Molecular reprogramming in tomato pollen during development and heat stress

---

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim

Fachbereich Biowissenschaften (FB15) der

Goethe-Universität Frankfurt am Main

von

Mario Keller

geboren in Frankfurt am Main (Hessen)

Frankfurt am Main 2018

(D30)

Vom Fachbereich Biowissenschaften (FB15) der Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Sven Klimpel

Gutachter: Prof. Dr. Enrico Schleiff, Jun. Prof. Dr. Michaela Müller-McNicoll

Datum der Disputation: 25.03.2019

Index of contents

Index of figures

Index of tables

Index of supplemental material

Index of supplemental figures

Index of supplemental tables

Abbreviations

2D-DIGE	difference gel electrophoresis
2-DE	two-dimensional gel electrophoresis
ADF	actin depolymerization factor
AGP	arabinogalactan protein
Aha1	activator of 90 kDA heat shock ATPase
ANOVA	analysis of variance
AP2	ApetalA2
ARF	auxin response factor
ATP-S	ATP-sulfurylase
BAG6	BAG family molecular chaperone regulator 6
BBP	basic blue protein
cDNA	complementary DNA
CS	citrate synthase
CSD	copper/zinc superoxide dismutases
Cy3/5	cyanine 3/5
DCL	Dicer-like
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
e.g.	for example (*exempli gratia*)
eIF	eukaryotic translation initiation factor
EST	expressed sequence tag
FPKM	fragments per kilobase of transcript per million fragments mapped
GAE	UDP-glucuronate 4-epimerase
GFF	generic feature format
Hsf	heat stress transcription factor
Hsp	heat shock protein
HSR	heat stress response
IDH_{reg}	regulatory subunit isocitrate dehydrogenase
LC-MS/MS	liquid chromatography tandem-mass spectrometry
LEA	late embryogenesis abundant protein
LFQ	label free quantification
LRR	leucine-rich repeat receptor-like protein
MACE	massive analysis of cDNA ends
MALDI-Tof-MS	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
miRNA	micro RNA
mRNA	messenger RNA
MS	mass spectrometry
MTI	miRNA-target interaction
NGS	next-generation sequencing
NMD	nonsense-mediated mRNA decay
ORF	open reading frame
PCA	principal component analysis
PCR	polymerase chain reaction
PDH_{β}	beta subunit of E1 component of the pyruvate dehydrogenase complex
PME	pectin methylesterase
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
qRT-PCR	quantitative real-time polymerase chain reaction
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNA-seq	RNA sequencing
ROS	reactive oxygen species
RP	ribosomal protein
RPKM	reads per kilobase of exon model per million mapped reads
rRNA	ribosomal RNA
SAM	Sequence Alignment/Map
SCS_{β}	beta subunit of succinyl-CoA-synthase
$SDH_{Fe\text{-}S}$	iron-sulphur subunit of succinate dehydrogenase
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPL	squamosa promoter binding protein-like
TCA	tricarboxylic acid
TF	transcription factor
TPM	transcripts per million
tRNA	transfer RNA
uORF	upstream ORF
USP	UDP-sugar pyrophosphorylase
UXS	UDP-glucuronate decarboxylases

Zusammenfassung

In der Regel ist der Lebenszyklus von Pflanzen ein sich wiederholender Wechsel zwischen einer haploiden Gametophytengeneration und einer diploiden Sporophytengeneration. Innerhalb der Gametophytengeneration reifen weibliche sowie männliche Gametophyten heran, welche sich durch eine erfolgreiche Befruchtung zu einer diploiden Zygote vereinen. Die entstandene Zygote markiert den Beginn einer neuen Sporophytengeneration. Der männliche Gametophyt, besser bekannt als Pollen, unterzieht sich während seiner Reifung mehreren strukturellen und morphologischen Veränderungen und ist deshalb ein interessantes Modellsystem für die Analyse von Zellwachstum und -entwicklung. Weiterhin sind Pflanzenpollen sehr empfindlich gegenüber abiotischen Stressen, wie etwa Hitze. Aus diesem Grund ist die Analyse von Pollen auch hinsichtlich der globalen Erwärmung und den damit verbundenen Ernteausfällen von besonderer Bedeutung.

Die Entwicklung und Hitzestressreaktion von Pollen beruht auf der feinabgestimmten Steuerung von biologischen Prozessen und Regulationsnetzwerken, was sowohl auf der Transkriptomebene als auch auf der Proteomebene erfolgt. Bisher haben sich die meisten Studien hauptsächlich mit dem Transkriptom von reifen Pollen beschäftigt, wohingegen die Analyse von früheren Entwicklungsstadien und des Proteoms von Pollen weitestgehend vernachlässigt wurde. Deshalb war es das Ziel dieser Arbeit das Transkriptom und Proteom von sich entwickelnden und wärmebelasteten Tomatenpollen genauer zu untersuchen, um Rückschlüsse über entwicklungsrelevante Prozesse und die Hitzestressantwort der analysierten Entwicklungsstadien ziehen zu können.

Analysiert wurden drei Pollenentwicklungsstadien, nämlich Tetraden, post-meiotischer Pollen und reifer Pollen. Die Pollen wurden aus Tomatenpflanzen isoliert, welche entweder unter Kontrollbedingungen bei 25°C gehalten wurden oder einem einstündigen Hitzestress von 38°C ausgesetzt waren, gefolgt von einer Erholung für anderethalb Stunden bei 25°C (wärmebelastet). Aus den nicht belasteten und wärmebelasteten Pollen wurden im Anschluss RNA und Proteine isoliert, um mittels Hochdurchsatzmethoden vorhandene „messenger RNAs“ (mRNAs), „micro RNAs“ (miRNAs) und Proteine zu identifizieren und quantifizieren.

Insgesamt lässt sich die Arbeit in drei Abschnitte unterteilen. Der erste Abschnitt befasst sich mit der Analyse der Pollen-Transkriptome mit dem Ziel entwicklungsrelevante Prozesse sowie die Hitzestressreaktion von Pollen besser zu verstehen. Im zweiten Abschnitt wird die Identifizierung und Charakterisierung von miRNAs beschrieben, welche einen Einfluss auf die Entwicklung und

Hitzestressantwort von Pollen haben. Der letzte Abschnitt beschäftigt sich mit dem Verhältnis der Pollen-Transkriptome und -Proteome und der Frage inwiefern Rückschlüsse vom Transkriptom auf das Proteome in sich entwickelnden und wärmebelasteten Pollen möglich sind.

Das Transkriptom von sich entwickelnden und wärmebelasteten Pollen

Um erste Einblicke in die Transkriptome von sich entwickelnden und wärmebelasteten Pollen zu erhalten, wurde eine Hauptkomponentenanalyse durchgeführt. Diese Analyse ergab, dass die Transkriptome weitestgehend stadienspezifisch sind, es jedoch auch eine Anpassung der Transkriptome als Reaktion auf den Hitzestress gibt.

Um entwicklungsrelevante Proteinfamilien und Prozesse zu detektieren, wurden zunächst alle mRNAs identifiziert, welche in einem oder zwei aufeinanderfolgenden Entwicklungsstadien akkumuliert sind. Diese Vorgehensweise lieferte fünf Gruppen von mRNAs, welche entweder in einem oder zwei aufeinanderfolgenden Entwicklungsstadien akkumuliert sind. Eine im Anschluss ausgeführte funktionale Anreicherungsanalyse ermöglichte es wichtige Prozesse und Proteinfamilien innerhalb der fünf Gruppen zu identifizieren. So konnte zum Beispiel gezeigt werden, dass in Tetraden so genannte „squamosa promoter binding proteins" (SPLs), Brassinosteroide und die abiotische Stressantwort eine wichtige Rolle spielen. Unter den der abiotischen Stressantwort zugehörigen mRNAs wurden viele mRNAs beobachtet, welche für Hitzestresstranskriptionsfaktoren (Hsfs) und Hitzeschockproteine (Hsps) kodieren. Diese Beobachtung ist im Einklang mit dem Konzept des „developmental priming", welches die Anhäufung von stressinduzierten Proteinen in frühen Entwicklungsstadien unter normalen Bedingungen als Schutzmaßnahme gegenüber plötzlich auftretenden Stressen beschreibt (Chaturvedi et al., 2013). Im Gegensatz dazu konnte anhand der in Tetraden und post-meiotischen Pollen akkumulierten mRNAs gezeigt werden, dass in diesen beiden Entwicklungsstadien die Proteinsynthese und Aminosäureaktivierung eine wichtige Rolle spielen. Bei den von den mRNAs kodierten Proteinen handelt es sich unteranderem um ribosomale Proteine der kleinen und großen Untereinheit, eukaryotische Initiationsfaktoren und Aminoacyl-tRNA-Synthetasen. Weiterhin konnte gezeigt werden, dass in post-meiotischen und reifen Pollen akkumulierte mRNAs vorrangig für Proteine der Glykolyse, des Citratzyklus, der Atmungskette und der ATP-Synthase kodieren. All diese Prozesse sind Teil der Zellatmung, was auf eine Vorbereitung auf den erhöhten Energieaufwand während der Keimung und des Pollenschlauchwachstums hindeutet. Anhand der nur in reifen Pollen akkumulierten mRNAs wurde sichtbar, dass in diesem Entwicklungsstadium die Vorbereitung für die Verlängerung und den Umbau der Zellwand und des Zytoskeletts während des

Pollenschlauchwachstums eine wichtige Rolle spielt. Die von den mRNAs kodierten Proteine umfassen zum Beispiel Cellulasen und Pektinmethylesterasen für Veränderungen der Zellwand sowie Aktin und verschiedene Aktin-bindende Proteine für Veränderungen der Aktinfilamente.

Der Anteil von mRNAs die aufgrund von Hitzestress differentiell reguliert sind, reicht in vegetativen Geweben von 4.2% bis 20% (Mangelsen et al., 2011; Li et al., 2013b). Ein vergleichbares Verhalten wurde auch für post-meiotischen und reifen Pollen beobachtet, bei denen 5.7% bzw. 4.5% der detektierten mRNAs differentiell reguliert sind. In Tetraden hingegen sind lediglich 1% der detektierten mRNAs differentiell reguliert. Ein Vergleich der in den drei Stadien differentiell regulierten mRNAs zeigte, dass es einen Kernsatz von 49 mRNAs gibt, welche in allen drei Entwicklungsstadien hochreguliert werden. Eine funktionale Analyse der 49 mRNAs zeigte, dass 19 der mRNAs für Mitglieder der Hsf und Hsp Familien kodieren. Die Mitglieder umfassen HsfA2 und HsfB1 von der Hsf Familie sowie 13 Mitglieder der „small Hsp“ (sHsp) Familie, ein Hsp70, ein Hsp90 und zwei Hsp100. Neben den zu erwartenden Hsf und Hsp kodierenden mRNAs wurde zum Beispiel auch eine mRNA identifiziert, welche für eine Ascorbat-Peroxidase kodiert. Ascorbat-Peroxidasen spielen eine wichtige Rolle bei der Entgiftung von reaktiven Sauerstoffspezies, welche unteranderem unter Hitzestress auftreten (Caverzan et al., 2012). Da ein Großteil der 49 mRNAs für Hsfs und Hsps kodiert, wurde das Hitzestressverhalten dieser Proteinfamilien genauer untersucht. Die Analyse ergab, dass in Tetraden zwar die meisten Hsf und Hsp mRNAs identifiziert wurden, jedoch doch nur halb so viele Hsf und Hsp mRNAs wie in post-meiotischen und reifen Pollen hochreguliert werden. Ein genauerer Blick auf die in post-meiotischen und reifen Pollen aber nicht in Tetraden hochregulierten Hsf und Hsp mRNAs zeigte, dass viele von diesen mRNAs bereits in nicht wärmebelasteten Tetraden akkumuliert sind und eine Hochregulation unter Hitzestress höchstwahrscheinlich nicht erforderlich ist. Diese Beobachtung lieferte ein weiteres Indiz für ein „developmental priming“ von Tetraden.

Die regulatorische Rolle von miRNAs in der Entwicklung und Hitzestressantwort von Pollen

Da bekannt ist, dass miRNAs eine wichtige Rolle in der Entwicklung und Stressantwort von Pflanzen spielen (Jones-Rhoades et al., 2006; Shukla et al., 2008), war es von Interesse herauszufinden inwieweit dies auch auf die Entwicklung und Hitzestressantwort von Pollen zutrifft. Hierzu wurden in einem ersten Schritt miRNAs in nicht belasteten und wärmebelasteten Pollen vorhergesagt. Um zu überprüfen ob die Vorhersagen verlässlich sind, wurde als erstes überprüft wie viele der insgesamt 793 detektierten miRNAs bereits bekannt sind. Es konnte gezeigt werden, dass 24 der

vorhergesagten miRNAs bereits in Tomate und weitere 14 in anderen Pflanzenspezies beschrieben sind, was die Verlässlichkeit der Vorhersage bestätigte.

Um den Einfluss von miRNAs auf die Abundanz von mRNAs überprüfen zu können, wurde als Nächstes für jede der 793 miRNAs eine Zielvorhersage gemacht um mRNAs zu identifizieren, welche durch die miRNA reguliert werden. Insgesamt konnte für etwas mehr als die Hälfte der miRNAs mindestens eine zu regulierende mRNA identifiziert werden, wobei im Mittel etwa 6 mRNAs von einer miRNA reguliert werden. Die Interaktion zwischen jedem Paar von miRNA und mRNA wurde im Anschluss als sogenannte „miRNA-target interaction“ (MTI) gespeichert.

Nach der Identifizierung von miRNAs, welche in einem oder zwei aufeinanderfolgenden Entwicklungsstadien akkumuliert sind, wurden im Anschluss alle entwicklungsrelevanten MTIs ermittelt. Diese MTIs haben miteinander gemein, dass sowohl die miRNA als auch die mRNA in ein oder zwei Entwicklungsstadien akkumuliert sind, und dass Veränderungen in der miRNA-Abundanz mit entgegengesetzten Veränderungen in der mRNA-Abundanz einhergehen. Insgesamt wurden 207 entwicklungsrelevante MTIs identifiziert. Neben vielen bisher unbekannten miRNAs, setzten sich einige der MTIs aus miRNAs zusammen, die bereits in der Literatur beschrieben sind. Eine dieser miRNAs ist miR395, welche in dieser Studie wie auch in der Literatur eine ATP-Sulfurylase kodierende mRNA reguliert. Die Regulation einer ATP-Sulfurylase kodierenden mRNA durch miR395 war bis dato nicht für Pollen beschrieben, bestätigt jedoch die Annahme, dass die Regulation der Schwefelassimilation für die Entwicklung von Pollen von großer Bedeutung ist (Birke et al., 2013). Weiterhin konnte beobachtet werden, dass die mRNAs von 34 MTIs für Transkriptionsfaktoren kodieren. Eine miRNA, welche die mRNAs von Transkriptionsfaktoren reguliert ist miR156. Basierend auf den Ergebnissen dieser Studie konnte erstmals gezeigt werden, dass miR156 für die Herunterregulation von fünf SPL kodierenden mRNAs im Übergang von Tetraden zu post-meiotischen Pollen verantwortlich ist.

Neben ihres Einflusses auf die Pollenentwicklung wurden die miRNAs auch hinsichtlich ihres Einflusses auf das Hitzestressverhalten der einzelnen Entwicklungsstadien untersucht. Insgesamt konnten 24 MTIs identifiziert werden, die einen Einfluss auf das Hitzestressverhalten von Pollen haben könnten. Bei dem Großteil dieser MTIs führt die Herunterregulation der miRNA zu einer Hochregulation der mRNA. Ein Beispiel für solch eine MTI ist die Interaktion zwischen miR408 und einer mRNA, welche für ein Plantacyanin kodiert. Die Herunterregulation von miR408 und die damit

verbundene Hochregulation von Plantacyanin kodierenden mRNAs konnte bisher nur für andere Stresse wie etwa Trockenstress gezeigt werden (Ma et al., 2015). Die in dieser Studie präsentierten Ergebnisse zeigen, dass die Herunterregulation von miR408 ebenfalls eine wichtige Rolle in der Hitzestressantwort von Pollen spielt.

Regulation von Proteinen während der Entwicklung und Hitzestressantwort von Pollen

Die bisherigen Resultate hatten gezeigt, dass es im Verlauf der Entwicklung und als Antwort auf Hitzestress zu Veränderungen des Transkriptoms kommt. Um zu überprüfen ob sich Veränderungen des Transkriptoms auch auf das Proteom übertragen lassen, sollte im letzten Abschnitt das Proteom von sich entwickelnden und wärmebelasteten Pollen genauer untersucht werden. Um einen ersten Einblick in die Beziehung von Transkriptom und Proteome zu erlangen, wurden die in einem Stadium detektierten mRNAs und Proteine miteinander hinsichtlich ihrer Abundanz korreliert. Sowohl für nicht belastete als auch für wärmebelastete Pollen konnte nur eine sehr geringe Korrelation festgestellt werden. Interessanterweise, erhöhte sich der Korrelationskoeffizient für mRNAs aus Tetraden und post-meiotischen Pollen wenn diese mit Proteinen des darauffolgenden Entwicklungsstadiums korreliert wurden. So lag der Korrelationskoeffizient von mRNAs aus post-meiotischen Pollen und Proteinen aus reifen Pollen bei 0,49, wohingegen der Korrelationskoeffizient mit Proteinen aus post-meiotisichen Pollen lediglich bei 0,23 lag.

Da die Transkriptome und Proteome teils nur sehr schwach korreliert sind, war es von Interesse herauszufinden wie sich die in den Entwicklungsstadien akkumulierten mRNAs auf Proteinebene verhalten. Hierzu wurden als Erstes alle Proteine identifiziert, die in einem oder zwei aufeinanderfolgenden Entwicklungsstadien akkumuliert sind. Im Anschluss wurde für alle Paare von akkumulierten mRNAs und Proteinen überprüft ob deren Akkumulation im gleichen Entwicklungsstadium oder in unterschiedlichen Entwicklungsstadium auftritt. Hieraus resultierend, konnten zwei Translationsmodi abgeleitet werden, welche während der Pollenentwicklung auftreten. Der erste Modus beschreibt eine direkte Translation, bei der die erstmalige Akkumulation von mRNA und Protein im gleichen Entwicklungsstadium auftritt. Im Gegensatz dazu beschreibt der zweite Modus eine zeitverzögerte Translation, da die erstmalige Akkumulation des Proteins im Vergleich zur erstmaligen Akkumulation der mRNA um ein Entwicklungsstadium verzögert ist. Weiterhin konnte eine mögliche Erklärung für die zeitverzögerte Translation von mRNAs gefunden werden. Eine kürzlich für die Tabakpflanze veröffentlichte Studie konnte zeigen, dass es während der Entwicklung von Pollen

zu einer Kurz- und Langzeitspeicherung von mRNAs in sogenannten „EDTA/puromycin-resistant particles" (EPPs) kommt (Hafidh et al., 2018). Bei EPPs handelt es sich um Ribonukleoproteinkomplexe, welche die Translation der gespeicherten mRNAs verhindern. Anhand der in der Studie angegebenen Tomaten-Orthologen von in EPPs gespeicherten mRNAs konnte verifiziert werden, dass etwa 60% der mRNAs, welche eine verzögerte Translation aufweisen, in EPPs gespeichert werden.

Da der Vergleich der Transkriptome und Proteome von sich entwickelnden Pollen auf eine partielle Entkopplung von mRNAs und den dazugehörigen Proteinen hindeutete, sollte im finalen Abschnitt das Hitzestressverhalten von Proteinen und ihrer zugrundeliegenden mRNAs verglichen werden. Die Identifizierung von Proteinen, welche als Reaktion auf den Hitzestress herunter- oder hochreguliert werden, zeigte, dass die Proteome der einzelnen Entwicklungsstadien weitaus stärker auf den Hitzestress reagieren als die Transkriptome. So wiesen zum Beispiel in Tetraden etwa 40% der detektierten Proteine eine veränderte Abundanz auf, wohingegen lediglich 1% der mRNAs verändert war. Ein direkter Vergleich der Regulation von Proteinen und ihren mRNAs zeigte, dass die Proteine zum Großteil unabhängig von ihren mRNAs reguliert werden. Ein genauerer Blick auf die Proteine, welche unabhängig von ihren mRNAs reguliert werden, zeigte, dass es sich bei vielen dieser Proteine um ribosomale Proteine der kleinen und großen Untereinheit handelt. Die Hitzestressregulation der ribosomalen Proteine erwies sich als weitestgehend stadienspezifisch. Eine solche Herunter- und Hochregulation von ribosomalen Proteinen könnte zu einer veränderten Zusammensetzung von Ribosomen führen, was wiederum einen Einfluss darauf haben könnte wie effizient und welche mRNAs translatiert werden (Xue and Barna, 2012). So konnte zum Beispiel in einer früheren Studie im Menschen gezeigt werden, dass Ribosomen, welche RPS25 beinhalten, bevorzugt eine gewisse Gruppe von mRNAs translatieren (Shi et al., 2017). Interessanterweise konnten in reifen Pollen zwei RPS25 identifiziert werden, die unter Hitzestress hochreguliert sind. Diese RPS25 könnten ähnlich wie beim Menschen zu einer bevorzugten Translation bestimmter mRNAs führen und so post-transkriptionell die Hitzestressantwort des reifen Pollens beeinflussen.

Abstract

In times of a growing world population and the associated demand for high crop yield, the understanding and improvement of plant reproduction is of central importance. One key step of plant reproduction is the development of the male gametophyte, which is better known as pollen. In addition, the development of pollen was shown to be very sensitive to abiotic stresses, such as heat, which can cause crop damage and yield loss. To obtain new insights in the development and heat stress response of pollen, a combined transcriptome and proteome analysis was performed for three pollen developmental stages of non- and heat-stressed tomato plants.

The analysis of the transcriptomes of non-stressed pollen developmental stages enabled the determination of mRNAs accumulated in certain developmental stages. The functional analysis of these mRNAs led to the identification of protein families and functional processes that are important at different times of pollen development. A subsequent comparison of the transcriptomes of non- and heat-stressed pollen revealed a core set of 49 mRNAs, which are upregulated in all three developmental stages. The encoded proteins include among other things different heat stress transcription factors and heat shock proteins, which are known key players of the plant heat stress response.

Furthermore, 793 potential miRNAs could be identified in the transcriptome of non- and heat-stressed pollen. Interestingly, 38 out of the 793 miRNAs have already been identified in plants. For more than half of these miRNAs potential target mRNAs were identified and the interactions between miRNAs and mRNAs linked to the development and heat stress response of pollen. In total, 207 developmentally relevant interactions could be determined, out of which 34 have an effect on transcriptional-networks. In addition, 24 of the interactions contribute the heat stress response of pollen, whereby this mainly affects post-meiotic pollen.

An initial correlation of the proteome and transcriptome of the developmental stages revealed that transcriptome analyses are not sufficient to draw exact conclusions about the state of the proteome. A closer look on the relationship of the transcriptome and proteome during pollen development revealed two translational modes that are active during the development of pollen. One mode leads to a direct translation of mRNAs, while the second mode leads a delayed translation at a later point in time. Regarding the delayed translation, it could be shown that this is likely due to a short-term storage of mRNAs in so-called EPPs. The comparison of the proteome and transcriptome response to heat

stress revealed that the proteome reacts much stronger and that the reaction is mainly independent from the transcriptome. Finally, the comparison of the proteome of non- and heat-stressed pollen provided first indications for changes in the ribosome composition in response to heat stress, as 57 ribosomal proteins are differentially regulated in at least one developmental stage.

1 Introduction

1.1 Cellular transcriptome and proteome dynamic

1.1.1 From gene to protein – basic regulatory principles

The central dogma of molecular biology describes the flow of genetic information as a sequential transfer of information between the biopolymers deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and protein (Crick, 1958). Initially, the information is encoded in the DNA, in the form of genes. During transcription, the information is transferred from DNA to RNA and in the case of messenger RNA (mRNA) subsequently to protein via translation (Fu et al., 2014; McManus et al., 2015). However, the dogma is very simplified and there are important aspects to consider when analyzing expression data. On the one hand, different synthesis and degradation rates can lead to an uncoupling and low correlation of mRNA and protein levels (Tebaldi et al., 2012; McManus et al., 2015). On the other hand, processes like pre-mRNA splicing (Piras et al., 2012) as well as post-transcriptional (Nachtergaele and He, 2017) and post-translational modifications (Khoury et al., 2011) lead to a higher complexity and offer an additional layer of regulation and functionality.

As mentioned above, transcription is the first step of the information flow and describes the synthesis of RNA molecules from a DNA template. This process relies on the fine-tuned interplay of an RNA polymerase and general transcription factors (TFs). In eukaryotes there exist three types of RNA polymerases, which are required for the synthesis of different classes of RNA (Hahn, 2004). For instance, RNA polymerase II is required for the synthesis of mRNAs (Cramer et al., 2001). Next to the transcription, the maturation of mRNAs relies on at least three processing steps, namely capping, splicing and polyadenylation (Proudfoot, 2000). After their maturation, mRNAs have three distinct fates. They are either translated into a protein, stored for later translation or degraded. Two examples for mRNA degradation mechanisms are the nonsense-mediated mRNA decay (NMD) and micro RNA (miRNA)-mediated silencing (Shyu et al., 2008). NMD is a surveillance mechanism, which is required for the degradation of aberrant mRNAs. These mRNAs harbor a premature termination codon, which, for instance, derives from mistakes during splicing. To prevent the translation of truncated proteins, these mRNAs are degraded via the NMD pathway (Kurosaki and Maquat, 2016). In contrast to the surveillance function of the NMD, miRNA-mediated silencing is a general mechanism for targeted regulation of mRNA abundance. miRNAs are short RNAs that typically have a length between 20 and 24 nucleotides and are part of the RNA-induced silencing complex (RISC). The miRNAs guide the

RISC to mRNAs by binding highly complementary regions on the mRNAs. The binding leads in most cases to the cleavage and degradation of the mRNA (Zhang et al., 2006; Ivashuta et al., 2011). Typically, miRNAs are transcribed from intergenic loci, which produces a primary miRNA transcript (pri-miRNA). In the next step, a member of the Dicer-like (DCL) family cleaves the pri-miRNA into a precursor-miRNA (pre-miRNA), which has a characteristic hairpin loop structure. The same DCL protein subsequently cleaves the pre-miRNA, which releases a miRNA-miRNA* duplex. After export into the cytoplasm, the duplex separates and the mature miRNA is incorporated in the RISC (Rogers and Chen, 2013; Budak and Akpinar, 2015).

In addition to mRNA cleavage, central mismatches between the miRNA and mRNA may promote translational inhibition instead of mRNA cleavage (Brodersen et al., 2008). Other mechanisms affecting translation are the storage of mRNAs in ribonucleoprotein complexes, such as EDTA/puromycin-resistant particles (EPPs) or processing bodies (Hafidh et al., 2018; Standart and Weil, 2018), as well as the translation of so-called upstream open reading frames (uORFs). The translation of uORFs, which are located upstream of the main ORF, typically inhibits the translation of the main ORF and by this directly has an impact on the protein abundance (Roy and Arnim, 2013).

The actual translation of mRNAs relies on the interplay of key components of the translation machinery, such as ribosomes, consisting of ribosomal RNA (rRNA) and ribosomal proteins (RPs), eukaryotic translation initiation factors (eIFs) and transfer RNAs (tRNAs). Ribosomes are required for the decoding of the mRNA as well as for catalyzing the peptidyl transferase reaction, eIFs for the initiation of translation and tRNAs for the delivery of amino acids to the ribosome (Merchante et al., 2017).

As pointed out, there are multiple regulatory mechanisms that can affect mRNA and protein abundance and offer the possibility for a massive reprogramming of the transcriptome and proteome during important biological processes, such as development (Chaturvedi et al., 2013; Palumbo et al., 2014) or the response to external stimuli (Trivellini et al., 2012; Kumar et al., 2017). For a complete understanding of these processes, it is therefore necessary to monitor the entire transcriptome and proteome dynamics. The large-scale monitoring of thousands of transcripts and proteins was enabled with the introduction of so-called high-throughput technologies like RNA sequencing (RNA-seq) and mass spectrometry (Angel et al., 2012; Lowe et al., 2017).

1.1.2 Measurement of transcriptome composition and transcript levels with high-throughput methods

Over the last 40 years several methods for the analysis of transcripts have been developed. The individual methods can be distinguished by the number of transcripts that can be analyzed in parallel (Lowe et al., 2017), also known as the throughput, as well as by the need for prior knowledge about the underlying nucleotide sequence of the transcripts (Smith and Osborn, 2009; Fondevilla et al., 2011).

One method for expression analyses of a predefined set of genes is the quantitative real-time polymerase chain reaction (qRT-PCR) technique (Adamski et al., 2014). The qRT-PCR is the standard for abundance analyses and often used as independent technique for the validation of high-throughput methods (Lowe et al., 2017). However, qRT-PCR analyses are limited in their throughput due to the size of the microtiter plates, which typically carry 96 or 384 wells (Love et al., 2004). Three methods that have higher throughput and overcome the restriction to a predefined set of genes are expressed sequence tags (ESTs), microarrays and RNA-seq, which all enjoyed great popularity over the past decades (Figure 1). First insights in the transcriptome without *a priori* knowledge were made possible with the introduction of ESTs, which were used for the first time in 1991 (Adams et al., 1991). Since then, EST analyses have been conducted in a variety of model plants like *Arabidopsis thaliana* (White et al., 2000; Zhu et al., 2003), *Medicago truncatula* (Györgyey et al., 2000; Journet et al., 2002) and maize (Verza et al., 2005; Yang et al., 2006). Here, they have contributed to novel insights into the transcriptomes of different tissues and developmental stages as well as to the refinement of gene annotations. The first two steps for the generation of ESTs are similar to those of qRT-PCR and include the isolation of RNA and subsequent reverse transcription of mRNA into complementary DNA (cDNA). In the next step, a cDNA library is generated by inserting the cDNAs into plasmid vectors that are subsequently introduced into *Escherichia coli* cells. After amplification of the cDNA containing plasmids through replication in the *E. coli* cells, either the 5' or 3' end of the cDNAs is sequenced (Bouck and Vision, 2007). The sequencing itself is based on the chain-termination method introduced

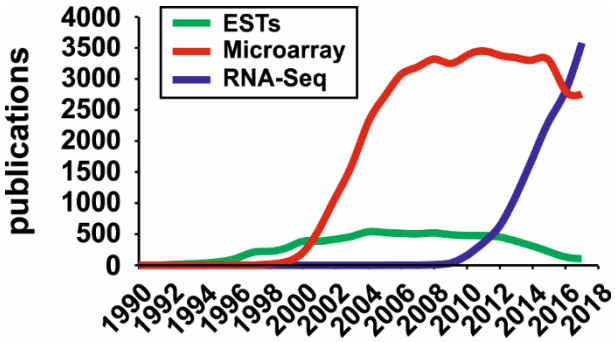


Figure 1: Publication rates of transcriptomics technologies
Shown is the number of publications referring to ESTs (green), Microarray (red) and RNA-Seq (blue) for transcriptome analysis.

by Sanger et al. (1977). The nucleotide sequences obtained from the sequencing are the final ESTs, which represent the 5' or 3' end of the initially isolated RNAs. After rejection of low-quality ESTs and the removal of contaminations, the first step of the downstream analysis is the clustering of ESTs and generation of consensus sequences, which aims to remove redundancy by merging ESTs that are sequenced from the same transcript. Afterwards, the consensus sequences can be searched against biological databases or aligned to a reference genome, if available, to functionally categorize and annotate them (Nagaraj et al., 2007). Further, it is assumed that the frequency of the observed ESTs is proportional to the abundance of the corresponding mRNA, which enabled expression profiling of known and novel transcripts (Park et al., 2006). However, it is suggested that at low sampling depths, EST analyses are rather qualitative than quantitative (Lee et al., 1995; Milla et al., 2002), which is especially a problem for low abundant transcripts, which are known to be underrepresented in EST collections (Alba et al., 2004).

To overcome this limitation and extend the throughput, ESTs are used for the construction of microarrays, which allow expression profiling of more than ten thousands of genes also if an organisms lacks genomic information (Chen et al., 2004). For instance, EST collections were used for the microarray design for a variety of plants, including citrus (Martinez-Godoy et al., 2008), apple (Janssen et al., 2008), peanut (Payton et al., 2009) and three *Brassica* species (Xiang et al., 2008). The history of microarrays dates back to the mid-1990s (Schena et al., 1995), but the large-scale use only really began with the beginning of the 21th century (Figure 1 red line). A microarray is generally a solid surface, such as a glass slide, that is spotted with thousands of cDNA probes (Yang and Speed, 2002), whereby today's cDNA microarrays carry more than 50,000 probes (Guo and Peddada, 2008). These probes are clusters of identical oligonucleotides that are complementary to defined regions of annotated genes. The aim of most microarray studies is the identification of genes that are differentially regulated between two samples (e.g. treatment and control). For this purpose, mRNA is isolated from both samples and reverse transcribed into cDNA, which is afterwards fluorescently labeled with a dye. To discriminate between the analyzed samples, two different dyes like cyanine 3 and 5 (Cy3 and Cy5) are used. The labeled cDNAs of both samples are then mixed and put on the same microarray for competitive hybridization with the probes that are attached to the surface of the microarray. After the hybridization, the microarray is washed, scanned with a laser and the emitted fluorescence signals of the Cy3 and Cy5 dyes captured for all probes. After normalization, the Cy3 and Cy5 signals can be used for the calculation of $\log_2$ ratios, which allows the comparison of

expression levels between the samples. This procedure is based on the assumption that if a gene is higher expressed in one of the samples, there is also more labeled cDNA of the gene that can hybridize on the microarray, which would lead to a higher fluorescence signal (Malone and Oliver, 2011). Although microarrays have enabled the possibility of large-scale transcriptome analyzes, also this technique has certain limitations. For example, the normal cDNA microarrays are limited to known transcripts and therefore the detection of unknown or alternatively spliced transcripts is not possible. Further, there can be a high background noise due to cross-hybridization, which leads to false positive signals. Moreover, the background noise together with the saturation of signals lead to a limited detection range, which makes it difficult to detect low abundant transcripts if their signal approaches the noise level (Mortazavi et al., 2008; Wang et al., 2009; Malone and Oliver, 2011).

One method overcoming the limitations of microarrays is RNA-seq, which uses next-generation sequencing (NGS) for the identification and quantification of expressed transcripts (Kukurba and Montgomery, 2015). In the first years, NGS was used for the (re)sequencing of genomes but quickly it was also applied to the sequencing of cDNAs, which enabled the identification and quantification of transcripts (Nowrousian, 2010). Publications using RNA-seq as the method of choice for transcriptome analyses emerged in 2008 (Mortazavi et al., 2008; Nagalakshmi et al., 2008; Wilhelm et al., 2008) and their number has exceeded that of microarray related publications within the last two years (Figure 1 blue line). In general, RNA-seq analyses can be separated into three parts, namely the library preparation, sequencing and downstream analyses.

The library preparation starts with the isolation of the RNA and a subsequent reverse transcription into cDNA. Next, the cDNA is fragmented, followed by a gel size selection to obtain cDNA fragments of a predefined length. After that, adapter sequences are ligated to the 5’ and 3’ end of the cDNA fragments, which harbor primer binding sites for the upcoming PCR amplification and the sequencing (Head et al., 2014; Kukurba and Montgomery, 2015). Depending on the biological question, there are many alternative protocols for the library preparation. For example, there exist many different strand-specific protocols that aim to preserve the information from which strand the observed transcripts originated, which is among other things important in the case of antisense transcripts or the discrimination of transcripts whose gene loci overlap but are located on opposite strands (Levin et al., 2010). Other frequent steps in protocols are the depletion of rRNA, which would otherwise diminish the sequencing information of other RNA types, as it constitutes more than 80% of the transcriptome (O’Neil et al., 2013) or the enrichment of mRNAs via oligo(dT)-beads (Cui et al., 2010).

After PCR amplification, the RNA-seq library is sequenced, which is nowadays typically done on Illumina platforms, such as HiSeq, NextSeq or MiSeq (Ambardar et al., 2016). The sequencing is performed on a flow cell, which is a glass slide that is spotted with millions of primers that are complementary to a region of the adapters that were ligated to the cDNA fragments. After the library is given on the flow cell, the cDNA fragments hybridize to the primers. Next, the cDNA fragments are amplified via bridge amplification, which leads to the formation of millions of dense clusters, each of them containing identical copies of the initial cDNA fragments. The amplified cDNA fragments in the clusters serve as sequencing templates. During the sequencing, a DNA strand is synthesized, whereby the incorporated nucleotides are monitored, which gives direct information about the sequence of the templates. Therefore, in a first step, a sequencing primer is hybridized to the templates, which serves as starting point for the sequencing. The sequencing itself is a repetition of sequencing cycles. During each cycle, a DNA polymerase incorporates one of four modified deoxynucleoside triphosphates (dNTPs) in the growing DNA strands. Each of the four dNTPs is labeled with a different reversible dye-terminator, which allows the identification of the incorporated dNTP and ensures that only a single dNTP is incorporated in each cycle. After the incorporation, the flow cell is imaged to identify the incorporated dNTP for each cluster, followed by the cleavage of the dye-terminators from the dNTPs, which allows the incorporation of a new dNTP in the next cycle. Typically, only 100 to 300 cycles are performed as the signal quality of the clusters is decreasing with an increasing number of cycles. The final output of the sequencing are millions of so-called reads, which represent the monitored nucleotide sequences of each cluster (Metzker, 2010; Buermans and den Dunnen, 2014). Nowadays, the templates are typically sequenced from both ends, which is called paired-end sequencing and leads to a higher coverage as well as to more reliable downstream analyses (Ambardar et al., 2016).

The analysis of the sequencing data typically starts with a quality control of the reads, which can be done with tools such as FastQC or HTQC. The obtained results give information about the read qualities as well as possible contaminations like sequenced adapter sequences (Yang and Kim, 2015). For the removal of low quality regions and adapter sequences, tools like Trimmomatic (Bolger et al., 2014) can be used. In the presence of an annotated reference genome, the next step of the analysis is the alignment of the millions of reads to the reference genome. This allows the identification of the genomic positions from which the reads originated. For this task, specialized tools have been developed, which are known as aligners. The choice of the appropriate aligner depends on the

analyzed organism. While for prokaryotes the reads can be aligned continuously, for most eukaryotes a splice-aware aligner is more appropriate, due to the presence of introns in the reference genome, which are not present in the sequenced reads (Kim et al., 2013). Bowtie2 (Langmead and Salzberg, 2012) and NextGenMap (Sedlazeck et al., 2013), for example, can be used for the continuous alignment of reads, whereas TopHat2 (Kim et al., 2013) and HISAT (Kim et al., 2015) are splice-aware. The output of the aligners is a file in the Sequence Alignment/Map (SAM) format, which provides for each read information about its alignment position like the chromosome, strand and alignment start as well as information about mismatches, insertions or deletions in the alignment (Li et al., 2009). The SAM file together with the gene annotation can afterwards be used as input for tools, such as the High-Throughput Sequencing python framework (HTSeq; Anders et al., 2015), which reports the number of reads aligned to each annotated gene, better known as read counts. To make read counts between genes and samples comparable, a normalization step is required. Here, typically the measures reads per kilobase of exon model per million mapped reads (RPKM; Mortazavi et al., 2008) and fragments per kilobase of transcript per million fragments mapped (FPKM; Trapnell et al., 2010) are used for single-end and paired-end reads, respectively. Both methods normalize for differences in the gene length, which makes genes within one sample comparable, and for differences in the number of sequenced reads, which allows the comparison of a gene between two samples (Li et al., 2017). However, in the presence of a small fraction of genes accumulating a large portion of the sequenced reads in one of the analyzed samples, both methods lead to skewed expression levels for the remaining genes (Rapaport et al., 2013). To solve this problem and enable a workflow for differential expression analyses, specialized R packages like DESeq2 (Love et al., 2014) or edgeR (Robinson et al., 2009) have been developed. The normalization techniques of both packages are based on the calculation of a single normalization factor for each analyzed sample with the basic assumption that the majority of the analyzed genes is not differentially expressed (Schurch et al., 2016). The subsequent testing for differential expression is done by modeling the read counts to a negative binomial distribution, followed by the testing with either a Wald (DESeq2) or an exact test (edgeR). One drawback of RNA-seq is that the read coverage across the genes is not uniform due to bias during the library preparation (Tauber and Von Haeseler, 2013). Steps that introduce such bias can be: (i) the enrichment of mRNAs via the polyA-tail, which leads to a higher coverage towards the 3’ end, (ii) reverse transcription via random hexamer primers, which leads to the under-representation of the 3’ end, or (iii) the fragmentation of the cDNA as some restriction enzymes show a preference to

specific sequences and are therefore not leading to a random fragmentation pattern (Wu et al., 2011; Finotello et al., 2014).

Two NGS-based methods that overcome the problem of a non-uniform coverage are QuantSeq (Moll et al., 2014) and massive analysis of cDNA ends (MACE; Zawada et al., 2014). Both methods generate only a single read per transcript, which is sufficient for analyzing gene and differential expression (Moll et al., 2014). In the MACE protocol, polyadenylated mRNA is first isolated and afterwards reverse transcribed into cDNA. Next, a biotinylated adapter is ligated to the end of the cDNA, which is afterwards fragmented. The biotinylated 3' ends are afterwards captured by streptavidin beads, ligated to 5' and 3' adapters and sent for sequencing (Zawada et al., 2014). The resulting reads can afterwards be quantified and the resulting read counts either normalized to the number of sequenced reads, which results in the measure of transcripts per million (TPM; Simm et al., 2015; Fragkostefanakis et al., 2016), or directly used as input for DESeq2 or edgeR.

1.1.3 Detection and quantification of proteins via mass spectrometry

Although transcriptomic approaches are widely used to analyze cellular dynamics, proteins are the functionally active molecules in the cell (Martínez-Gómez et al., 2012). Therefore, proteomics studies are essential to obtain a complete picture of active processes in the cell (Vij and Tyagi, 2007; Yates et al., 2009). There exist different methods for the large-scale quantification of proteins, which are either gel-based or gel-free (Abdallah et al., 2012). One gel-based approach for the quantification of proteins is two-dimensional gel electrophoresis (2-DE), which was developed in the mid-1970s (O'Farrell, 1975). In the first dimension, proteins are separated based on their charge by isoelectric focusing and in the second dimension according to their molecular weight by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Beranova-Giorgianni, 2003). This procedure enables the separation of complex protein mixtures into over 10,000 spots corresponding to over 1,000 proteins that can afterwards be identified via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-ToF-MS) (Abdallah et al., 2012). However, 2-DE shows a low gel-to-gel reproducibility, when identical spots on different gels are compared (Lilley et al., 2002). This problem can be overcome with difference gel electrophoresis (2D-DIGE), which enables the comparison of up to three samples on a single gel. For this purpose, the samples are labeled with different fluorescent dyes (e.g. Cy2, Cy3 or Cy5) before they are separated on the gel. After the gel is scanned with the wavelength of each dye, the measured intensities for each spot can be used for the comparison of protein abundance between the samples (Lilley et al., 2002; Abdallah et al., 2012). MALDI-ToF-MS analyses can afterwards

identify the proteins of spots that show changes in their abundance between the analyzed samples (Hoffert et al., 2004).

An alternative to the in-gel quantification approaches is the peptide-based quantification of proteins by mass spectrometry (MS) (Deracinois et al., 2013). This is often done via a coupling of a SDS-PAGE and liquid chromatography tandem-mass spectrometry (LC-MS/MS). Here, in a first step, proteins are separated by SDS-PAGE, followed by the excision of the gel lane and division into slices, which allows the pre-fractionation and thus a reduction of the complexity of the sample. Next, the proteins in the slices are proteolytically digested, followed by a separation of the peptides via LC and subsequent MS/MS measurements. The obtained spectra are afterwards searched against a database with theoretical spectra, which allows the identification as well as quantification of peptides and subsequent assignment of the peptides to the protein from which they derived (Dzieciatkowska et al., 2014). For the comparison of protein levels between multiple samples, either a label-based or label-free approach can be applied. An example for a label-based approach is the use of isobaric tags for relative and absolute quantification (iTRAQ). With this approach, it is possible to analyze up to eight samples in parallel. For this purpose, after the SDS-PAGE separation and in-gel digestion for each sample a different isobaric tag is attached to the peptides, followed by the pooling of the labeled peptides and the LC-MS/MS measurement. As the tags are isobaric, identical peptides from different samples have the same mass to charge ratio (*m*/*z*) and appear as a single precursor ion in the first MS scan (MS1). After fragmentation of the tagged peptides, uniquely identifiable reporter ions of the tags appear in the low *m*/*z* range of the second MS scan (MS2) spectrum and enable a relative quantification, whereas the fragmentation pattern of the peptide is used for the identification of the peptide (Rauniyar and Yates, 2014). However, label-based approaches are often cost intensive due to the expensive labels and limited to a small number of samples that can be analyzed in parallel (Neilson et al., 2011). In contrast, label-free approaches work without the modification of peptides and quantify proteins based either on spectral counting or on the measurement of peak intensities. Spectral counting uses for the identification and quantification solely the MS2 spectrum and estimates the abundance of a protein by the number of MS2 spectra detected for the peptides of a protein (Abdallah et al., 2012). An alternative is offered by the MaxQuant software package, which determines the intensities of peptide peaks at the MS1 level in the *m*/*z*-retention time plane. The intensity of a peptide is either the full peak volume or the intensity maximum over the retention time profile. The identification of the peptides is performed at the MS2 level by searching the MS2 spectra against a sequence database. The

intensities of a protein's peptides are afterwards used for the calculation of label free quantification (LFQ) values, which can be compared between the analyzed samples (Cox and Mann, 2008; Cox et al., 2014).

However, mass spectrometric approaches have certain limitations. For instance, membrane proteins and low-abundant proteins are underrepresented in mass spectrometry analyses (Han et al., 2008). Further, not all peptides of a protein can be observed or correctly identified, which is due to certain features of the peptides, such as modifications and hydrophobicity (Eichacker et al., 2004; Han et al., 2008). In addition, the identification of peptides is dependent on the quality of the protein database. Unknown proteins or isoforms as well as unknown modifications are not detectable in mass spectrometry analyses (Wang and Wilson, 2013).

1.2 Development of pollen in flowering plants

Plants play an essential role in the maintenance of today's atmosphere through the production of oxygen and simultaneous consumption of carbon dioxide (Igamberdiev and Lea, 2006). Further, flowering plants are essential for the food supply of the growing world population and therefore of great interest (Hill and Li, 2016). In general, flowering plants undergo a repetitive transition between a haploid gametophyte and a diploid sporophyte generation during their life cycle (Borg et al., 2009). Main function of the gametophyte generation is the development of female and male haploid gametophytes, which give rise to a diploid zygote after successful fertilization, marking the beginning of a new sporophyte generation (Yadegari and Drews, 2004). The development of the male gametophyte, better known as pollen, takes place in the anthers of stamina, which are specialized reproductive organs of the sporophyte (Scott et al., 2004; Yadegari and Drews, 2004). Pollen provides a model system for the analysis of cell growth and differentiation (Becker et al., 2003; Honys and Twell, 2004) and was focus of multiple studies in a variety of plant species, such as *A. thaliana* (Becker et al., 2003; Honys and Twell, 2003; Honys and Twell, 2004) and *Solanum lycopersicum* (hereinafter referred to as tomato; Chaturvedi et al., 2013; Giorno et al., 2013; Keller et al., 2017).

The development of pollen can be separated in two consecutive phases, namely microsporogenesis and microgametogenesis (Figure 2). During microsporogenesis, the diploid pollen mother cell, also known as microsporocyte, undergoes meiotic division to give rise to a tetrad of four haploid microspores. The phase is completed when the post-meiotic microspores are released from the tetrads (Borg et al., 2009). The release of the microspores is achieved by the activity of enzymes

secreted by the tapetum, which is the inner layer of the stamina and in addition responsible for the nutrient supply of the developing pollen (Scott et al., 2004). During microgametogenesis, the post-meiotic microspores enlarge and produce a large vacuole, which leads to the migration of the nucleus towards the cell wall. A final asymmetric cell division, better known as pollen mitosis I, leads to the characteristic cell-within-a-cell structure of the pollen grain with a generative cell that is embedded in the cytoplasm of a larger vegetative cell (Yamamoto et al., 2003).

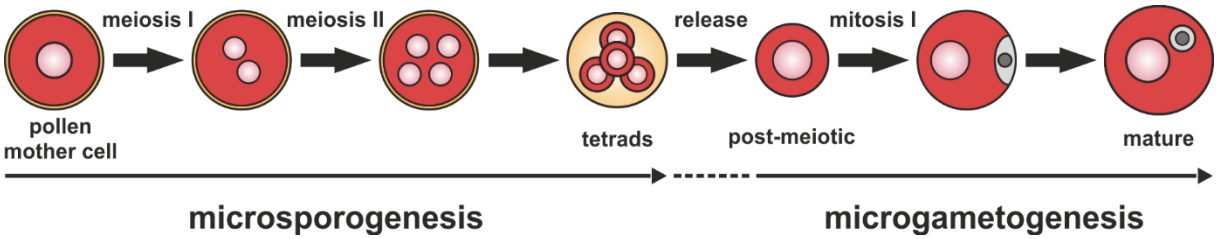


Figure 2: Development of pollen in tomato
During microsporogenesis, the diploid pollen mother cell undergoes meiotic division, which produces a tetrad of four haploid microspores. After release, the post-meiotic microspores enlarge and undergo an asymmetric mitotic division, which leads to a cell-within-a-cell structure. The mature pollen is bicellular and composed of a generative cell engulfed within the cytoplasm of a larger vegetative cell.

The engulfed generative cell, which is representing the male germline, and the surrounding vegetative cell have distinct fates (Borg et al., 2009). The vegetative cell is responsible for the nutrient supply of the generative cell and gives rise to the pollen tube after successful pollination. In contrast, the generative cell gives rise to a pair of sperm cells by a second mitotic division, also known as pollen mitosis II (Giorno et al., 2013; Rutley and Twell, 2015). In plant species with mature bicellular pollen grains (e.g. tomato) pollen mitosis II occurs during pollen tube growth, whereas in species with mature tricellular pollen grains (e.g. *A. thaliana*) it already occurs before the end of maturation (Giorno et al., 2013). After maturation, pollen grains dehydrate, followed by the opening of the anthers, which allows the dispersal of the pollen grains in the environment. Upon adhesion to the stigma, pollen grains rehydrate, start germination and produce the pollen tube. The pollen tube then intrudes into the stigma and grows towards the ovary, where it comes to a double fertilization of the female gametophyte with the two sperm cells (Firon et al., 2012).

1.2.1 Transcriptome and proteome dynamics during pollen development

The development of pollen is accompanied by dynamic changes in the composition and quantity of mRNAs and proteins. Most of the pollen studies published so far are based on mature pollen, as it is one of the most accessible stages (Ischebeck et al., 2014). Initial large-scale transcriptomic studies were performed in 2003 by Honys and Twell (2003) and Becker et al. (2003). In both studies, the authors used Arabidopsis GeneChip arrays, which allowed the detection of up to 8,000 of the 27,000

annotated protein-encoding genes. In total, the authors were able to identify 992 (Honys and Twell, 2003) and 1,584 (Becker et al., 2003) genes as expressed in mature pollen. Based on their findings, Honys and Twell (2003) estimated the number of expressed genes in mature *A*. *thaliana* pollen to be higher than 3,500. Subsequent functional classifications of genes exclusively expressed in pollen revealed an enrichment of signal transduction, cell wall metabolism, metabolic processes and cytoskeleton in both studies. Only one year later Honys and Twell (2004) expanded their analyses to four pollen developmental stages, namely microspores, bicellular pollen, immature tricellular pollen and mature pollen grains. Further, they utilized with the ATH1 GeneChip an array that carried probe sets for around 22,500 annotated genes, which increased the amount of detectable genes by a factor of about three. The transcriptome diversity of the developmental stages revealed a decrease from earlier to later developmental stages with 11,565 detected genes in microspores and only 7,235 genes in mature pollen grains. A similar decrease in transcriptome diversity was observed in rice, where the number of expressed genes decreased from 14,590 in microspores to only 5,945 in mature pollen grains (Wei et al., 2010). Today, the estimated number of expressed genes in mature *A*. *thaliana* pollen ranges from 3,945 to 7,235 across different studies with an average of 6,044 expressed genes (Rutley and Twell, 2015).

In contrast to the relatively well-characterized transcriptome of developing pollen, information about the pollen proteome is much more limited. Early proteomic pollen studies were based on 2-DE and rather limited in their output with only 110 to 135 detected proteins (Holmes-Davis et al., 2005; Noir et al., 2005; Sheoran et al., 2006). The first large-scale analysis in *A*. *thaliana* revealed the presence of at least 3,465 proteins in mature pollen grains (Grobei et al., 2009). The authors could further show an overrepresentation of proteins related to metabolism, energy, protein fate, protein synthesis, cellular transport and development.

Changes of the proteome along the course of pollen development was so far analyzed in tomato and tobacco based on five and eight developmental stages, respectively (Chaturvedi et al., 2013; Ischebeck et al., 2014). In total, in developing tomato and tobacco pollen 1,821 and 3,817 proteins, respectively, could be identified. The authors of both studies could further show that early developmental stages tend to accumulate proteins related to heat stress (HS), such as heat shock proteins (Hsps). In contrast, late developmental stages accumulate proteins required for germination and pollen tube growth, like those of cell wall and lipid metabolism as well as vesicle trafficking and the tricarboxylic acid (TCA) cycle.

When comparing the transcriptome and proteome of pollen, it turns out that the overrepresentation of processes like cell wall metabolism and signal transduction in the transcriptome of mature pollen (Becker et al., 2003; Honys and Twell, 2003) is also apparent in the proteome (Dai et al., 2006; Grobei et al., 2009; Chaturvedi et al., 2013). However, other processes like carbon and energy metabolism, which are overrepresented in the proteome of mature pollen (Dai et al., 2006) showed no overrepresentation in one of the transcriptomic studies. Further, a comparison of tobacco mRNA and *A. thaliana* protein levels of a phosphoglycerate kinase and a pyruvate decarboxylase revealed different abundance patterns along the course of pollen development. Both enzymes had most abundant protein levels in dehydrated pollen grains, whereas transcript levels peaked in earlier stages and were strongly diminished in mature pollen (Ischebeck et al., 2014).

1.3 Response of pollen to elevated temperature

The observed change in abundance and composition of the pollen transcriptome and proteome during development is accompanied by different sensitivities of the individual stages to abiotic stresses (De Storme and Geelen, 2014). One of the abiotic conditions with the biggest impact on the development of pollen is heat (Rieu et al., 2017). Application of HS to tomato plants during pollen development revealed a drastic reduction of fruit set as well as produced and germinating pollen (Sato et al., 2000). Regarding the sensitivity of the individual stages it was shown that pollen is most sensitive to high temperature from meiosis to pollen mitosis I and relatively heat tolerant thereafter (Bokszczanin et al., 2013; Rieu et al., 2017). Both meiosis and pollen mitosis I were shown to be disrupted by stresses like cold and heat (Zinn et al., 2010; Draeger and Moore, 2017), which might explain the sensitivity of the early stages to stresses.

In general, exposure of plants to elevated temperature can cause proteins to denature and form aggregates as well as increase the fluidity of the membrane (Lima et al., 2013). As plants are sessile organisms, they cannot escape high temperature and therefore have developed a complex stress response network (Guo et al., 2016). Key players of this stress response network are specialized Hsps that contribute to maintenance of protein homeostasis in the cytosol and organelles. Hsps are classified according to their molecular weight into the following families: Hsp100, Hsp90, Hsp70, Hsp60, Hsp40, small Hsp (sHsp / Hsp20) and Hsp10 (Feder and Hofmann, 1999; Baniwal et al., 2004; Fragkostefanakis et al., 2015; Zhang et al., 2015a). The number of members varies between the different Hsp families. For instance, in *A. thaliana* the size of the families ranges from only five (Hsp10 family) up to 129 members (Hsp40 family) (Fragkostefanakis et al., 2015). Despite their name, not all

Hsps are induced by HS. Many Hsps are constitutively expressed under normal conditions, whereas others are exclusively expressed upon HS (Rhoads et al., 2005; Ferradini et al., 2015). Expression of Hsps is mainly under the control of Hsfs. Hsfs induce the transcription of Hsps and other HS-related genes by binding so-called heat shock elements, which are palindromic motifs located upstream of the transcriptions start sites (Nover et al., 2001; Scharf et al., 2012). In contrast to vegetative tissues, mature and germinating pollen of most plant species lack the general HS response (HSR). Here, either no or only a subset of the HS inducible Hsps are synthesized, whereby the missing response was apparent at both the transcriptional and translational level (Cooper et al., 1984; Hopf et al., 1992; Mascarenhas and Crone, 1996). However, early pollen developmental stages, like microspores, clearly show an accumulation of various Hsfs and Hsps in response to HS at transcript and protein level (Frova et al., 1989; Frank et al., 2009; Rieu et al., 2017). Further, although early developmental stages were shown to be most sensitive to high temperature (Bokszczanin et al., 2013; Rieu et al., 2017), it could be shown that certain Hsfs and Hsps are accumulated in early developmental stages under normal conditions (Chaturvedi et al., 2013; Ischebeck et al., 2014; Fragkostefanakis et al., 2016). This phenomenon is called developmental priming and thought to protect the early sensitive stages in the process of meiosis and mitotic division against sudden stresses (Chaturvedi et al., 2013; Zhang et al., 2017b).

Large-scale transcriptomic and proteomic studies about the HSR of pollen are rather rare. However, it could be shown that 30 genes are upregulated in response to elevated temperature in microspores (Frank et al., 2009). Most of the upregulated genes are encoding for Hsps and proteins of reactive oxygen species (ROS) scavenging pathways, which is comparable to the other transcriptomic studies (Qin et al., 2008; Mangelsen et al., 2011; Li et al., 2013b). Furthermore, Chaturvedi et al. (2015) identified in post-meiotic and mature pollen 43 and eight heat-treatment responsive proteins, respectively. The responsive proteins in post-meiotic pollen included among other things HSP20 and HSP22 of the sHsp family as well as chaperone protein htpG of the Hsp90 family, whereas the responsive proteins in mature pollen included proteins related to glycolysis and the TCA cycle.

Despite the limited information available for pollen, numerous studies have analyzed the transcriptome response to elevated temperature in vegetative tissues of a variety of species, such as *A. thaliana* (Busch et al., 2005) and important crop plants like wheat (Qin et al., 2008; Kumar et al., 2015), barley (Mangelsen et al., 2011), switchgrass (Li et al., 2013b) and grape (Liu et al., 2012a). However, a general statement about the extent of the transcriptome response in plants cannot be made. The

effect of elevated temperature ranged from 4.2% differentially regulated transcripts in switchgrass (Li et al., 2013b) up to 20% differentially regulated transcripts across three time points in barley (Mangelsen et al., 2011). Reasons for this difference are diverse and include the use of sensitive and tolerant genotypes, different tissues (e.g. leaves, seedlings, caryopses or pooled tissues) as well as different temperature regimes in the mentioned studies.

Similar to the transcriptome also the response of the proteome to elevated temperature has been analyzed in important crop and model plants, including rice (Lee et al., 2007; Liao et al., 2014), soybean (Ahsan et al., 2010), alfalfa (Li et al., 2013a), wheat (Zhang et al., 2017a), grape (Jiang et al., 2017) and *A*. *thaliana* (Echevarría-Zomeño et al., 2016). The amount of differentially regulated proteins detected in the studies was strongly dependent on the used technique. Studies that were based on 2-DE were able to detect between 27 and 81 differentially regulated proteins (Lee et al., 2007; Ahsan et al., 2010; Li et al., 2013a; Liao et al., 2014). In contrast, recent studies using isobaric labeling revealed the differential regulation of 221 to 808 proteins in response to high temperature (Echevarría-Zomeño et al., 2016; Jiang et al., 2017; Zhang et al., 2017a). By considering all detected proteins, the proportion of differentially regulated proteins was 6.1, 10.3 and 14.3% in *A*. *thaliana* seedlings (Echevarría-Zomeño et al., 2016), wheat grains (Zhang et al., 2017a) and grape leaves (Jiang et al., 2017), respectively.

A look on the functional classification of the differentially regulated transcripts and differentially regulated proteins reveals, in part, similarities in the proteins and processes affected by high temperature. For instance, Hsps, ribosomal proteins, proteins involved in sugar signaling and proteins of ROS scavenging pathways are upregulated in both the transcriptome (Qin et al., 2008; Mangelsen et al., 2011; Kumar et al., 2015) and proteome (Ahsan et al., 2010; Echevarría-Zomeño et al., 2016; Jiang et al., 2017). Further, several photosynthesis-related proteins and proteins associated with amino acid metabolism are downregulated in response to high temperature at transcriptome and proteome level. Other processes, such as abscisic acid and calcium signaling, are induced in the transcriptome (Qin et al., 2008; Mangelsen et al., 2011; Li et al., 2013b) but show no changes in the proteome.

1.4 Objectives of this study

Pollen is an interesting model system for analyzing cell growth and development in plants as it undergoes several structural and morphological changes during its maturation. In addition, pollen is extremely sensitive to abiotic stresses like high temperature and therefore of major interest in the agricultural aspect of crop damage and yield loss due to global warming. So far, most studies analyzed pollen mainly at the transcriptome level with a strong focus on mature pollen. In contrast, only a small number of studies was dealing with the transcriptional characterization of early developmental stages and the proteome of pollen. Therefore, this study should focus on the analysis of the development and HSR of pollen at both the transcriptome and proteome level with an additional focus on regulatory mechanisms affecting these two levels.

The first part of this study addresses the transcriptome composition of developing and heat-stressed pollen. For this purpose, transcriptomic data from MACE experiments should be analyzed to identify mRNAs accumulated in certain developmental stages as well as mRNAs differentially regulated in response to HS. Subsequent functional analyses should enable the detection of important functional processes and protein families for the development and HSR of pollen. In the second part, the role of miRNAs in the development and HSR of pollen should be examined, as these are known to play an important role in plant development and stress response. Therefore, in a first step miRNAs have to be predicted by means of small RNA-seq libraries, followed by expression profiling of the predicted miRNAs. In the second step, a target prediction should be performed to identify potential target mRNAs of the predicted miRNAs, which should allow the determination of miRNAs required for the regulation of mRNAs during development and in response to HS. In the final part, the relationship between the transcriptome and proteome of developing and heat-stressed pollen should be examined. Based on proteomic data from LC-MS/MS experiments, proteins should be identified, quantified and correlated with their underlying mRNAs. As for pollen it is known that the transcriptome and proteome are weakly correlated (Grobei et al., 2009), it would be of interest to check if the weak correlation results from a delay between mRNAs and protein accumulation as, for instance, described for *Plasmodium falciparum* (Le Roch et al., 2004). Further, the regulation of proteins and mRNAs in heat-stressed pollen should be compared to find out if changes in protein abundance result from changes in mRNA abundance or if there is an mRNA-independent regulation.

2 Methods

2.1 Biological methods

Members of the Solanaceae Pollen Thermotolerance Initial Training Network (SPOT-ITN; http://www.spot-itn.eu/) conducted the biological methods. For transparency of the origin of the datasets analyzed in this study, the most important steps of the biological methods are briefly explained.

2.1.1 Growth conditions, stress treatment and pollen isolation

Plant growth, stress treatment and pollen isolation were carried out as described by Bokszczanin et al. (2015). In brief, tomato plants of cultivar Red Setter were grown under controlled conditions in a glasshouse facility of ALSIA – Research Center Metapontum Agrobios (Metaponto, Italy). The growth conditions included a 12-hour light-dark cycle with a day temperature of 24-26°C and a night temperature of 18-20°C at a relative humidity of 65-70%. For the heat treatment, plants were transferred in a preheated growth chamber and exposed to 38°C for 1h, followed by a decrease to 25°C within 30 minutes and a recovery at this temperature for an additional hour. The untreated plants were kept at 25°C for the same period. The isolation of the different pollen developmental stages was based on the knowledge about the correlation between the flower bud size and the developmental stage. Therefore, flower buds were sorted with respect to the developmental stages as follows: a length of 4-6mm for tetrads, 6-8mm for post-meiotic pollen and ≥10mm for mature pollen. Subsequently, the anthers were carefully dissected out of the flower buds, followed by the isolation of the pollen from the anthers. Non- and heat-stressed pollen were collected from three independent experiments that were performed on three consecutive days. Samples collected on the same day were treated as one biological replicate.

2.1.2 RNA isolation, library preparation and sequencing

RNA isolation as well as MACE and small RNA library preparation and sequencing (GenXPro Frankfurt am Main, Germany) were prepared as described by Bokszczanin et al. (2015).

RNA isolation – RNA was isolated from the three biological replicates for each developmental stage and condition by using the Macherey-Nagel NucleoSpin miRNA isolation kit according to the manufacturer's protocol, which resulted in a small (<200nt) and large fraction (>200nt).

MACE library preparation and sequencing – For the preparation of the MACE libraries, in a first step polyadenylated RNA was isolated from the large RNA fraction (>200nt) via the Dynabeads mRNA

Purification Kit (Life Technologies) and afterwards reverse transcribed into cDNA with the SuperScript Double-Stranded cDNA Synthesis Kit (Life Technologies) using biotinylated poly(dT) primers. Next, the cDNA was fragmented to an average size of 250bp with Bioruptor (Diagenode). Biotinylated cDNA ends were captured by Dynabeads M-270 Streptavidin Beads (Life Technologies) and ligated with T4 DNA Ligase 1 (NEB) to modified adapters (TrueQuant, GenXPro). PCR amplification was performed with KAPA HiFi Hot-Start Polymerase (KAPA Biosystems), followed by purification with Agencourt AMPure XP beads (Beckman Coulter) and sequencing on a HiSeq2000 (Illumina).

small RNA-seq library preparation and sequencing – For the preparation of the small RNA libraries, modified 3' and 5' adapters (TrueQuant, GenXPro) were ligated to the RNAs from the small fraction (<200nt) using T4 RNA Ligase 2 and T4 RNA Ligase 1 (NEB), respectively. The products were subsequently reverse transcribed into cDNA with SuperScript III (Life Technologies) and amplified by PCR with KAPA HiFi Hot-Start Polymerase (KAPA Biosystems). After size-selection by PAGE, the final libraries were sequenced on a HiSeq2000 (Illumina).

2.1.3 Protein extraction and LC-MS/MS measurements

The extraction of proteins and LC-MS/MS measurements were performed as described by Keller et al. (2018) and are briefly explained in the following.

Protein extraction – Proteins were extracted from three biological replicates for each developmental stage and condition. In a first step, pollen samples were freeze-dried and grounded, followed by the resuspension of the homogenized pollen samples in a protein extraction buffer. After phenol extraction, proteins were pre-fractionated by SDS-PAGE and in-gel digested with trypsin.

LC-MS/MS – Digested peptides were separated by one dimensional nanoflow liquid chromatography before entering the mass spectrometer. The MS measurement was performed on an Orbitrap LTQ XL mass spectrometer (Thermo). After each full MS1 scan, ten MS2 scans were performed on the 10 most abundant peptides.

2.2 Databases

2.2.1 miRBase

The microRNA database (miRBase; http://www.mirbase.org/; Kozomara and Griffiths-Jones, 2014) is a repository containing all published miRNA sequences and their associated annotation. In addition, for each miRNA literature references and secondary structure predictions of the corresponding pre-miRNA are provided. The nucleotide sequences of all deposited mature miRNAs were downloaded and used for the determination of known miRNAs among the predicted pollen miRNAs.

2.2.2 PubMed – NCBI

The PubMed database of the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/pubmed/; Agarwala et al., 2016) comprises millions of article abstracts and links to full journal articles. Further, PubMed allows the search for articles based on a simple or an advance search with a search builder tool. The latter one was used for the generation of the overview about EST, microarray and RNA-seq related publications over the past 30 years (Figure 1).

2.2.3 Rfam

The RNA family database (Rfam; http://rfam.xfam.org/; Kalvari et al., 2018) represents the families of different RNA types (e.g. snRNAs or snoRNAs) as multiple sequence alignments, consensus secondary structures and covariance models. In addition, the Rfam offers the opportunity to download the individual sequences of the multiple sequence alignments, which were subsequently used for the filtering of the potential pollen miRNAs.

2.2.4 SGN

The Sol Genomics Network (SGN; https://solgenomics.net/; Fernandez-Pozo et al., 2015) is a database comprising different tools and genomic data for the analysis of multiple members of the Solanaceae family and close relatives. The reference genome of tomato cultivar Heinz (SL3.0), the cDNA sequences (ITAG3.2), protein sequences (ITAG3.2) and the corresponding annotation (ITAG3.2) were downloaded from the FTP site of SGN.

2.2.5 TAIR

The Arabidopsis Information Resource (TAIR; https://www.arabidopsis.org/; Berardini et al., 2015) is a repository containing genomic data of Arabidopsis and further supporting information. The proteome of *A*. *thaliana* (TAIR10) was downloaded from the FTP server and used for the prediction of orthologs between *A*. *thaliana* and tomato.

2.2.6 PlantTFDB

The plant transcription factor database (PlantTFDB; http://planttfdb.cbi.pku.edu.cn/; Guo et al., 2008) is a database comprising predicted TFs of multiple plant species. All predicted TFs of tomato were downloaded from PlantTFDB and used for the detection of TFs and their family among the analyzed transcripts and proteins.

2.3 Programs

2.3.1 Bowtie2

Bowtie2 (version 2.3.4; http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; Langmead and Salzberg, 2012) is a fast gapped-read alignment program, which uses a full-text minute index to align sequencing reads to a reference genome. It was used for the alignment of the small RNA-seq libraries to the reference genome of tomato as well as the potential miRNAs against the nucleotide sequences downloaded from the Rfam.

2.3.2 CorelDraw

CorelDraw (version X7; https://www.coreldraw.com/) is a vector graphics editor that was used for the generation of figures in this study.

2.3.3 Cytoscape

Cytoscape (version 3.4.0; http://www.cytoscape.org/; Shannon et al., 2003) is a software package for the integration and visualization of biological interaction networks. It was used for the generation of miRNA-target networks.

2.3.4 DESeq2

DESeq2 (version 1.16.1; https://bioconductor.org/packages/release/bioc/html/DESeq2.html; Love et al., 2014) is an R package for differential expression analysis of count-based expression data. It normalizes read counts and models them by a negative binomial distribution. Further, DESeq2 provides different statistical tests for the detection of differentially expressed genes between two or more samples. DESeq2 was used for the detection of differentially regulated mRNAs either between the developmental stages or between non- and heat-stressed pollen stages.

2.3.5 Eclipse

Eclipse (version 4.4.2; https://www.eclipse.org/) and its plugin PyDev (http://www.pydev.org/) are integrated development environments, which were used for the implementation and execution of custom python scripts.

2.3.6 FactoMineR

FactoMineR (version 1.36; http://factominer.free.fr/; Lê et al., 2008) is an R package for multivariate exploratory data analysis. FactoMineR was used for the performance of principal component analyses (PCAs), which allowed the exploration and evaluation of the transcriptomic and proteomic data.

2.3.7 HTSeq

HTSeq is a python package (version 0.6.0; http://htseq.readthedocs.io/en/release_0.10.0/; Anders et al., 2015) for the processing of data from high-throughput experiments. The htseq-count script was used for the generation of read counts for each annotated gene.

2.3.8 InParanoid

InParanoid (version 4.1, http://inparanoid.sbc.su.se/cgi-bin/index.cgi; Remm et al., 2001) is a program for the prediction of orthologs and in-paralogs between two species by means of their proteomes. It was used for the detection of orthologs and co-orthologs between *A*. *thaliana* and tomato. The reported ortholog groups were afterwards screened for groups containing known *A*. *thaliana* Hsps, which in turn enabled the identification of Hsp orthologs in tomato.

2.3.9 MaxQuant

MaxQuant (version 1.6.0.16; http://www.biochem.mpg.de/5111795/maxquant; Cox et al., 2014) is a quantitative proteomics software package for the analysis of label-based (e.g. SILAC or iTRAQ) and label-free proteomics experiments. It was used for the label-free quantification of the LC-MS/MS libraries.

2.3.10 Mendeley

Mendeley (version 1.17.12; https://www.mendeley.com) is a reference management software, which was used for the management of references in this study.

2.3.11 Mercator

Mercator (https://mapman.gabipd.org/de/app/mercator; Lohse et al., 2014) is a web server enabling functional annotation of proteins by the assignment of MapMan terms to the query proteins. The MapMan ontology contains 35 main biological categories that are organized in a hierarchically structured tree. Mercator was used for the functional annotation of the tomato protein sequences.

2.3.12 NextGenMap

NextGenMap (version 0.4.12; http://cibiv.github.io/NextGenMap/; Sedlazeck et al., 2013) is a flexible and fast read aligner, which is hash-table based and has the ability to align reads even to highly

polymorphic genomes. It was used for the alignment of the MACE libraries to the reference genome of tomato.

2.3.13 Python scripts and libraries

Within this study, custom scripts were written in python (version 2.7; https://www.python.org/) to process and analyze the transcriptomic and proteomic data. This included the generation and normalization of expression values, statistical testing, the generation of input files for other tools, as well as the parsing of output files and the preparation of data for figures. In this context, the site-packages numpy (http://www.numpy.org/) and scipy (https://www.scipy.org/) were used for mathematical and statistical calculations, HTSeq (see 2.3.7) for the generation of read counts and Biopython (https://biopython.org/) for the handling and manipulation of nucleotide sequences.

2.3.14 R scripts and libraries

Custom R scripts were written and served the execution of DESeq2 (see 2.3.4) and FactoMineR (see 2.3.6) as well as the clustering of the miRNAs via the built in kmeans function.

2.3.15 Sigmaplot

Sigmaplot (version 13.0; https://systatsoftware.com/) is a software package for the statistical and graphical analysis of scientific data. It was used for the generation of figures in this study.

2.3.16 TargetFinder

TargetFinder (https://github.com/carringtonlab/TargetFinder) is a tool for the prediction of small RNA binding sites on target mRNAs based on Smith-Waterman alignments and a position-weighted scoring matrix. It was used for the prediction of miRNA targets among the cDNA sequences of tomato.

2.3.17 ViennaRNA Package

The ViennaRNA package (version 2.4.3; https://www.tbi.univie.ac.at/RNA/; Lorenz et al., 2011) consists of several stand-alone programs for the prediction and comparison of RNA secondary structures. The RNAfold program is part of the ViennaRNA package and determines an optimal secondary structure by computing the minimum free energy for each input sequence. RNAfold was used for the structure prediction of potential pollen pre-miRNAs.

2.4 Determination of publication rates of transcriptomics technologies

The publication rates of ESTs, Microarrays and RNA-seq were determined by PubMed searches (May 17, 2018; see 2.2.2) and used for sthe preparation of Figure 1. The searches were performed with the following user-defined queries, whereby the *XXXX* in the queries was replaced by the respective year (ranging from 1991 to 2017). (i) ESTs: *(expressed sequence tag[Title/Abstract] OR expressed sequence tags[Title/Abstract]) AND ("XXXXX"[Date - Publication] : "XXXX"[Date - Publication])*; (ii) Microarrays: *(Microarray[Title/Abstract] OR Gene Chip[Title/Abstract]) AND (mRNA[Title/Abstract] OR messenger RNA[Title/Abstract] OR cDNA[Title/Abstract] OR complementary DNA[Title/Abstract] OR transcript[Title/Abstract] OR transcriptome[Title/Abstract] OR gene expression[Title/Abstract]) AND ("XXXX"[Date - Publication] : "XXXX"[Date - Publication])*; (iii) RNA-seq: *(RNA-seq[Title/Abstract]) AND ("XXXX"[Date - Publication] : "XXXX"[Date - Publication])*.

2.5 Transcriptome analysis

2.5.1 Read alignment and quantification

Read alignment to the reference genome – In total, there were six samples (tetrads CO, tetrads HS, post-meiotic CO, post-meiotic HS, mature CO and mature HS) that were run in biological triplicates, resulting in a total number of 18 MACE libraries. The reads of the 18 MACE libraries were aligned to the reference genome of tomato (SL3.0; cultivar Heinz), which was downloaded from the SGN (see 2.2.4). The alignment was performed with NextGenMap (see 2.3.12), which was executed with default parameters except for the following modifications: --kmer-skip 1, --silent-clip and --no-unal. Subsequently, all alignments with an edit distance (sum of insertions, deletions and mismatches) greater than two were eliminated from the reported SAM file.

Quantification and normalization – As in tomato only a single isoform is annotated for each protein-encoding gene in the generic feature format (GFF) file, no discrimination between different mRNAs from the same gene took place. Therefore, the quantification of the mRNA level of a gene was based on the total number of reads aligned to the gene, which was done with the htseq-count script of HTSeq (see 2.3.7). The input of the htseq-count script were the SAM files with the alignments and the GFF file of tomato (ITAG3.2; see 2.2.4). Due to differences in the number of aligned reads between the different MACE libraries, the read counts of all mRNAs were library-wise normalized by dividing the read count of an mRNA by the total number of aligned reads, followed by the multiplication by one million, which led to TPM values for all mRNAs.

2.5.2 Threshold estimation and data quality control via PCA

Threshold estimation – To determine a threshold, at which it is assumed that an mRNA was really detected in a MACE library, all mRNAs not detected in every biological replicate of a sample were extracted. Next, these mRNAs were used to generate a distribution of replicate averaged TPM values for each of the six samples, whereby only replicates in which the mRNA was detected were taken into account. Based on the 95th percentiles of each of the six distributions a detection threshold of 1 TPM could be determined. In the end, all mRNAs detected (≥1 TPM) in at least two out of the three replicates were considered as detected for a sample. The TPM values of mRNAs not detected in a sample were set to 0 for all three replicates.

PCA – To obtain a first overview about the behavior of the replicates and samples, a PCA was performed with FactoMineR (see 2.3.6). As input served $\log_{10}$ transformed TPM values of the 18 MACE libraries.

2.5.3 Analysis of stage-accumulated mRNAs

Identification of stage-accumulated mRNAs – To identify mRNAs accumulated (stage-accumulated mRNAs) in a single or two consecutive developmental stages, a differential regulation analysis was performed with DESeq2 (see 2.3.4) by taking the read counts of the replicates of all CO samples as input. In total, four statistical tests were performed independently. These included a likelihood ratio test for the identification of mRNAs with significantly different levels between any two stages (adjusted p-value <0.05) and three pairwise Wald tests (tetrads vs post-meiotic, tetrads vs mature and post-meiotic vs mature) to figure out between which stages the difference is (adjusted p-value <0.05 and a $\log_2$ fold change <-1 or >1). The outcome of this procedure was used for the determination of stage-accumulated mRNAs. mRNAs were considered as accumulated in a single stage if they had significantly higher levels in this stage than in the other two stages. Further, mRNAs were considered as accumulated in two consecutive stages if they had no significant difference between the two stages, but in both stages significantly higher levels than in the remaining third stage. This approach resulted in five groups of stage-accumulated mRNAs, namely mRNAs accumulated in tetrads, tetrads + post-meiotic, post-meiotic, post-meiotic + mature and mature.

Functional enrichment analysis – For the functional characterization of the stage-accumulated mRNAs, all mRNAs of tomato were functionally annotated based on the MapMan ontology. For this purpose, the protein sequences of the mRNAs were submitted to the Mercator web server, which

assigned MapMan terms to each protein sequence, if possible (see 2.3.11). In the next step, a functional enrichment analysis was performed to identify functional categories that are enriched among the stage-accumulated mRNAs. For this purpose, a python script (see 2.3.13) was written, which performs a Fisher's exact test and a subsequent correction for multiple hypothesis testing via the false discovery rate (FDR) method of Benjamini and Hochberg. The test was performed for all MapMan terms of the second hierarchy level (e.g. protein.synthesis), as this level is available for the majority of mRNAs and sufficient for a clear identification of important processes. Each MapMan term was tested for the null hypothesis that there is no dependency between the term and the stage-accumulated mRNAs, whereby all annotated mRNAs served as background. A rejection of the null hypothesis means that there is a dependency and thus an enrichment of the term among the stage-accumulated mRNAs. The reported p-value of each term was afterwards corrected for multiple hypothesis testing and terms with an adjusted p-value <0.05 considered as enriched.

2.5.4 Collection of of Hsfs and identification of Hsps in tomato

The analysis of Hsfs and Hsps was based on known tomato Hsfs, described by Scharf et al. (2012), and predicted tomato orthologs of *A*. *thaliana* Hsps. The ortholog prediction was performed with InParanoid (see 2.3.8) by taking the proteomes of *A*. *thaliana* (TAIR10; see 2.2.5) and tomato (ITAG3.2; see 2.2.4) as input. The reported ortholog groups were afterwards screened for groups containing *A*. *thaliana* Hsps, which were described by Fragkostefanakis et al. (2015). Tomato genes assigned to the same group as *A*. *thaliana* Hsps were considered as tomato Hsps.

2.5.5 Analysis of mRNAs differentially regulated in response to HS

Differentially regulated mRNAs upon HS – For the identification of mRNAs with differentially regulated levels in response to HS, for each developmental stage a Wald test was performed with DESeq2, whereby the read counts of the replicates of the non- and heat-stressed samples were used as input. mRNAs with an adjusted p-value <0.05 and a $\log_2$ fold change <-1 and >1 were considered as down- and upregulated, respectively.

Hsf and Hsp analysis –The HS response of the seven Hsp families (Hsp100, Hsp90, Hsp70, Hsp60, Hsp40, sHsp and Hsp10) and of the Hsf family was analyzed based on the differential expression results of each developmental stage. For this purpose, all members of the Hsp and Hsf families were categorized as downregulated, not regulated or upregulated. Afterwards, the percentage each category constitutes in a family was determined.

2.6 Analysis of miRNAs in pollen

2.6.1 miRNA prediction, quantification and characterization

miRNA prediction (read alignment) – The reads obtained from the sequencing of the small RNA-seq libraries were in a first step filtered for reads with a length between 18 and 24 nucleotides and identical reads collapsed. The collapsing was done by representing reads with an identical sequence as a read stack, whereby the original number of reads with this sequence was represented as the height of the stack. Only read stacks with a height of at least two (the sequence occurred at least twice in the library) were further considered. The read stacks were afterwards aligned to the reference genome of tomato using bowtie2 (see 2.3.1), which was executed in the end-to-end mode with default parameters except for the following modifications: -D 20, -R 3, -L 10, -i S,1,0.50 and -k 20. Read stacks with an edit distance higher than 1 were removed directly afterwards.

miRNA prediction (filtering) – After the alignment, several filtering steps were applied to identify the most likely mature miRNAs among the aligned read stacks. First, if more than one alignment was reported for a read stack, only the alignments with the highest alignment score were kept. Second, all alignments that overlap with an alignment of a higher read stack were removed. The remaining read stacks were afterwards aligned against all sequences of the Rfam (see 2.2.3) with bowtie2 using the same parameters as for the read stack alignments but with the prohibition of alignments to the reverse complement by means of the --norc parameter. After rejection of alignments that didn't have the highest alignment score for a read stack or an edit distance higher than 1, only those read stacks were kept that had either no reported alignments or were aligned to a miRNA sequence. Next, based on the assumption that miRNAs originate either from intronic or intergenic regions of the genome, all alignments overlapping with an exon of the tomato GFF file were removed. In the final filtering step, all alignments that were not detected in all three biological replicates of a sample were excluded.

miRNA prediction (secondary structure prediction) – Since during its maturation the precursor of the miRNA forms a characteristic hairpin loop structure, the final step for the detection of miRNAs was a secondary structure prediction with RNAfold of the ViennaRNA package (see 2.3.17). For this purpose, for each alignment of a read stack two potential precursor regions were excised out of the genome. In accordance with Friedländer et al. (2012) the first region ranges from 20 nucleotides upstream to 70 nucleotides downstream of the alignment and the second region from 70 nucleotides upstream to 20 nucleotides downstream. The excision of two regions was necessary, as it is not clear

if the potential miRNA originated from the 5' or 3' arm of the pre-miRNA. All excised regions were subsequently used as input for RNAfold, which was run with default parameters. The reported secondary structures were afterwards examined for the existence of a valid hairpin loop structure using the following 5 criteria: (i) the miRNA had to be positioned in a hairpin loop structure, (ii) the hairpin loop has to cover at least 75% of the excised region, (iii) the miRNA must not be located in the loop, (iv) the loop has a minimal length of 4 nucleotides and (v) the minimal free energy of the predicted structure is lower than -35 (kcal/mol). All read stacks with at least one secondary structure fulfilling the five criteria were considered as mature miRNAs. If multiple alignments of a read stack fulfilled the criteria, the genomic locations were stored but only a single mature miRNA was taken into account for the subsequent analyses. For the naming of the miRNAs, all detected miRNAs were sorted in lexicographical order based on their nucleotide sequence. The final name of a miRNA consisted of "solyc-miR" as a prefix and the position of the miRNA in the sorted list as suffix.

miRNA quantification – Similarly to the quantification of the mRNA levels of genes, also for the quantification of the miRNAs the TPM measure was used. For this purpose, the height of the read stack was divided by the sum of all read stack heights in a library, followed by the multiplication with one million.

Determination of known and novel miRNAs – To find out which of the predicted miRNAs are already known and which are novel, the miRNAs were searched against the mature miRNAs downloaded from the miRBase (see 2.2.1). A miRNA was considered as known if there was a perfect match with a miRNA from the miRBase. The predicted miRNAs that are already known were further classified based on the closest related organism with a perfectly matching miRNA. The closest related organism could be either tomato itself, a member of the Solanaceae family (*Nicotiana tabacum* or *Solanum tuberosum*) or a more distant species within the Viridiplantae kingdom.

Prediction of miRNA targets – The prediction of miRNA targets was performed with TargetFinder (see 2.3.16), which was run with default parameters. As input served the nucleotide sequences of the miRNAs and the cDNA sequences of tomato (ITAG3.2; see 2.2.4) that were downloaded from the SGN. Each pair of miRNA and predicted target were stored as miRNA-target interaction (MTI), whereby multiple binding sites of a miRNA on a target were stored as independent MTIs.

2.6.2 Analysis of the relationship between stage-accumulated miRNAs and mRNAs

Detection of stage-accumulated miRNAs – For the detection of stage-accumulated miRNAs, the miRNAs were clustered based on their relative abundance in the non-stressed pollen stages. For this purpose, the TPM values of the three biological replicates of each stage were averaged, followed by the division by the sum of the average TPM values of all three stages, which resulted in a relative abundance value between 0 and 1 for each stage. The relative abundance profiles of all miRNAs were used as input for the kmeans function of R (see 2.3.14). In total, nine clusters were created (k=9) by performing 1000 independent clusterings with a maximum of 100 iterations, whereby the clustering with the lowest residual sum of squares was reported. Afterwards, the clusters and their miRNAs were classified based on the average relative abundance across the three stages. A cluster and its miRNAs were considered as accumulated in a single stage if the average relative abundance of this stage was at least two times higher than the average of the other two stages. Further, clusters and miRNAs were considered as accumulated in two consecutive stages if their average relative abundance differed less than twofold, but were at least two times higher than the cluster average of the remaining third stage.

Developmentally relevant MTIs – For the identification of MTIs that might play a role in the development of pollen, all MTIs consisting of a stage-accumulated miRNA and a stage-accumulated mRNA (see 2.5.3) were extracted in a first step. Afterwards, those MTIs were filtered, where a change in the miRNA abundance is accompanied by a change of the mRNA abundance in the opposite direction. After that, the MTIs were classified into categories based on the abundance profile of their mRNAs and miRNAs. In the final step, the MTIs containing TF mRNAs (see 2.2.6) were used for the construction of an interaction network with Cytoscape (see 2.3.3).

2.6.3 Analysis of the relationship between miRNAs and mRNAs differentially regulated in response to HS

Detection of HS-relevant miRNAs – The miRNAs detected in a developmental stage were clustered based on their relative abundance in the non- and heat-stressed samples. In total seven clusters (k=7) were created using the kmeans function of R, whereby 1000 independent clusterings with a maximum of 100 iterations were performed. As for each developmental stage only the two clusters with miRNAs exclusively detected in non- (c1) and heat-stressed samples (c7) showed a clear difference (at least twofold increase or decrease), the miRNAs of the cluster c1 were considered as downregulated and the miRNAs of cluster c7 as upregulated in response to HS for all three stages.

MTIs with a role in HSR – The MTIs made up by down- or upregulated miRNAs and down- or upregulated mRNAs (see 2.5.5 – *Differentially regulated mRNAs upon HS*) were gathered and those pairs of miRNAs and mRNAs with an opposite HS regulation considered as MTIs with a potential role in HSR.

2.7 Analysis of the relationship between the transcriptome and proteome of pollen

2.7.1 Quantitative proteomics

Label-free quantification – The 18 LC-MS/MS libraries were used as input for MaxQuant (see 2.3.9) by importing them as independent experiments. The proteome of tomato (ITAG3.2; see 2.2.4), which was downloaded from the SGN, served as protein database. MaxQuant was executed with default parameters, with the addition of the calculation of LFQ intensities, whereby phospho (STY) modifications were additionally taken into account for quantification. For further analyses the "proteinGroups.txt" output file was used, which contains the detected protein groups (proteins not distinguishable via the detected peptides) and their LFQ intensities in the 18 libraries as well as further information that can be used for filtering of the data.

2.7.2 Pre-processing and correlation

Filtering of detected protein groups – In a first step, proteins representing contaminants or reverse database sequences were excluded, which was based on their "CON__" and "REV__" prefixes. Further, all protein groups only identified by a modification site were excluded. Subsequently, a protein group was determined as detected in a sample if it had LFQ intensities for at least two of the three biological replicates, otherwise the LFQ intensities of all replicates were set to 0.

Reduction of multi protein groups – As a small proportion of the detected protein groups harbor more than one protein, the question was whether all proteins in the protein group were expressed or if they should be excluded from further analyses. MaxQuant indicates in its output the so-called "majority proteins" of a group, which are proteins that have at least half of the peptides that were identified for the leading protein. Therefore, the proteins of the multi protein groups were classified as either majority or minority proteins, followed by the calculation of their average mRNA levels across all MACE libraries. Based on a histogram representing the mRNA levels of the minority and majority proteins, the decision was to exclude the minority proteins from the protein groups as a substantial amount of the minority proteins was not detected in the transcriptome of pollen.

Correlation of mRNA and protein levels – The direct correlation of mRNA and protein levels was based on genes with measured mRNA and protein levels, whereby only protein groups containing a single protein were taken into account. In total, six correlations were performed, whereby the transcriptome of each stage was correlated with its own proteome and the proteome of the later stages. The correlation was based on the replicate averaged and afterwards $\log_2$ transformed TPM values and LFQ intensities.

2.7.3 Analysis of the relationship between stage-accumulated proteins and mRNAs

Identification of stage-accumulated protein groups – For the detection of stage-accumulated protein groups, first an analysis of variance (ANOVA) was performed to detect protein groups with significantly different LFQ intensities between the CO samples of the stages, followed by a Tukey post-hoc test to identify between which stages the difference is. As for most statistical analyses normally distributed data is beneficial, the LFQ intensities of all protein groups were in a first step $\log_2$ transformed after a pseudo count of 1 was added to the LFQ intensities. Next, for each of the detected protein groups the ANOVA was performed on the $\log_2$ transformed LFQ intensities of the three replicates of each non-stressed developmental stage. However, as the proteome coverage was very low, replicates with a $\log_2$ value of 0 (LFQ intensity + pseudo count was 1) would distort the statistical test. Therefore, all replicates with a value of 1 were removed before the ANOVA, except if all three replicates had a value of 1. The p-values of all protein groups were afterwards corrected for multiple hypothesis testing with the FDR method of Benjamini and Hochberg and all protein groups with and adjusted p-value <0.05 were subsequently used for three pairwise Tukey post-hoc tests that were run with a significance level of 0.05. Based on the results of the three tests, stage-accumulated protein groups were determined. A protein group was considered as accumulated in a single stage if the group had significantly higher levels in this stage than in the other two stages. Further, protein groups were considered as accumulated in two consecutive stages if they had no significant difference between the two stages, but in both stages significantly higher levels than in the remaining third stage.

Detection of mRNAs with direct and delayed translation in developing pollen – In a first step, the overlap between each group of stage-accumulated mRNAs (see 2.5.3 – *Identification of stage-accumulated mRNAs*) and each group of stage-accumulated protein groups was determined. The overlap between a group of stage-accumulated mRNAs and a group of stage-accumulated protein groups was defined as the genes with an mRNA and a protein in the two compared groups. Based on

the calculated overlaps, two translational modes could be derived that were defined as direct and delayed translation. Direct translation is presumed when the first stage with accumulated levels is the same for the mRNA and protein (e.g. mRNA is accumulated in tetrads and protein in tetrads + post-meiotic). In contrast, delayed translation is apparent when the accumulated level of the protein is delayed by one stage in comparison to the mRNA (e.g. mRNA is accumulated in tetrads and protein in post-meiotic + mature). In the end, five groups of mRNAs with direct or delayed translation could be determined, namely tetrads direct, tetrads delay, post-meiotic direct, post-meiotic delay and mature direct.

2.7.4 Analysis of the relationship between proteins and mRNAs differentially regulated in response to HS

Detection and analysis of protein groups differentially regulated in response to HS – For the identification of down- and upregulated protein groups, for each of the three stages an unpaired Student's *t*-test was performed. As input served the log_2 transformed LFQ intensities of the replicates of the non- and heat-stressed samples, whereby again log_2 values of 0 were excluded in the test of a protein group. Next, the reported p-values of a stage's *t*-tests were corrected for multiple hypothesis testing (FDR method of Benjamini and Hochberg) and all protein groups with an adjusted p-value <0.05 considered as down- and upregulated, respectively. Afterwards, the HS regulation of the mRNAs of the down- and upregulated proteins was analyzed to examine if the protein regulation is dependent or independent of the mRNA regulation.

3 Results

The aim of this study was the molecular biological analysis of non- and heat-stressed pollen developmental stages to gain insights in the development and HSR of pollen. For this purpose, the three developmental stages tetrad, post-meiotic pollen and mature pollen were isolated from tomato plants (cultivar Red Setter) that were either kept under control conditions at 25°C (non-stressed) or exposed to HS at 38°C for one hour, followed by a recovery phase at 25°C (heat-stressed). RNA and protein samples were taken from each sample (combination of developmental stage and temperature) in biological triplicates and used for the preparation of 18 MACE, small RNA-seq and LC-MS/MS libraries, respectively.

3.1 The transcriptome of developing and heat-stressed pollen

There are different molecular levels at which the behavior of an organism during development or in the response to external stimuli can be analyzed. One of the most accessible and best-studied levels is the transcriptome, which was analyzed in the first part of this study for developing and heat-stressed pollen. For this purpose, the reads of the 18 MACE libraries were aligned to the reference genome of tomato (Table 1).

Table 1: Alignment statistics of the MACE libraries
Shown are the alignment statistics for the 18 MACE libraries. For each non- (CO) and heat-stressed stage (HS), the total number of reads in each biological replicate as well as the percentage of aligned reads and reads aligned to annotated mRNAs are shown.

sample		replicate	total number of reads	percentage of aligned reads	percentage of aligned reads to mRNAs
tetrads	CO	1	4,220,434	73.6%	52.6%
		2	4,548,389	76.6%	55.1%
		3	5,398,343	82.5%	60.2%
	HS	1	6,917,504	76.6%	53.0%
		2	5,690,789	82.2%	58.9%
		3	4,775,683	79.9%	58.0%
post-meiotic	CO	1	7,221,666	86.5%	65.6%
		2	8,131,060	84.7%	64.3%
		3	5,115,229	87.3%	65.7%
	HS	1	7,283,901	86.8%	67.0%
		2	6,345,970	87.4%	67.2%
		3	2,311,857	85.8%	63.3%
mature	CO	1	7,423,230	86.0%	63.9%
		2	5,387,727	88.6%	66.6%
		3	6,847,578	87.7%	64.6%
	HS	1	3,493,375	83.5%	61.4%
		2	4,738,745	84.0%	62.0%
		3	3,205,951	87.2%	64.6%

The total number of reads across the different libraries ranged from 2.3 (post-meiotic HS replicate 3) up to 8.1 million reads (post-meiotic CO replicate 2). Regarding the alignment rates of the libraries, the

lowest percentage of aligned reads was observed for the first CO replicate of tetrads with 73.6% of aligned reads, whereby 52.6% were aligned to mRNAs. In contrast, for the second CO replicate of mature pollen 88.6% and 66.6% of the reads were aligned to the genome and annotated mRNAs, respectively.

After the alignment of the MACE libraries, the next step was the quantification of mRNAs. For this purpose, read counts were created for all protein-coding genes, followed by a library-wise normalization, which resulted in TPM values for all mRNAs. In the next step, a threshold was determined, at which an mRNA is considered as detected. For this purpose, for each sample a distribution of replicate averaged TPM values was generated, whereby only mRNAs not detected in all three replicates of a sample were taken into account (Figure 3). The determination of an appropriate threshold was based on the 95th percentile of the distributions. The 95th percentile of the six distributions ranged from 0.47 TPM (mature CO) up to 0.91 TPM (post-meiotic HS), which led to the determination of a threshold of 1 TPM. An mRNA was in the end considered as detected in a sample if it was detected (≥1 TPM) in at least two of the three biological replicates.

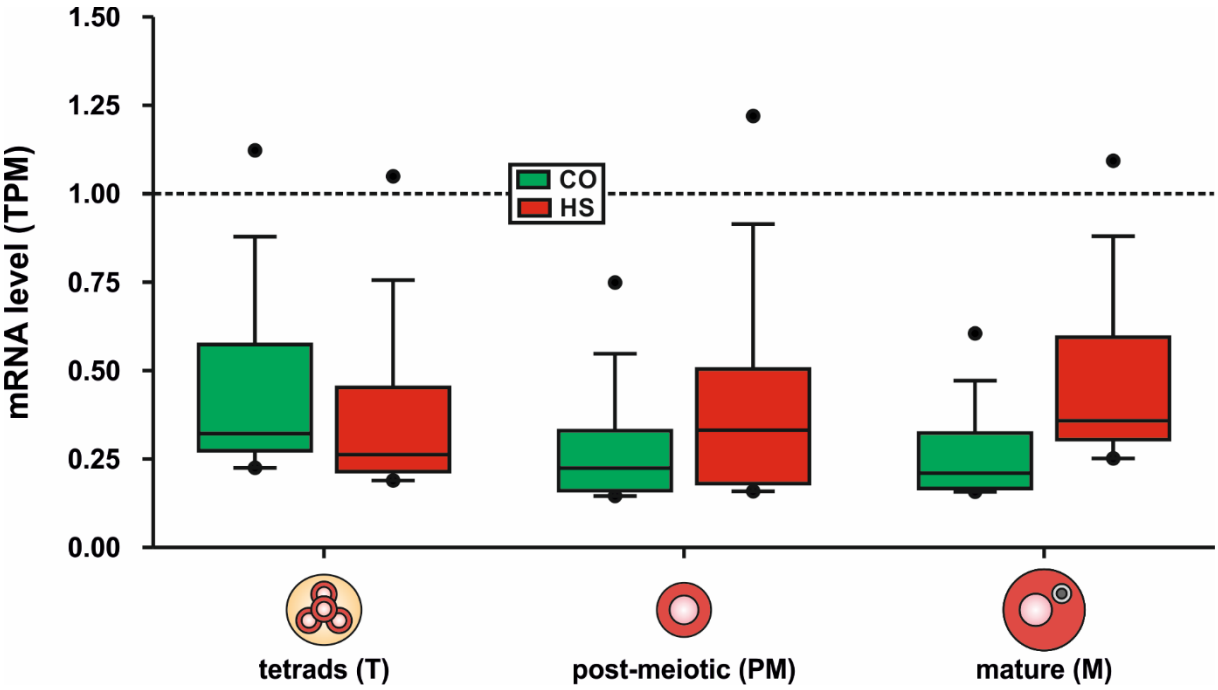


Figure 3: Determination of an mRNA detection threshold
For each sample, the mRNAs not detected in all three biological replicates were used for the generation of a TPM value distribution. The analyzed samples included non- (CO) and heat-stressed (HS) tetrads (T), post-meiotic (PM) and mature pollen (M). The horizontal line in each box represents the median and the whiskers the 10th and 90th percentile, respectively. The dashed line indicates the inferred threshold of 1 TPM.

For a first overview of the behavior of the replicates and samples, a PCA was performed (Figure 4). The PCA revealed that principal component 1 (PC1) explains around 60% of the variance of the

transcriptomic data and is already sufficient to separate the three developmental stages under both conditions (Figure 4A and B). The inclusion of PC2, which explains 20.5% variance, leads to an additional separation of the post-meiotic samples (Figure 4A), whereas PC3 clearly separates HS from CO samples but accounts for only 3.6% of the variance (Figure 4B).

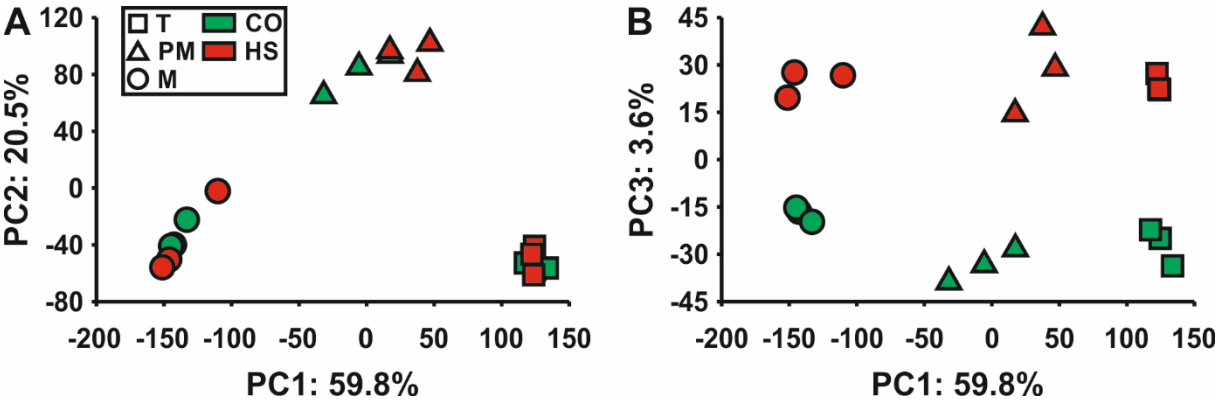


Figure 4: PCA on the MACE libraries of non- and heat-stressed pollen
(A and B) Shown are PCA plots for the analyzed MACE libraries of tetrads (T; square), post-meiotic (PM; triangle) and mature pollen (M; circle) either non- (CO; green) or heat-stressed (HS; red). Either PC1 against PC2 **(A)** or PC1 against PC3 **(B)** is plotted. Axes are labeled with the variance explained by the respective PC.

The PCA provided first indications of extensive changes in the transcriptome of developing pollen as well as evidence of minor alterations in the pollen transcriptomes in response to HS. Therefore, the transcriptome of developing and heat-stressed pollen was analyzed in more detail in the following to identify key processes of pollen development and the pollen HSR.

3.1.1 Stage-dependent induction of protein families and functional processes in developing pollen

As differences in the transcriptome seemed to be strongest between the developmental stages, the comparison of the transcriptomes of the developmental stages was the first step of this study. For a first qualitative overview of the transcriptomes, a venn diagram for detected mRNAs in the pollen transcriptomes was created (Figure 5A). In total, 19,080 different mRNAs were identified across all three developmental stages of which slightly above the half (10,387 mRNAs) were detected in all three stages. The individual mRNA diversity revealed a decrease from early to late stages with 18,212 mRNAs detected in tetrads, followed by 14,923 mRNAs in post-meiotic and 11,068 mRNAs in mature pollen. Similarly, the number of mRNAs detected exclusively in two consecutive stages was nearly ten times higher for tetrads and post-meiotic pollen (3,846 mRNAs) than for post-meiotic and mature pollen (415 mRNAs).

In the next step, the 19,080 mRNAs detected in developing pollen were examined for mRNAs accumulated in one or two consecutive developmental stages. For this purpose, a differential regulation analysis was performed, which resulted in five groups of stage-accumulated mRNAs (Figure 5B), namely mRNAs accumulated in only one of the three stages (tetrads, post-meiotic and mature) as well as mRNAs accumulated in two consecutive stages (tetrads + post-meiotic and post-meiotic + mature).

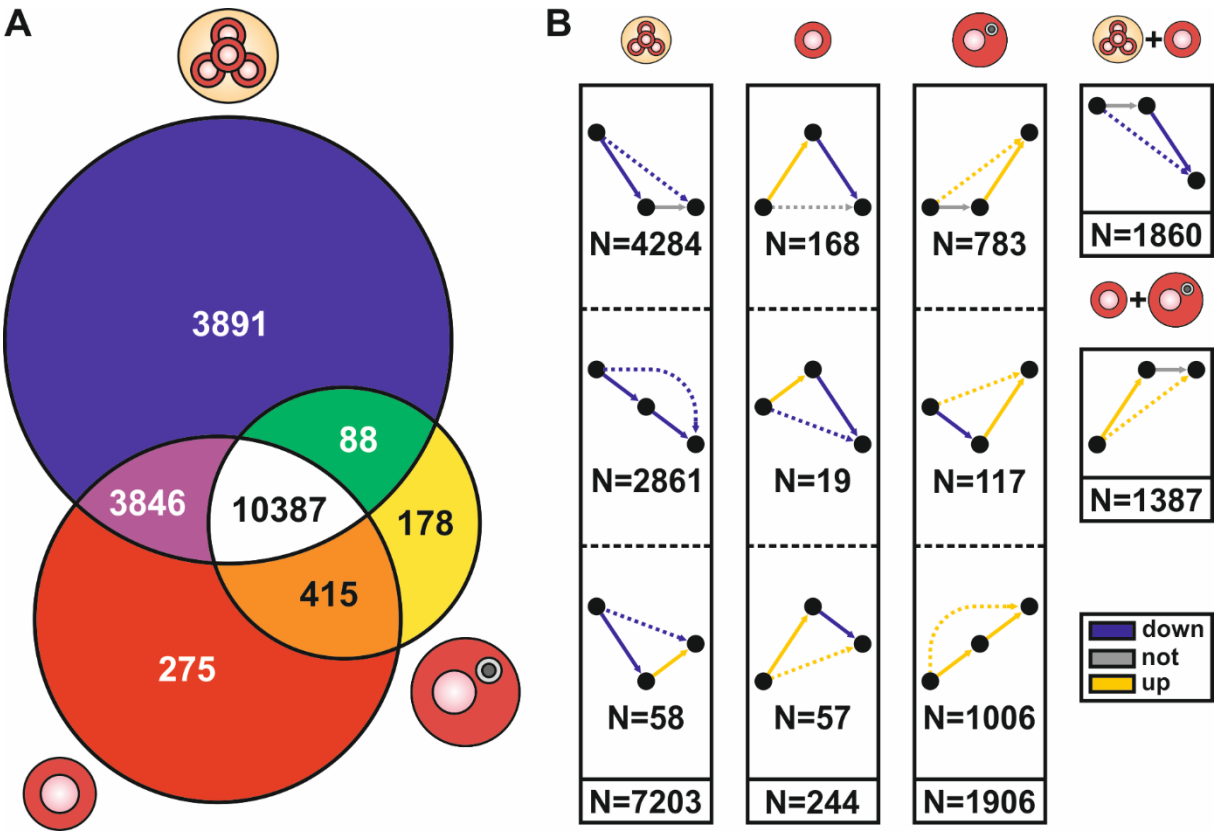


Figure 5: Detected and stage-accumulated mRNAs in developing pollen
(A) Shown is a venn diagram indicating the overlap of detected mRNAs between tetrads (blue), post-meiotic (red) and mature pollen (yellow). **(B)** mRNAs were categorized as either tetrads, post-meiotic, mature, tetrads + post-meiotic or post-meiotic + mature accumulated. Categorization was based on information about downregulation (down; blue arrows), no regulation (not; grey arrows) and upregulation (up; yellow arrows) between stages. The solid arrow indicates the regulation between two adjacent developmental stages and the dashed arrow the regulation between tetrads and mature pollen.

The highest number of stage-accumulated mRNAs was observed for tetrads with 7,203 mRNAs, followed by mature pollen with 1,906 mRNAs. With only 244 mRNAs, post-meiotic pollen shows the least number of accumulated mRNAs. However, post-meiotic pollen has a substantial amount of mRNAs showing a co-accumulation with tetrads (tetrads + post-meiotic; 1,860 mRNAs) and mature pollen (post-meiotic + mature; 1,387 mRNAs).

After the determination of stage-accumulated mRNAs, it was of interest whether these mRNAs encode for proteins of particular functional processes, which might play a role in the development of pollen.

Therefore, the proteins encoded by the mRNAs were first functionally annotated based on the MapMan ontology, followed by a functional enrichment analysis, which revealed overrepresented functional terms for each group of stage-accumulated mRNAs (Figure 6; Supplemental Table 1 highlighted in bold). For nearly all groups multiple enriched terms could be identified.

For tetrads (7,203 accumulated mRNAs) 28 enriched terms were identified (Figure 6 I), whereby the terms 'development.squamosa promoter binding like (SPL)', 'hormone metabolism.brassinosteroid' and 'stress.abiotic' were examined in more detail as these are known to play important roles in plant development (Chen et al., 2010; Yang et al., 2011; Chaturvedi et al., 2016). The term 'development.squamosa promoter binding like (SPL)' comprises squamosa promoter binding protein-like proteins (SPLs), which are plant-specific TFs (Chen et al., 2010). Out of the 15 mRNAs belonging to this term, eight are accumulated in tetrads. The term 'hormone metabolism.brassinosteroid' includes 40 of the mRNAs accumulated in tetrads. To get deeper insights in the proteins encoded by the 40 mRNAs, the proteins were searched among the tomato proteins involved in brassinosteroid synthesis and signaling, which were recently described by Simm et al. (2016). Interestingly, the proteins comprise nearly all proteins required for the synthesis of brassinolide from campesterol, namely CPD, DET2, CYP90D1 and Br6ox. Further, the proteins include the receptor kinases BRI1 and BAK1, the cytoplasmic kinase BIN2 and the transcription factor BES1/BZR2, which are all required for brassinosteroid signaling. The third enriched term 'stress.abiotic' includes 154 of the mRNAs accumulated in tetrads, whereby 64 of them are related to heat, nine to cold, 23 to drought/salt, 10 to touch/wounding, three to light stress and 45 not directly linked to a specific stress. Among the 64 mRNAs related to heat, multiple mRNAs encoding for members of the Hsf (Supplemental Table 2) and Hsp families (Supplemental Table 3) were identified. The Hsfs include HsfA1c, HsfA2, HsfA4a, HsfA9 as well as HsfB2a and the Hsps six members of the sHsp family, 19 of the Hsp40 family as well as two of the Hsp70, Hsp90 and Hsp100 family.

Out of the 11 enriched terms in tetrads and post-meiotic pollen (1,860 accumulated mRNAs), 'protein.synthesis' was most enriched and, together with 'protein.aa activation', comprises mRNAs encoding for components of the translation machinery (Figure 6 II). 162 of the mRNAs accumulated in tetrads and post-meiotic pollen belong to 'protein.synthesis' and encode, among other things, for 26 RPs of the small subunit, 41 RPs of the large subunit and 11 eIFs or subunits of eIFs. The term 'protein.aa activation' comprises 24 of the mRNAs accumulated in tetrads and post-meiotic pollen. The encoded proteins of these mRNAs include the aminoacyl-tRNA synthetases required for the transfer of

alanine, arginine, glutamine, glycine, histidine, leucine, lysine, methionine, proline, serine, threonine, tyrosine and valine onto tRNAs.

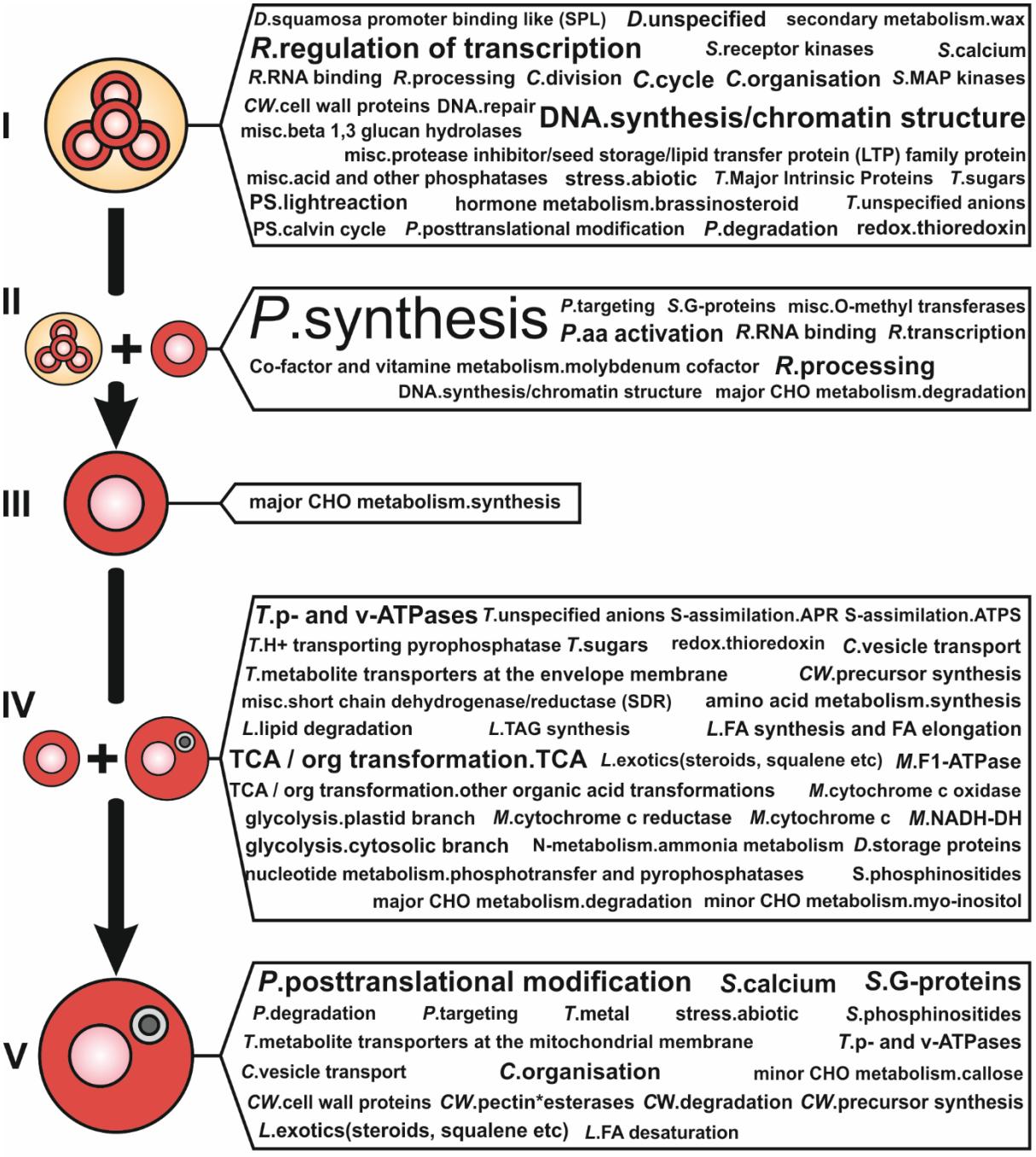


Figure 6: Enriched MapMan terms within the groups of stage-accumulated mRNAs
Shown are terms of the MapMan hierarchy that are enriched among stage-accumulated mRNAs. mRNAs are accumulated in tetrads (I), tetrads + post-meiotic (II), post-meiotic (III), post-meiotic + mature (IV) or mature (V). The font size of the enriched terms is inverse proportional to the adjusted p-value obtained from the functional enrichment analysis.

Only a single enriched term was observed for post-meiotic pollen (244 accumulated mRNAs), which is 'major CHO metabolism.synthesis' (Figure 6 III). This term covers five of the mRNAs accumulated in

post-meiotic pollen, which encode for a small and large subunit of the ADP-glucose pyrophosphorylase, a sucrose phosphate phosphatase, an isoamylase and a glucose-1-phosphate adenylyltransferase.

With 31 enriched terms, the highest number of enriched terms was observed for post-meiotic and mature pollen (1,387 accumulated mRNAs) (Figure 6 IV). One of the most enriched terms is 'transport.p- and v-ATPases', which covers 19 of the mRNAs accumulated in post-meiotic and mature pollen. The proteins encoded by the 19 mRNAs include three out of four V_0 subunits (a, c and e) and seven out of eight V_1 subunits (A, B, C, D, E, F and H) of the V-type proton ATPase. Out of the remaining enriched terms, many are related to cellular respiration. These terms include for example 'glycolysis.cytosolic branch', which covers 13 of the accumulated mRNAs. These mRNAs encode for the proteins required for step 2 (glucose-6-phosphate isomerase), 3 (phosphofructokinase), 4 (fructose-bisphosphate aldolase), 6 (glyceraldehyde 3-phosphate dehydrogenase), 8 (phosphoglycerate mutase), 9 (enolase) and 10 (pyruvate kinase) of the cytosolic glycolysis. A second term related to cellular respiration is 'TCA / org transformation.TCA', which comprises 18 of the accumulated mRNAs. The proteins encoded by the 18 mRNAs include subunits of all three enzymes of the pyruvate dehydrogenase complex (E1, E2 and E3) as well as at least one subunit of all TCA cycle enzymes. Further enriched terms related to cellular respiration belong to the superior hierarchy level 'mitochondrial electron transport / ATP synthesis'. The accumulated mRNAs included in these terms encode for subunits and assembly factors of complex I (NADH-DH), III (cytochrome c reductase) and IV (cytochrome c oxidase) of the mitochondrial respiratory chain, cytochrome c and subunits of the F1-ATPase.

Out of the 19 enriched terms of mature pollen (1,906 accumulated mRNAs), four are related to the cell wall (Figure 6 V). The first of these terms is 'cell wall.precursor synthesis', which covers 13 of the mRNAs accumulated in mature pollen. The encoded proteins are involved in the biosynthesis of UDP-xylose, UDP-D-galacturonate and DP-4-dehydro-6-deoxy-D-mannose. The second term is 'cell wall.cell wall proteins', which comprises nine of the accumulated mRNAs. These mRNAs encode for arabinogalactan proteins (AGPs), leucine-rich repeat receptor-like protein kinase family proteins (LRR) and alpha-1,4-glucan-protein synthases. The last two terms related to the cell wall are 'cell wall.degradation' and 'cell wall.pectin*esterases'. The term 'cell wall.degradation' comprise 24 of the accumulated mRNAs, which encode for cellulases, pectate lyases and polygalacturonases. Regarding 'cell wall.pectin*esterases', 18 of the accumulated mRNAs are included in this term and encode for

pectin methylesterases (PMEs) and pectinacetylesterases. One enriched term not related to the cell wall is ‘cell.organisation’, which covers 60 of the mRNAs accumulated in mature pollen. The encoded proteins include actin and actin-related proteins, such as actin depolymerization factors (ADFs), profilin and fimbrin, as well as alpha and beta tubulin chains, which are components of the microtubules.

So far, the study has revealed strong changes in the mRNA transcriptome of developing pollen, which was apparent on the one hand in the mRNA diversity (the number of detected mRNAs) and on the other hand in the regulation of mRNA levels during development. Further, the functional analysis of mRNAs accumulated in specific developmental stages has revealed processes and protein families that are of importance at different points in time during pollen development.

3.1.2 Identifications of the pollen HSR core set and the effect of heat stress on stage-accumulated mRNAs

After the examination of transcriptome alterations during pollen development, the next step was to analyze the effect of HS on the transcriptome homeostasis of the developmental stages. To obtain a first overview of the diversity of the transcriptome of non- and heat-stressed pollen stages, the number of detected mRNAs between non- and heat-stressed samples of each stage were compared (Figure 7). For all three stages, the vast majority of detected mRNAs is shared between non- and heat-stressed samples. For instance, out of the 18,887 mRNAs detected in tetrads 93% (17,571 mRNAs) are detected in the non- and heat-stressed sample (Figure 7A), while these are 89% (14,656 out of 16,432 mRNAs) for post-meiotic (Figure 7B) and 88% (10,583 out of 12,048 mRNAs) for mature pollen (Figure 7C). Further, for tetrads the number of exclusively detected mRNAs was nearly identical for the non- (641 mRNAs) and heat-stressed sample (675 mRNAs). In contrast, for post-meiotic pollen, with 1,509 mRNAs, more than five times as many mRNAs were only detected in the heat-stressed sample than in the non-stressed sample (267 mRNAs). For mature pollen the number of mRNAs only detected in the heat-stressed sample (980 mRNAs) was two times higher than the number of mRNAs detected in the non-stressed sample (485 mRNAs).

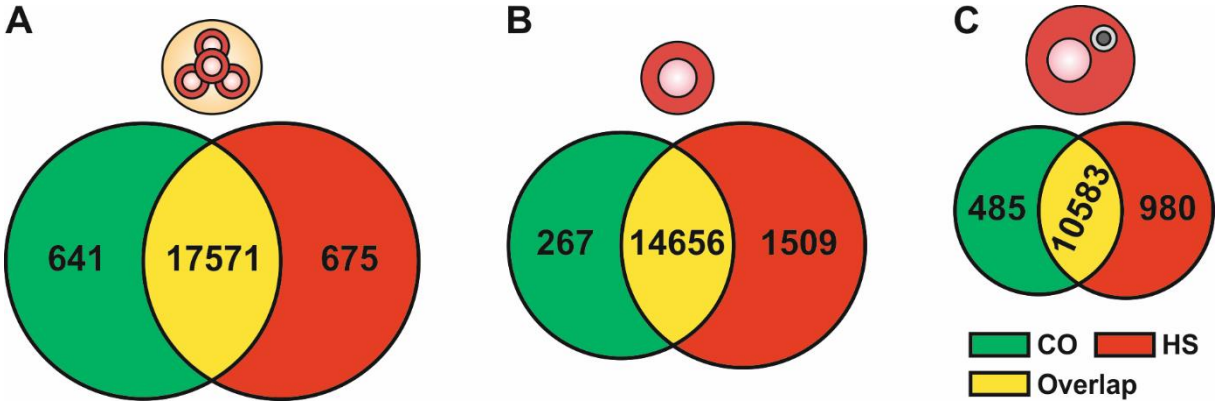


Figure 7: Detected mRNAs between non- and heat-stressed pollen stages
Venn diagrams indicate the number of detected mRNAs between non- (CO) and heat-stressed (HS) tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. mRNAs were either exclusively detected in CO (green) or HS (red) or detected in both conditions (Overlap, yellow).

After this first qualitative overview about detected mRNAs, it was of interest to identify those mRNAs, which are significantly down- and upregulated after HS. For this purpose, a differential regulation analysis was performed for the mRNAs detected in tetrads (18,887 mRNAs), post-meiotic (16,432 mRNAs) and mature pollen (12,048 mRNAs) (Figure 8). At this point it is important to point out that not all of the mRNAs only detected in non- or heat-stressed samples are significantly down- and upregulated, respectively. For instance, of the 641 mRNAs only detected in non-stressed tetrads, only 67 are significantly downregulated. The remaining mRNAs were too low abundant to be considered as significantly downregulated.

By far the lowest number of differentially regulated mRNAs was observed, with only 1%, for tetrads, whereby 82 mRNAs are down- and 104 upregulated (Figure 8A). In contrast, post-meiotic and mature pollen have with 5.7% and 4.5%, respectively, a much higher fraction of differentially regulated mRNAs (Figure 8B and C). In total, for post-meiotic pollen 256 down- and 684 upregulated mRNAs were identified and for mature pollen 113 down- and 432 upregulated mRNAs. Despite differences in the extent of the transcriptomic response, for all stages the number of upregulated mRNAs exceeded the number of downregulated mRNAs, which was most apparent for mature pollen where the number of upregulated mRNAs was nearly four times higher.

A comparison of the mRNAs downregulated in the three stages revealed that the downregulation is strongly stage-specific and that there is no common downregulation of mRNAs (Supplemental Figure 1A). However, the comparison of the upregulated mRNAs revealed a pollen HSR core set of 49 mRNAs, which showed in all three developmental stages an upregulation (Supplemental Figure 1B). Out of these 49 mRNAs, 19 belong to the MapMan term 'stress.abiotic.heat' and encode for members

of the Hsf and Hsp families (Supplemental Table 4). The detected Hsfs and Hsps include HsfA2 and HsfB1 as well as 13 members of the sHsp family, one Hsp70, one Hsp90 and two Hsp100. The remaining 30 mRNAs encode for other proteins than Hsfs or Hsps, whereby five of them do also belong to 'stress.abiotic.heat'.

As many of the mRNAs of the pollen HSR core set encode for Hsfs and Hsps, the HS regulation of these protein families should be analyzed in more detail. For this purpose, for each family the percentage of downregulated, not regulated and upregulated members was determined (Figure 9; Supplemental Table 5). For all three stages the lowest percentage of detected members was observed for the Hsf family (66.7% for tetrads, 59.3% for post-meiotic and 37% for mature pollen) and the highest percentage of detected members for the Hsp70 and Hsp90 family (100% for all stages). Furthermore, all families had the highest percentage of detected members in tetrads, while at the same time tetrads was the stage with the lowest percentage of upregulated members. The Hsf family can illustrate these findings. The highest percentage of Hsfs was, with 66.7%, detected in tetrads, which is followed by post-meiotic pollen with 59.3% and mature pollen with 37.0%. Despite having the highest number of

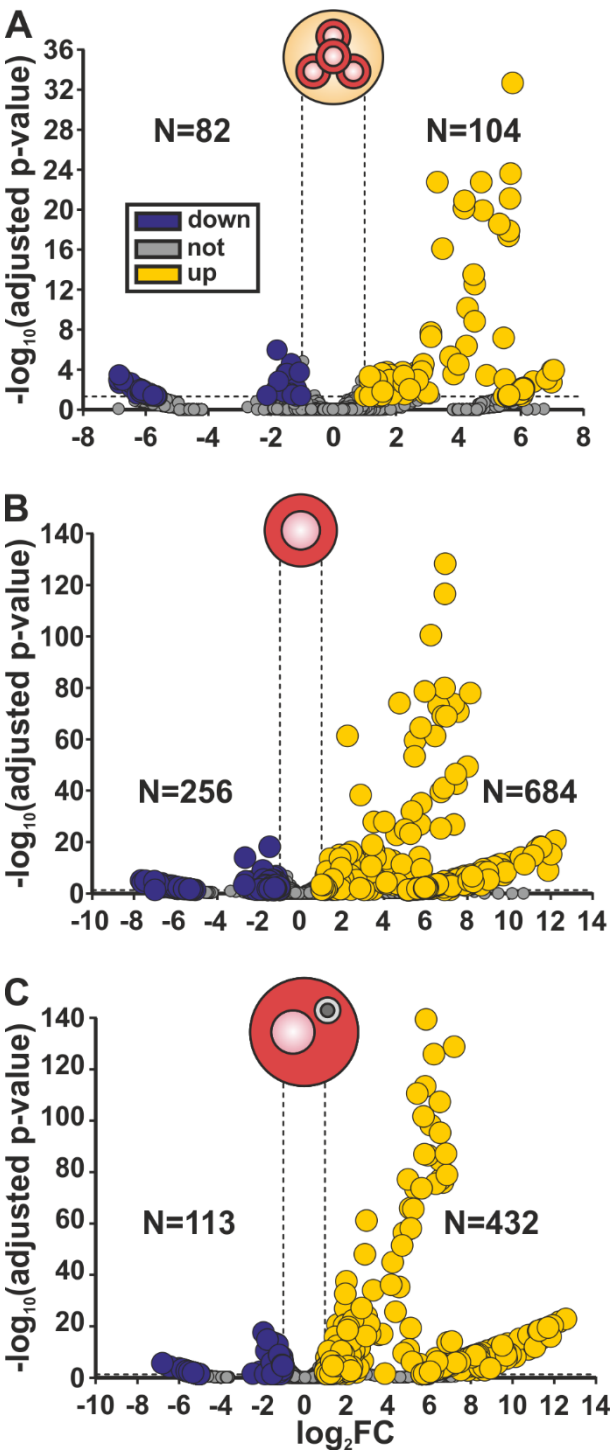


Figure 8: Response of pollen transcriptomes to HS
Shown are volcano plots indicating the differential regulation results of tetrads **(A)**, post-meioitc **(B)** and mature pollen **(C)**. Plotted is the log_2FC against the negative decadic logarithm of the adjusted p-value. mRNAs are either downregulated (blue circles), not regulated (grey circles) or upregulated (yellow circles) in response to HS. mRNAs were defined as differentially regulated if they had an adjusted p-value below 0.05 (horizontal line) and a |log2FC| > 1(vertical lines).

detected Hsfs, tetrads had with 11.1% the lowest number of upregulated Hsfs. In post-meiotic and mature pollen on the other hand 18.5% of the Hsfs are upregulated. Out of the upregulated Hsfs, HsfA2 and HsfB1 are upregulated in all stages, HsfB2b in tetrads and mature pollen, HsfA7 in post-meiotic and mature pollen, HsfB2a and HsfC1 only in post-meiotic and HsfA1b only in mature pollen.

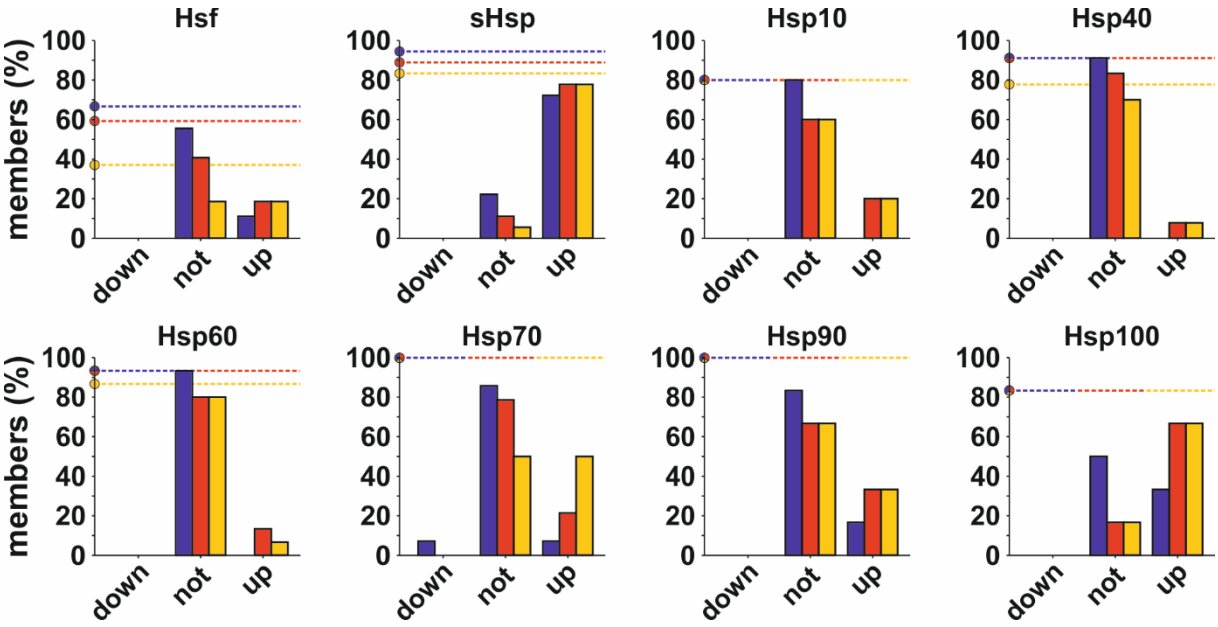


Figure 9: HS regulation of Hsf and Hsp families in pollen developmental stages
For each family the percentage of downregulated, not regulated and upregulated members in tetrads (blue), post-meiotic (red) and mature pollen (yellow) is shown. Dashed lines with circles indicate the percentage of family members detected in a developmental stage.

Interestingly, the total number of Hsf and Hsp members that are upregulated in post-meiotic (38 members) and mature pollen (41 members) is two times higher than the number of upregulated members in tetrads (20 members). Interestingly, many of the Hsfs and Hsps upregulated in post-meiotic and/or mature pollen but not in tetrads belong to the mRNAs accumulated in non-stressed tetrads. (Supplemental Table 5 last column), which matches the concept of developmental priming. Hsfs and Hsps with this behavior include HsfB2a as well as one Hsp10, Hsp70 and Hsp100 (Figure 10). HsfB2a is upregulated in post-meiotic pollen, not present in mature pollen and not differentially regulated in tetrads. Despite a missing upregulation in response to HS, the mRNA level of HsfB2a in heat-stressed tetrads is slightly higher than its level in heat-stressed post-meiotic pollen. The same holds true for the Hsp10, Hsp70 and Hsp100, which are all upregulated in post-meiotic and mature pollen but lack changes of their mRNA levels in tetrads. All three Hsps have in non-stressed tetrads already high mRNA levels that exceed the levels in at least one of the other heat-stressed stages.

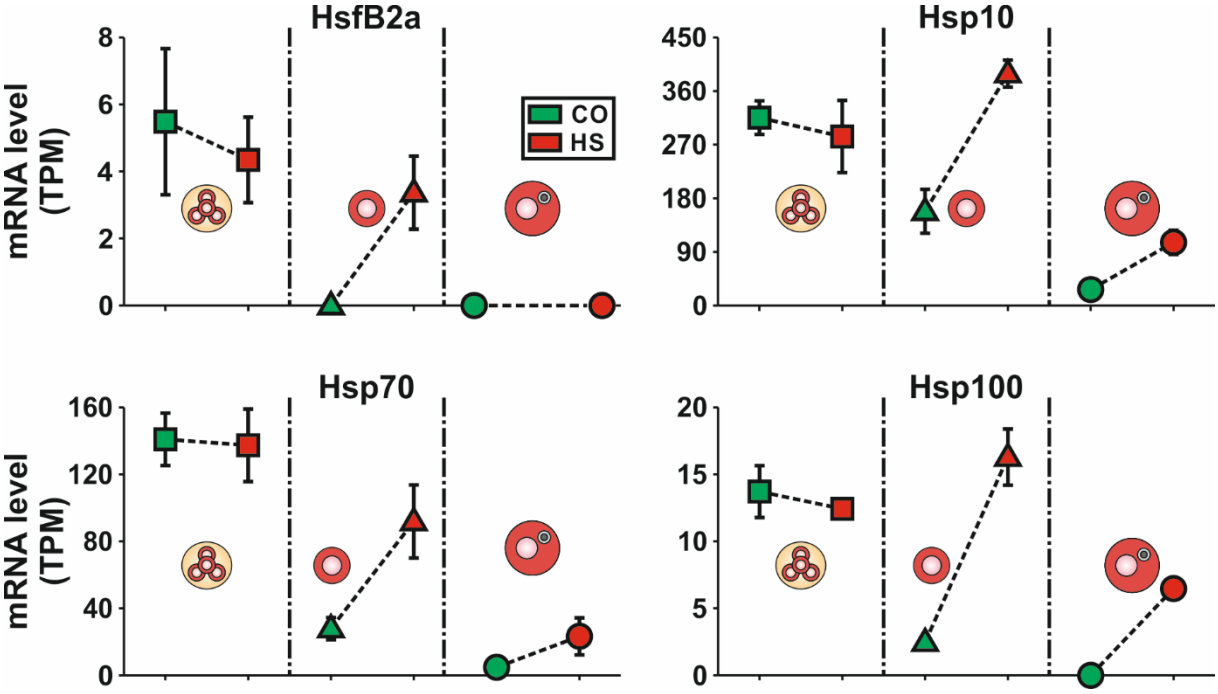


Figure 10: Hsf and Hsp family members pointing towards a developmental priming of tetrads
Shown are the mRNA levels (TPM) of HsfB2a (Solyc03g026020), an Hsp10 (Solyc07g042250), an Hsp70 (Solyc03g082920) and an Hsp100 (Solyc03g117950) in non- (green) and heat-stressed (red) tetrads (square), post-meiotic (triangle) and mature pollen (circle). Shapes indicate the mean of three biological replicates and error bars are the standard error.

After the observance that stage-accumulated mRNAs play a potential role in the HS response of pollen, the next step was to go in the opposite direction and analyze the effect of HS on the five groups of stage-accumulated mRNAs (Figure 5B). In accordance with the general HS response of the pollen transcriptomes (Figure 8), the analysis of the HS regulation of the stage-accumulated mRNAs revealed that the vast majority of these mRNAs is not differentially regulated in the stage in which they are accumulated (Figure 11). However, a small fraction of the stage-accumulated mRNAs is differentially regulated. Except for mRNAs accumulated in tetrads and post-meiotic pollen, the majority of the stage-accumulated mRNAs is downregulated, which is most apparent for mRNAs accumulated in post-meiotic and mature pollen out of which 119 are downregulated in post-meiotic pollen. Related to this, the HS regulation of the mRNAs accumulated in post-meiotic and mature pollen differs between the two stages. While in heat-stressed post-meiotic pollen 119 mRNAs are down- and three upregulated, these are only 25 down- and 16 upregulated in mature pollen.

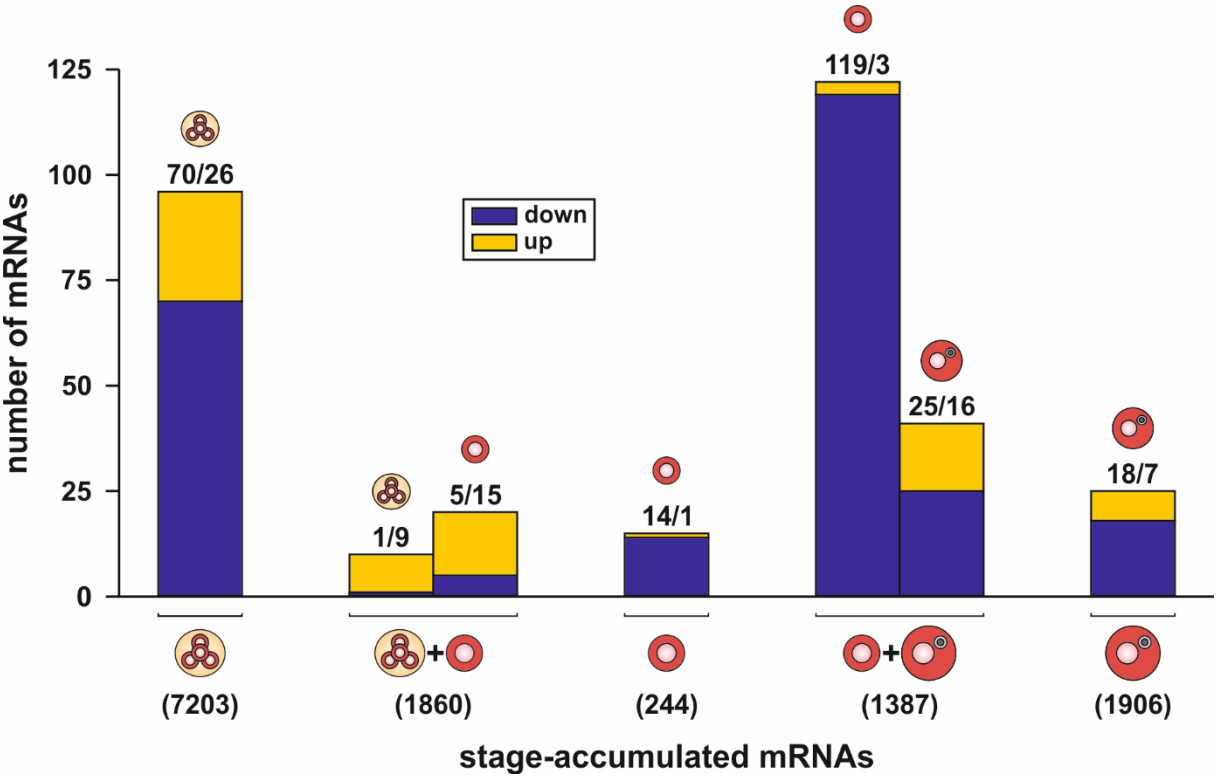


Figure 11: Stage-accumulated mRNAs differentially regulated in response to HS
mRNAs of the five groups of stage-accumulated mRNAs were searched among the mRNAs that are either down- (blue) or upregulated (yellow) in response to HS in the respective stage or stages they are accumulated. Below the bars, the total number of stage-accumulated mRNAs is indicated for each of the five groups. The bars indicate the number of stage-accumulated mRNAs that are differentially regulated. The first number above the bars represents the number of downregulated mRNAs and the second the number of upregulated mRNAs.

Although the majority of the stage-accumulated mRNAs was not affected by HS, the small number of differentially regulated mRNAs might have an influence on important functional processes of the developmental stages. For this purpose, the functional role of the stage-accumulated mRNAs that are differentially regulated in response to HS was further examined (Supplemental Table 6 to Supplemental Table 10). Among the stage-accumulated mRNAs that are upregulated in response to HS, 24 members of the pollen HSR core set (Supplemental Table 4) were identified. With 19 members most of them are accumulated and upregulated in tetrads, while four additional members are accumulated and upregulated in tetrads and post-meiotic pollen as well as one member in post-meiotic and mature pollen and another member only in mature pollen.

In contrast to the HS-related role of the upregulated mRNAs, the stage-accumulated mRNAs with a HS-induced downregulation possess very diverse functionalities. Nevertheless, among the 70 mRNAs accumulated and downregulated in tetrads, eight mRNAs encode for members of the G2-like, CAMT, bHLH, Dof, B3, M-type_MADS, and WRKY TF family. Further, two of the mRNAs accumulated and downregulated in post-meiotic pollen encode for members of the ERF and bZIP TF family. Among the 119 mRNAs accumulated in post-meiotic and mature pollen and downregulated in post-meiotic pollen,

multiple mRNAs encoding for potassium, sugar and amino acid transporters could be identified as well as developmentally relevant proteins like late embryogenesis abundant (LEA) proteins, vicilin-like and a patatin.

The comparison of non- and heat-stressed pollen transcriptomes led to the identification of a pollen HSR core set of 49 mRNAs. Further, Hsf and Hsp family members were identified that might contribute to a developmental priming of tetrads as they are HS responsive in at least one of the other two stages but already accumulated in non-stressed tetrads. The impact of HS on stage-accumulated mRNAs is rather small and revealed only a small number of differentially regulated mRNAs. However, among the stage-accumulated mRNAs that are upregulated, multiple members of the pollen HSR core set were identified, which indicates an accumulation of HS-relevant mRNAs already in non-stressed stages. Further, the downregulation of stage-accumulated mRNAs, like those encoding for transporters and TFs, might have a negative effect on pollen development.

3.2 The regulatory role of miRNAs in the development and heat stress response of pollen

The results so far have shown intensive transcriptome changes during the development of pollen and minor changes of the transcriptomes in response to HS. As it is known that miRNAs play a role in the development (Jones-Rhoades et al., 2006) and stress response of plants (Shukla et al., 2008), it was of interest whether this also applies to the development and HSR of pollen. The prediction of miRNAs was performed with a self-developed workflow that resulted in 793 potential miRNAs across all analyzed samples (Supplemental Table 11). The highest number of miRNAs was detected for non-stressed tetrads (494 miRNAs), followed by heat-stressed tetrads (360 miRNAs), non- (291 miRNAs) and heat-stressed post-meiotic pollen (181 miRNAs) and non- (127 miRNAs) and heat-stressed mature pollen with (112 miRNAs). The gradually decrease in the number of predicted miRNAs from early to late stages is similar to the decrease of the mRNA transcriptome during pollen development.

To find out if the predicted miRNAs are known or novel miRNAs, they were searched against the mature miRNA sequences deposited in the miRBase, whereby only perfectly matching miRNAs were considered as known. The known miRNAs were further classified based on their taxonomically closest hit in the miRBase as known in: (i) tomato, (ii) a species of the Solanaceae family other than tomato or (iii) in a more distant species of the Viridiplantae kingdom (Table 2). Although the vast majority of the predicted miRNAs had no perfect hits in the miRBase, for each sample between 28 and 31 of the

predicted miRNAs are present in the miRBase. The number of predicted miRNAs already known in tomato is in a range between 21 and 23 miRNAs. Out of these miRNAs, 18 have been identified in all six samples, while six are only present in a subset of the samples, like solyc-miR530 (miRBase entry: sly-miR156d-3p), which was not identified in non- and heat-stressed tetrads (Supplemental Table 12). Not yet known in tomato but in another species of the Solanaceae family are five to eight of the identified miRNAs. Two of them were identified in all samples, while nine are missing in at least one of the samples, such as solyc-miR625 (closest miRBase entry: stu-miR393-5p), which was not identified in heat-stressed mature pollen. Further, one or two of the identified miRNAs are known in distant species of the Viridiplantae kingdom. Two of them were exclusively detected in a single sample, while solyc-miR660 (closest miRBase entry: gma-miR172k) was detected in all samples except for heat-stressed tetrads.

Table 2: Known and novel miRNAs among the predicted miRNAs
The predicted miRNAs of each non- (CO) and heat-stressed (HS) developmental stage were searched against the miRBase. Based on the search results, the miRNAs were classified in the following order as detected in: (i) tomato, (ii) the Solanaceae family, (iii) Viridiplantae kingdom. If there were no matches in the miRBase, the miRNAs were considered as novel. The total number of predicted miRNAs is given in the last line.

classification	tetrads		post-meiotic		mature	
	CO	HS	CO	HS	CO	HS
tomato	23	22	22	22	21	22
Solanaceae	5	5	8	7	7	5
Viridiplantae	2	1	1	1	1	1
novel	464	332	260	151	98	84
total	494	360	291	181	127	112

To obtain information about the impact of the predicted miRNAs on the development and HSR of pollen, a target prediction was performed to identify potential target mRNAs. For 422 out of the 793 predicted miRNAs at least one potential target mRNA was predicted (Supplemental Figure 2), which is a coverage of about 53%. The average number of predicted targets per miRNA is 5.9, whereby 41% (172 miRNAs) of the 422 miRNAs have only a single predicted target mRNA. In addition, 89% of the 422 miRNAs do not have more than 10 target mRNAs. The combinations of miRNA and target mRNA were stored as MTIs.

3.2.1 Identification and impact of developmentally relevant miRNAs

After a first overview about the predicted miRNAs, the next step was the identification of miRNAs that might play a role during the development of pollen. Out of the 793 predicted miRNAs, 641 were identified in at least one of the non-stressed stages, whereby around 12.3% (79 miRNAs) are present in all of them (Figure 12). Similar to the mRNAs, also for the miRNAs a decrease in their diversity from tetrads (494 miRNAs) towards mature pollen (127 miRNAs) was observed.

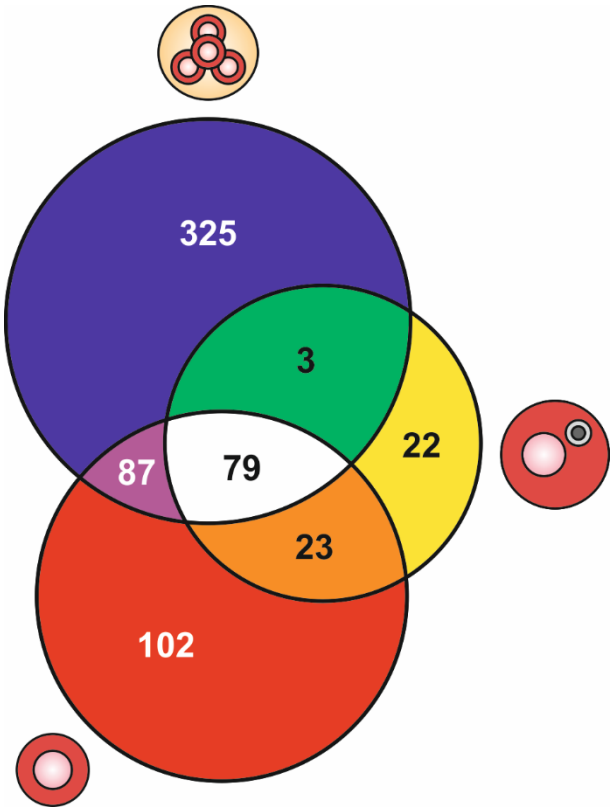


Figure 12: Detected miRNAs in developing pollen
Shown is a venn diagram indicating the overlap of detected miRNAs between non-stressed tetrads (blue), post-meiotic (red) and mature pollen (yellow).

For the identification of developmentally relevant miRNAs, the 641 miRNAs were clustered based on their relative abundance in the non-stressed stages (Supplemental Figure 3). Out of the nine generated miRNA clusters, eight comprised miRNAs accumulated in a single or two consecutive stages (Figure 13). Cluster one and two comprise miRNAs accumulated in tetrads, clusters three and four miRNAs accumulated in tetrads and post-meiotic pollen, cluster five miRNAs accumulated in post-meiotic pollen, cluster six miRNAs accumulated in post-meiotic and mature pollen and cluster seven and eight those miRNAs accumulated solely in mature pollen. Overall, between 21 (post-meiotic and mature pollen) and 355 stage-accumulated miRNAs (tetrads) were identified.

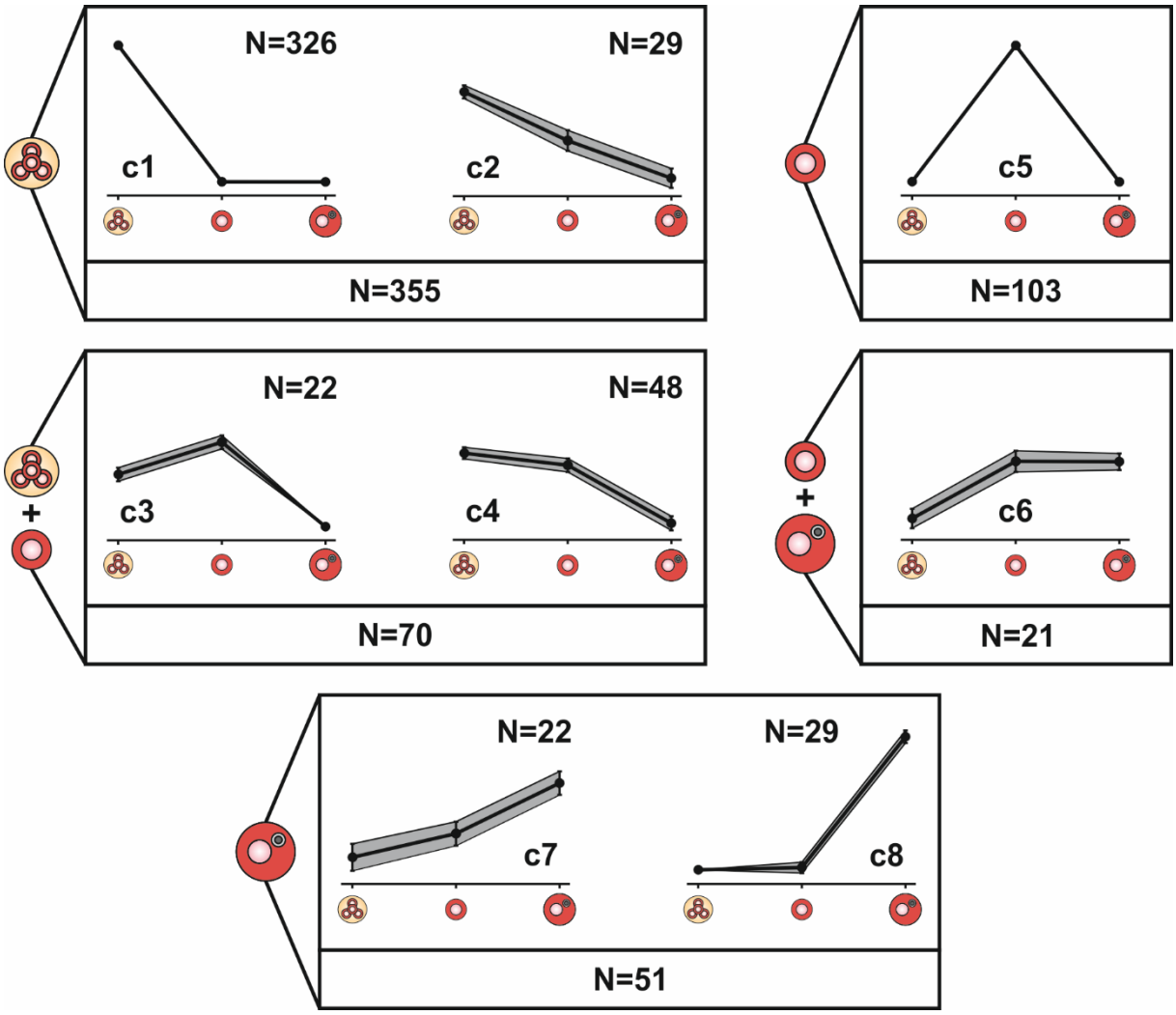


Figure 13: Clusters of stage-accumulated miRNAs
Shown are miRNA clusters, which comprise miRNAs accumulated in one or two consecutive stages. Clusters were categorized as either tetrads, post-meiotic, mature, tetrads + post-meiotic or post-meiotic + mature accumulated.

After the identification of stage-accumulated miRNAs, it was of interest whether they are involved in the regulation of mRNAs, which had been shown to be relevant for the development of pollen. For this purpose, the previously identified MTIs were searched for those consisting of a stage-accumulated miRNA and a stage-accumulated mRNA, whereby the miRNA and mRNA had to have an inverse behavior in their abundance during the development of pollen. In total, eight categories (A to H) of such developmentally relevant MTIs were detected (Figure 14), which comprise between one and 73 MTIs. The 73 MTIs belong to category A and have in common that their mRNA is downregulated during the transition from tetrads to post-meiotic pollen, while the miRNA behaves in the opposite direction.

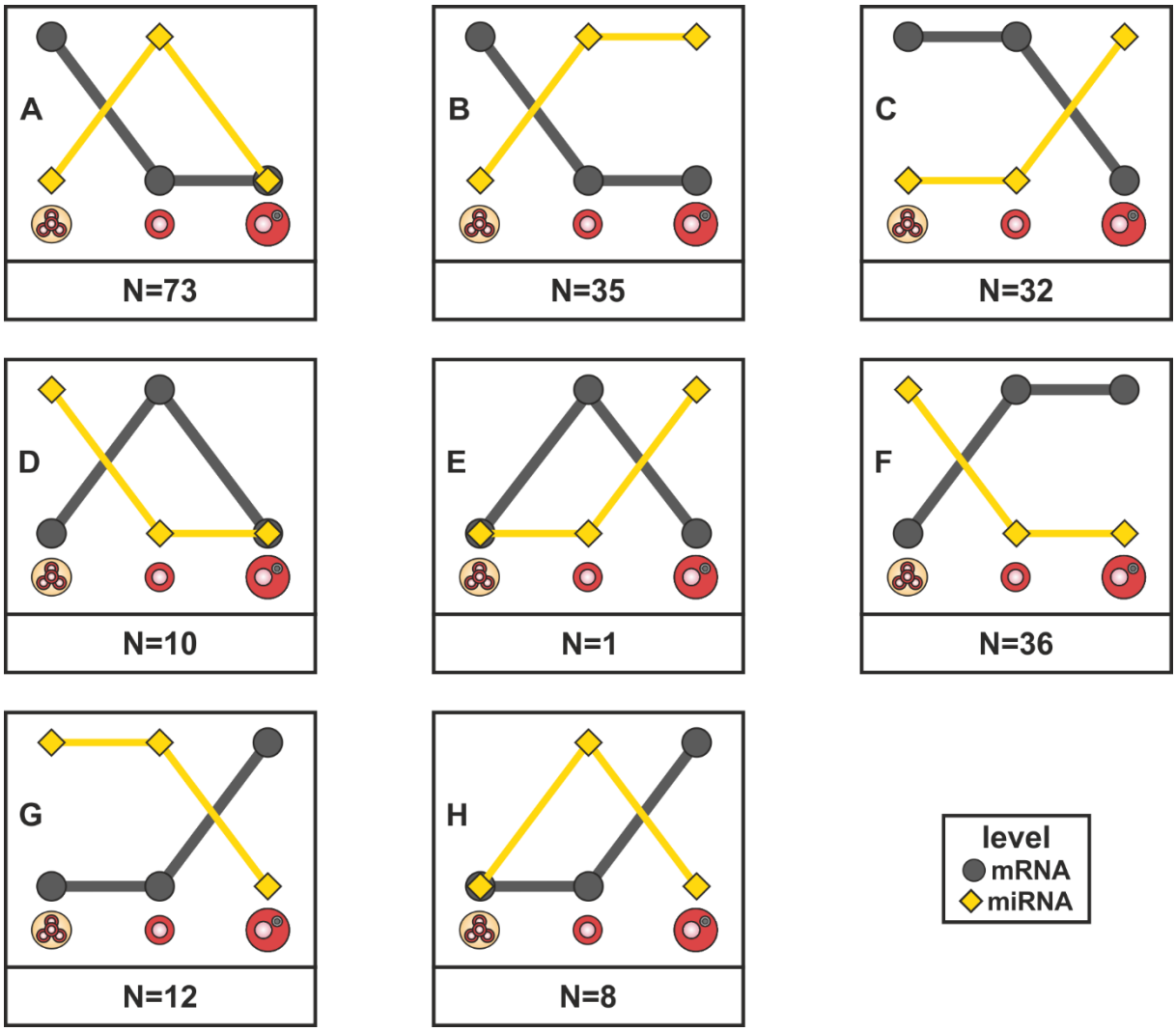


Figure 14: Developmentally relevant MTIs
Shown are eight categories (A to H) of developmentally relevant MTIs, consisting of pairs of stage-accumulated miRNAs and stage-accumulated mRNAs. The information in which developmental stage or stages the mRNA and miRNA are accumulated is represented by simplified mRNA (grey circles) and miRNA levels (yellow diamonds) along the developmental stages. The criterion to define an MTI as developmentally relevant is an inverse behavior between the miRNA and mRNA level.

In total, 207 MTIs are distributed across the eight categories. To assess the impact of these MTIs, their mRNAs were functionally analyzed (Supplemental Table 13). The mRNAs of the MTIs encode for proteins of very different functionalities. For instance, one of the mRNAs in category A encodes a copper/zinc superoxide dismutase (CSD) , which is targeted by a miRNA (solyc-miR756) that was so far only known in potato (miRBase entry: stu-miR398b-3p). Another example is an MTI belonging to category F, which consists of solyc-miR658 and an mRNA encoding for an ATP-sulfurylase (ATP-S). So far, soylc-miR658 was not identified in the Solanaceae family but in species that are more distant (e.g. miRBase entry: tae-miR395b).

However, despite a functional diversity, 34 out of the 207 MTIs contain mRNAs encoding TFs out of 15 different families (Figure 15). Out of the different miRNAs targeting TFs, three have already been

identified in plants. One of them is solyc-miR664, which is already known in tomato (miRBase entry: sly-miR156d-5p). This miRNA targets five members of the SBP family, which are all SPL proteins. Interestingly, these proteins are also targeted by the so far unknown miRNA solyc-miR663, which belongs to category A. The second known miRNA is solyc-miR684, which is also known in tomato (miRBase entry: sly-miR160a). This miRNA belongs to category C and targets two mRNAs encoding for TFs of the auxin response factor (ARF) family. The same mRNAs are also targeted by a second miRNA of category C, which is the so far unknown solyc-mIR683. The last known miRNA is solyc-miR660, which was so far only described in soybean (miRBase entry: gma-miR172k). This miRNA targets, among other things, six mRNAs encoding for members of the ApetalA2 (AP2) family.

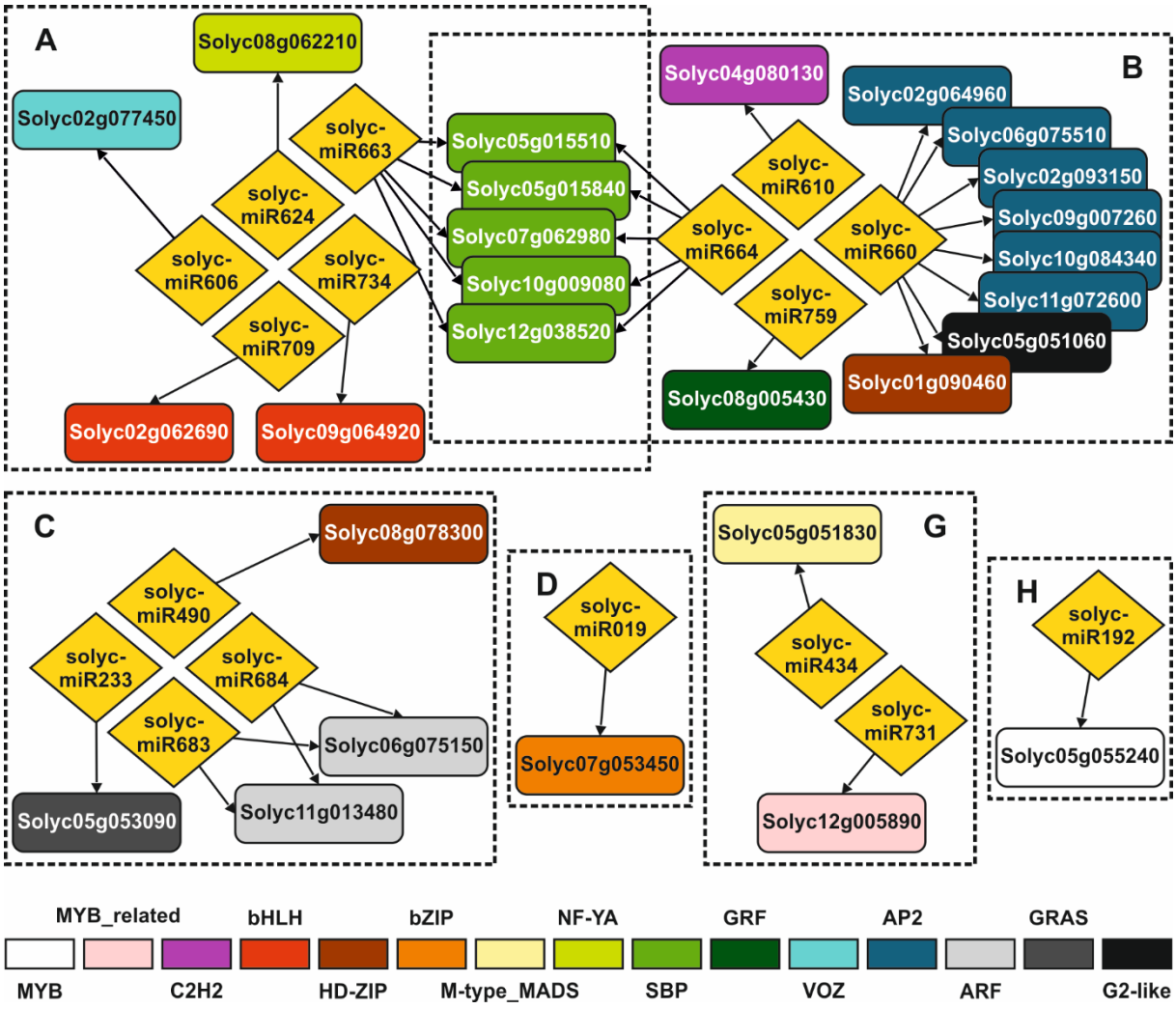


Figure 15: TFs regulated by miRNAs in a developmental context
Shown are developmentally relevant MTIs (category A, B, C, D, G and H) consisting of miRNAs (diamonds) and mRNAs that encode for TFs (rounded rectangles). Arrows indicate interactions and the color of the TFs represents the family affiliation.

The analysis of the miRNAs has so far revealed qualitative as well as quantitative differences between the developmental stages. Similar to the number of detected mRNAs also the number of detected miRNAs showed a decrease from tetrads towards mature pollen, which was also apparent in the number of stage-accumulated miRNAs. Further, the analysis of stage-accumulated miRNAs in combination with stage-accumulated mRNAs led to the identification of 207 developmentally relevant MTIs, of which 34 seem to influence the mRNA levels of TFs.

3.2.2 Detection of miRNAs involved in the heat stress response of pollen

After the identification of miRNAs with a potential role in pollen development, the next step was the identification of miRNAs involved in the HSR of pollen. A first overview of the detected miRNAs between the non- and heat-stressed stages revealed 633 detected miRNAs in tetrads, 337 in post-meiotic and 156 in mature pollen (Figure 16). Unlike for the mRNAs, with 67% and 60% the majority of the miRNAs detected in tetrads and post-meiotic pollen, respectively, was either only detected in the non-stressed samples or only in the heat-stressed samples. For mature pollen, the percentage of miRNAs only detected in one of the samples was with 47% much lower than for the other two stages.

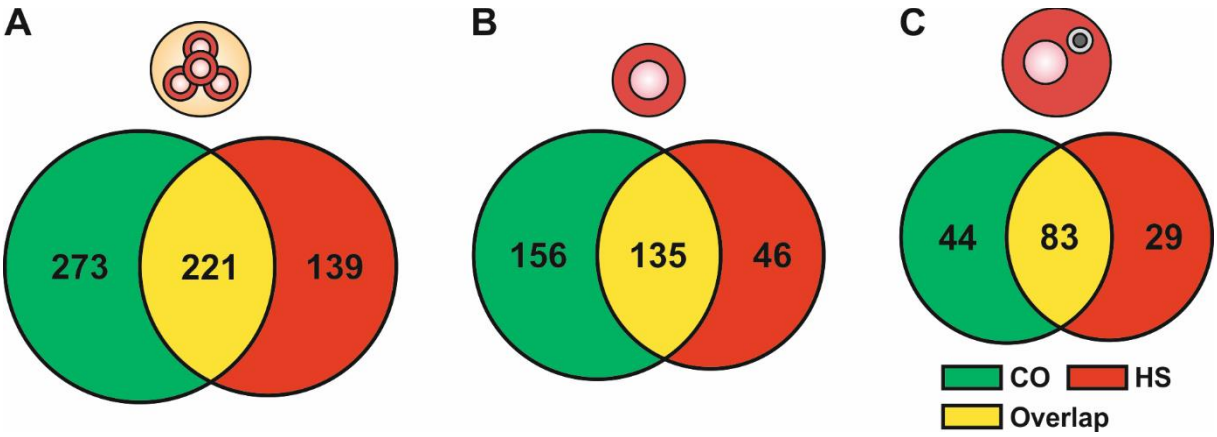


Figure 16: Detected miRNAs between non- and heat-stressed pollen stages
Venn diagrams indicate the number of detected miRNAs between non- (CO) and heat-stressed (HS) tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. miRNAs were either solely detected under CO (green), after HS (red) or under both conditions (Overlap, yellow).

A subsequent clustering of the miRNAs, based on their relative abundance in the non- and heat-stressed samples, led to seven clusters for each developmental stage (Supplemental Figure 4 to Supplemental Figure 6). Out of these clusters for each developmental stage one cluster comprised downregulated miRNAs (c1 clusters) and one cluster upregulated miRNAs (c7 clusters). Interestingly, the downregulated miRNAs of each stage are exactly the miRNAs only detectable in the non-stressed sample, while the upregulated miRNAs are exactly those only detectable in the heat-stressed

samples. The miRNAs of the remaining clusters (c2 to c6) showed no or only minor changes in response to HS.

Based on the known MTIs, the target mRNAs of the down- and upregulated miRNAs were searched for those also differentially regulated in response to HS (Supplemental Figure 7). Next to MTIs, whose miRNAs and mRNAs are regulated the same, also 24 MTIs comprising miRNAs and mRNAs with an opposite regulation were identified. 17 of these 24 MTIs are repressed by HS, which leads to an upregulation of the mRNA (Table 3). One of the HS-relevant MTIs in tetrads contains the unknown solyc-miR708 and an mRNA encoding for an activator of 90 kDA heat shock ATPase (Aha1). In post-meiotic pollen, one of the HS-relevant MTIs contains solyc-miR678, which is already known in the Solanaceae family (e.g. miRBase entry: nta-miR408), and an mRNA encoding for a basic blue protein (BBP). One of the HS-relevant MTIs in mature pollen contains the unknown solyc-miR461, which targets an mRNA encoding for a BAG family molecular chaperone regulator 6 (BAG6). Additionally, one of the HS-relevant MTIs was detected in post-meiotic and in mature pollen. This MTI contains the unknown solyc-miR749, which targets an mRNA encoding for a cytochrome b561-related family protein (CYB561). Next to MTIs repressed in response to HS, also HS-induced MTIs were identified, which are leading to a downregulation of the mRNAs. One example for such an MTI is the interaction of solyc-miR544 and an mRNA encoding profilin in post-meiotic pollen. In response to HS, solyc-miR544 is upregulated, which leads to the downregulation of the profilin mRNA. Interestingly, four additional MTIs in post-meiotic pollen also include solyc-miR544. The mRNAs of these MTIs encode, among other things, for a peptidyl-prolyl cis-trans isomerase and a LEA protein.

In contrast to 207 developmentally relevant MTIs, only 24 HS-relevant MTIs could be identified, which suggests that miRNAs rather play a role in pollen development than in the HSR of pollen. Nevertheless, the 24 MTIs might play a crucial for the fine-tuning of the pollen HSR, whereby most of them are active in post-meiotic pollen. Further, 75% (18 MTIs) of the HS-relevant MTIs lead to an upregulation of mRNAs, whereby the mRNAs encode for other proteins than the typical Hsfs or Hsps, which makes them interesting candidates for further analyses.

Table 3: HS-relevant MTIs with opposite regulation between miRNA and target mRNA

Listed are MTIs consisting of miRNAs (column 3) and mRNAs (column 5) that are down- (↓) or upregulated (↑) in response to HS (column 2 and 4) in the different developmental stages (column 1). For each mRNA the second hierarchy level MapMan term (column 6) and the functional description (ITAG3.2; column 7) is given.

stage	miRNA		mRNA		MapMan term	description
	reg.	ID	reg.	ID		
tetrads	↓	solyc-miR544	↑	Solyc01g110630	hormone metabolism.auxin	small auxin up-regulated RNA9
	↓	solyc-miR708	↑	Solyc10g078930	not assigned.no ontology	activator of 90 kDa heat shock ATPase
	↓	solyc-miR544	↑	Solyc02g079790	not assigned.unknown	transmembrane protein
post-meiotic	↓	solyc-miR749	↑	Solyc02g070680	redox.ascorbate and glutathione	cytochrome b561-related family protein
	↓	solyc-miR678	↑	Solyc01g104400	misc.plastocyanin-like	basic blue protein
	↓	solyc-miR593	↑	Solyc11g006490	RNA.regulation of transcription	basic-leucine zipper (bZIP) transcription factor family protein
	↓	solyc-miR163	↑	Solyc01g028987	DNA.synthesis/chromatin structure	Dead box ATP-dependent RNA helicase, putative
	↓	solyc-miR257	↑	Solyc04g072360	protein.degradation	F-box/RNI-like superfamily protein
	↓	solyc-miR037	↑	Solyc04g076527	RNA.regulation of transcription	GATA transcription factor, putative
	↓	solyc-miR192	↑	Solyc01g009160	not assigned.no ontology	harpin-induced1-like
	↓	solyc-miR776	↑	Solyc08g068770	misc.GCN5-related N-acetyltransferase	N-hydroxycinnamoyl-CoA:tyramine N-hydroxycinnamoyl transferase
	↓	solyc-miR490	↑	Solyc02g079380	not assigned.no ontology	pentatricopeptide repeat-containing protein
	↓	solyc-miR792	↑	Solyc02g092420	not assigned.no ontology	phototropic-responsive NPH3 family protein
	↓	solyc-miR792	↑	Solyc04g082510	protein.postranslational modification	protein kinase
	↓	solyc-miR490	↑	Solyc09g019970	protein.degradation	ubiquitin carboxyl-terminal hydrolase
	↑	solyc-miR677	↓	Solyc04g056560	protein.postranslational modification	protein phosphatase-2c
	↑	solyc-miR544	↓	Solyc12g044630	cell.organisation	profilin
	↑	solyc-miR544	↓	Solyc01g097960	development.unspecified	late embryogenesis abundant D-like protein
	↑	solyc-miR544	↓	Solyc07g006040	not assigned.no ontology	DNA-directed RNA polymerase subunit beta
	↑	solyc-miR544	↓	Solyc02g083440	not assigned.unknown	F-box
	↑	solyc-miR544	↓	Solyc02g086910	protein.folding	peptidyl-prolyl cis-trans isomerase
mature	↓	solyc-miR749	↑	Solyc02g070680	redox.ascorbate and glutathione	cytochrome b561-related family protein
	↓	solyc-miR461	↑	Solyc01g095320	stress.abiotic	BAG family molecular chaperone regulator 6
	↑	solyc-miR522	↓	Solyc02g075620	cell wall.pectin*esterases	pectinesterase

3.3 Regulation of proteins during the development and heat stress response of pollen

The analysis of the development and HSR of pollen was so far only examined at the transcriptome level. Therefore, to identify similarities and differences between the regulation of pollen mRNAs and proteins, LC-MS/MS libraries were constructed to identify and quantify proteins isolated from non- and heat-stressed pollen stages with MaxQuant. The identified proteins are reported in so-called protein groups, which comprise proteins that are not distinguishable by detected peptides. Further, for each protein group so-called LFQ intensities are reported, which are a measure of relative protein abundance and allow a comparison of protein levels between the libraries.

Before further downstream analyses, the composition of the protein groups was analyzed to check how many groups contain more than one protein (multi-protein groups) and if the multi-protein groups contain proteins, which are likely to be false positives. A first look on the size of the protein groups revealed that, with 879, more than 77% of the 1,136 detected protein groups contain only a single protein (Figure 17 pie chart). Out of the remaining 257 multi-protein groups the majority comprises two proteins and only 33 more than three proteins. The problem of the 257 multi-protein groups is that some of the proteins of these groups might not have been present and were detected by peptides of other proteins in the group. In this context, MaxQuant reports for each protein group so-called majority proteins, which are those proteins that have at least half of the peptides that the leading protein has. The removal of all proteins that are not majority proteins from the 257 multi-protein groups seems to be an appropriate way to reduce the size of these protein groups and remove false positive proteins. To verify this procedure, the mRNA levels of the proteins should serve as an estimator about the reliability that the proteins were really detected in one of the libraries. For this purpose, the proteins of the multi-protein groups were classified as either majority (424 proteins) or minority proteins (272 proteins) with respect to the number of peptides by which they were identified. Afterwards for each majority and minority protein the average mRNA level across all analyzed MACE libraries was determined, which should provide information on the extent to which the mRNAs are on average present in non- and heat-stressed pollen stages. (Figure 17 bar chart). Interestingly, 88 out of the 272 minority proteins were not detected at the mRNA level in any of the analyzed MACE libraries, while with 37 out of 424 this fraction is much lower for the majority proteins. Further, the mRNA level of the minority proteins is mainly in a range between 2 and 16 TPM, while the average mRNA level of majority proteins is mainly greater than 16 TPM with a peak at 128 TPM. The not existing or rather low

mRNA levels of the minority proteins and the high mRNA levels of the majority proteins justify the exclusion of minority proteins from further analyses as they are likely to introduce bias in the analyses.

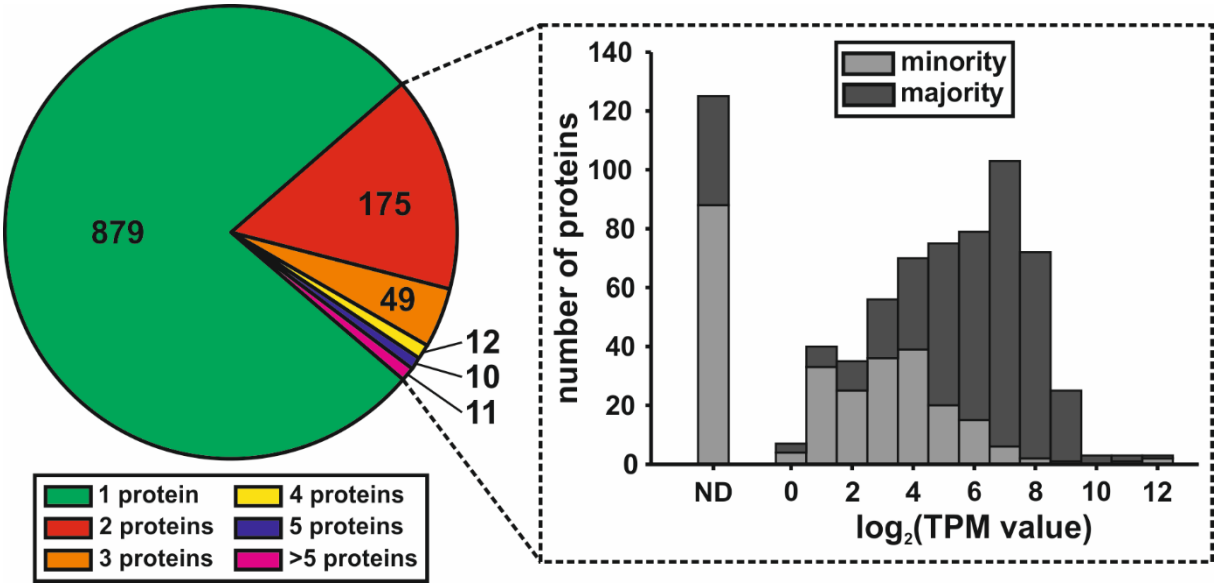


Figure 17: Protein group sizes and mRNA levels of majority and minority proteins
The pie chart indicates the number of protein groups with a certain number of proteins. Further, the bar charts shows the average mRNA level (TPM value) across all 18 MACE libraries for the majority (dark grey) and minority proteins (light grey) of multi-protein groups. Proteins with no detected mRNA in any MACE library are marked as not detected (ND).

For a first global overview about the relationship of the pollen mRNAs and proteins a correlation analysis was performed, whereby only proteins from single-protein groups (no other proteins in the group) were used. The correlation of transcriptomes and proteomes of the same non-stressed stage revealed only a low correlation for all three stages (Figure 18A to C). The highest correlation was observed for mature pollen with 0.46, followed by tetrads with 0.3 and post-meiotic pollen with 0.23. Further, to take into account a possible lagging of the protein levels behind the mRNA levels, the transcriptome of tetrads and post-meiotic pollen was also correlated with the proteome of later stages (Figure 18D). Interestingly, for tetrads and post-meiotic pollen the correlation coefficient increases when their transcriptomes are correlated with the proteome of the subsequent stage. For tetrads, the coefficient increased from 0.3 to 0.35 and for post-meiotic pollen it doubled from 0.23 to 0.49. The low correlation of the transcriptome and proteome of the same stage was also apparent for the heat-stressed stages (Supplemental Figure 8), whereby mature pollen shows with a coefficient of 0.44 the highest correlation, followed by tetrads with 0.37 and post-meiotic pollen with 0.29.

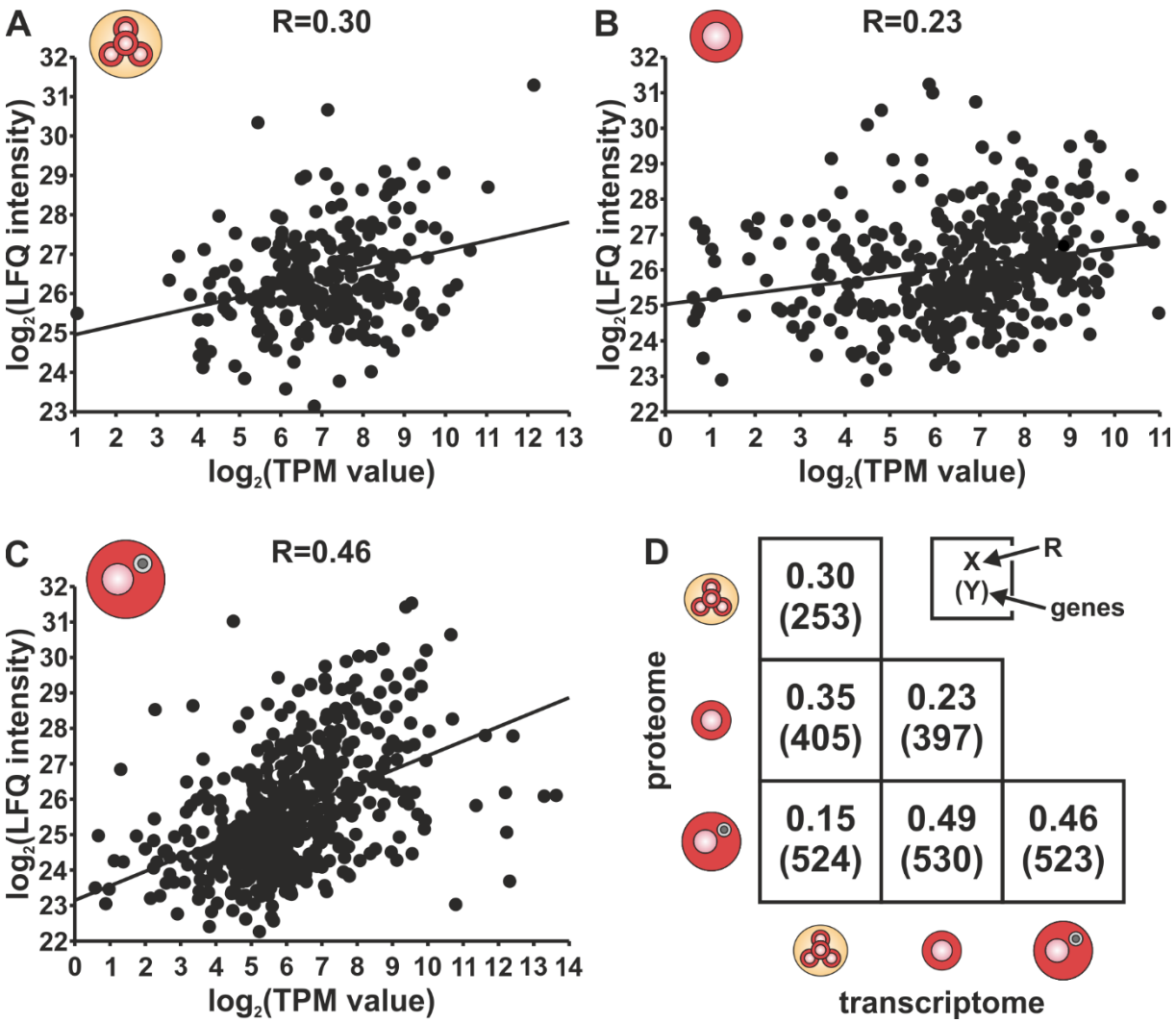


Figure 18: Correlation of non-stressed pollen transcriptomes and proteomes
(A, B and C) Shown are scatterplots indicating the correlation between mRNA (TPM value) and protein levels (LFQ intensity) in tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. Only genes with measured mRNA and measured protein were considered. In addition, the pearson correlation coefficient (R) is shown. **(D)** Correlation matrix between different combinations of transcriptomes and proteomes of pollen stages. The transcriptome of a developmental stage was only compared to its own proteome or the proteome of a later stage. The first value of a cell indicates the pearson correlation coefficient and the value in parentheses shows the number of genes used for the calculation.

So far, the comparison of the transcriptomes and proteomes of non- and heat-stressed pollen has revealed only a low correlation. This finding emphasizes need for a combined transcriptome and proteome analysis to obtain a more complete picture about the regulation of pollen development and the pollen HSR. Further, the removal of minority proteins from multi-protein groups has turned out as an appropriate way to improve the quality and reliability of the proteomic data.

3.3.1 Translational regulation of protein levels during pollen development

After first global insights in the correlation of the transcriptomes and proteomes, the regulation of mRNAs and their encoded proteins should be analyzed in a developmental context. A first glimpse in the number of detected protein groups between the non-stressed stages revealed the existence of 933 protein groups across all three stages, whereby with 218 groups around 23% of the groups are common to all stages (Figure 19A). Further, in contrast to the mRNAs and miRNAs, an increase in the protein diversity from tetrads (342 groups) towards mature pollen (706 groups) was observed.

In the next step, protein groups accumulated in one or two consecutive stages were determined among the 933 groups detected in developing pollen based on a differential regulation analysis, which resulted in five groups of stage-accumulated protein groups (Figure 19B).

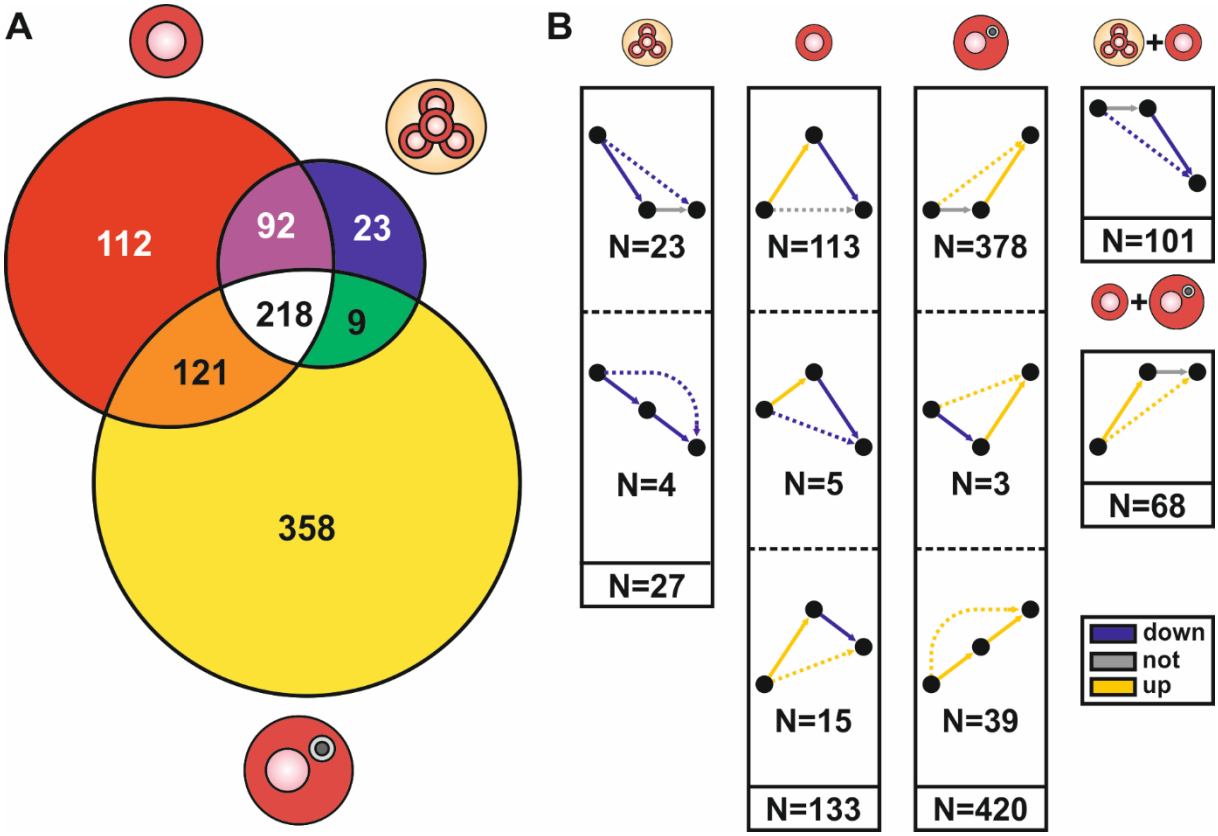


Figure 19: Detected and stage-accumulated protein groups in developing pollen
(A) Shown is a venn diagram indicating the overlap of detected protein groups between tetrads (blue), post-meiotic (red) and mature pollen (yellow). **(B)** Protein groups were categorized as either tetrads, post-meiotic, mature, tetrads + post-meiotic or post-meiotic + mature accumulated. Categorization was based on information about downregulation (down; blue arrows), no regulation (not; grey arrows) and upregulation (up; yellow arrows) between stages. The solid arrow indicates the regulation between two adjacent developmental stages and the dashed arrow the regulation between tetrads and mature pollen.

The highest number of stage-accumulated protein groups was observed for mature pollen with 420 groups, followed by post-meiotic pollen with 133 and tetrads with only 27 groups. Further, there are 101 protein groups that are accumulated in tetrads and post-meiotic pollen as well as 68 in post-

meiotic and mature pollen. After the identification of the stage-accumulated protein groups, the majority proteins of these protein groups were treated independently.

After the identification of the five groups of stage-accumulated proteins, these groups and the five groups of stage-accumulated mRNAs (Figure 5B) were used to analyze the translational regulation of proteins during pollen development, whereby only pairs of mRNAs and proteins were considered where the mRNA and the protein are stage-accumulated. The mRNA and protein of a pair can belong to one of the five groups of stage-accumulated mRNAs and five groups of stage-accumulated proteins, respectively, which in total leads to 25 possible combinations. Out of the 25 combinations, 15 combinations were used to analyze translational regulation during pollen development (Figure 20).

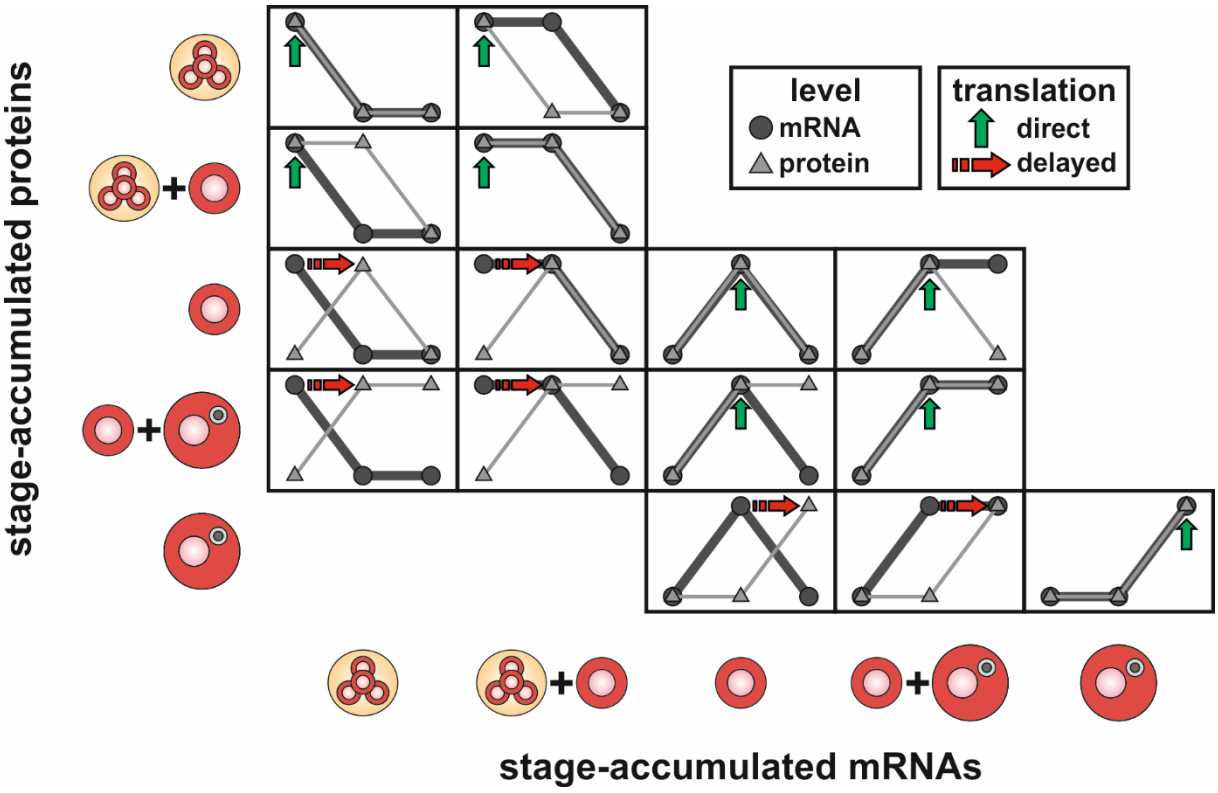


Figure 20: Two possible translation modes in developing pollen
Shown are simplified mRNA (dark grey circle) and protein level profiles (light grey triangle) of stage-accumulated mRNAs and proteins. Based on these profiles, mRNAs were classified as either being directly translated (green arrows) or showing a delay in their translation (red arrows).

The 15 combinations have in common that the accumulation of the proteins occurs in the same stage as the accumulation of the mRNAs or at maximum one stage after. Based on the 15 combinations, two translational modes active in developing pollen could be determined. The first mode is a direct translation and implies that the first stage with an accumulation is the same for the mRNA and the encoded protein (Figure 20 green arrows). The second mode is defined as a delayed translation, as the increase of the protein is one stage behind the stage where the mRNA is accumulated (Figure 20

red arrows), which indicates that the mRNA is translated during the transition from one stage to another. In total, nine of the 15 combinations can be assigned to a direct translation as well as six combinations to a delayed translation.

After the determination of the two translational modes, the number of mRNA and protein pairs belonging to each of the 15 combinations was determined (Figure 21). In total, there are 503 pairs of mRNAs and proteins distributed across the 15 combinations, out of which 262 show a direct translation and 241 a delayed translation. When further subdividing the pairs based on the stage in which the mRNA is accumulated for the first time, there are 112 pairs with a direct translation in tetrads (tetrads direct), 85 with a translation during the transition from tetrads to post-meiotic pollen (tetrads delay), 34 with a direct translation in post-meiotic pollen (post-meiotic direct), 156 with a translation during the transition from post-meiotic to mature pollen (post-meiotic delay) and 116 with a direct translation in mature pollen (mature direct).

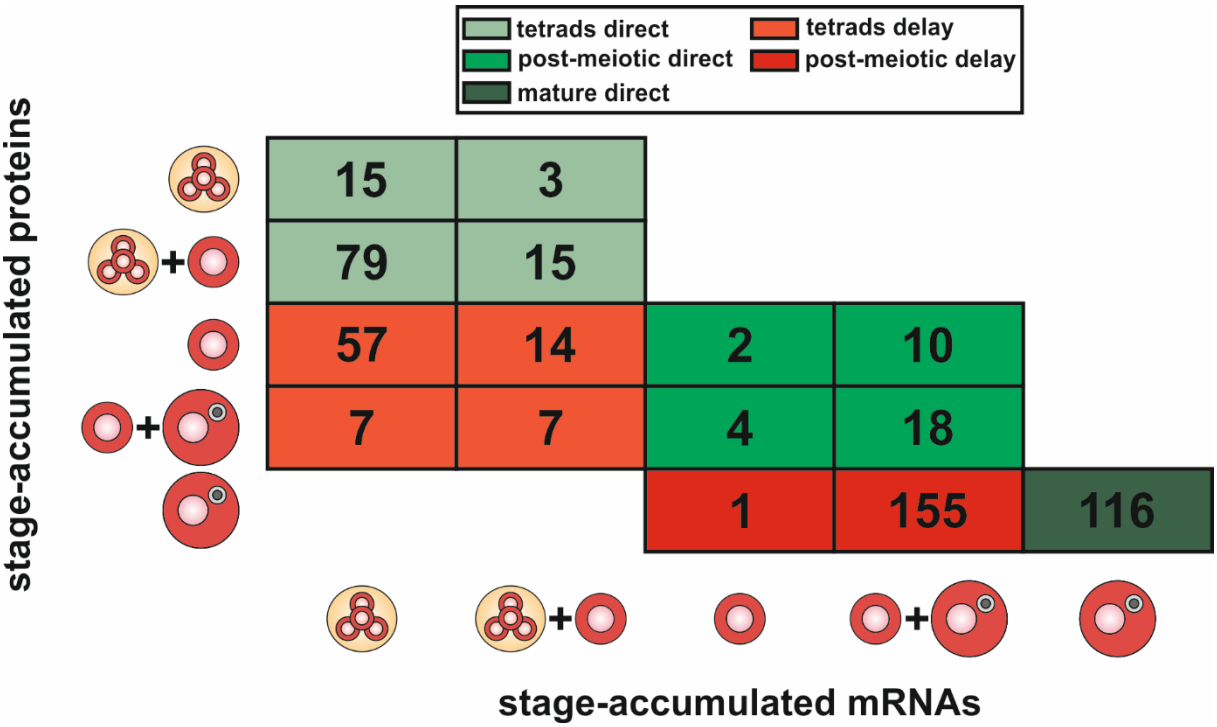


Figure 21: mRNAs with direct and delayed translation during pollen development
Shown is the number of pairs of mRNAs and proteins belonging to the 15 analyzed combinations. The combinations were further classified as showing direct (green) or delayed (translation) in tetrads (light colors), post-meiotic (medium colors) or mature pollen (dark color).

To understand the effect of the direct and delayed translation on biological processes, the proteins regulated by direct and delayed translation were functionally analyzed based on the second hierarchy terms of the MapMan ontology (Supplemental Table 14). With 15 proteins, many of the 112 proteins with direct translation in tetrads belong to ‘PS.lightreaction’, whereby 11 of them are chlorophyll a-b

binding proteins. In addition, a substantial amount of proteins with direct translation in tetrads belongs to ‘protein.synthesis’ and ‘DNA.synthesis/chromatin structure’. The respective proteins are two RPs of the small and nine RPs of the large subunit as well as ten members of the H4 histone family. A smaller amount of proteins belongs to ‘stress.abiotic’, including two members of Hsp90 and one member of the sHsp family. Similar to proteins with direct translation also many of the 85 proteins with delayed translation in tetrads belong to ‘protein.synthesis’. These proteins are two RPs of the small and three RPs of the large subunit as well as four eIFs or subunits of eIFs. Further, six proteins with delayed translation in tetrads belong to ‘DNA.synthesis/chromatin structure’, which are all members of the H2B histone family. A smaller number of proteins belongs to ‘misc.gluco-, galacto- and mannosidases’, which include two beta-glucosidases, one beta-galactosidase and a glycosyl hydrolase family protein. Regarding the 34 proteins with a direct translation in post-meiotic pollen, most of them do not share a functionality. However, three of the proteins belong to ‘TCA / org transformation.TCA’, whereby two of them are subunits of two different enzymes of the pyruvate dehydrogenase complex (E1 and E3), whereas one is a fumarase.

As the proteome of mature pollen is the best-studied pollen proteome (Zhang et al., 2017b), the proteins accumulated in mature pollen were analyzed in more detail with a focus on the 272 proteins either translated during the transition from post-meiotic to mature pollen (post-meiotic delay; 156 proteins) or directly in mature pollen (mature direct; 116 proteins) (Figure 22). Out of the 272 proteins, many proteins are involved in the synthesis of cell wall precursors. These proteins include a UDP-sugar pyrophosphorylase (USP), which is a product of a delayed translation, as well as two UDP-glucuronate 4-epimerases (GAEs) that are directly translated in mature pollen. The other proteins are four UDP-glucuronate decarboxylases (UXSs) and two MUR4s, whereby these proteins are either the product of a delayed or direct translation. In addition, to proteins involved in the synthesis of cell wall precursors, also cell wall proteins were identified. These proteins include two AGPs as well as a LRR, which are all directly translated in mature pollen. Also accumulated in mature pollen are many proteins involved in vesicle transport. These proteins include four COPI subunits (γ, δ, ε and ζ), the SNARE proteins YKT61 and VAP33-like and the SNARE-interacting protein KEULE, whereby, except for KEULE, all proteins are a product of a delayed translation. In addition, proteins required for the formation and modification of the cytoskeleton were identified, which are related to either the actin filaments or microtubules. The actin filaments-related proteins include five actins, a profilin, two ADFs as well as the actin-binding proteins fimbrin and villin, whereby all proteins, except for villin, are directly

translated in mature pollen. The microtubules-related proteins are two tubulin alpha chains and one tubulin beta chain, which are all directly translated. Also accumulated in mature pollen are four subunits of the catalytic V_1 complex (C, D, E and H) and subunit d of the membrane-embedded V_0 complex of the V-ATPase. The four subunits of the V_1 complex are all a product of a delayed translation, whereas the subunit of the V_0 complex is directly translated in mature pollen. Out of the remaining proteins accumulated in mature pollen, many are located in the mitochondrial matrix or the mitochondrial inner membrane and are involved in either the TCA cycle or subunits of the F-ATPase. The proteins involved in the TCA cycle include a beta subunit of the E1 component of the pyruvate dehydrogenase complex (PDH_{β}), a citrate synthase (CS), two regulatory subunits of the isocitrate dehydrogenase (IDH_{reg}), a beta subunit of the succinyl-CoA-synthase (SCS_{β}) as well as an iron-sulphur subunit of the succinate dehydrogenase ($SDH_{FE\text{-}S}$). Interestingly, although they participate in the same pathway, these proteins show a mixture of direct and delayed translation. In contrast the three subunits of the F_1 complex (β, δ and ε) of the F-ATPase and are all a product of a delayed translation.

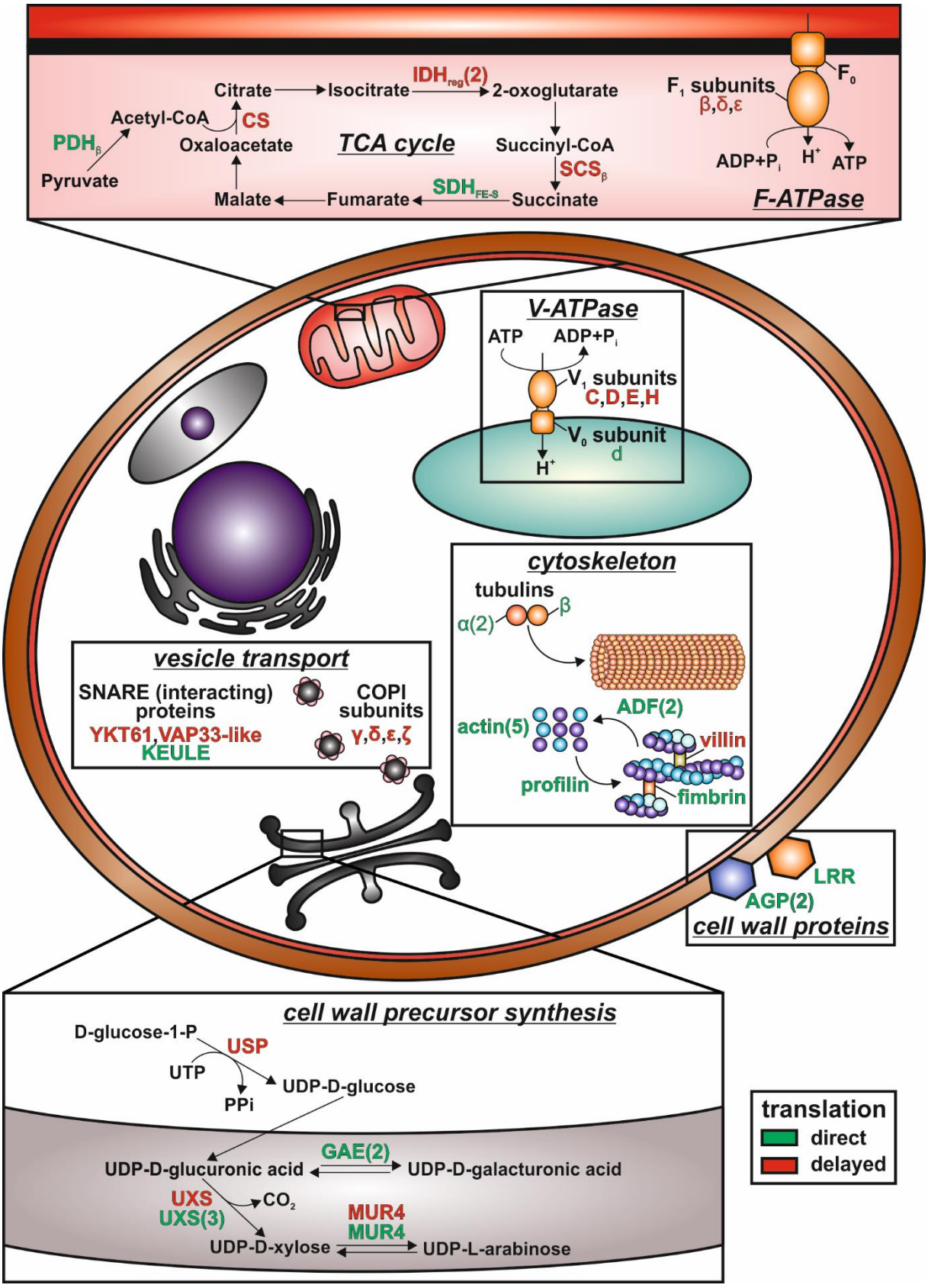


Figure 22: Proteins accumulated in mature pollen as a result of direct and delayed translation

Shown are proteins accumulated in mature pollen and are the product of a direct (green letters) or delayed translation (red letters). Redundancy of proteins is indicated in parentheses.

The combination of transcriptomic and proteomic data enabled the identification of the modes of direct and delayed translation. The two modes seem to be present throughout all pollen developmental stages and enable either a direct translation of synthesized mRNAs or a storage of the synthesized mRNAs for a translation in the transition from one stage to another. Further, the proteins controlled by the two modes were functionally analyzed, which enabled insights in the translational regulation of proteins accumulated in mature pollen.

3.3.2 Uncoupling of the transcriptome and proteome during the heat stress response of pollen

The comparison of the developmental transcriptome and proteome has revealed that the analysis of transcriptomic data is not sufficient to draw direct conclusions about the proteome. To examine if this also holds true for the HSR of pollen developmental stages, the regulation of mRNAs and proteins in response to HS was compared.

A first overview about the detected protein groups between non- and heat-stressed developmental stages showed that, similar to the mRNAs, the majority of protein groups detected for a developmental stage is shared between the non- and heat-stressed samples (Figure 23). However, the percentage of shared protein groups differs between the developmental stages. While in mature pollen 81.5% (628 of 771 protein groups) of the protein groups are shared between the non- and heat-stressed sample, these are only 61.4% in tetrads (297 of 484 protein groups) and 57.1% in post-meiotic pollen (400 of 700 protein groups). Further, the extent of protein groups exclusively detected in non- and heat-stressed samples differed between the stages. For tetrads, 142 protein groups are exclusively detected in the heat-stressed sample, which are three times more groups than exclusively detected in the non-stressed sample (45 protein groups). In contrast, for post-meiotic and mature pollen the number of protein groups exclusively detected in the non- or heat-stressed sample are more or less identical.

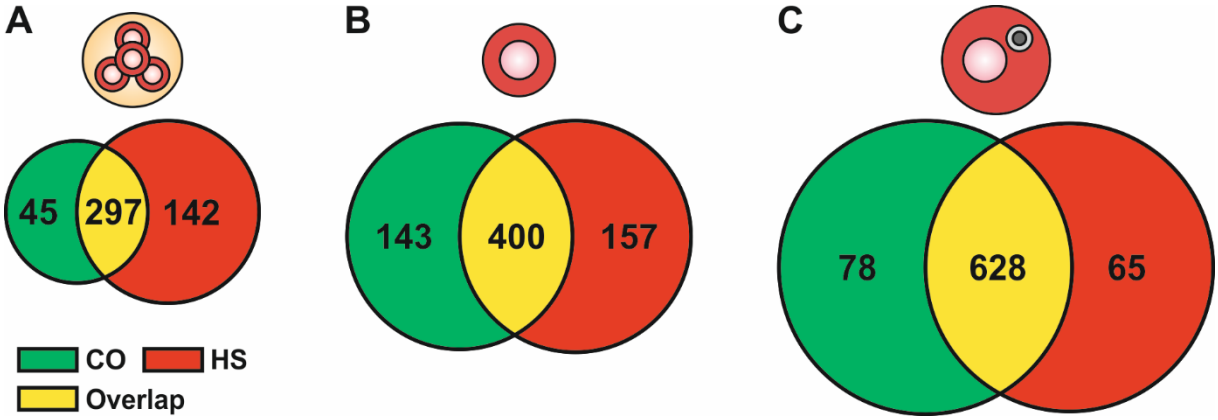


Figure 23: Detected protein groups between non- and heat-stressed pollen stages
Venn diagrams indicate the number of detected protein groups between non- (CO) and heat-stressed (HS) tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. Groups were either solely detected under CO (green), after HS (red) or under both conditions (Overlap, yellow).

After the qualitative overview, a differential regulation analysis was performed to identify those protein groups with significantly altered levels after HS. In total, 46 down- and 146 upregulated protein groups were detected in tetrads, 145 down- and 161 upregulated groups in post-meiotic pollen and 78 down- and 71 upregulated groups in mature pollen. Similar to the mRNAs, also the HS regulation of protein groups is highly stage-specific (Supplemental Figure 9). Out of the upregulated protein groups, no group was upregulated in all three stages. However, there are two protein groups downregulated in all three stages. The first protein group contains a single mitochondrial glycoprotein, while the second protein group contains proteins annotated as RPL8s.

Comparing the percentage of mRNAs and proteins differentially regulated in response to HS, it becomes apparent that HS has a much stronger impact on the proteome than on the transcriptome. For instance, in tetrads only 1% of the mRNAs are differentially regulated but 39.7% of the protein groups. Similar to this, also for post-meiotic and mature pollen a stronger effect on the proteome was observed, whereby the effect in post-meiotic pollen was with 43.7% of differentially regulated protein groups (5.7% differentially regulated mRNAs) much higher than in mature pollen with 19.3% (4.5% differentially regulated mRNAs). To compare differentially regulated proteins and their coding mRNAs in more detail, the majority proteins of down- and upregulated protein groups were treated independently in the following.

Overall, the comparison of the regulation of differentially regulated proteins and their mRNAs revealed only a very low degree of co-regulation and also inverse regulation between proteins and their mRNAs (Figure 24). Most of the co-regulated pairs of mRNAs and proteins are upregulated, whereby eight are upregulated in tetrads, 13 in post-meiotic and two in mature pollen. Interestingly, these upregulated

pairs include multiple members of Hsp families (Supplemental Table 15). In tetrads, five sHsps and one Hsp100 are co-upregulated, while it are six sHsps and one Hsp40 in post-meiotic pollen and a single Hsp10 in mature pollen. Regarding a co-downregulation, only one pair each was co-downregulated in post-meiotic and mature pollen. The protein downregulated in post-meiotic pollen is an adenosylhomocysteinase and the protein downregulated in mature pollen an S-adenosyl-L-methionine-dependent methyltransferases superfamily protein.

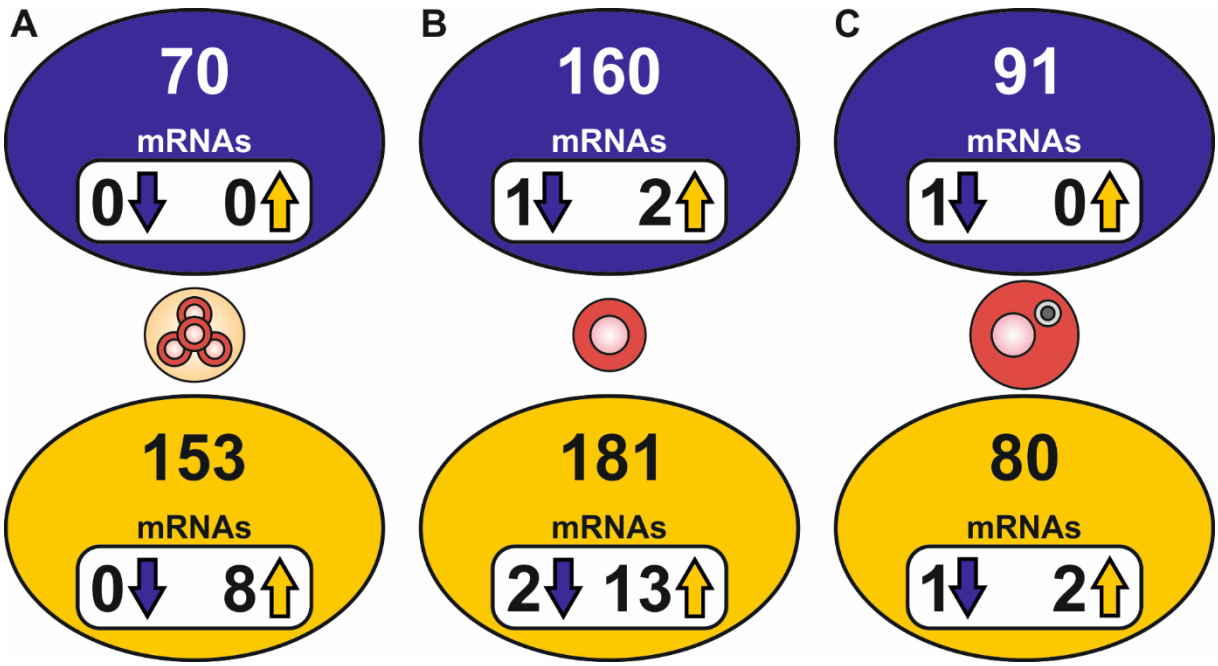


Figure 24: HS regulation of mRNAs encoding down- and upregulated proteins
Shown is the HS regulation of mRNAs encoding proteins that are down- (blue ellipses) or upregulated (yellow ellipses) in response to HS in tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. The number of down- and upregulated proteins is shown in the upper part of the ellipses and the number of down- (blue arrow) and upregulated mRNAs (yellow arrow) in the white rectangle.

A closer look on the proteins differentially regulated independently of their mRNAs revealed for all three stages that many of these proteins are RPs (Figure 25; Supplemental Table 16). In total, 57 RPs are differentially regulated in at least one stage, whereby the highest number of differentially regulated RPs was observed for post-meiotic pollen (30 regulated RPs), followed by tetrads (21 regulated RPs) and mature pollen (15 regulated RPs). Out of the 57 RPs, 21 are part of the small subunit and 36 of the large subunit.

The regulation of the RPs seems to be highly stage-specific as 50 out of the 57 RPs are differentially regulated in only one stage. The RPs that are downregulated in two or three stages, include one RPS8, two RPL8, one RPL21 as well as one RPL18, while the RPs that are upregulated in two or three stages include one RPL5 and one RPL28. Interestingly, for four proteins of the RPS8 family a completely different regulation was observed. While one of the RPS8s is downregulated in tetrads and mature pollen, in each developmental stage one of the remaining RPS8s is upregulated.

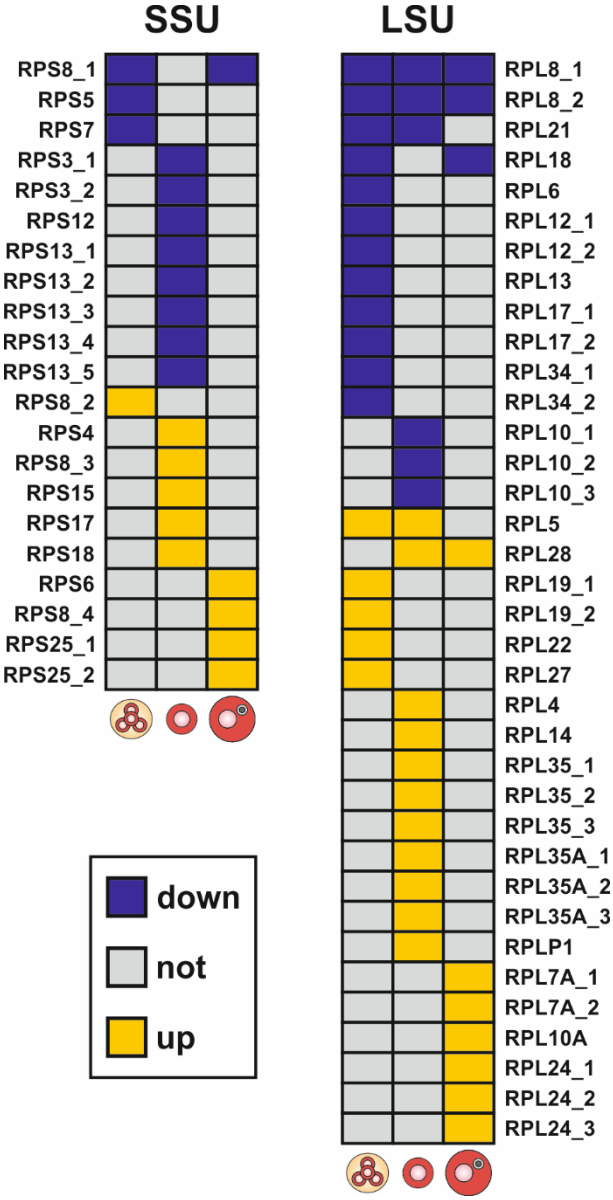


Figure 25: Alteration of RP levels in response to HS
Shown are RPs of the small (SSU) and large subunit (LSU) differentially regulated in at least one pollen developmental stage. For each RP the family as well its regulation in tetrads (upper row), post-meiotic (middle row) and mature pollen (lower row) is indicated. RPs are either downregulated (blue), not regulated (grey) or upregulated (yellow) in a developmental stage.

The analysis of the heat-stressed pollen proteomes, revealed a much stronger effect of HS on pollen proteomes than one pollen transcriptomes. For instance, in post-meiotic pollen 43.7% of the observed protein groups had altered levels after the application of HS, while this was only observed for 5.7% of the mRNAs. When further comparing the regulation of differentially regulated proteins and their mRNAs, only a small percentage of pairs of proteins and mRNAs show a co-regulation. This finding, suggests that the HS regulation of proteins is only to a small extent dependent on the regulation of their mRNAs. The subsequent analysis of proteins that are mRNA-independently regulated revealed that many of these proteins are RPs, whereby their regulation is highly stage-specific.

4 Discussion

In this study, transcriptomic and proteomic approaches were combined to gain new molecular biological insights about the development and HSR of pollen. The analysis of the transcriptome and proteome revealed a shared increase in the mRNA and miRNA diversity from tetrads to mature pollen, while the protein showed the opposite effect. Further, mRNAs accumulated in certain developmental stages enabled the identification of functional processes with a potential role in pollen development. The regulation of these mRNAs takes in part place via miRNAs, which have a preference for the regulation of mRNAs encoding for TFs. In contrast, the regulation of the encoded proteins occurs via two translational modes, which either lead to a direct or delayed translation of the mRNAs. A comparison of the HSR of the different stages led, among other things, to the identification of a pollen HSR core set consisting of 49 mRNAs. In addition, a small set of miRNAs involved in the regulation of HS-responsive mRNAs was identified. The analysis of the HSR of the pollen proteomes showed that the regulation of the proteome is largely independent of the transcriptome and also much more pronounced. The response of the proteomes was especially apparent for RPs, which showed a mainly stage-specific down- and upregulation.

4.1 New insights in the development of pollen by transcriptome and proteome analyses

4.1.1 Diversity of mRNAs, miRNAs and proteins in developing pollen

The initial PCA analysis of the MACE libraries revealed that the mRNA transcriptome of developing pollen is highly stage-specific, as there is a clear separation of the developmental stages by the first principal component (Figure 4A). In total, about 19,000 different mRNAs were identified in developing pollen (Figure 5A), which is much higher than the 13,977 expressed genes during the development of *A. thaliana* pollen (Honys and Twell, 2004). One reason for this difference is the inclusion of the tetrad stage in this study, which was not analyzed in *A. thaliana* but had roughly 3,900 mRNAs not detected in the other stages. A second reason for the lower number of expressed genes in *A. thaliana* might be the use of microarrays in the *A. thaliana* study, which are known to have a lower sensitivity and lack probes for some of the mRNAs (Wang et al., 2009; Malone and Oliver, 2011). A comparison of the number of different mRNAs detected in a stage revealed a decrease in the mRNA diversity from early to late developmental stages (Figure 5A), which seems to be a general pollen phenomenon, as it has also been described in other species, like *A. thaliana* and rice (Honys and Twell, 2004; Wei et al., 2010). Similarly, also for miRNAs a decrease in the diversity was observed (Figure 12), which

supports the idea that the repertoire of miRNAs scales with the size of the transcriptome that is to be regulated (Lenz et al., 2011). In contrast, the protein diversity increases from tetrads towards mature pollen (Figure 19A), what has already been described for tomato (Chaturvedi et al., 2013).

4.1.2 Functional processes required during pollen development

Next to dynamic changes of the mRNA diversity, it has also been shown that mRNAs accumulated in one or two consecutive stages (Figure 5B) encode for proteins of specific functional processes (Figure 6). The identified processes and protein families provide new insights into the development of pollen and the molecular requirements for proper pollen development. The identification of functional processes with an important role in pollen developmental stages was in most studies based on the identification of an interesting set of genes (e.g. accumulated mRNAs or proteins in a developmental stage), which was afterwards functionally categorized, followed by the summation of all genes assigned to a functional category (Honys and Twell, 2003; Honys and Twell, 2004; Dai et al., 2006; Chaturvedi et al., 2013; Ischebeck et al., 2014). In contrast, a functional enrichment analysis on a set of genes enables the identification of statistically overrepresented functional processes, which increases the likelihood of identifying processes that are most relevant (Huang et al., 2009; Tipney and Hunter, 2010).

Functional processes in tetrads – Based on the enrichment analysis of mRNAs accumulated in tetrads (Figure 6 I) a potential role of brassinosteroids, SPLs and the abiotic stress response in the early developmental phase of pollen was determined. Brassinosteroids are phytohormones that play an important role in different plant developmental processes, including root, fruit and flower development (Bao et al., 2004; Fu et al., 2008; Gonzalez-Garcia et al., 2011; Manzano et al., 2011). Further, brassinosteroids were also shown to be of importance for the development of pollen, as the mutation of brassinosteroid synthesis and signaling genes led to a reduced pollen number, viability and release (Ye et al., 2010). Recently, brassinosteroids were also shown to promote pollen germination and pollen tube growth (Vogler et al., 2014), while the results of this study suggest an additional role in the early developmental phase of pollen. Similar to brassinosteroids, also SPLs are important key players for plant growth and developmental processes, including leaf, flower, trichome and fruit development (Chen et al., 2010; Preston and Hileman, 2013). In addition, it was shown that members of the SPL family are required for proper pollen sac development and pollen production (Unte et al., 2003; Wang et al., 2016). Although, it is known that SPLs are of importance for the development of pollen, little is known about their regulation in developing pollen. According to the results of this study, the mRNAs of

eight of the 15 SPLs detected in the genome of tomato (Supplemental Table 1) are accumulated in tetrads, which indicates a role of SPLs in the early developmental phase of pollen. The enrichment of the abiotic stress response in tetrads was mainly based on the accumulation of HS-responsive mRNAs like those of Hsfs and Hsps. The accumulation of stress-responsive proteins during the early pollen development was termed developmental priming and is thought to protect early developmental stages against sudden stresses (Chaturvedi et al., 2013; Zhang et al., 2017b). So far, developmental priming was only described at the proteome level, like for example in tomato and tobacco, where members of the sHsp, Hsp70 and Hsp90 family accumulated in early developmental stages (Chaturvedi et al., 2013; Ischebeck et al., 2014). The observation of accumulated Hsf and Hsp mRNAs in tetrads indicates that developmental priming is also regulated at the transcriptome level and, next to Hsps, also relies on Hsfs. Further, some of the Hsf and Hsp mRNAs accumulated in tetrads are upregulated in post-meiotic and/or mature pollen in response to HS but lack an upregulation in tetrads (Figure 10). Despite a missing upregulation in tetrads, the levels of these mRNAs in the heat-stressed samples are comparable between tetrads and the other two stages, which further supports the idea of a developmental priming of tetrads with HS-responsive Hsf and Hsp mRNAs.

Functional processes in tetrads and post-meiotic pollen – The enrichment analysis of mRNAs accumulated in tetrads and post-meiotic pollen indicates a preparation of a higher translational activity in these two stages, as the two enriched terms ‘protein.synthesis’ and ‘protein.aa activation’ (Figure 6 II) comprise key components of the translation machinery, such as RPs, eIFs and aminoacyl-tRNA synthetases. The observed accumulation of mRNAs encoding such translational components has also been described in other species, like tobacco (Bokvaj et al., 2015; Hafidh et al., 2018) and *A*. *thaliana* (Honys and Twell, 2004). Further, knockout of RPs, eIFs and aminoacyl-tRNA synthetases may lead to defects in pollen development, germination or pollen tube growth (Berg et al., 2005; Xia et al., 2010; Yan et al., 2016), which emphasizes the developmental importance of these protein families.

Functional processes in post-meiotic and mature pollen – Within the group of mRNAs accumulated in post-meiotic and mature pollen an enrichment of V-ATPases and processes required for aerobic respiration, such as glycolysis, the TCA cycle and the mitochondrial electron transport chain, was observed (Figure 6 IV). The accumulated mRNAs belonging to V-ATPases encode, among other things, for the subunits A and E of the catalytic V_1 complex. In *A*. *thaliana* it was shown that a T-DNA insertion allele of the single copy gene encoding subunit A (VHA-A; AT1G78900) leads to complete male gametophytic lethality (Dettmer et al., 2005). Further, subunit E of the V_1 complex is in *A*.

thaliana encoded by the three gene copies VHA-E1 (AT4G11150), VHA-E2 (AT3G08560) and VHA-E3 (AT1G64200), whereby VHA-E2 was proposed to be pollen-specific (Strompen et al., 2005). According to the ortholog prediction of this study, two of the mRNAs accumulated in post-meiotic and mature pollen are orthologs of the *A*. *thaliana* E subunits, whereby one of them (Solyc09g048990) is an ortholog of the pollen-specific VHA-E2. In addition, the tomato ortholog of VHA-E2 was, together with four other V-ATPase subunits, accumulated in the proteome of mature pollen (Figure 22). These results highlight that V-ATPases play a role in the late phase of pollen development, whereby the accumulation of the mRNAs is initiated in post-meiotic pollen. The observed enrichment of aerobic respiration indicates a preparation for the upcoming energy demand during germination and pollen tube growth (Selinski and Scheibe, 2014). So far, the preparation for a high-energy demand was only described for the proteome of mature pollen, where an accumulation of proteins involved in energy production was observed (Dai et al., 2006; Grobei et al., 2009; Chaturvedi et al., 2013). In contrast to the general assumption that energy metabolism is under-represented in transcriptomic studies of mature pollen (Honys and Twell, 2003; Chaturvedi et al., 2013), the results of this study clearly show an accumulation of energy-related mRNAs in post-meiotic and mature pollen. Further, also in this study energy-related proteins were detected in the proteome of mature pollen (Figure 22; Supplemental Table 14), which further supports the idea that pollen germination depends on presynthesized proteins stored in mature pollen (Chaturvedi et al., 2013).

Functional processes in mature pollen – Terms enriched in the group of mRNAs accumulated in mature pollen, included multiple terms related to the cell wall, such as cell wall precursor synthesis and pectin esterases, as well as the term ‘cell.organisation’ (Figure 6 V), which comprises mRNAs encoding for proteins related to actin filaments and microtubules. Many of the proteins encoded by the mRNAs belonging to the mentioned terms are also accumulated in the proteome of mature pollen (Figure 22; Supplemental Table 14). The charging of mature pollen with mRNAs and proteins related to the cell wall and actin is a known phenomenon and thought to be a preparation for cell wall synthesis and the establishment of an actin-rich cytoskeleton during germination and pollen tube growth (Honys and Twell, 2003; Zou et al., 2009). Out of the accumulated mRNAs and proteins involved in cell wall precursor synthesis, two UXS and one MUR4 proteins had already been identified in the membrane proteome of tomato pollen (Paul et al., 2016), while one additional UXS and two GAE proteins were only identified in this study. GAE proteins are required for the synthesis of UDP-D-galacturonic acid, which is one of the activated precursors necessary for the synthesis of pectins

(Usadel et al., 2004; Bethke et al., 2016). Pectins are thought to be the principal components of the apical cell wall of growing pollen tubes (Bosch et al., 2005), which suggests an important role of the two identified GAEs in the preparation of mature pollen for the upcoming pollen tube growth. Also related to pectins are PMEs, out of which three have accumulated mRNAs and proteins in mature pollen. After the secretion of pectins into the apoplastic space, PMEs demethylesterifiy the pectins, which either leads to a loosening or stiffening of the wall in dependence on the apoplastic pH (Bosch et al., 2005; Tian et al., 2006). Mutation of VGD1, a PME in *A*. *thaliana*, resulted in unstable and poorly growing pollen tubes, which highlights the importance of PMEs during pollen tube growth (Jiang et al., 2005)). The observed accumulation of mRNAs and proteins, required for the formation and remodeling of actin filaments, was also described for mature *A*. *thaliana* pollen (Sheoran et al., 2006) and is thought to be a preparation for the upcoming germination and pollen tube growth (Honys and Twell, 2003; Zou et al., 2009). During pollen tube growth, the actin cytoskeleton appears in three distinct states, which are involved in important processes, such as cytoplasmic streaming, tip growth and vesicular transport of cell wall and membrane components (Qu et al., 2015; Zhu et al., 2017). The accumulated proteins related to actin filaments are nearly all a result of a direct translation and include actins as well as ADFs, profilin and fimbrin. The accumulation of these proteins is of major importance as the mutation of actin-related genes, such as ADFs or fimbrin, leads to a delay in pollen germination and inhibition or retardation of pollen tube growth (Wu et al., 2010; Zhu et al., 2017).

4.1.3 The regulatory role of miRNAs during pollen development

Within this study, it could be shown that the abundance of mRNAs undergoes dynamic changes during the development of pollen according to the special needs of the individual stage. Based on the combination of the mRNA and miRNA data, it has been shown that the changes in mRNA abundance are, among other things, regulated by miRNAs. Further, it has been shown that miRNAs also have an indirect effect on the expression of non-target genes via the regulation of mRNAs encoding for TFs.

In total, 641 miRNAs were detected in developing pollen (Figure 12), out of which 35 are known in tomato or other plant species (Supplemental Table 12). The observed number of miRNAs in developing tomato pollen is slightly higher than the 486 miRNAs detected in developing pollen of diploid and autotetraploid rice (Li et al., 2016b), whereby the analyzed developmental stages of the studies did not overlap completely. However, the observed 641 miRNAs are comparable to the 670 detected miRNAs in non-stressed leaves of wild tomato (Zhou et al., 2016). The percentage of known miRNAs in this study (35 of 641 miRNAs) is much lower than the percentage of known miRNAs in the

rice (192 of 486 miRNAs) and wild tomato study (576 of 670 miRNAs), which is likely due to the different criteria for the determination of known miRNAs. While in this study, miRNAs needed a perfect match to a mature miRNA of the miRBase, in the other studies also miRNAs with non-perfect matches to pre-miRNAs were considered as known.

The prediction of miRNA targets was performed with TargetFinder, which resulted in an average of 5.9 targets per miRNA. This number is comparable to the average of 4.8 miRNA targets in an *A*. *thaliana* study (Wang et al., 2004), whereby it has to be mentioned that the number of predicted miRNAs is strongly dependent on the prediction tool used. For instance, another *A*. *thaliana* study estimated the average number of miRNA targets to 31.1 (Lenz et al., 2011), which is more than six times higher than the previously reported 4.8 targets. This finding is likely due to the use of RNAhybrid as target prediction tool, which is mainly used in animals and in comparison to plant miRNA target prediction tools overestimates the number of targets in plants (Srivastava et al., 2014). The same study also showed that TargetFinder has the highest recall and precision rate among available plant miRNA target prediction tools, which supports the reliability of the target prediction in this study.

The connection of stage-accumulated mRNAs and miRNAs based on the predicted MTIs, led to the identification of 207 developmentally relevant MTIs in pollen, which were further assigned into eight categories based on the abundance profile of their mRNAs and miRNAs (Figure 14). One of the MTIs belonging to category A consists of solyc-miR756, which is a known miRNA of the miR398 family in potato (miRBase entry: stu-miR398b-3p), and an mRNA encoding for a CSD (Supplemental Table 13). The regulation of CSDs by miR398 was also shown in *A*. *thaliana* (Guan et al., 2013). Further, the interaction between miR398 and CSDs is part of various abiotic stress responses in *A*. *thaliana* (Sunkar et al., 2006; Yamasaki et al., 2007; Guan et al., 2013). For instance, under oxidative stress miR398 is downregulated, which leads to an accumulation of CSD1 and CSD2 mRNAs and an associated detoxification of superoxide radicals (Sunkar et al., 2006). Regarding the role of miR398 in the development of tomato pollen, the assignment of miR398 in category A indicates that pollen requires high levels of CSD mRNAs in the early developmental phase to protect tetrads against oxidative stress, whereas in post-meiotic and mature pollen the demand is diminished. Therefore, to counteract the high levels of CSD mRNAs in tetrads, miR398 is accumulated during the transition from tetrads to post-meiotic pollen, leading to a CSD mRNA reduction (Figure 14 category A). A second interesting MTI was identified in category F and is made up by solyc-miR658 and an mRNA encoding for an ATP-S. So far, solyc-miR658 was only identified in distant species of the Viridiplantae kingdom

(e.g. miRBase entry: tae-miR395b) where it is part of the miR395 family. The observed ATP-S catalyzes the first step of primary S-assimilation by activating sulfate (Anjum et al., 2015). Regulation of ATP-S by miR395 was also described in *A*. *thaliana*, where accumulation of miR395 led to a reduction of ATP-S mRNAs and as a result to an accumulation of sulfate (Davidian and Kopriva, 2010; Liang et al., 2010). According to its assignment to category F, the accumulation of miR395 in tetrads should lead to an accumulation of sulfate, which would be accompanied by a decrease of reduced sulfur. Upon the transition from tetrads to post-meiotic pollen, S-assimilation should be induced by the downregulation of miR395 and the associated accumulation of ATP-S mRNAs (Figure 14 category A). The idea of an impaired S-assimilation in tetrads is supported by a second MTI from category F. This MTI consists of a novel miRNA (solyc-miR237) and an mRNA encoding a sulfate transporter. The suppression of the sulfate transporter in tetrads is likely leading to a diminished sulfate uptake and thus also to a reduced S-assimilation. These findings support the idea that the regulation of S-assimilation is of high importance for the development of pollen (Birke et al., 2013).

In addition to the three mentioned MTIs, 34 of the remaining developmentally relevant MTIs have been shown to affect the mRNA levels of TFs (Figure 15). The regulation of TFs by miRNAs is a known phenomenon in plants and of great importance for plant growth and development (Li and Zhang, 2016; Samad et al., 2017). In total, 17 miRNAs with an effect on TFs were identified, out of which three are already known. The two miRNAs solyc-miR664 and solyc-miR684 have already been identified in tomato, where they belong to the miR156 (miRBase entry: sly-miR156d-5p) and miR160 family (miRBase entry: sly-miR160a), respectively. The third known miRNA is solyc-miR660, which has a perfect match to a soybean miRNA of the miR172 family (miRBase entry: gma-miR172k). The identified member of the miR156 family (solyc-miR664) targets five SPLs, which belong to the SBP TF family (Figure 15 category B). The importance of SPLs in the early developmental phase of pollen was already pointed out by their enrichment among the mRNAs accumulated in tetrads (Figure 6). The regulation of SPLs by miR156 seems to be conserved in plants and was shown in a variety of plant species, such as *A*. *thaliana* (Xing et al., 2010; Wang et al., 2016), barley (Tripathi et al., 2018) and pear (Qian et al., 2017). Overexpression of miR156 in *A*. *thaliana* led to a significant downregulation of members of the SPL family and also to a strong decrease in the number of produced pollen grains (Wang et al., 2016), while overexpression of miR156 in a *spl8* mutant background resulted in fully sterile plants that completely lack pollen sacs (Xing et al., 2010). According to its assignment to category B, the identified miR156 is most active in post-meiotic and mature pollen leading to reduced

mRNA levels of the five targeted SPLs, while in tetrads miR156 levels are much lower, which results in higher mRNA levels of the five SPLs (Figure 14 category B). Interestingly, the five SPLs are also targeted by the so far unknown solyc-miR663 of category A. This miRNA is, except for the lack of the last two nucleotides, identical to the identified miR156 (solyc-miR664) and thus also a likely member of the miR156 family. The second known miRNA identified in tomato pollen belongs to the miR160 family (solyc-miR684) and targets two members of the ARF TF family (Figure 15 category C). ARFs control different developmental processes by the activation or repression of the transcription of auxin-regulated genes in dependence on the cellular auxin level (Li et al., 2016a). Regulation of ARFs by miR160 has been described for different plant species, such as *A. thaliana* and rice. In addition, transgenic *A. thaliana* and rice plants, carrying miR160-resistant ARFs, revealed dramatic growth and developmental defects, which points to an important role of miR160 in auxin dependent developmental processes (Mallory et al., 2005; Huang et al., 2016). In *A. thaliana*, it was shown that ARF17 is expressed from meiosis to the bicellular microspore stage but not in mature pollen, whereby *arf17* mutant plants showed a male-sterile phenotype with pollen wall-patterning defects and pollen degradation (Yang et al., 2013). Interestingly, one of the ARFs regulated by miR160 in tomato pollen (Solyc11g013480) is a predicted ortholog to *A. thaliana* ARF17, which indicates that the activity of miR160 is essential for tomato pollen development. According to its assignment into category C, tomato miR160 is only active in mature pollen, while its reduced levels in tetrads and post-meiotic pollen allow accumulation of the ARF mRNAs (Figure 14 category C), which is in agreement with the finding in *A. thaliana*. The second miRNA targeting the two ARFs (solyc-miR683) shares a high sequence similarity to the identified miR160 (solyc-miR684) but has a substitution at the 15th position and an additional adenine at the 3' end. Nevertheless, it is most likely also a member of the miR160 family. The identified member of the miR172 family (solyc-miR660) targets, among other things, six members of the AP2 TF family (Figure 15 category B). In *A. thaliana*, miR172 was also shown to target AP2 mRNAs, whereby the regulation was mainly accomplished by translational inhibition and only to a small extent by mRNA cleavage. Elevated accumulation of miR172 resulted in floral organ identity defects, while the disruption of the base pairing between miR172 and AP2 resulted in floral patterning defects, which emphasizes an important role of miR172 in flower development (Chen, 2004). In contrast to the findings in *A. thaliana*, in maize and *Nicotiana benthamiana* miR172 regulates AP2-like and AP2 proteins via mRNA cleavage, which is in agreement with the findings of this study (Lauter et al., 2005; Mlotshwa et al., 2006). Despite a difference in the regulation, also *Nicotiana*

benthamiana lines expressing a miR172-resistant *AP2* mutant showed floral patterning defects (Mlotshwa et al., 2006). Regarding the role of miR172 in the development of pollen little is known. Nevertheless, miR172 was shown to target at least two members of the AP2 family in mature pollen of *Brassica campestris* (Jiang et al., 2014). The finding that the miR172 family member identified in tomato pollen belongs to MTI category B, provides evidence that miR172 is already active in post-meiotic pollen, whereas the miR172 levels in tetrads are much lower, which explains the accumulation of AP2 mRNAs in tetrads (Figure 14 category B). The remaining 12 miRNAs, which are targeting mRNAs of TFs, represent novel candidates for the fine-tuning of transcriptional-networks and should be examined in future.

4.1.4 Regulation of protein levels by direct and delayed translation

The analysis of the transcriptome revealed an accumulation of mRNAs of a specific functionality in certain developmental stages. As the translation efficiency of mRNAs is dependent on multiple factors, it was of interest whether the observed accumulation of mRNAs also applies to the encoded proteins. The analysis of the proteins revealed that the translation of the underlying mRNAs is under the control of two distinct translational modes, which were termed direct and delayed translation.

The initial correlation analysis of the transcriptomes and proteomes of the pollen developmental stages revealed only a low correlation, ranging from 0.23 to 0.46 (Figure 18A to C). A similar low correlation was observed for the transcriptome and proteome of mature *A*. *thaliana* pollen, which had a pearson correlation coefficient of 0.31 (Grobei et al., 2009). Reasons for an uncoupling of the transcriptome and proteome are diverse and can include differences in translation rates as well as differences in the half-life of mRNAs and proteins (Schwanhüusser et al., 2011). Interestingly, the correlation of the transcriptome of tetrads and post-meiotic pollen with the proteome of the subsequent stage revealed especially for post-meiotic pollen an improvement of the correlation coefficient (Figure 18D). A similar improvement of the correlation was reported for transcriptomes and proteomes of different developmental stages of the unicellular parasite *Plasmodium falciparum* (Le Roch et al., 2004). Also here, the transcriptome of developmental stages correlated best with the proteome of later stages. One explanation for the observed improvement of the correlation is provided by the identified mode of delayed translation (Figure 20), which implies that proteins accumulate one stage later than their mRNAs. A delayed translation of mRNAs was also described in *Medicago truncatula* seeds, where especially mRNAs of LEA proteins were stored and translated with a delay (Verdier et al., 2013). Interestingly, three annotated LEA mRNAs were identified in the group of mRNAs with delayed

translation in post-meiotic pollen (Figure 21 post-meiotic delay), which indicates that the delay in LEA mRNA translation is also existent in developing pollen. One possibility for the storage of mRNAs in developing pollen offer EPPs, which were recently examined in tobacco (Hafidh et al., 2018). The authors analyzed four pollen developmental stages (uninucleate microspore, early bicellular pollen, late bicellular pollen and mature pollen) as well as pollen during pollen tube growth and could show that a subset of the mRNAs, present in the transcriptome of developing pollen, is stored in EPPs. EPPs are ribonucleoprotein complexes and enable either a short- or long-term storage of mRNAs, whose translation is required in a later developmental stage. This storage of mRNAs in EPPs is a possible explanation for the delayed translation of mRNAs accumulated in tetrads (Figure 21 tetrads delay) and post-meiotic pollen (post-meiotic delay). As the authors provide tomato homologs for all mRNAs identified in the EPPs of the different tobacco pollen developmental stages, a possible relationship between the delayed translation and EPPs could be verified. For this purpose, the 156 mRNAs with delayed translation in post-meiotic pollen were searched in EPPs of uninucleate microspores, which are more or less similar to post-meiotic pollen. Remarkably, around 60% of the 156 mRNAs have a homolog in EPPs of uninucleate microspores, which is a strong evidence that the observed delay in translation is, in part, the result of a short-term storage of mRNAs in EPPs.

4.2 Response of the pollen transcriptome and proteome to heat stress

4.2.1 The heat stress response core set and effect of heat stress on stage-accumulated mRNAs

The analysis of the pollen transcriptome during development revealed strong changes in mRNA abundance in dependence of the demand of the respective developmental stage. This finding together with the knowledge about the high sensitivity of pollen to HS, rose the question about the HSR of pollen and if mRNAs with an important role in pollen development are affected by HS.

The qualitative (Figure 7) and quantitative comparison (Figure 8) of non- and heat-stressed pollen developmental stages revealed differences in the HSR of the different stages. The least transcriptomic response was observed for tetrads, in which only 1% of the quantified mRNAs are differentially regulated in response to HS. In contrast, the observed 5.7% and 4.5% of differentially regulated mRNAs for post-meiotic and mature pollen, respectively, are comparable to the 4.2% in switchgrass and 3.7% to 5.2% in Chinese cabbage (Li et al., 2013b; Dong et al., 2015). One of the few large-scale transcriptomic studies that addressed the transcriptome response of pollen to HS was recently published and focused on the HSR of mature *A. thaliana* pollen (Rahmati Ishka et al., 2018). The

authors could show that 15% of the quantified mRNAs were differentially regulated, which is much higher than the 4.5% observed for mature tomato pollen in this study. One reason for the higher percentage of differentially regulated mRNAs in *A. thaliana* could be the applied HS regime, which included a diurnal cycle of hot and cold temperature for a period of one week and the isolation of pollen at the peak of the HS treatment at 40°C. In contrast, the HS regime of this study included a HS of 38°C for one hour, followed by a recovery phase before isolation of pollen. Despite the differences in the HS regime, the differentially regulated mRNAs in *A. thaliana* may provide insights in the general HS response of mature pollen.

The comparison of the differentially regulated mRNAs between the developmental stages revealed that the downregulation of mRNAs is highly stage-specific and that no mRNA is downregulated in all three stages (Supplemental Figure 1A). In contrast, a core set of 49 mRNAs with an upregulation in all three developmental stages was identified (Supplemental Figure 1B; Supplemental Table 4). A search for orthologs of the 49 mRNAs in the group of upregulated mRNAs in mature *A. thaliana* pollen (Rahmati Ishka et al., 2018), led to the identification of 25 orthologs that are also upregulated in response to HS. Next to the expected Hsf and Hsp mRNAs, eight mRNAs encode for proteins with a different functionality, which are interesting candidates for new insights in the pollen HSR. The first mRNA encodes for a BAG6, whose HS accumulation was also described in *A*. *thaliana* seedlings (Kang et al., 2006), whereby HsfA2 induces the expression of BAG6 (Nishizawa-Yokoi et al., 2009). Interestingly, *A*. *thaliana AtBAG6* knockdown lines show a moderate increase in thermotolerance with a simultaneous increase in the expression of *HSP18.2* and *HSP25.3*, which are both members of the sHsp family (Echevarría-Zomeño et al., 2016). The reason for the negative effect of BAG6 on *HSP18.2* and *HSP25.3* is its interaction with CaM3 (Kang et al., 2006), which functions as transcriptional activator of the two Hsps (Zhang et al., 2009). Therefore, BAG6 is likely to act as a limiter of the plant HSR (Echevarría-Zomeño et al., 2016), which also appears to apply to pollen. The second of the eight mRNAs encodes for an ascorbate peroxides, which is one of the two major ROS scavenging enzymes (Ozyigit et al., 2016). Ascorbate peroxidases play an important role in the detoxification of H_2O_2, which can cause significant damage to cells (Caverzan et al., 2012). Due to the accumulation of intracellular H_2O_2 under HS, ascorbate peroxidases are of particular importance in the HSR of plants (Volkov et al., 2006). Further, one mRNA encodes for a member of the glutathione S-transferase family. Glutathione S-transferases are involved in the detoxification of different substrates and were shown to be implicated in biotic and abiotic stress responses (Marrs, 1996). Interestingly,

two mRNAs encode for proteins annotated as co-chaperones of Hsp90. The first of these mRNAs encodes Aha1. In plants and humans, Aha1 functions as a co-chaperone and activator of Hsp90 by stimulating the ATPase activity of Hsp90 (Kadota et al., 2008; Li et al., 2012). Recently, it was also proposed that Aha1 is able to act as an autonomous chaperone that prevents stress-denatured proteins from aggregation (Tripathi et al., 2014). The second mRNA encodes a putative calcyclin-binding protein, which was recently shown to bind to Hsp90 and proposed to regulate the activity of Hsp90 by dephosphorylation (Góral et al., 2016). Another mRNA encodes for a RNA-binding protein, whose *A*. *thaliana* ortholog is SR45a. So far, it has been shown that the expression of SR45a is increased in response to stresses, such as high light irradiation, drought and HS (Tanabe et al., 2007; Gulledge et al., 2012). Further, there exist two splice variants of SR45a. The first variant has an exon-skipping event that leads to the full-length isoform, whereas the second variant retains the exon, which leads to a premature stop codon and the loss of the C-terminal RS motif. Interestingly, under HS the proportion of the full-length isoform dramatically increases (Gulledge et al., 2012), which offers an additional layer for the plant HSR. Regarding the last two mRNAs encoding a beta-1,3-glucanase and a not annotated protein, not much about their role in the plant HSR is known. Nevertheless, the upregulation of a beta-1,3-glucanase was also observed in switchgrass and the authors speculated that, similar to biotic stresses, the modification of the cell wall might be an important component of the plant HSR (Li et al., 2013b).

The analysis of the Hsf and Hsp families revealed that for all three stages the strongest HS effect was observed for the sHsp family (Figure 9; Supplemental Table 5). For each developmental stage, at least 72% of the annotated sHsps are upregulated in response to HS. Similar to this observation, also in heat-stressed *A*. *thaliana* roots and heat-stressed tomato leaves the sHsp family show a very strong HS induction (Swindell et al., 2007; Fragkostefanakis et al., 2015). Further, in all three stages an HS-induced upregulation of HsfA2 and HsfB1 was observed, which is in agreement with previous findings in heat-stressed *A*. *thaliana* seedlings (Dong et al., 2011). The importance of HsfA2 and HsfB1 in the HSR of tomato becomes also clear by their interaction with HsfA1, which is the master regulator of the HSR in tomato (Hahn et al., 2011). In addition, one Hsp90 and two Hsp100s are upregulated in all three stages. Similar to this, the *A*. *thaliana* orthologs of these Hsps, namely AtHsp90-1, ClpB1 and ClpB4, are also upregulated in response to HS (Yamada and Nishimura, 2008; Kissen et al., 2016). Interestingly, HsfA2 and the three Hsps belong to the mRNAs upregulated in heat-stressed mature *A*.

thaliana pollen (Rahmati Ishka et al., 2018), which emphasizes the importance of these proteins in the HSR of pollen.

The effect of HS on the mRNAs accumulated in one or two consecutive stages showed that the vast majority of these mRNAs is not affected. Nevertheless, a small number of these mRNAs was differentially regulated, whereby the majority was downregulated (Figure 11). Among the mRNAs accumulated in non-stressed tetrads and downregulated in heat-stressed tetrads (Supplemental Table 6) as well as those accumulated in non-stressed post-meiotic pollen and downregulated in heat-stressed post-meiotic pollen (Supplemental Table 8) many TFs of different families were identified. The observed TF families, such as CAMTA, bZIP or WRKY, are known to be of importance for the proper development of pollen (Mitsuda et al., 2003; Iven et al., 2010; Lei et al., 2017). Therefore, the downregulation of these TF families is likely to impair the development of pollen. By far the highest percentage of downregulated mRNAs was observed for the mRNAs accumulated in post-meiotic and mature pollen, out of which 8.6% (119 of 1,387 mRNAs) show a HS-induced downregulation in post-meiotic pollen (Supplemental Table 9). Interestingly, one of the 119 mRNAs encodes the villin protein accumulated in the proteome of mature pollen (Figure 22). Villin proteins are required for the bundling of actin filaments and the regulation of actin dynamics during pollen tube growth. Further, it could be shown that *A*. *thaliana VLN5* loss-of-function mutant plants have retarded pollen tube growth (Zhang et al., 2010). Therefore, the downregulation of villin mRNAs in response to HS is likely leading to an impairment in pollen tube growth. Next to villin, also multiple mRNAs encoding for sugar transporters were downregulated in response to HS. Sugar transporters reside, among other things, on the pollen surface and are required for the uptake of carbohydrates in the developing pollen. Perturbations in the sugar supply can lead to an impairment of pollen development and male sterility (Slewinski, 2011). Hence, the HS-induced downregulation of sugar transporter mRNAs will result in a shortage in the sugar supply, which in turn leads to impaired pollen development. Further, the 119 mRNAs include four mRNAs encoding for LEA proteins. These proteins are required for the acquisition of desiccation tolerance as it is required in dehydrated seeds and pollen (Goyal et al., 2005; Firon et al., 2012). The suppression of LEA proteins in response to HS will therefore have a negative effect on the desiccation tolerance of the dehydrated mature pollen and possibly lead to a reduced number of viable and germinating pollen.

4.2.2 The effect of miRNAs on the heat stress response of pollen

The clustering of the miRNAs based on their relative abundance in non- and heat-stressed samples led to the identification of down- (c1 clusters) and upregulated miRNAs (c7 clusters) for each developmental stage (Supplemental Figure 4 to Supplemental Figure 6). Based on the down- and upregulated miRNAs, 24 HS-relevant MTIs with an opposite HS regulation between their miRNA and mRNA could be identified (Table 3; Supplemental Figure 7). One of these HS-relevant MTIs in tetrads comprises the mRNA of the pollen HSR core set encoding for Aha1. This mRNA is targeted by the so far unknown solyc-miR708, whose downregulation is likely leading to the upregulation of the mRNA. The regulation of mRNAs encoding Hsp90 co-chaperones by miRNAs has so far been described only in humans (Liu et al., 2012b; Zheng et al., 2016), but the observed MTI provides evidence that this also applies to plants. One of the HS-relevant MTIs observed in post-meiotic pollen includes solyc-miR678, which was already identified in other species of the Solanaceae family (e.g. miRBase entry: nta-miR408), where it belongs to the miR408 family. The identified miR408 targets an mRNA encoding for a BBP, whereby in response to HS the miRNA is down- and the mRNA upregulated. The targeting of BBP mRNAs, also known as plantacyanins, by miR408 has also been described in other plant species, such as *A. thaliana* (Sunkar and Zhu, 2004) and wheat (Yao et al., 2007). Further, the importance of miR408 in abiotic stress response was examined in *A. thaliana* and it could be shown that miR408 is downregulated in response drought and osmotic stress, whereas overexpression of miR408 leads to a higher sensitivity to drought and osmotic stress (Ma et al., 2015). Therefore, it is likely that the downregulation of miR408 and upregulation of the mRNA encoding BBP will have a positive effect on the tolerance of post-meiotic pollen to HS. One of the HS-relevant MTI in mature pollen includes the downregulated solyc-miR461 and an upregulated mRNA of the pollen HSR core set encoding for BAG6. The regulation of BAG6 by a miRNA was also described during the development of Siberian apricot seeds (Niu et al., 2016), although the miRNA shares no sequence similarity to solyc-miR461. Interestingly, one of the HS-relevant MTIs was detected in post-meiotic and mature pollen. The MTI includes the so far unknown solyc-miR749 and an mRNA encoding a CYB561. Although, the role of CYB561s in stress responses is not clear, there is evidence that CYB561s support stress defense (Asard et al., 2013) and thus the regulation via solyc-miR749 is likely to contribute to the HSR of post-meiotic and mature pollen. Next to the mentioned MTIs leading to an upregulation of mRNAs, also seven MTIs leading to the downregulation of their mRNAs were identified. Interestingly, five of the seven MTIs include solyc-miR544, which belongs to the novel

miRNAs. The upregulation of this miRNA in post-meiotic pollen leads, among other things, to the simultaneous downregulation of two mRNAs encoding for the actin-related profilin and a LEA protein. The regulation of profilin and LEA encoding mRNAs by miRNAs has already been described in humans (Liu et al., 2013) and cotton (Magwanga et al., 2018), respectively. As already mentioned LEA proteins are required for the desiccation tolerance of dehydrated pollen (Goyal et al., 2005; Firon et al., 2012), whereas profilin is required for pollen germination and pollen tube growth (Yu and Parthasarathy, 2014). The miRNA-induced downregulation of the mRNAs encoding these proteins might therefore have a negative effect on pollen viability and tube growth.

4.2.3 Regulation of the pollen proteome in response to heat stress

Similar to the transcriptome and proteome of developing pollen, also for the heat-stressed pollen developmental stages only a low correlation ranging from 0.29 to 0.44 (Supplemental Figure 8) was observed. Further, it could be shown that the HS regulation of proteins is mainly independent from the regulation of the underlying mRNAs (Figure 24; Supplemental Table 15). The low correlation and different regulation of mRNAs and proteins in response to HS was also described in other species, such as rice (Batista et al., 2017), soybean (Valdés-López et al., 2016), *Ulva prolifera* (Fan et al., 2017) and *Aiptasia pallida* (Cziesielski et al., 2018). This finding, indicates, that the regulation of protein levels in response to HS is controlled at the post-transcriptional level and therefore mRNA measurements provide only limited insights in the regulation of proteins.

Although the HS regulation of proteins is mainly independent from the regulation of the underlying mRNAs, 25 times a pair of mRNAs and proteins showed a co-regulation in response to HS (Supplemental Table 15). Seven of these pairs showed a co-upregulation in tetrads and post-meiotic pollen. Interestingly, the mRNAs of these pairs do all belong to the pollen HSR core set, whereby five of the mRNAs encode for sHsps. Although upregulated in tetrads and post-meiotic pollen, the encoded proteins of the seven mRNAs are not upregulated in mature pollen. This finding further supports the idea that the typical accumulation of Hsps in response to HS does not take place in mature pollen (Hopf et al., 1992). Further, one Hsp100 and a further sHsps showed a co-upregulation exclusively in tetrads and post-meiotic pollen, respectively, which further confirms the accumulation of Hsps in early developmental stages (Frova et al., 1989; Rieu et al., 2017).

Within the group of proteins that are regulated independently of their underlying mRNA, 57 RPs were identified (Figure 25). The observed differential regulation of RPs in response to HS might lead to an

alternative ribosome composition, which is known to have an impact on how and which mRNAs are being translated (Xue and Barna, 2012). For instance, in humans it could be shown that under HS an isoform of the mitochondrially localized MRPL18 is generated that remains in the cytosol, where it is incorporated into the 80S ribosome complex and facilitates the translation of stress-responsive proteins such as Hsp70 (Zhang et al., 2015b). One of the RP families with a different regulation of its members is the RPS8 family. One of the RP8s is downregulated in tetrads and mature pollen, while in each developmental stage a different RPS8 is upregulated. Results in human suggest that RPS8 functions as a rate-limiting factor in translational regulation and further is involved in the inhibition of cap- and IRES-mediated translation (Hao et al., 2011). If this role of RPS8 also applies to plants, the RPS8s upregulated in the different developmental stages contribute to a stage-specific translation regulation during HS. Further, in mouse it could be shown that ribosomes containing RPS25 preferentially translate a subpool of mRNAs (Shi et al., 2017). Interestingly, two RPS25s are exclusively upregulated in mature pollen, which suggests that also in the case of mature pollen a subpool of mRNAs might be preferentially translate in response to HS. Next to translation of a specific subpool of mRNAs, RPs can also regulate mRNA translation by the translation of uORFs, which prevents the translation of the main ORF (Merchante et al., 2017). For instance, RPL4 and RPL10 were shown to be involved in the translation of uORFs, which had a negative effect on the translation of the main ORF (Imai et al., 2008; Kakehi et al., 2015). Interestingly, members of the RPL4 and RPL10 family are differentially regulated in heat-stressed post-meiotic pollen. Three RPL10s are downregulated in post-meiotic pollen, while one RPL4 is upregulated in post-meiotic pollen. The differential regulation of these proteins could contribute to changes in the translation of uORFs and by this influence the synthesis of the actual protein.

5 Conclusion and outlook

The results of this study enabled new insights in the development and HSR of pollen at the mRNA, miRNA and protein level (Figure 26). The analysis of stage-accumulated mRNAs enabled the identification of functional processes of particular importance in certain developmental stages. For instance, in tetrads a special role of brassinosteroids and SPLs as well as of the abiotic stress response was observed, while in tetrads and post-meiotic pollen the synthesis of proteins and activation of amino acids is of importance. In contrast, post-meiotic and mature pollen share a preference for V-ATPases and aerobic respiration, while solely in mature pollen cell wall and cytoskeleton related processes are important. These results provide especially for the less characterized tetrad stage new insights into protein families and processes, which are required for the early phase of pollen development.

Next to functional processes, also an important role of miRNAs in the development of pollen was observed. For instance, in post-meiotic and mature pollen miR156 and miR172 are required for the suppression of mRNAs encoding for SPL and AP2 TFs, while miR395, miR398 and miR160 are required for mRNA suppression in tetrads, post-meiotic and mature pollen, respectively. The observed preferential regulation of TFs enables the direct control of transcriptional-networks, which is likely supporting the molecular reprogramming during pollen development. As the presented results do all rely on *in silico* analyses, the miRNAs and their interactions with target mRNAs must be verified by future experiments. The verification of the miRNAs could be done using stem-loop RT-qPCR, which unlike the standard qRT-PCR can identify and quantify small RNAs by extending the miRNA with a stem-loop RT primer (Kramer, 2011). The cleavage of the target mRNAs could be verified using degradome sequencing, which enables the identification and quantification of cleavage sites by sequencing the 5’ ends of uncapped RNAs (Li et al., 2010). The uncapped RNAs correspond in this case to the 3’ fragments of cleaved mRNAs and are captured via a free monophosphate located at the 5’ end of the fragments. The identified cleavage sites can afterwards be compared to the predicted binding sites of the miRNAs.

Further, the translation of mRNAs in developing pollen has been shown to be under the control of two translational modes. The first mode results in a direct translation of the mRNAs, while the second mode leads to a delayed translation. The observed delay in mRNA translation is likely, in part, the result of a storage of mRNAs in EPPs, which keep the mRNAs translationally silenced. Here it would

be interesting to identify the RNA-binding proteins required for the storage of the mRNAs. One way to identify these proteins would be the RNA-binding protein purification and identification technique. This technique enables the affinity purification of tagged mRNAs and of the proteins bound to them. The co-purified proteins can afterwards be identified via MS approaches (Slobodin and Gerst, 2010).

Regarding the HSR of developing pollen, a HSR core set of 49 mRNAs, which are upregulated in all analyzed stages, could be identified. Next to the classical Hsf and Hsp encoding mRNAs, also mRNAs encoding for other proteins, such as BAG6 or an ascorbate peroxidase, are part of the pollen HSR core set. The typical approach to verify the importance of proteins in the HSR is the generation of overexpression and knockout lines, followed by HS application and phenotyping (Liu and Charng, 2013). A similar approach could be applied for the so far not characterized members of the pollen HSR core set to analyze their impact on the HSR of pollen.

In addition, it could be shown that the HS regulation of mRNAs is to a small extent regulated via the induction or repression of miRNAs. For instance, the HS-induced repression of the known miRNA miR408 and the novel miRNAs solyc-miR708, solyc-miR461 and solyc-miR749 leads to an accumulation of their target mRNAs, which encode for proteins, such as BBP, CYB561, Aha1 or BAG6. The importance of the miRNAs in the HSR of pollen could be analyzed by the knockdown of the miRNAs using CRISPR/cas9 (Chang et al., 2016), which should have an effect on the thermotolerance.

The analysis of the proteomes of the heat-stressed pollen stages revealed that the HS regulation of proteins is largely independent of the underlying mRNAs. Especially affected by HS are RPs of the small and large subunit. For instance in post-meiotic pollen 13 RPs of the small and 17 RPs of the large subunit are differentially regulated, which is likely leading to an altered ribosome composition. The alteration of the ribosome composition represents a new layer in the HSR of pollen, which is completely independent from the general induction of Hsfs and Hsps. To verify the effect of an altered ribosome composition on the translation of mRNAs, a similar approach as described by Shi et al. (2017) could be applied. In this approach, ribosomes containing an endogenously tagged RP are isolated, followed by ribosome profiling, which enables the identification and quantification of mRNAs bound by the ribosomes. By performing a second ribosome profiling on all ribosomes of a cell and comparing the abundance of an mRNA between the two ribosome profiling experiments, mRNAs preferentially bound by ribosomes containing the tagged RP can be identified.

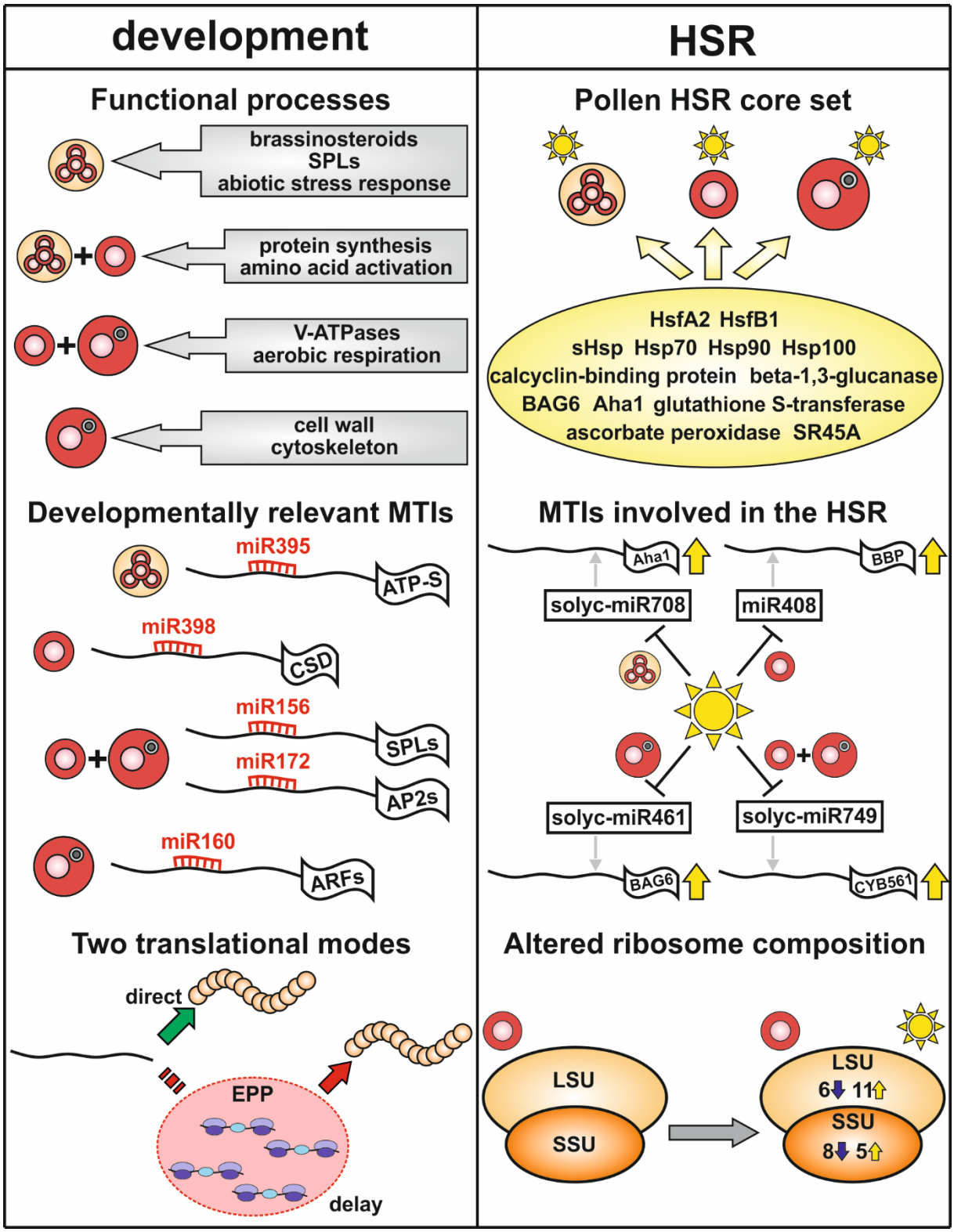


Figure 26: New molecular insights in the development and HSR of pollen
Shown are the key results for the development and HSR of pollen. Regarding the development of pollen, functional processes, MTIs and two translational modes could be identified, which are important in the individual developmental stages. The analysis of the heat-stressed pollen stages enabled the identification of a pollen HSR core set, MTIs and differentially regulated RPs, which are part of the HSR of developing pollen.

6 References

Abdallah C, Dumas-Gaudot E, Renaut J, Sergeant K (2012) Gel-based and gel-free quantitative proteomics approaches at a glance. Int J Plant Genomics **2012**: 494572

Adams M, Kelley J, Gocayne J, Dubnick M, Polymeropoulos M, Xiao H, Merril C, Wu A, Olde B, Moreno R, et al (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. Science (80-) **252**: 1651–1656

Adamski MG, Gumann P, Baird AE (2014) A method for quantitative analysis of standard and high-throughput qPCR expression data based on input sample quantity. PLoS One **9**: e103917

Agarwala R, Barrett T, Beck J, Benson DA, Bollin C, Bolton E, Bourexis D, Brister JR, Bryant SH, Canese K, et al (2016) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res **44**: D7–D19

Ahsan N, Donnart T, Nouri MZ, Komatsu S (2010) Tissue-specific defense and thermo-adaptive mechanisms of soybean seedlings under heat stress revealed by proteomic approach. J Proteome Res **9**: 4189–4204

Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JKC, et al (2004) ESTs, cDNA microarrays, and gene expression profiling: Tools for dissecting plant physiology and development. Plant J **39**: 697–714

Ambardar S, Gupta R, Trakroo D, Lal R, Vakhlu J (2016) High Throughput Sequencing: An Overview of Sequencing Chemistry. Indian J Microbiol **56**: 394–404

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

Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD (2012) Mass spectrometry-based proteomics: existing capabilities and future directions. Chem Soc Rev **41**: 3912–28

Anjum NA, Gill R, Kaushik M, Hasanuzzaman M, Pereira E, Ahmad I, Tuteja N, Gill SS (2015) ATP-sulfurylase, sulfur-compounds, and plant stress tolerance. Front Plant Sci **6**: 210

Asard H, Barbaro R, Trost P, Bérczi A (2013) Cytochromes *b* 561: Ascorbate-Mediated Trans-Membrane Electron Transport. Antioxid Redox Signal **19**: 1026–1035

Baniwal SK, Bharti K, Chan KY, Fauth M, Ganguli A, Kotak S, Mishra SK, Nover L, Port M, Scharf KD, et al (2004) Heat stress response in plants: A complex game with chaperones and more than twenty heat stress transcription factors. J Biosci **29**: 471–487

Bao F, Shen J, Brady SR, Muday GK, Asami T, Yang Z (2004) Brassinosteroids interact with auxin to promote lateral root development in Arabidopsis. Plant Physiol **134**: 1624–31

Batista R, Fonseca C, Planchon S, Negrão S, Renaut J, Oliveira MM (2017) Environmental stress is the major cause of transcriptomic and proteomic changes in GM and non-GM plants. Sci Rep **7**: 10624

Becker JD, Boavida LC, Carneiro J, Haury M, Feijó JA (2003) Transcriptional profiling of Arabidopsis tissues reveals the unique characteristics of the pollen transcriptome. Plant Physiol **133**: 713–25

Beranova-Giorgianni S (2003) Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: Strengths and limitations. TrAC - Trends Anal Chem **22**: 273–281

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E (2015) The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. Genesis **53**: 474–485

Berg M, Rogers R, Muralla R, Meinke D (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J **44**: 866–878

Bethke G, Thao A, Xiong G, Li B, Soltis NE, Hatsugai N, Hillmer RA, Katagiri F, Kliebenstein DJ, Pauly M, et al (2016) Pectin Biosynthesis Is Critical for Cell Wall Integrity and Immunity in *Arabidopsis thaliana*. Plant Cell **28**: 537–556

Birke H, Heeg C, Wirtz M, Hell R (2013) Successful Fertilization Requires the Presence of at Least One Major O-Acetylserine(thiol)lyase for Cysteine Synthesis in Pollen of Arabidopsis. Plant Physiol **163**: 959–972

Bokszczanin KL, Krezdorn N, Fragkostefanakis S, Müller S, Rycak L, Chen Y, Hoffmeier K, Kreutz J, Paupière MJ, Chaturvedi P, et al (2015) Identification of novel small ncRNAs in pollen of tomato. BMC Genomics **16**: 714

Bokszczanin KL, Solanaceae Pollen Thermotolerance Initial Training Network (SPOT-ITN) Consortium, Fragkostefanakis S (2013) Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. Front Plant Sci **4**: 315

Bokvaj P, Hafidh S, Honys D (2015) Transcriptome profiling of male gametophyte development in Nicotiana tabacum. Genomics Data **3**: 106–111

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Borg M, Brownfield L, Twell D (2009) Male gametophyte development: A molecular perspective. J Exp Bot **60**: 1465–1478

Bosch M, Cheung AY, Hepler PK (2005) Pectin methylesterase, a regulator of pollen tube growth. Plant Physiol **138**: 1334–46

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Mol Ecol **16**: 907–924

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science (80-) **320**: 1185–1190

Budak H, Akpinar BA (2015) Plant miRNAs: biogenesis, organization and origins. Funct Integr Genomics **15**: 523–531

Buermans HPJ, den Dunnen JT (2014) Next generation sequencing technology: Advances and applications. Biochim Biophys Acta - Mol Basis Dis **1842**: 1932–1941

Busch W, Wunderlich M, Schöffl F (2005) Identification of novel heat shock factor-dependent genes and biochemical pathways in Arabidopsis thaliana. Plant J **41**: 1–14

Caverzan A, Passaia G, Rosa SB, Ribeiro CW, Lazzarotto F, Margis-Pinheiro M (2012) Plant responses to stresses: Role of ascorbate peroxidase in the antioxidant protection. Genet Mol Biol **35**: 1011–1019

Chang H, Yi B, Ma R, Zhang X, Zhao H, Xi Y (2016) CRISPR/cas9, a novel genomic tool to knock down microRNA in vitro and in vivo. Sci Rep **6**: 22312

Chaturvedi P, Doerfler H, Jegadeesan S, Ghatak A, Pressman E, Castillejo MA, Wienkoop S, Egelhofer V, Firon N, Weckwerth W (2015) Heat-treatment-responsive proteins in different developmental stages of tomato pollen detected by targeted mass accuracy precursor alignment (tMAPA). J Proteome Res **14**: 4463–4471

Chaturvedi P, Ghatak A, Weckwerth W (2016) Pollen proteomics: from stress physiology to developmental priming. Plant Reprod **29**: 119–132

Chaturvedi P, Ischebeck T, Egelhofer V, Lichtscheidl I, Weckwerth W (2013) Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming. J Proteome Res **12**: 4892–903

Chen X (2004) A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower

Development. Science (80-) **303**: 2022–2025

Chen X, Zhang Z, Liu D, Zhang K, Li A, Mao L (2010) SQUAMOSA promoter-binding protein-like transcription factors: Star players for plant growth and development. J Integr Plant Biol **52**: 946–951

Chen YA, Mckillen DJ, Wu S, Jenny MJ, Chapman R, Gross PS, Warr GW, Almeida JS (2004) Optimal cDNA microarray design using expressed sequence tags for organisms with limited genomic information. BMC Bioinformatics **5**: 191

Cooper P, David Ho TH, Hauptmann RM (1984) Tissue specificity of the heat-shock response in maize. Plant Physiol **75**: 431–441

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M (2014) Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol Cell Proteomics **13**: 2513–2526

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol **26**: 1367–1372

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 ångstrom resolution. Science (80-) **292**: 1863–1876

Crick F (1958) On protein synthesis. Symp Soc Exp Biol **12**: 138–63

Cui P, Lin Q, Ding F, Xin C, Gong W, Zhang L, Geng J, Zhang B, Yu X, Yang J, et al (2010) A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. Genomics **96**: 259–265

Cziesielski MJ, Liew YJ, Cui G, Schmidt-Roach S, Campana S, Marondedze C, Aranda M (2018) Multi-omics analysis of thermal stress response in a zooxanthellate cnidarian reveals the importance of associating with thermotolerant symbionts. Proceedings Biol Sci. doi: 10.1098/rspb.2017.2654

Dai S, Li L, Chen T, Chong K, Xue Y, Wang T (2006) Proteomic analyses of Oryza sativa mature pollen reveal novel proteins associated with pollen germination and tube growth. Proteomics **6**: 2504–2529

Davidian JC, Kopriva S (2010) Regulation of sulfate uptake and assimilation - The same or not the same? Mol Plant **3**: 314–325

Deracinois B, Flahaut C, Duban-Deweer S, Karamanos Y (2013) Comparative and Quantitative Global Proteomics Approaches: An Overview. Proteomes **1**: 180–218

Dettmer J, Schubert D, Calvo-Weimar O, Stierhof YD, Schmidt R, Schumacher K (2005) Essential role of the V-ATPase in male gametophyte development. Plant J **41**: 117–124

Dong H, Zhen Z, Peng J, Chang L, Gong Q, Wang NN (2011) Loss of ACS7 confers abiotic stress tolerance by modulating ABA sensitivity and accumulation in Arabidopsis. J Exp Bot **62**: 4875–4887

Dong X, Yi H, Lee J, Nou I-S, Han C-T, Hur Y (2015) Global Gene-Expression Analysis to Identify Differentially Expressed Genes Critical for the Heat Stress Response in Brassica rapa. PLoS One **10**: e0130451

Draeger T, Moore G (2017) Short periods of high temperature during meiosis prevent normal meiotic progression and reduce grain number in hexaploid wheat (Triticum aestivum L.). Theor Appl Genet **130**: 1785–1800

Dzieciatkowska M, Hill R, Hansen KC (2014) GeLC-MS/MS analysis of complex protein mixtures. Methods Mol Biol **1156**: 53–66

Echevarría-Zomeño S, Fernández-Calvino L, Castro-Sanz AB, López JA, Vázquez J, Castellano MM (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

Eichacker LA, Granvogl B, Mirus O, Müller BC, Miess C, Schleiff E (2004) Hiding behind hydrophobicity: Transmembrane segments in mass spectrometry. J Biol Chem **279**: 50915–50922

Fan M, Sun X, Xu N, Liao Z, Li Y, Wang J, Fan Y, Cui D, Li P, Miao Z (2017) Integration of deep transcriptome and proteome analyses of salicylic acid regulation high temperature stress in Ulva prolifera. Sci Rep **7**: 11052

Feder ME, Hofmann GE (1999) HEAT-SHOCK PROTEINS, MOLECULAR CHAPERONES, AND THE STRESS RESPONSE: Evolutionary and Ecological Physiology. Annu Rev Physiol **61**: 243–282

Fernandez-Pozo N, Menda N, Edwards JD, Saha S, Tecle IY, Strickler SR, Bombarely A, Fisher-York T, Pujar A, Foerster H, et al (2015) The Sol Genomics Network (SGN)-from genotype to phenotype to breeding. Nucleic Acids Res **43**: D1036–D1041

Ferradini N, Iannacone R, Capomaccio S, Metelli A, Armentano N, Semeraro L, Cellini F, Veronesi F, Rosellini D (2015) Assessment of heat shock protein 70 induction by heat in alfalfa varieties and constitutive overexpression in transgenic plants. PLoS One **10**: e0126051

Finotello F, Lavezzo E, Bianco L, Barzon L, Mazzon P, Fontana P, Toppo S, Di Camillo B (2014) Reducing bias in RNA sequencing data: a novel approach to compute counts. BMC Bioinformatics **15 Suppl 1**: S7

Firon N, Nepi M, Pacini E (2012) Water status and associated processes mark critical stages in pollen development and functioning. Ann Bot **109**: 1201–14

Fondevilla S, Küster H, Krajinski F, Cubero JI, Rubiales D (2011) Identification of genes differentially expressed in a resistant reaction to Mycosphaerella pinodes in pea using microarray technology. BMC Genomics **12**: 28

Fragkostefanakis S, Mesihovic A, Simm S, Paupière MJ, Hu Y, Paul P, Mishra SK, Tschiersch B, Theres K, Bovy A, et al (2016) HsfA2 Controls the Activity of Developmentally and Stress-Regulated Heat Stress Protection Mechanisms in Tomato Male Reproductive Tissues. Plant Physiol **170**: 2461–2477

Fragkostefanakis S, Simm S, Paul P, Bublak D, Scharf KD, Schleiff E (2015) Chaperone network composition in Solanum lycopersicum explored by transcriptome profiling and microarray meta-analysis. Plant, Cell Environ **38**: 693–709

Frank G, Pressman E, Ophir R, Althan L, Shaked R, Freedman M, Shen S, Firon N (2009) Transcriptional profiling of maturing tomato (Solanum lycopersicum L.) microspores reveals the involvement of heat shock proteins, ROS scavengers, hormones, and sugars in the heat stress response. J Exp Bot **60**: 3891–3908

Friedländer MR, MacKowiak SD, Li N, Chen W, Rajewsky N (2012) MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res **40**: 37–52

Frova C, Taramino G, Binelli G (1989) Heat- shock proteins during pollen development in maize. Dev Genet **10**: 324–332

Fu FQ, Mao WH, Shi K, Zhou YH, Asami T, Yu JQ (2008) A role of brassinosteroids in early fruit development in cucumber. J Exp Bot **59**: 2299–2308

Fu Y, Dominissini D, Rechavi G, He C (2014) Gene expression regulation mediated through reversible m 6 A RNA methylation. Nat Rev Genet **15**: 293–306

Giorno F, Wolters-Arts M, Mariani C, Rieu I (2013) Ensuring Reproduction at High Temperatures: The Heat Stress Response during Anther and Pollen Development. Plants **2**: 489–506

Gonzalez-Garcia M-P, Vilarrasa-Blasi J, Zhiponova M, Divol F, Mora-Garcia S, Russinova E,

Cano-Delgado AI (2011) Brassinosteroids control meristem size by promoting cell cycle progression in Arabidopsis roots. Development **138**: 849–859

Góral A, Bieganowski P, Prus W, Krzemieñ-Ojak Ł, Dziołka BK, Fabczak H, Filipek A (2016) Calcyclin binding protein/siah-1 interacting protein is a hsp90 binding chaperone. PLoS One **11**: e0156507

Goyal K, Walton LJ, Tunnacliffe A (2005) LEA proteins prevent protein aggregation due to water stress. Biochem J **388**: 151–7

Grobei MA, Qeli E, Brunner E, Rehrauer H, Zhang R, Roschitzki B, Basler K, Ahrens CH, Grossniklaus U (2009) Deterministic protein inference for shotgun proteomics data provides new insights into Arabidopsis pollen development and function. Genome Res **19**: 1786–1800

Guan Q, Lu X, Zeng H, Zhang Y, Zhu J (2013) Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in Arabidopsis. Plant J **74**: 840–851

Gulledge AA, Roberts AD, Vora H, Patel K, Loraine AE (2012) Mining Arabidopsis thaliana RNA-seq data with integrated genome browser reveals stress-induced alternative splicing of the putative splicing regulator SR45A. Am J Bot **99**: 219–231

Guo AY, Chen X, Gao G, Zhang H, Zhu QH, Liu XC, Zhong YF, Gu X, He K, Luo J (2008) PlantTFDB: A comprehensive plant transcription factor database. Nucleic Acids Res **36**: D966-9

Guo M, Liu J-H, Ma X, Luo D-X, Gong Z-H, 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 W, Peddada S (2008) Adaptive choice of the number of bootstrap samples in large scale multiple testing. Stat Appl Genet Mol Biol **7**: Article13

Györgyey J, Vaubert D, Jiménez-Zurdo JI, Charon C, Troussard L, Kondorosi Á, Kondorosi É (2000) Analysis of *Medicago truncatula* Nodule Expressed Sequence Tags. Mol Plant-Microbe Interact **13**: 62–71

Hafidh S, Potěšil D, Müller K, Fíla J, Michailidis C, Herrmannová A, Feciková J, Ischebeck T, Valášek LS, Zdráhal Z, et al (2018) Dynamics of the Pollen Sequestrome Defined by Subcellular Coupled Omics. Plant Physiol **178**: 258–282

Hahn A, Bublak D, Schleiff E, Scharf K-D (2011) Crosstalk between Hsp90 and Hsp70 Chaperones and Heat Stress Transcription Factors in Tomato. Plant Cell **23**: 741–755

Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol **11**: 394–403

Han X, Aslanian A, Yates JR (2008) Mass spectrometry for proteomics. Curr Opin Chem Biol **12**: 483–490

Hao Y, Kong X, Ruan Y, Gan H, Chen H, Zhang C, Ren S, Gu J (2011) CDK11p46 and RPS8 associate with each other and suppress translation in a synergistic manner. Biochem Biophys Res Commun **407**: 169–174

Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, Ordoukhanian P (2014) Library construction for next-generation sequencing: overviews and challenges. Biotechniques **56**: 61–4, 66, 68, passim

Hill CB, Li C (2016) Genetic Architecture of Flowering Phenology in Cereals and Opportunities for Crop Improvement. Front Plant Sci **7**: 1906

Hoffert JD, van Balkom BWM, Chou C-L, Knepper M a (2004) Application of difference gel electrophoresis to the identification of inner medullary collecting duct proteins. Am J Physiol Renal Physiol **286**: F170-9

Holmes-Davis R, Tanaka CK, Vensel WH, Hurkman WJ, McCormick S (2005) Proteome mapping

of mature pollen of Arabidopsis thaliana. Proteomics **5**: 4864–4884

Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol **5**: R85.1-R.85.13

Honys D, Twell D (2003) Comparative analysis of the Arabidopsis pollen transcriptome. Plant Physiol **132**: 640–52

Hopf N, Plesofsky-Vig N, Brambl R (1992) The heat shock response of pollen and other tissues of maize. Plant Mol Biol **19**: 623–630

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc **4**: 44–57

Huang J, Li Z, Zhao D (2016) Deregulation of the OsmiR160 target gene OsARF18 causes growth and developmental defects with an alteration of auxin signaling in rice. Sci Rep **6**: 29938

Igamberdiev AU, Lea PJ (2006) Land plants equilibrate O2 and CO2 concentrations in the atmosphere. Photosynth Res **87**: 177–194

Imai A, Komura M, Kawano E, Kuwashiro Y, Takahashi T (2008) A semi-dominant mutation in the ribosomal protein L10 gene suppresses the dwarf phenotype of the acl5 mutant in Arabidopsis thaliana. Plant J **56**: 881–890

Ischebeck T, Valledor L, Lyon D, Gingl S, Nagler M, Meijón M, Egelhofer V, Weckwerth W (2014) Comprehensive Cell-specific Protein Analysis in Early and Late Pollen Development from Diploid Microsporocytes to Pollen Tube Growth. Mol Cell Proteomics **13**: 295–310

Ivashuta S, Banks IR, Wiggins BE, Zhang Y, Ziegler TE, Roberts JK, Heck GR (2011) Regulation of gene expression in plants through miRNA inactivation. PLoS One **6**: e21330

Iven T, Strathmann A, Böttner S, Zwafink T, Heinekamp T, Guivarc'h A, Roitsch T, Dröge-Laser W (2010) Homo- and heterodimers of tobacco bZIP proteins counteract as positive or negative regulators of transcription during pollen development. Plant J **63**: 155–166

Janssen BJ, Thodey K, Schaffer RJ, Alba R, Balakrishnan L, Bishop R, Bowen JH, Crowhurst RN, Gleave AP, Ledger S, et al (2008) Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. BMC Plant Biol **8**: 16

Jiang J, Liu X, Liu C, Liu G, Li S, Wang L (2017) Integrating Omics and Alternative Splicing Reveals Insights into Grape Response to High Temperature. Plant Physiol **173**: 1502–1518

Jiang J, Lv M, Liang Y, Ma Z, Cao J (2014) Identification of novel and conserved miRNAs involved in pollen development in Brassica campestris ssp. chinensis by high-throughput sequencing and degradome analysis. BMC Genomics **15**: 146

Jiang L, Yang S-L, Xie L-F, Puah CS, Zhang X-Q, Yang W-C, Sundaresan V, Ye D (2005) VANGUARD1 encodes a pectin methylesterase that enhances pollen tube growth in the Arabidopsis style and transmitting tract. Plant Cell **17**: 584–96

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs AND THEIR REGULATORY ROLES IN PLANTS. Annu Rev Plant Biol **57**: 19–53

Journet EP, van Tuinen D, Gouzy J, Crespeau H, Carreau V, Farmer MJ, Niebel A, Schiex T, Jaillon O, Chatagnier O, et al (2002) Exploring root symbiotic programs in the model legume Medicago truncatula using EST analysis. Nucleic Acids Res **30**: 5579–5592

Kadota Y, Amigues B, Ducassou L, Madaoui H, Ochsenbein F, Guerois R, Shirasu K (2008) Structural and functional analysis of SGT1-HSP90 core complex required for innate immunity in plants. EMBO Rep **9**: 1209–1215

Kakehi JI, Kawano E, Yoshimoto K, Cai Q, Imai A, Takahashi T (2015) Mutations in ribosomal proteins, RPL4 and RACK1, suppress the phenotype of a thermospermine-deficient mutant of arabidopsis thaliana. PLoS One **10**: e0117309

Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI (2018) Rfam 13.0: Shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res **46**: D335–D342

Kang CH, Jung WY, Kang YH, Kim JY, Kim DG, Jeong JC, Baek DW, Jin JB, Lee JY, Kim MO, et al (2006) AtBAG6, a novel calmodulin-binding protein, induces programmed cell death in yeast and plants. Cell Death Differ **13**: 84–95

Keller M, Hu Y, Mesihovic A, Fragkostefanakis S, Schleiff E, Simm S (2017) Alternative splicing in tomato pollen in response to heat stress. DNA Res **24**: 205–217

Keller M, SPOT-ITN Consortium, Simm S (2018) The coupling of transcriptome and proteome adaptation during development and heat stress response of tomato pollen. BMC Genomics **19**: 447

Khoury GA, Baliban RC, Floudas CA (2011) Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. Sci Rep **1**: 90

Kim D, Langmead B, Salzberg SL (2015) HISAT: A fast spliced aligner with low memory requirements. Nat Methods **12**: 357–360

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

Kissen R, Øverby A, Winge P, Bones AM (2016) Allyl-isothiocyanate treatment induces a complex transcriptional reprogramming including heat stress, oxidative stress and plant defence responses in Arabidopsis thaliana. BMC Genomics **17**: 740

Kozomara A, Griffiths-Jones S (2014) MiRBase: Annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res **42**: D68–D73

Kramer MF (2011) Stem-loop RT-qPCR for miRNAs. Curr Protoc Mol Biol **2011**: Unit 15.10

Kukurba KR, Montgomery SB (2015) RNA sequencing and analysis. Cold Spring Harb Protoc **2015**: 951–969

Kumar M, Padula MP, Davey P, Pernice M, Jiang Z, Sablok G, Contreras-Porcia L, Ralph PJ (2017) Proteome Analysis Reveals Extensive Light Stress-Response Reprogramming in the Seagrass Zostera muelleri (Alismatales, Zosteraceae) Metabolism. Front Plant Sci **7**: 2023

Kumar RR, Goswami S, Sharma SK, Kala YK, Rai GK, Mishra DC, Grover M, Singh GP, Pathak H, Rai A, et al (2015) Harnessing Next Generation Sequencing in Climate Change: RNA-Seq Analysis of Heat Stress-Responsive Genes in Wheat (*Triticum aestivum* L.). Omi A J Integr Biol **19**: 632–647

Kurosaki T, Maquat LE (2016) Nonsense-mediated mRNA decay in humans at a glance. J Cell Sci **129**: 461–467

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods **9**: 357–359

Lauter N, Kampani A, Carlson S, Goebel M, Moose SP (2005) microRNA172 down-regulates glossy15 to promote vegetative phase change in maize. Proc Natl Acad Sci **102**: 9412–9417

Lê S, Josse J, Husson F (2008) **FactoMineR** : An *R* Package for Multivariate Analysis. J Stat Softw **25**: 1–18

Lee DG, Ahsan N, Lee SH, Kyu YK, Jeong DB, Lee IJ, Lee BH (2007) A proteomic approach in analyzing heat-responsive proteins in rice leaves. Proteomics **7**: 3369–3383

Lee NH, Weinstock KG, Kirkness EF, Earle-Hughes JA, Fuldner RA, Marmaros S, Glodek A, Gocayne JD, Adams MD, Kerlavage AR (1995) Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor

treatment. Proc Natl Acad Sci **92**: 8303–8307

Lei R, Li X, Ma Z, Lv Y, Hu Y, Yu D (2017) Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. Plant J **91**: 962–976

Lenz D, May P, Walther D (2011) Comparative analysis of miRNAs and their targets across four plant species. BMC Res Notes **4**: 483

Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, Friedman N, Gnirke A, Regev A (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat Methods **7**: 709–715

Li C, Zhang B (2016) MicroRNAs in Control of Plant Development. J Cell Physiol **231**: 303–313

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078–2079

Li J, Soroka J, Buchner J (2012) The Hsp90 chaperone machinery: Conformational dynamics and regulation by co-chaperones. Biochim Biophys Acta - Mol Cell Res **1823**: 624–635

Li S-B, Xie Z-Z, Hu C-G, Zhang J-Z (2016a) A Review of Auxin Response Factors (ARFs) in Plants. Front Plant Sci **7**: 47

Li W, Wei Z, Qiao Z, Wu Z, Cheng L, Wang Y (2013a) Proteomics analysis of alfalfa response to heat stress. PLoS One **8**: e82725

Li X, Brock GN, Rouchka EC, Cooper NGF, Wu D, O'Toole TE, Gill RS, Eteleeb AM, O'Brien L, Rai SN (2017) A comparison of per sample global scaling and per gene normalization methods for differential expression analysis of RNA-seq data. PLoS One **12**: e0176185

Li X, Shahid M, Wu J, Wang L, Liu X, Lu Y (2016b) Comparative Small RNA Analysis of Pollen Development in Autotetraploid and Diploid Rice. Int J Mol Sci **17**: 499

Li Y-F, Wang Y, Tang Y, Kakani V, Mahalingam R (2013b) Transcriptome analysis of heat stress response in switchgrass (Panicum virgatum L.). BMC Plant Biol **13**: 153

Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, Axtell MJ, Zhang W, Sunkar R (2010) Transcriptome-wide identification of microRNA targets in rice. Plant J **62**: 742–759

Liang G, Yang F, Yu D (2010) MicroRNA395 mediates regulation of sulfate accumulation and allocation in Arabidopsis thaliana. Plant J **62**: 1046–1057

Liao J-L, Zhou H-W, Zhang H-Y, Zhong P-A, Huang Y-J (2014) Comparative proteomic analysis of differentially expressed proteins in the early milky stage of rice grains during high temperature stress. J Exp Bot **65**: 655–671

Lilley KS, Razzaq A, Dupree P (2002) Two-dimensional gel electrophoresis: Recent advances in sample preparation, detection and quantitation. Curr Opin Chem Biol **6**: 46–50

Lima RB, Dos Santos TB, Vieira LGE, Ferrarese MDLL, Ferrarese-Filho O, Donatti L, Boeger MRT, Petkowicz CLDO (2013) Heat stress causes alterations in the cell-wall polymers and anatomy of coffee leaves (Coffea arabica L.). Carbohydr Polym **93**: 135–143

Liu G-T, Wang J-F, Cramer G, Dai Z-W, Duan W, Xu H-G, Wu B-H, Fan P-G, Wang L-J, Li S-H (2012a) Transcriptomic analysis of grape (Vitis vinifera L.) leaves during and after recovery from heat stress. BMC Plant Biol **12**: 174

Liu H -c., Charng Y -y. (2013) Common and Distinct Functions of Arabidopsis Class A1 and A2 Heat Shock Factors in Diverse Abiotic Stress Responses and Development. Plant Physiol **163**: 276–290

Liu H, Wang Y, Li X, Zhang Y, Li J, Zheng Y, Liu M, Song X, Li X (2013) Expression and regulatory function of miRNA-182 in triple-negative breast cancer cells through its targeting of profilin 1.

Tumour Biol **34**: 1713–22

Liu X, Zou L, Zhu L, Zhang H, Du C, Li Z, Gao C, Zhao X, Bao S, Zheng H (2012b) MiRNA mediated up-regulation of cochaperone p23 acts as an anti-apoptotic factor in childhood acute lymphoblastic leukemia. Leuk Res **36**: 1098–1104

Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: A fast and simple web server for genome scale functional annotation of plant sequence data. Plant, Cell Environ **37**: 1250–1258

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA Package 2.0. Algorithms Mol Biol **6**: 26

Love DR, Pichler FB, Dodd A, Copp BR, Greenwood DR (2004) Technology for high-throughput screens: The present and future using zebrafish. Curr Opin Biotechnol **15**: 564–571

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**: 550

Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T (2017) Transcriptomics technologies. PLoS Comput Biol **13**: e1005457

Ma C, Burd S, Lers A (2015) MiR408 is involved in abiotic stress responses in Arabidopsis. Plant J **84**: 169–187

Magwanga RO, Lu P, Kirungu JN, Dong Q, Hu Y, Zhou Z, Cai X, Wang X, Hou Y, Wang K, et al (2018) Cotton Late Embryogenesis Abundant (LEA2) Genes Promote Root Growth and Confer Drought Stress Tolerance in Transgenic Arabidopsis thaliana. G3 (Bethesda) **8**: 2781–2803

Mallory AC, Bartel DP, Bartel B (2005) MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. Plant Cell **17**: 1360–75

Malone JH, Oliver B (2011) Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol **9**: 34

Mangelsen E, Kilian J, Harter K, Jansson C, Wanke D, Sundberg E (2011) Transcriptome analysis of high-temperature stress in developing barley caryopses: Early stress responses and effects on storage compound biosynthesis. Mol Plant **4**: 97–115

Manzano S, Martínez C, Megías Z, Gómez P, Garrido D, Jamilena M (2011) The role of ethylene and brassinosteroids in the control of sex expression and flower development in Cucurbita pepo. Plant Growth Regul **65**: 213–221

Marrs KA (1996) THE FUNCTIONS AND REGULATION OF GLUTATHIONE S-TRANSFERASES IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol **47**: 127–158

Martinez-Godoy MA, Mauri N, Juarez J, Marques MC, Santiago J, Forment J, Gadea J (2008) A genome-wide 20 K citrus microarray for gene expression analysis. BMC Genomics **9**: 318

Martínez-Gómez P, Sánchez-Pérez R, Rubio M (2012) Clarifying Omics Concepts, Challenges, and Opportunities for *Prunus* Breeding in the Postgenomic Era. Omi A J Integr Biol **16**: 268–283

Mascarenhas JP, Crone DE (1996) Pollen and the heat shock response. Sex Plant Reprod **9**: 370–374

McManus J, Cheng Z, Vogel C (2015) Next-generation analysis of gene expression regulation – comparing the roles of synthesis and degradation. Mol BioSyst **11**: 2680–2689

Merchante C, Stepanova AN, Alonso JM (2017) Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J **90**: 628–653

Metzker ML (2010) Sequencing technologies the next generation. Nat Rev Genet **11**: 31–46

Milla MAR, Butler E, Huete AR, Wilson CF, Anderson O, Gustafson JP (2002) Expressed sequence tag-based gene expression analysis under aluminum stress in rye. Plant Physiol **130**: 1706–16

Mitsuda N, Isono T, Sato MH (2003) Arabidopsis CAMTA Family Proteins Enhance V-PPase Expression in Pollen. Plant Cell Physiol **44**: 975–981

Mlotshwa S, Yang Z, Kim Y, Chen X (2006) Floral patterning defects induced by Arabidopsis APETALA2 and microRNA172 expression in Nicotiana benthamiana. Plant Mol Biol **61**: 781–793

Moll P, Ante M, Seitz A, Reda T (2014) QuantSeq 3′ mRNA sequencing for RNA quantification. Nat Methods **11**: i–iii

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods **5**: 621–628

Nachtergaele S, He C (2017) The emerging biology of RNA post-transcriptional modifications. RNA Biol **14**: 156–163

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science (80-) **320**: 1344–1349

Nagaraj SH, Gasser RB, Ranganathan S (2007) A hitchhiker's guide to expressed sequence tag (EST) analysis. Brief Bioinform **8**: 6–21

Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, Van Sluyter SC, Haynes PA (2011) Less label, more free: Approaches in label-free quantitative mass spectrometry. Proteomics **11**: 535–553

Nishizawa-Yokoi A, Yoshida E, Yabuta Y, Shigeoka S (2009) Analysis of the regulation of target genes by an Arabidopsis heat shock transcription factor, HsfA2. Biosci Biotechnol Biochem **73**: 890–5

Niu J, Wang J, An J, Liu L, Lin Z, Wang R, Wang L, Ma C, Shi L, Lin S (2016) Integrated mRNA and miRNA transcriptome reveal a cross-talk between developing response and hormone signaling for the seed kernels of Siberian apricot. Sci Rep **6**: 35675

Noir S, Bräutigam A, Colby T, Schmidt J, Panstruga R (2005) A reference map of the Arabidopsis thaliana mature pollen proteome. Biochem Biophys Res Commun **337**: 1257–1266

Nover L, Bharti K, Döring P, Mishra SK, Ganguli A, Scharf KD (2001) Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? Cell Stress Chaperones **6**: 177–89

Nowrousian M (2010) Next-generation sequencing techniques for eukaryotic microorganisms: Sequencing-based solutions to biological problems. Eukaryot Cell **9**: 1300–1310

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem **250**: 4007–21

O'Neil D, Glowatz H, Schlumpberge M (2013) Ribosomal RNA depletion for efficient use of RNA-seq capacity. Curr Protoc Mol Biol **Chapter 4**: Unit 4.19

Ozyigit II, Filiz E, Vatansever R, Kurtoglu KY, Koc I, Öztürk MX, Anjum NA (2016) Identification and Comparative Analysis of H2O2-Scavenging Enzymes (Ascorbate Peroxidase and Glutathione Peroxidase) in Selected Plants Employing Bioinformatics Approaches. Front Plant Sci **7**: 301

Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, Pezzotti M, Paci P (2014) Integrated Network Analysis Identifies Fight-Club Nodes as a Class of Hubs Encompassing Key Putative Switch Genes That Induce Major Transcriptome Reprogramming during Grapevine Development. Plant Cell Online **26**: 4617–4635

Park S, Sugimoto N, Larson MD, Beaudry R, van Nocker S (2006) Identification of genes with potential roles in apple fruit development and biochemistry through large-scale statistical analysis of expressed sequence tags. Plant Physiol **141**: 811–24

Paul P, Chaturvedi P, Selymesi M, Ghatak A, Mesihovic A, Scharf KD, Weckwerth W, Simm S, Schleiff E (2016) The membrane proteome of male gametophyte in Solanum lycopersicum. J Proteomics **131**: 48–60

Payton P, Kottapalli KR, Rowland D, Faircloth W, Guo B, Burow M, Puppala N, Gallo M (2009) Gene expression profiling in peanut using high density oligonucleotide microarrays. BMC Genomics **10**: 265

Piras V, Tomita M, Selvarajoo K (2012) Is central dogma a global property of cellular information flow? Front Physiol **3**: 439

Preston JC, Hileman LC (2013) Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. Front Plant Sci **4**: 80

Proudfoot NJ (2000) Connecting transcription to messenger RNA processing. Trends Biochem Sci **25**: 290–293

Qian M, Ni J, Niu Q, Bai S, Bao L, Li J, Sun Y, Zhang D, Teng Y (2017) Response of miR156-SPL Module during the Red Peel Coloration of Bagging-Treated Chinese Sand Pear (Pyrus pyrifolia Nakai). Front Physiol **8**: 550

Qin D, Wu H, Peng H, Yao Y, Ni Z, Li Z, Zhou C, Sun Q (2008) Heat stress-responsive transcriptome analysis in heat susceptible and tolerant wheat (Triticum aestivum L.) by using Wheat Genome Array. BMC Genomics **9**: 432

Qu X, Jiang Y, Chang M, Liu X, Zhang R, Huang S (2015) Organization and regulation of the actin cytoskeleton in the pollen tube. Front Plant Sci **5**: 786

Rahmati Ishka M, Brown E, Weigand C, Tillett RL, Schlauch KA, Miller G, Harper JF (2018) A comparison of heat-stress transcriptome changes between wild-type Arabidopsis pollen and a heat-sensitive mutant harboring a knockout of cyclic nucleotide-gated cation channel 16 (cngc16). BMC Genomics **19**: 549

Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol **14**: R95

Rauniyar N, Yates JR (2014) Isobaric labeling-based relative quantification in shotgun proteomics. J Proteome Res **13**: 5293–5309

Remm M, Storm CEV, Sonnhammer ELL (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. J Mol Biol **314**: 1041–1052

Rhoads DM, White SJ, Zhou Y, Muralidharan M, Elthon TE (2005) Altered gene expression in plants with constitutive expression of a mitochondrial small heat shock protein suggests the involvement of retrograde regulation in the heat stress response. Physiol Plant **123**: 435–444

Rieu I, Twell D, Firon N (2017) Pollen Development at High Temperature: From Acclimation to Collapse. Plant Physiol **173**: 1967–1976

Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**: 139–140

Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, Yan SF, Williamson KC, Holder AA, Carucci DJ, et al (2004) Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. Genome Res **14**: 2308–2318

Rogers K, Chen X (2013) Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. Plant Cell **25**: 2383–2399

Roy B, Arnim AG von (2013) Translational Regulation of Cytoplasmic mRNAs. Arab B **11**: e0165

Rutley N, Twell D (2015) A decade of pollen transcriptomics. Plant Reprod **28**: 73–89

Samad AFA, Sajad M, Nazaruddin N, Fauzi IA, Murad AMA, Zainal Z, Ismail I (2017) MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. Front Plant Sci **8**: 565

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci **74**: 5463–5467

Sato S, Peet MM, Thomas JF (2000) Physiological factors limit fruit set of tomato (Lycopersicon esculentum Mill.) under chronic, mild heat stress. Plant, Cell Environ **23**: 719–726

Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. Biochim Biophys Acta - Gene Regul Mech **1819**: 104–119

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. Science (80-) **270**: 467–470

Schurch NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, Wrobel N, Gharbi K, Simpson GG, Owen-Hughes T, et al (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? Rna **22**: 839–851

Schwanhüusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature **473**: 337–342

Scott RJ, Spielman M, Dickinson HG (2004) Stamen structure and function. Plant Cell **16**: S46-60

Sedlazeck FJ, Rescheneder P, Von Haeseler A (2013) NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. Bioinformatics **29**: 2790–2791

Selinski J, Scheibe R (2014) Pollen tube growth: where does the energy come from? Plant Signal Behav **9**: e977200

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res **13**: 2498–2504

Sheoran IS, Sproule KA, Olson DJH, Ross ARS, Sawhney VK (2006) Proteome profile and functional classification of proteins in Arabidopsis thaliana (Landsberg erecta) mature pollen. Sex Plant Reprod **19**: 185–196

Shi Z, Fujii K, Kovary KM, Genuth NR, Röst HL, Teruel MN, Barna M (2017) Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. Mol Cell **67**: 71–83.e7

Shukla LI, Chinnusamy V, Sunkar R (2008) The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochim Biophys Acta - Gene Regul Mech **1779**: 743–748

Shyu A Bin, Wilkinson MF, Van Hoof A (2008) Messenger RNA regulation: To translate or to degrade. EMBO J **27**: 471–481

Simm S, Fragkostefanakis S, Paul P, Keller M, Einloft J, Scharf KD, Schleiff E (2015) Identification and expression analysis of ribosome biogenesis factor co-orthologs in solanum lycopersicum. Bioinform Biol Insights **9**: 1–17

Simm S, Scharf KD, Jegadeesan S, Chiusano ML, Firon N, Schleiff E (2016) Survey of genes involved in biosynthesis, transport, and signaling of phytohormones with focus on solanum lycopersicum. Bioinform Biol Insights **10**: 185–207

Slewinski TL (2011) Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: A physiological perspective. Mol Plant **4**: 641–662

Slobodin B, Gerst JE (2010) A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. Rna **16**: 2277–2290

Smith CJ, Osborn AM (2009) Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. FEMS Microbiol Ecol **67**: 6–20

Srivastava PK, Moturu TR, Pandey P, Baldwin IT, Pandey SP (2014) A comparison of performance of plant miRNA target prediction tools and the characterization of features for genome-wide target prediction. BMC Genomics **15**: 348

Standart N, Weil D (2018) P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. Trends Genet **34**: 612–626

De Storme N, Geelen D (2014) The impact of environmental stress on male reproductive development in plants: biological processes and molecular mechanisms. Plant Cell Environ **37**: 1–18

Strompen G, Dettmer J, Stierhof YD, Schumacher K, Jürgens G, Mayer U (2005) Arabidopsis vacuolar H+-ATPase subunit E isoform 1 is required for Golgi organization and vacuole function in embryogenesis. Plant J **41**: 125–132

Sunkar R, Kapoor A, Zhu J-K (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell **18**: 2051–65

Sunkar R, Zhu J-K (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell **16**: 2001–19

Swindell WR, Huebner M, Weber AP (2007) Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. BMC Genomics **8**: 125

Tanabe N, Yoshimura K, Kimura A, Yabuta Y, Shigeoka S (2007) Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 and atSR45a, in response to environmental stress. Plant Cell Physiol **48**: 1036–1049

Tauber S, Von Haeseler A (2013) Exploring the sampling universe of RNA-seq. Stat Appl Genet Mol Biol **12**: 175–188

Tebaldi T, Re A, Viero G, Pegoretti I, Passerini A, Blanzieri E, Quattrone A (2012) Widespread uncoupling between transcriptome and translatome variations after a stimulus in mammalian cells. BMC Genomics **13**: 220

Tian GW, Chen MH, Zaltsman A, Citovsky V (2006) Pollen-specific pectin methylesterase involved in pollen tube growth. Dev Biol **294**: 83–91

Tipney H, Hunter L (2010) An introduction to effective use of enrichment analysis software. Hum Genomics **4**: 202–206

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol **28**: 511–515

Tripathi RK, Bregitzer P, Singh J (2018) Genome-wide analysis of the SPL/miR156 module and its interaction with the AP2/miR172 unit in barley. Sci Rep **8**: 7085

Tripathi V, Darnauer S, Hartwig NR, Obermann WMJ (2014) Aha1 can act as an autonomous chaperone to prevent aggregation of stressed proteins. J Biol Chem **289**: 36220–36228

Trivellini A, Jibran R, Watson LM, O'Donoghue EM, Ferrante A, Sullivan KL, Dijkwel PP, Hunter DA (2012) Carbon Deprivation-Driven Transcriptome Reprogramming in Detached Developmentally Arresting Arabidopsis Inflorescences. Plant Physiol **160**: 1357–1372

Unte US, Sorensen A-M, Pesaresi P, Gandikota M, Leister D, Saedler H, Huijser P (2003) SPL8,

an SBP-box gene that affects pollen sac development in Arabidopsis. Plant Cell **15**: 1009–19

Usadel B, Schlüter U, Mølhøj M, Gipmans M, Verma R, Kossmann J, Reiter WD, Pauly M (2004) Identification and characterization of a UDP-D-glucuronate 4-epimerase in Arabidopsis. FEBS Lett **569**: 327–331

Valdés-López O, Batek J, Gomez-Hernandez N, Nguyen CT, Isidra-Arellano MC, Zhang N, Joshi T, Xu D, Hixson KK, Weitz KK, et al (2016) Soybean Roots Grown under Heat Stress Show Global Changes in Their Transcriptional and Proteomic Profiles. Front Plant Sci **7**: 517

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, et al (2013) A Regulatory Network-Based Approach Dissects Late Maturation Processes Related to the Acquisition of Desiccation Tolerance and Longevity of Medicago truncatula Seeds. Plant Physiol **163**: 757–774

Verza NC, Rezende E Silva T, Neto GC, Nogueira FTS, Fisch PH, De Rosa VE, Rebello MM, Vettore AL, Da Silva FR, Arruda P (2005) Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags. Plant Mol Biol **59**: 363–374

Vij S, Tyagi AK (2007) Emerging trends in the functional genomics of the abiotic stress response in crop plants. Plant Biotechnol J **5**: 361–380

Vogler F, Schmalzl C, Englhart M, Bircheneder M, Sprunck S (2014) Brassinosteroids promote Arabidopsis pollen germination and growth. Plant Reprod **27**: 153–167

Volkov RA, Panchuk II, Mullineaux PM, Schöffl F (2006) Heat stress-induced H2O2 is required for effective expression of heat shock genes in Arabidopsis. Plant Mol Biol **61**: 733–746

Wang P, Wilson SR (2013) Mass spectrometry-based protein identification by integrating de novo sequencing with database searching. BMC Bioinformatics **14 Suppl 2**: S24

Wang X-J, Reyes JL, Chua N-H, Gaasterland T (2004) Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. Genome Biol **5**: R65

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. Nat Rev Genet **10**: 57–63

Wang Z, Wang Y, Kohalmi SE, Amyot L, Hannoufa A (2016) SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2 controls floral organ development and plant fertility by activating ASYMMETRIC LEAVES 2 in Arabidopsis thaliana. Plant Mol Biol **92**: 661–674

Wei LQ, Xu WY, Deng ZY, Su Z, Xue Y, Wang T (2010) Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in Oryza sativa. BMC Genomics **11**: 338

White JA, Todd J, Newman T, Focks N, Girke T, de Ilárduya OM, Jaworski JG, Ohlrogge JB, Benning C (2000) A New Set of Arabidopsis Expressed Sequence Tags from Developing Seeds. The Metabolic Pathway from Carbohydrates to Seed Oil. Plant Physiol **124**: 1582–1594

Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bähler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature **453**: 1239–1243

Wu Y, Yan J, Zhang R, Qu X, Ren S, Chen N, Huang S (2010) Arabidopsis FIMBRIN5, an Actin Bundling Factor, Is Required for Pollen Germination and Pollen Tube Growth. Plant Cell Online **22**: 3745–3763

Wu Z, Wang X, Zhang X (2011) Using non-uniform read distribution models to improve isoform expression inference in RNA-Seq. Bioinformatics **27**: 502–508

Xia C, Wang YJ, Li WQ, Chen YR, Deng Y, Zhang XQ, Chen LQ, Ye D (2010) The Arabidopsis eukaryotic translation initiation factor 3, subunit F (AteIF3f), is required for pollen germination and embryogenesis. Plant J **63**: 189–202

Xiang D, Datla R, Li F, Cutler A, Malik MR, Krochko JE, Sharma N, Fobert P, Georges F, Selvaraj G, et al (2008) Development of a *Brassica* seed cDNA microarray. Genome **51**: 236–242

Xing S, Salinas M, Höhmann S, Berndtgen R, Huijser P (2010) miR156-Targeted and Nontargeted SBP-Box Transcription Factors Act in Concert to Secure Male Fertility in *Arabidopsis*. Plant Cell **22**: 3935–3950

Xue S, Barna M (2012) Specialized ribosomes: A new frontier in gene regulation and organismal biology. Nat Rev Mol Cell Biol **13**: 355–369

Yadegari R, Drews GN (2004) Female gametophyte development. Plant Cell **16**: S133-41

Yamada K, Nishimura M (2008) Cytosolic heat shock protein 90 regulates heat shock transcription factor in Arabidopsis thaliana. Plant Signal Behav **3**: 660–662

Yamamoto Y, Nishimura M, Hara-Nishimura I, Noguchi T (2003) Behavior of Vacuoles during Microspore and Pollen Development in Arabidopsis thaliana. Plant Cell Physiol **44**: 1192–1201

Yamasaki H, Abdel-Ghany SE, Cohu CM, Kobayashi Y, Shikanai T, Pilon M (2007) Regulation of copper homeostasis by micro-RNA in Arabidopsis. J Biol Chem **282**: 16369–16378

Yan H, Chen D, Wang Y, Sun Y, Zhao J, Sun M, Peng X (2016) Ribosomal protein L18aB is required for both male gametophyte function and embryo development in Arabidopsis. Sci Rep **6**: 31195

Yang CJ, Zhang C, Lu YN, Jin JQ, Wang XL (2011) The mechanisms of brassinosteroids' action: From signal transduction to plant development. Mol Plant **4**: 588–600

Yang H, Kaur N, Kiriakopolos S, McCormick S (2006) EST generation and analyses towards identifying female gametophyte-specific genes in Zea mays L. Planta **224**: 1004–1014

Yang IS, Kim S (2015) Analysis of Whole Transcriptome Sequencing Data: Workflow and Software. Genomics Inform **13**: 119–25

Yang J, Tian L, Sun M-X, Huang X-Y, Zhu J, Guan Y-F, Jia Q-S, Yang Z-N (2013) AUXIN RESPONSE FACTOR17 Is Essential for Pollen Wall Pattern Formation in Arabidopsis. Plant Physiol **162**: 720–731

Yang YH, Speed T (2002) Design issues for cDNA microarray experiments. Nat Rev Genet **3**: 579–588

Yao Y, Guo G, Ni Z, Sunkar R, Du J, Zhu JK, Sun Q (2007) Cloning and characterization of microRNAs from wheat (Triticum aestivum L.). Genome Biol **8**: R96

Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. Annu Rev Biomed Eng **11**: 49–79

Ye Q, Zhu W, Li L, Zhang S, Yin Y, Ma H, Wang X (2010) Brassinosteroids control male fertility by regulating the expression of key genes involved in Arabidopsis anther and pollen development. Proc Natl Acad Sci **107**: 6100–6105

Yu LX, Parthasarathy M V. (2014) Molecular and cellular characterization of the tomato pollen profilin, LePro1. PLoS One **9**: e86505

Zawada AM, Rogacev KS, Müller S, Rotter B, Winter P, Fliser D, Heine GH (2014) Massive analysis of cDNA Ends (MACE) and miRNA expression profiling identifies proatherogenic pathways in chronic kidney disease. Epigenetics **9**: 161–172

Zhang B, Pan X, Cobb GP, Anderson TA (2006) Plant microRNA: A small regulatory molecule with big impact. Dev Biol **289**: 3–16

Zhang H, Qu X, Bao C, Khurana P, Wang Q, Xie Y, Zheng Y, Chen N, Blanchoin L, Staiger CJ, et al (2010) Arabidopsis VILLIN5, an Actin Filament Bundling and Severing Protein, Is Necessary for Normal Pollen Tube Growth. Plant Cell **22**: 2749–2767

Zhang J, Liu B, Li J, Zhang L, Wang Y, Zheng H, Lu M, Chen J (2015a) Hsf and Hsp gene families in Populus: Genome-wide identification, organization and correlated expression during development and in stress responses. BMC Genomics **16**: 181

Zhang W, Zhou R-G, Gao Y-J, Zheng S-Z, Xu P, Zhang S-Q, Sun D-Y (2009) Molecular and Genetic Evidence for the Key Role of AtCaM3 in Heat-Shock Signal Transduction in Arabidopsis. Plant Physiol **149**: 1773–1784

Zhang X, Gao X, Coots RA, Conn CS, Liu B, Qian SB (2015b) Translational control of the cytosolic stress response by mitochondrial ribosomal protein L18. Nat Struct Mol Biol **22**: 404–410

Zhang Y, Pan J, Huang X, Guo D, Lou H, Hou Z, Su M, Liang R, Xie C, You M, et al (2017a) Differential effects of a post-anthesis heat stress on wheat (Triticum aestivum L.) grain proteome determined by iTRAQ. Sci Rep **7**: 3468

Zhang Z, Hu M, Feng X, Gong A, Cheng L, Yuan H (2017b) Proteomes and Phosphoproteomes of Anther and Pollen: Availability and Progress. Proteomics **17**: 1600458

Zheng D, Sabbagh JJ, Blair LJ, Darling AL, Wen X, Dickey CA (2016) MicroRNA-511 binds to FKBP5 mRNA, which encodes a chaperone protein, and regulates neuronal differentiation. J Biol Chem **291**: 17897–17906

Zhou R, Wang Q, Jiang F, Cao X, Sun M, Liu M, Wu Z (2016) Identification of miRNAs and their targets in wild tomato at moderately and acutely elevated temperatures by high-throughput sequencing and degradome analysis. Sci Rep **6**: 33777

Zhu J, Nan Q, Qin T, Qian D, Mao T, Yuan S, Wu X, Niu Y, Bai Q, An L, et al (2017) Higher-Ordered Actin Structures Remodeled by Arabidopsis ACTIN-DEPOLYMERIZING FACTOR5 Are Important for Pollen Germination and Pollen Tube Growth. Mol Plant **10**: 1065–1081

Zhu W, Schlueter SD, Brendel V (2003) Refined Annotation of the Arabidopsis Genome by Complete Expressed Sequence Tag Mapping. Plant Physiol **132**: 469–484

Zinn KE, Tunc-Ozdemir M, Harper JF (2010) Temperature stress and plant sexual reproduction: Uncovering the weakest links. J Exp Bot **61**: 1959–1968

Zou J, Song L, Zhang W, Wang Y, Ruan S, Wu WH (2009) Comparative proteomic analysis of arabidopsis mature pollen and germinated pollen. J Integr Plant Biol **51**: 438–455

7 Supplemental Material

Supplemental Table 1: Distribution of stage-accumulated mRNAs among MapMan terms
Shown is he distribution of mRNAs accumulated in tetrads (T, column 3), tetrads and post-meiotic (T + PM, column 4), post-meiotic (PM, column 5), post-meiotic and mature (PM + M, column 6), mature (M, column 7) as well as all annotated genes (Genome, column 8) among the MapMan terms of the second hierarchy level (column 2). Terms shown to be significantly enriched (adjusted p-value less than 0.05 in a Fisher exact test) are highlighted in bold.

1st level	2nd level	T (7,203)	T + PM (1,860)	PM (244)	PM + M (1,387))	M (1,906	Genome (35,768)
no prediction		21 (0.29%)	9 (0.48%)	3 (1.23%)	8 (0.58%)	11 (0.58%)	206 (0.58%)
PS	lightreaction	**92 (1.28%)**	9 (0.48%)	0 (0%)	2 (0.14%)	2 (0.1%)	279 (0.78%)
	photorespiration	9 (0.12%)	3 (0.16%)	1 (0.41%)	2 (0.14%)	2 (0.1%)	32 (0.09%)
	calvin cycle	**22 (0.31%)**	7 (0.38%)	0 (0%)	1 (0.07%)	1 (0.05%)	61 (0.17%)
major CHO metabolism	synthesis	5 (0.07%)	3 (0.16%)	**5 (2.05%)**	5 (0.36%)	0 (0%)	39 (0.11%)
	degradation	19 (0.26%)	**11 (0.59%)**	1 (0.41%)	**10 (0.72%)**	9 (0.47%)	77 (0.22%)
minor CHO metabolism	raffinose family	5 (0.07%)	1 (0.05%)	0 (0%)	0 (0%)	3 (0.16%)	15 (0.04%)
	trehalose	8 (0.11%)	1 (0.05%)	0 (0%)	0 (0%)	1 (0.05%)	22 (0.06%)
	sugar alcohols	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	myo-inositol	4 (0.06%)	0 (0%)	1 (0.41%)	**5 (0.36%)**	5 (0.26%)	27 (0.08%)
	others	12 (0.17%)	4 (0.22%)	1 (0.41%)	3 (0.22%)	2 (0.1%)	46 (0.13%)
	callose	4 (0.06%)	0 (0%)	0 (0%)	0 (0%)	**4 (0.21%)**	13 (0.04%)
	sugar kinases	1 (0.01%)	0 (0%)	1 (0.41%)	0 (0%)	0 (0%)	3 (0.01%)
	galactose	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	6 (0.02%)
	misc	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
glycolysis	cytosolic branch	16 (0.22%)	2 (0.11%)	0 (0%)	**13 (0.94%)**	4 (0.21%)	68 (0.19%)
	plastid branch	3 (0.04%)	1 (0.05%)	0 (0%)	**6 (0.43%)**	2 (0.1%)	21 (0.06%)
	unclear/dually targeted	7 (0.1%)	0 (0%)	0 (0%)	3 (0.22%)	1 (0.05%)	21 (0.06%)
fermentation	LDH	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	2 (0.01%)
	PDC	3 (0.04%)	1 (0.05%)	0 (0%)	1 (0.07%)	0 (0%)	5 (0.01%)
	ADH	1 (0.01%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	7 (0.02%)
	aldehyde dehydrogenase	7 (0.1%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	15 (0.04%)
gluconeogenesis / glyoxylate cycle	citrate synthase	1 (0.01%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	2 (0.01%)
	malate synthase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	Malate DH	1 (0.01%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	2 (0.01%)
	PEPCK	1 (0.01%)	0 (0%)	1 (0.41%)	0 (0%)	0 (0%)	2 (0.01%)
	pyruvate dikinase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	isocitrate lyase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
OPP	oxidative PP	5 (0.07%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	17 (0.05%)

	non-reductive PP	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	10 (0.03%)
	electron transfer	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (0.01%)
TCA / org transformation.	TCA	4 (0.06%)	0 (0%)	0 (0%)	**18 (1.3%)**	2 (0.1%)	46 (0.13%)
	other organic acid transformations	5 (0.07%)	1 (0.05%)	0 (0%)	**6 (0.43%)**	0 (0%)	23 (0.06%)
	carbonic anhydrases	0 (0%)	0 (0%)	1 (0.41%)	1 (0.07%)	2 (0.1%)	14 (0.04%)
mitochondrial electron transport / ATP synthesis	NADH-DH	4 (0.06%)	3 (0.16%)	0 (0%)	**12 (0.87%)**	1 (0.05%)	73 (0.2%)
	NADH-DH	2 (0.03%)	0 (0%)	0 (0%)	2 (0.14%)	0 (0%)	8 (0.02%)
	electron transfer flavoprotein	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	alternative oxidase	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	1 (0.05%)	4 (0.01%)
	cytochrome c reductase	0 (0%)	0 (0%)	1 (0.41%)	**4 (0.29%)**	1 (0.05%)	10 (0.03%)
	cytochrome c	3 (0.04%)	1 (0.05%)	0 (0%)	**5 (0.36%)**	1 (0.05%)	23 (0.06%)
	cytochrome c oxidase	4 (0.06%)	3 (0.16%)	1 (0.41%)	**6 (0.43%)**	0 (0%)	40 (0.11%)
	uncoupling protein	2 (0.03%)	0 (0%)	0 (0%)	1 (0.07%)	1 (0.05%)	4 (0.01%)
	F1-ATPase	2 (0.03%)	1 (0.05%)	1 (0.41%)	**9 (0.65%)**	0 (0%)	29 (0.08%)
	unspecified	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	1 (0.05%)	4 (0.01%)
cell wall	precursor synthesis	12 (0.17%)	1 (0.05%)	0 (0%)	**11 (0.79%)**	**13 (0.68%)**	67 (0.19%)
	cellulose synthesis	19 (0.26%)	2 (0.11%)	0 (0%)	2 (0.14%)	8 (0.42%)	68 (0.19%)
	hemicellulose synthesis	8 (0.11%)	0 (0%)	0 (0%)	3 (0.22%)	2 (0.1%)	35 (0.1%)
	pectin synthesis	0 (0%)	0 (0%)	0 (0%)	2 (0.14%)	0 (0%)	10 (0.03%)
	cell wall proteins	**21 (0.29%)**	0 (0%)	1 (0.41%)	1 (0.07%)	**9 (0.47%)**	52 (0.15%)
	degradation	47 (0.65%)	8 (0.43%)	1 (0.41%)	4 (0.29%)	**24 (1.26%)**	173 (0.48%)
	modification	24 (0.33%)	1 (0.05%)	0 (0%)	2 (0.14%)	3 (0.16%)	83 (0.23%)
	pectin*esterases	27 (0.37%)	4 (0.22%)	0 (0%)	2 (0.14%)	**18 (0.94%)**	107 (0.3%)
lipid metabolism	FA synthesis and FA elongation	29 (0.4%)	10 (0.54%)	1 (0.41%)	**17 (1.23%)**	4 (0.21%)	128 (0.36%)
	FA desaturation	0 (0%)	1 (0.05%)	0 (0%)	1 (0.07%)	**5 (0.26%)**	19 (0.05%)
	Phospholipid synthesis	15 (0.21%)	0 (0%)	1 (0.41%)	6 (0.43%)	8 (0.42%)	56 (0.16%)
	TAG synthesis	2 (0.03%)	2 (0.11%)	0 (0%)	**3 (0.22%)**	1 (0.05%)	11 (0.03%)
	glyceral metabolism	2 (0.03%)	0 (0%)	0 (0%)	2 (0.14%)	2 (0.1%)	6 (0.02%)
	lipid transfer proteins etc	8 (0.11%)	1 (0.05%)	0 (0%)	1 (0.07%)	4 (0.21%)	16 (0.04%)
	unassigned	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	exotics(steroids, squalene etc)	17 (0.24%)	2 (0.11%)	0 (0%)	**11 (0.79%)**	**20 (1.05%)**	119 (0.33%)
	lipid degradation	44 (0.61%)	6 (0.32%)	1 (0.41%)	**20 (1.44%)**	16 (0.84%)	211 (0.59%)
	glycolipid synthesis	5 (0.07%)	1 (0.05%)	0 (0%)	0 (0%)	2 (0.1%)	11 (0.03%)
N-metabolism	nitrate metabolism	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	3 (0.01%)
	ammonia metabolism	3 (0.04%)	0 (0%)	1 (0.41%)	**4 (0.29%)**	0 (0%)	20 (0.06%)
	N-degradation	2 (0.03%)	2 (0.11%)	1 (0.41%)	0 (0%)	0 (0%)	5 (0.01%)
	misc	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	9 (0.03%)

amino acid metabolism	synthesis	32 (0.44%)	12 (0.65%)	3 (1.23%)	**21 (1.51%)**	13 (0.68%)	181 (0.51%)
	degradation	27 (0.37%)	5 (0.27%)	1 (0.41%)	6 (0.43%)	4 (0.21%)	127 (0.36%)
	misc	4 (0.06%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	9 (0.03%)
S-assimilation	ATPS	0 (0%)	0 (0%)	0 (0%)	**2 (0.14%)**	0 (0%)	3 (0.01%)
	APR	0 (0%)	0 (0%)	0 (0%)	**2 (0.14%)**	0 (0%)	3 (0.01%)
	sulfite redox	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	1 (0%)
	sulfite oxidase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	AKN	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (0.01%)
metal handling	acquisition	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	10 (0.03%)
	binding, chelation and storage	16 (0.22%)	0 (0%)	1 (0.41%)	2 (0.14%)	1 (0.05%)	48 (0.13%)
	regulation	2 (0.03%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	4 (0.01%)
secondary metabolism	isoprenoids	36 (0.5%)	3 (0.16%)	0 (0%)	6 (0.43%)	8 (0.42%)	138 (0.39%)
	phenylpropanoids	50 (0.69%)	2 (0.11%)	1 (0.41%)	6 (0.43%)	1 (0.05%)	201 (0.56%)
	N misc	14 (0.19%)	3 (0.16%)	0 (0%)	1 (0.07%)	1 (0.05%)	40 (0.11%)
	sulfur-containing	13 (0.18%)	2 (0.11%)	0 (0%)	4 (0.29%)	2 (0.1%)	104 (0.29%)
	wax	**16 (0.22%)**	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	40 (0.11%)
	flavonoids	54 (0.75%)	5 (0.27%)	1 (0.41%)	3 (0.22%)	4 (0.21%)	261 (0.73%)
	simple phenols	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	30 (0.08%)
	unspecified	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (0.01%)
hormone metabolim	abscisic acid	27 (0.37%)	3 (0.16%)	2 (0.82%)	4 (0.29%)	5 (0.26%)	100 (0.28%)
	auxin	48 (0.67%)	9 (0.48%)	1 (0.41%)	12 (0.87%)	16 (0.84%)	290 (0.81%)
	brassinosteroid	**40 (0.56%)**	3 (0.16%)	0 (0%)	4 (0.29%)	3 (0.16%)	112 (0.31%)
	cytokinin	14 (0.19%)	1 (0.05%)	1 (0.41%)	0 (0%)	1 (0.05%)	127 (0.36%)
	ethylene	38 (0.53%)	5 (0.27%)	1 (0.41%)	3 (0.22%)	8 (0.42%)	191 (0.53%)
	gibberelin	21 (0.29%)	0 (0%)	0 (0%)	3 (0.22%)	0 (0%)	107 (0.3%)
	jasmonate	20 (0.28%)	1 (0.05%)	2 (0.82%)	0 (0%)	1 (0.05%)	72 (0.2%)
	salicylic acid	4 (0.06%)	0 (0%)	1 (0.41%)	2 (0.14%)	2 (0.1%)	33 (0.09%)
Co-factor and vitamine metabolism	molybdenum cofactor	1 (0.01%)	**4 (0.22%)**	0 (0%)	0 (0%)	0 (0%)	8 (0.02%)
	thiamine	2 (0.03%)	1 (0.05%)	0 (0%)	1 (0.07%)	0 (0%)	5 (0.01%)
	riboflavin	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	9 (0.03%)
	pantothenate	6 (0.08%)	2 (0.11%)	0 (0%)	3 (0.22%)	1 (0.05%)	21 (0.06%)
	folate & vitamine K	3 (0.04%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	8 (0.02%)
	biotin	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (0.01%)
	iron-sulphur clusters	3 (0.04%)	1 (0.05%)	0 (0%)	0 (0%)	1 (0.05%)	13 (0.04%)
	ubiquinone	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	1 (0%)
	lipoic acid	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	3 (0.01%)
	NADH kinase	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	1 (0%)

	isochorismatase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
tetrapyrrole synthesis	glu-tRNA synthetase	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	glu-tRNA reductase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	GSA	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	ALA dehydratase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	porphobilinogen deaminase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	uroporphyrinogen III synthase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	uroporphyrinogen decarboxylase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	coproporphyrinogen III oxidase	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	protoporphyrin IX oxidase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	magnesium chelatase	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (0.01%)
	magnesium protoporphyrin IX methyltransferase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	magnesium-protoporphyrin IX monomethyl ester (oxidative) cyclase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	divinyl chlorophyllide-a 8-vinyl-reductase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	protochlorophyllide reductase	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	chlorophyll synthase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	chlorophyll b synthase	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	ferrochelatase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	6 (0.02%)
	heme oxygenase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	uroporphyrin-III C-methyltransferase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	sirohydrochlorin ferrochelatase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	regulation	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	unspecified	4 (0.06%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	9 (0.03%)
stress	biotic	118 (1.64%)	9 (0.48%)	3 (1.23%)	11 (0.79%)	27 (1.42%)	635 (1.78%)
	abiotic	**154 (2.14%)**	15 (0.81%)	2 (0.82%)	23 (1.66%)	**49 (2.57%)**	537 (1.5%)
redox	thioredoxin	**32 (0.44%)**	3 (0.16%)	0 (0%)	**8 (0.58%)**	6 (0.31%)	74 (0.21%)
	ascorbate and glutathione	22 (0.31%)	4 (0.22%)	2 (0.82%)	8 (0.58%)	5 (0.26%)	110 (0.31%)
	heme	3 (0.04%)	0 (0%)	1 (0.41%)	0 (0%)	0 (0%)	8 (0.02%)
	glutaredoxins	7 (0.1%)	2 (0.11%)	0 (0%)	2 (0.14%)	3 (0.16%)	54 (0.15%)
	peroxiredoxin	3 (0.04%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	7 (0.02%)
	dismutases and catalases	6 (0.08%)	1 (0.05%)	0 (0%)	1 (0.07%)	0 (0%)	15 (0.04%)
	misc	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	7 (0.02%)
polyamine metabolism	synthesis	5 (0.07%)	0 (0%)	0 (0%)	0 (0%)	2 (0.1%)	20 (0.06%)
	degradation	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
nucleotide metabolism	synthesis	11 (0.15%)	7 (0.38%)	0 (0%)	2 (0.14%)	1 (0.05%)	41 (0.11%)
	degradation	9 (0.12%)	2 (0.11%)	0 (0%)	6 (0.43%)	6 (0.31%)	52 (0.15%)
	salvage	16 (0.22%)	1 (0.05%)	0 (0%)	5 (0.36%)	8 (0.42%)	57 (0.16%)

	phosphotransfer and pyrophosphatases	6 (0.08%)	2 (0.11%)	0 (0%)	**6 (0.43%)**	3 (0.16%)	35 (0.1%)
	deoxynucleotide metabolism	7 (0.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	14 (0.04%)
Biodegradation of Xenobiotics	hydroxyacylglutathione hydrolase	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	4 (0.01%)
	lactoylglutathione lyase	5 (0.07%)	0 (0%)	0 (0%)	3 (0.22%)	0 (0%)	15 (0.04%)
	3-hydroxybutyryl-CoA dehydrogenase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
C1-metabolism	glycine hydroxymethyltransferase	0 (0%)	2 (0.11%)	0 (0%)	1 (0.07%)	1 (0.05%)	7 (0.02%)
	formate-tetrahydrofolate ligase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	1 (0%)
	dihydropteridine diphosphokinase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	5-formyltetrahydrofolate cyclo-ligase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	2 (0.01%)
	Methylenetetrahydrofolate dehydrogenase & Methenyltetrahydrofolate cyclohydrolase	0 (0%)	1 (0.05%)	0 (0%)	1 (0.07%)	0 (0%)	3 (0.01%)
	methylenetetrahydrofolate reductase	0 (0%)	0 (0%)	0 (0%)	1 (0.07%)	1 (0.05%)	2 (0.01%)
	GTP cyclohydrolase I	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	1 (0%)
	tetrahydrofolate synthase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (0.01%)
	dihydroneopterin aldolase	0 (0%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
	formate dehydrogenase	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0%)
misc	misc2	12 (0.17%)	1 (0.05%)	0 (0%)	1 (0.07%)	1 (0.05%)	52 (0.15%)
	UDP glucosyl and glucoronyl transferases	64 (0.89%)	8 (0.43%)	1 (0.41%)	14 (1.01%)	11 (0.58%)	286 (0.8%)
	gluco-, galacto- and mannosidases	30 (0.42%)	8 (0.43%)	3 (1.23%)	4 (0.29%)	5 (0.26%)	97 (0.27%)
	beta 1,3 glucan hydrolases	**25 (0.35%)**	3 (0.16%)	0 (0%)	2 (0.14%)	7 (0.37%)	72 (0.2%)
	acyl transferases	2 (0.03%)	1 (0.05%)	0 (0%)	1 (0.07%)	2 (0.1%)	14 (0.04%)
	O-methyl transferases	5 (0.07%)	**6 (0.32%)**	1 (0.41%)	4 (0.29%)	0 (0%)	29 (0.08%)
	oxidases - copper, flavone etc	34 (0.47%)	6 (0.32%)	2 (0.82%)	11 (0.79%)	10 (0.52%)	153 (0.43%)
	nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases	17 (0.24%)	0 (0%)	0 (0%)	0 (0%)	2 (0.1%)	86 (0.24%)
	glutathione S transferases	19 (0.26%)	2 (0.11%)	0 (0%)	0 (0%)	1 (0.05%)	98 (0.27%)
	cytochrome P450	50 (0.69%)	7 (0.38%)	3 (1.23%)	2 (0.14%)	1 (0.05%)	289 (0.81%)
	alcohol dehydrogenases	8 (0.11%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	20 (0.06%)
	peroxidases	14 (0.19%)	1 (0.05%)	2 (0.82%)	0 (0%)	5 (0.26%)	110 (0.31%)
	acid and other phosphatases	**31 (0.43%)**	7 (0.38%)	1 (0.41%)	8 (0.58%)	9 (0.47%)	94 (0.26%)
	oxygenases	2 (0.03%)	1 (0.05%)	0 (0%)	1 (0.07%)	0 (0%)	15 (0.04%)
	myrosinases-lectin-jacalin	7 (0.1%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	28 (0.08%)
	dynamin	7 (0.1%)	0 (0%)	0 (0%)	3 (0.22%)	0 (0%)	22 (0.06%)
	invertase/pectin methylesterase inhibitor family protein	10 (0.14%)	1 (0.05%)	1 (0.41%)	1 (0.07%)	0 (0%)	27 (0.08%)
	plastocyanin-like	13 (0.18%)	5 (0.27%)	3 (1.23%)	1 (0.07%)	0 (0%)	43 (0.12%)
	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	**20 (0.28%)**	1 (0.05%)	0 (0%)	2 (0.14%)	1 (0.05%)	52 (0.15%)
	short chain dehydrogenase/reductase (SDR)	18 (0.25%)	2 (0.11%)	1 (0.41%)	**8 (0.58%)**	2 (0.1%)	75 (0.21%)
	rhodanese	3 (0.04%)	2 (0.11%)	0 (0%)	0 (0%)	0 (0%)	7 (0.02%)
	GCN5-related N-acetyltransferase	7 (0.1%)	1 (0.05%)	0 (0%)	3 (0.22%)	2 (0.1%)	47 (0.13%)

	sulfotransferase	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	9 (0.03%)
	aminotransferases	3 (0.04%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (0.01%)
	calcineurin-like phosphoesterase family protein	3 (0.04%)	0 (0%)	0 (0%)	1 (0.07%)	1 (0.05%)	7 (0.02%)
	GDSL-motif lipase	31 (0.43%)	5 (0.27%)	0 (0%)	4 (0.29%)	5 (0.26%)	111 (0.31%)
	other Ferredoxins and Rieske domain	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (0.02%)
	zinc finger	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
RNA	processing	**106 (1.47%)**	**58 (3.12%)**	0 (0%)	6 (0.43%)	15 (0.79%)	372 (1.04%)
	transcription	39 (0.54%)	**21 (1.13%)**	1 (0.41%)	1 (0.07%)	4 (0.21%)	149 (0.42%)
	regulation of transcription	**717 (9.95%)**	166 (8.92%)	17 (6.97%)	59 (4.25%)	109 (5.72%)	2644 (7.39%)
	RNA binding	**66 (0.92%)**	**29 (1.56%)**	3 (1.23%)	6 (0.43%)	14 (0.73%)	215 (0.6%)
DNA	synthesis/chromatin structure	**140 (1.94%)**	**30 (1.61%)**	2 (0.82%)	6 (0.43%)	7 (0.37%)	340 (0.95%)
	repair	**34 (0.47%)**	11 (0.59%)	0 (0%)	2 (0.14%)	1 (0.05%)	96 (0.27%)
	unspecified	47 (0.65%)	15 (0.81%)	0 (0%)	11 (0.79%)	4 (0.21%)	241 (0.67%)
protein	aa activation	16 (0.22%)	**24 (1.29%)**	1 (0.41%)	2 (0.14%)	0 (0%)	94 (0.26%)
	synthesis	89 (1.24%)	**162 (8.71%)**	10 (4.1%)	18 (1.3%)	10 (0.52%)	700 (1.96%)
	targeting	70 (0.97%)	**32 (1.72%)**	5 (2.05%)	22 (1.59%)	**32 (1.68%)**	355 (0.99%)
	posttranslational modification	**233 (3.23%)**	36 (1.94%)	6 (2.46%)	38 (2.74%)	**122 (6.4%)**	976 (2.73%)
	degradation	**414 (5.75%)**	99 (5.32%)	9 (3.69%)	62 (4.47%)	**118 (6.19%)**	1680 (4.7%)
	folding	23 (0.32%)	11 (0.59%)	1 (0.41%)	5 (0.36%)	2 (0.1%)	86 (0.24%)
	glycosylation	19 (0.26%)	2 (0.11%)	0 (0%)	3 (0.22%)	5 (0.26%)	55 (0.15%)
	assembly and cofactor ligation	11 (0.15%)	1 (0.05%)	0 (0%)	0 (0%)	2 (0.1%)	139 (0.39%)
signaling	in sugar and nutrient physiology	16 (0.22%)	3 (0.16%)	0 (0%)	2 (0.14%)	6 (0.31%)	60 (0.17%)
	receptor kinases	**184 (2.55%)**	15 (0.81%)	1 (0.41%)	6 (0.43%)	30 (1.57%)	720 (2.01%)
	calcium	**82 (1.14%)**	13 (0.7%)	2 (0.82%)	17 (1.23%)	**49 (2.57%)**	284 (0.79%)
	phosphinositides	14 (0.19%)	2 (0.11%)	1 (0.41%)	**8 (0.58%)**	**12 (0.63%)**	65 (0.18%)
	G-proteins	65 (0.9%)	**24 (1.29%)**	2 (0.82%)	12 (0.87%)	**49 (2.57%)**	255 (0.71%)
	MAP kinases	**26 (0.36%)**	3 (0.16%)	0 (0%)	3 (0.22%)	8 (0.42%)	68 (0.19%)
	14-3-3 proteins	6 (0.08%)	0 (0%)	0 (0%)	3 (0.22%)	1 (0.05%)	13 (0.04%)
	misc	4 (0.06%)	1 (0.05%)	0 (0%)	0 (0%)	0 (0%)	10 (0.03%)
	lipids	1 (0.01%)	1 (0.05%)	0 (0%)	0 (0%)	2 (0.1%)	10 (0.03%)
	phosphorelay	4 (0.06%)	0 (0%)	1 (0.41%)	0 (0%)	0 (0%)	7 (0.02%)
	light	42 (0.58%)	13 (0.7%)	0 (0%)	3 (0.22%)	2 (0.1%)	149 (0.42%)
	gravity	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	unspecified	3 (0.04%)	0 (0%)	0 (0%)	1 (0.07%)	2 (0.1%)	8 (0.02%)
cell	organisation	**146 (2.03%)**	18 (0.97%)	0 (0%)	23 (1.66%)	**60 (3.15%)**	457 (1.28%)
	division	**50 (0.69%)**	14 (0.75%)	1 (0.41%)	5 (0.36%)	14 (0.73%)	134 (0.37%)
	cycle	**65 (0.9%)**	15 (0.81%)	0 (0%)	4 (0.29%)	5 (0.26%)	147 (0.41%)

	vesicle transport	37 (0.51%)	5 (0.27%)	0 (0%)	**21 (1.51%)**	**19 (1%)**	178 (0.5%)
	cell death	5 (0.07%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (0.02%)
	unspecified	0 (0%)	2 (0.11%)	0 (0%)	0 (0%)	0 (0%)	5 (0.01%)
development	storage proteins	8 (0.11%)	1 (0.05%)	0 (0%)	**8 (0.58%)**	3 (0.16%)	44 (0.12%)
	late embryogenesis abundant	3 (0.04%)	0 (0%)	1 (0.41%)	3 (0.22%)	1 (0.05%)	20 (0.06%)
	squamosa promoter binding like (SPL)	**8 (0.11%)**	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	15 (0.04%)
	organ development	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (0.01%)
	multitarget	2 (0.03%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (0.01%)
	unspecified	**227 (3.15%)**	59 (3.17%)	5 (2.05%)	22 (1.59%)	53 (2.78%)	817 (2.28%)
transport	p- and v-ATPases	8 (0.11%)	2 (0.11%)	1 (0.41%)	**19 (1.37%)**	**14 (0.73%)**	63 (0.18%)
	sugars	**27 (0.37%)**	5 (0.27%)	2 (0.82%)	**13 (0.94%)**	10 (0.52%)	81 (0.23%)
	amino acids	18 (0.25%)	5 (0.27%)	3 (1.23%)	7 (0.5%)	9 (0.47%)	95 (0.27%)
	nitrate	8 (0.11%)	0 (0%)	0 (0%)	1 (0.07%)	0 (0%)	25 (0.07%)
	ammonium	1 (0.01%)	0 (0%)	0 (0%)	1 (0.07%)	1 (0.05%)	12 (0.03%)
	sulphate	3 (0.04%)	1 (0.05%)	0 (0%)	3 (0.22%)	2 (0.1%)	16 (0.04%)
	phosphate	4 (0.06%)	2 (0.11%)	1 (0.41%)	4 (0.29%)	0 (0%)	27 (0.08%)
	metabolite transporters at the envelope membrane	7 (0.1%)	3 (0.16%)	1 (0.41%)	**8 (0.58%)**	3 (0.16%)	36 (0.1%)
	metabolite transporters at the mitochondrial membrane	10 (0.14%)	8 (0.43%)	0 (0%)	7 (0.5%)	**12 (0.63%)**	74 (0.21%)
	nucleotides	7 (0.1%)	0 (0%)	1 (0.41%)	0 (0%)	4 (0.21%)	33 (0.09%)
	NDP-sugars at the ER	4 (0.06%)	1 (0.05%)	0 (0%)	1 (0.07%)	2 (0.1%)	13 (0.04%)
	metal	18 (0.25%)	7 (0.38%)	1 (0.41%)	6 (0.43%)	**15 (0.79%)**	98 (0.27%)
	peptides and oligopeptides	22 (0.31%)	6 (0.32%)	1 (0.41%)	5 (0.36%)	5 (0.26%)	96 (0.27%)
	unspecified cations	5 (0.07%)	4 (0.22%)	1 (0.41%)	4 (0.29%)	6 (0.31%)	52 (0.15%)
	potassium	12 (0.17%)	2 (0.11%)	1 (0.41%)	6 (0.43%)	5 (0.26%)	65 (0.18%)
	ABC transporters and multidrug resistance systems	31 (0.43%)	9 (0.48%)	1 (0.41%)	8 (0.58%)	8 (0.42%)	180 (0.5%)
	peroxisomes	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (0.01%)
	unspecified anions	**11 (0.15%)**	1 (0.05%)	0 (0%)	**5 (0.36%)**	1 (0.05%)	25 (0.07%)
	Major Intrinsic Proteins	**18 (0.25%)**	1 (0.05%)	0 (0%)	1 (0.07%)	6 (0.31%)	47 (0.13%)
	porins	1 (0.01%)	1 (0.05%)	0 (0%)	2 (0.14%)	0 (0%)	6 (0.02%)
	calcium	6 (0.08%)	3 (0.16%)	0 (0%)	1 (0.07%)	4 (0.21%)	31 (0.09%)
	cyclic nucleotide or calcium regulated channels	8 (0.11%)	1 (0.05%)	0 (0%)	0 (0%)	3 (0.16%)	22 (0.06%)
	hormones	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	1 (0%)
	H+ transporting pyrophosphatase	1 (0.01%)	0 (0%)	0 (0%)	**3 (0.22%)**	0 (0%)	7 (0.02%)
	membrane system unknown	1 (0.01%)	0 (0%)	0 (0%)	0 (0%)	1 (0.05%)	4 (0.01%)
	misc	47 (0.65%)	8 (0.43%)	1 (0.41%)	8 (0.58%)	14 (0.73%)	204 (0.57%)
not assigned	no ontology	459 (6.37%)	177 (9.52%)	18 (7.38%)	96 (6.92%)	139 (7.29%)	2092 (5.85%)
	unknown	1,976 (27.43%)	528 (28.39%)	84 (34.43%)	426 (30.71%)	552 (28.96%)	14,407 (40.28%)

Supplemental Table 2: Known Hsfs in *S. lycopersicum*
Listed are name (column 1 and 3) and gene identifier (SolycID, column 2 and 4) of all members of the tomato Hsf family described by Scharf et al. (2012).

Hsf name	SolycID	Hsf name	SolycID
HsfA1a	Solyc08g005170	HsfA9	Solyc07g040680
HsfA1b	Solyc03g097120	HsfB1	Solyc02g090820
HsfA1c	Solyc08g076590	HsfB2a	Solyc03g026020
HsfA1e	Solyc06g072750	HsfB2b	Solyc08g080540
HsfA2	Solyc08g062960	HsfB3a	Solyc04g016000
HsfA3	Solyc09g009100	HsfB3b	Solyc10g079380
HsfA4a	Solyc03g006000	HsfB4a	Solyc04g078770
HsfA4b	Solyc07g055710	HsfB4b	Solyc11g064990
HsfA4c	Solyc02g072000	HsfB5	Solyc02g078340
HsfA5	Solyc12g098520	HsfC1	Solyc12g007070
HsfA6a	Solyc09g082670	SolycHsfl1	Solyc02g072060
HsfA6b	Solyc06g053960	SolycHsfl2	Solyc02g079180
HsfA7	Solyc09g065660	SolycHsfl3	Solyc11g008410
HsfA8	Solyc09g059520		

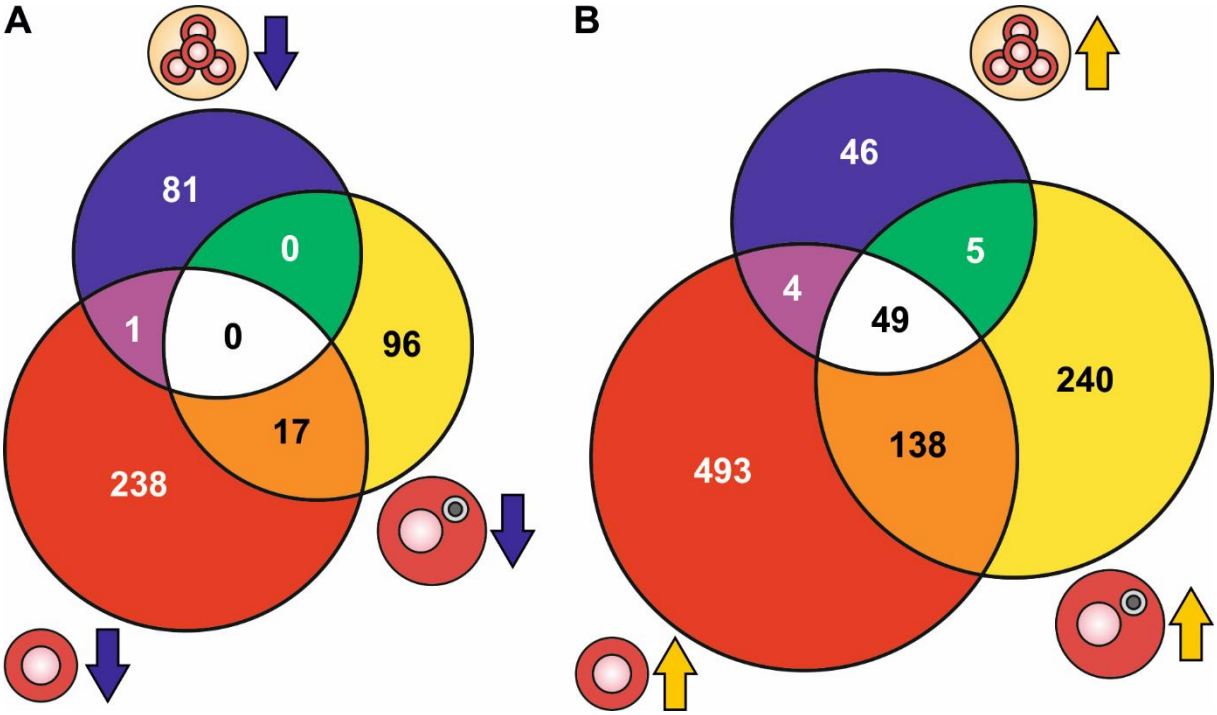
Supplemental Table 3: Putative Hsp orthologs identified in tomato Listed are putative tomato orthologs to *A. thaliana* Hsps identified via Inparanoid. For each identified ortholog the Hsp family (column 1), the gene identifier of the identified tomato ortholog (SolycID, column 2) as well as the gene identifiers (AGIs, column 3) and names (column 4) of the *A. thaliana* Hsp by which the ortholog was identified is given.

Hsp family	SolycID	Arabidopsis AGI	Arabidopsis Name
Hsp10	Solyc01g088610	AT3G60210	Plastid Cpn10 (1)
	Solyc05g056390	AT1G14980	Cpn10 (1)
	Solyc07g008800	AT1G14980	Cpn10 (1)
	Solyc07g042250	AT5G20720	Cpn21
	Solyc12g009250	AT5G20720	Cpn21
sHsp	Solyc02g080410	AT4G21870	AtHsp15.4-CV
	Solyc03g082420	AT4G27670	AtHsp25.3-P
	Solyc03g123540	AT1G54050	AtHsp17.4-CIII
	Solyc04g014480	AT5G37670	AtHsp15.7-Po
	Solyc06g076520	AT2G29500|AT1G07400|AT1G59860	AtHsp17.6B-CI|AtHsp17.8-CI|AtHsp17.6A-CI
	Solyc06g076540	AT2G29500|AT1G07400|AT1G59860	AtHsp17.6B-CI|AtHsp17.8-CI|AtHsp17.6A-CI
	Solyc06g076560	AT2G29500|AT1G07400|AT1G59860	AtHsp17.6B-CI|AtHsp17.8-CI|AtHsp17.6A-CI
	Solyc06g076570	AT2G29500|AT1G07400|AT1G59860	AtHsp17.6B-CI|AtHsp17.8-CI|AtHsp17.6A-CI
	Solyc07g064020	AT5G54660	AtHsp21.7-CVI
	Solyc08g062340	AT5G12020|AT5G12030	AtHsp17.6-CII|AtHsp17.7-CII
	Solyc08g062450	AT5G12020|AT5G12030	AtHsp17.6-CII|AtHsp17.7-CII
	Solyc08g078695	AT4G25200|AT5G51440	AtHsp23.6-M|AtHsp23.5-M
	Solyc08g078710	AT5G47600	AtHsp14.7-CVII
	Solyc08g078720	AT5G47600	AtHsp14.7-CVII
	Solyc09g015000	AT3G46230|AT5G59720|AT1G53540	AtHsp17.4-CI|AtHsp18.1-CI|AtHsp17.6C-CI
	Solyc09g015020	AT3G46230|AT5G59720|AT1G53540	AtHsp17.4-CI|AtHsp18.1-CI|AtHsp17.6C-CI
	Solyc11g020330	AT4G10250	AtHsp22.0-ER
	Solyc12g042830	AT1G52560	AtHsp26.5-MII
Hsp40	Solyc01g006860	AT5G06410	AtDjC17
	Solyc01g044270	AT2G25560	AtDjC44
	Solyc01g057750	AT5G18750	AtDjC47
	Solyc01g065570	AT3G06778	C
	Solyc01g066770	AT5G49060	AtDjC38
	Solyc01g067780	AT5G18750	AtDjC47
	Solyc01g079580	AT2G47440|AT3G62570|AT4G02100	D|D|D
	Solyc01g079610	AT3G62600	AtDjA8

	Solyc01g081330	AT1G28210	AtDjC23
	Solyc01g086740	AT5G48030	AtDjA3
	Solyc01g088730	AT5G23590	AtDjC37
	Solyc01g096700	AT5G06110|AT3G11450	AtDjC2|AtDjC1
	Solyc01g098250	AT3G04980|AT5G27240	AtDjC46|AtDjC45
	Solyc01g099260	AT3G05345	C
	Solyc01g100230	AT1G56300	AtDjC53
	Solyc01g105340	AT4G39960|AT2G22360	AtDjA5|AtDjA6
	Solyc01g105780	AT5G64360|AT5G09540	D|D
	Solyc01g109890	AT4G39150|AT2G21510	AtDjB7|AtDjC91
	Solyc02g014860	AT5G23240	AtDjC86
	Solyc02g062160	AT4G36040|AT2G17880	AtDjC26|AtDjC27
	Solyc02g064740	AT4G37480	AtDjC81
	Solyc02g077670	AT2G20560|AT4G28480	AtDjB3|AtDjB2
	Solyc02g080130	AT1G61770	AtDjC22
	Solyc02g083835	AT4G36040|AT2G17880	AtDjC26|AtDjC27
	Solyc03g007280	AT5G53150	AtDjC43
	Solyc03g019630	AT5G37440|AT5G37750|AT4G19570	AtDjC63|AtDjC62|AtDjC41
	Solyc03g019640	AT5G53150	AtDjC43
	Solyc03g063350	AT2G42080|AT3G58020	AtDjC70|AtDjC71
	Solyc03g097240	AT5G16650	AtDjC36
	Solyc03g115140	AT1G74250	AtDjC12
	Solyc03g116710	AT1G80120	AtDjC55
	Solyc03g116790	AT1G80030	AtDjA7
	Solyc03g117590	AT1G80920	AtDjC25
	Solyc03g118500	AT3G17830	AtDjA4
	Solyc03g123560	AT3G14200	AtDjC32
	Solyc04g005820	AT1G10350|AT1G59725	AtDjB5|AtDjB6
	Solyc04g007370	AT1G68370	AtDjB11
	Solyc04g009770	AT5G22060|AT3G44110	AtDjA2|AtDjA1
	Solyc04g063390	AT1G77020	AtDjC90
	Solyc04g076080	AT2G33735	AtDjC56
	Solyc04g077430	AT4G07990	AtDjC88
	Solyc04g080040	AT1G75100	AtDjC18
	Solyc04g080600	AT1G75310	AtDjC3
	Solyc05g006240	AT1G24120	AtDjB12
	Solyc05g006820	AT1G10350|AT1G59725	AtDjB5|AtDjB6
	Solyc05g008120	AT1G71000	AtDjC31
	Solyc05g009160	AT1G68370	AtDjB11
	Solyc05g011930	AT1G69060	AtDjC69
	Solyc05g018610	AT5G05750|AT3G57340	AtDjB15|AtDjB14
	Solyc05g040050	AT5G35753|AT2G05230|AT2G05250	|AtDjC78|C
	Solyc05g053760	AT4G13830	AtDjC29
	Solyc05g055160	AT5G22060|AT3G44110	AtDjA2|AtDjA1
	Solyc06g009670	AT2G47440|AT3G62570|AT4G02100	D|D|D
	Solyc06g068500	AT1G80920	AtDjC25
	Solyc06g068820	AT3G59280	AtDjID3B
	Solyc06g071110	AT3G47940	AtDjC57
	Solyc06g073950	AT5G22080	AtDjC35
	Solyc06g082840	AT4G12780|AT4G12770	AtDjC5|AtDjC9
	Solyc06g082970	AT5G21430	D
	Solyc07g007230	AT5G64360|AT5G09540	D|D
	Solyc07g026810	AT5G49580	AtDjB16
	Solyc07g055260	AT2G42750	AtDjC87
	Solyc07g065960	AT3G13310	AtDjC76
	Solyc07g065970	AT3G13310	AtDjC76
	Solyc07g066290	AT1G16680|AT1G79030	AtDjC39|AtDjC40
	Solyc08g006120	AT3G59280	AtDjID3B
	Solyc08g021920	AT2G26890	AtDjC30
	Solyc08g029220	AT3G04980|AT5G27240	AtDjC46|AtDjC45

	Solyc08g067795	AT5G25530	AtDjB10
	Solyc08g068300	AT5G11500	AtDjlD2
	Solyc08g078660	AT2G35720	AtDjC28
	Solyc09g005120	AT3G08970	AtDjB9
	Solyc09g005350	AT5G01390|AT3G08910	AtDjB1|AtDjB4
	Solyc09g007390	AT3G09700|AT5G03030	AtDjC16|AtDjC14
	Solyc09g007630	AT5G03160	AtDjC19
	Solyc09g009180	AT1G79940	AtDjC20
	Solyc09g015360	AT5G59610	AtDjC83
	Solyc09g074610	AT4G09350	AtDjC85
	Solyc09g092260	AT4G13830	AtDjC29
	Solyc10g078560	AT5G18140	AtDjC79
	Solyc10g079090	AT5G06910|AT3G12170	AtDjC24|AtDjC33
	Solyc10g081220	AT2G41000	AtDjC80
	Solyc10g086040	AT2G35795	AtDjC15
	Solyc11g005400	AT2G18465	AtDjC82
	Solyc11g018670	AT1G65280	AtDjC67
	Solyc11g065260	AT5G12430	AtDjC34
	Solyc11g071930	AT5G22080	AtDjC35
	Solyc12g056850	AT1G77930	AtDjC75
	Solyc12g088890	AT4G10130	AtDjC13
	Solyc12g095800	AT1G76700|AT1G21080	AtDjB8|AtDjC92
Hsp60	Solyc01g028810	AT1G55490|AT3G13470	Cpn60-β (3)|Cpn60-β (2)
	Solyc01g088080	AT3G03960	CCT-θ
	Solyc01g090750	AT3G20050	CCT-α
	Solyc02g063090	AT3G02530|AT5G16070	CCT-ζ (1)|CCT-ζ (2)
	Solyc02g068730	AT5G18820	Cpn60-α (2)
	Solyc02g085790	AT3G02530|AT5G16070	CCT-ζ (1)|CCT-ζ (2)
	Solyc03g118910	AT3G18190	CCT-δ
	Solyc03g121640	AT3G13860	Cpn60 (3)
	Solyc05g010240	AT1G26230	Cpn60-β (4)
	Solyc05g013990	AT1G24510	CCT-ε
	Solyc05g053470	AT3G23990	Cpn60 (2)
	Solyc05g056310	AT5G26360	CCT-γ
	Solyc06g065520	AT3G11830	CCT-η
	Solyc11g069000	AT5G20890	CCT-β
	Solyc11g069790	AT2G28000	Cpn60-α (1)
Hsp70	Solyc01g103450	AT5G49910|AT4G24280	AtHsp70-7|AtHsp70-6
	Solyc01g106210	AT5G09590	AtHsp70-10
	Solyc01g106260	AT5G09590	AtHsp70-10
	Solyc02g080470	AT1G11660	AtHsp70-16
	Solyc03g082920	AT5G28540|AT5G42020	AtHsp70-11|AtHsp70-12
	Solyc06g076020	AT3G12580	AtHsp70-4
	Solyc07g043560	AT4G16660	AtHsp70-17
	Solyc08g082820	AT5G28540|AT5G42020	AtHsp70-11|AtHsp70-12
	Solyc09g010630	AT3G12580	AtHsp70-4
	Solyc09g075950	AT2G32120	AtHsp70-8
	Solyc10g086410	AT3G12580	AtHsp70-4
	Solyc11g066060	AT3G12580	AtHsp70-4
	Solyc12g043110	AT1G79930|AT1G79920	AtHsp70-14|AtHsp70-15
	Solyc12g043120	AT1G79930|AT1G79920	AtHsp70-14|AtHsp70-15
Hsp90	Solyc03g007890	AT5G52640	AtHsp90-1
	Solyc04g081570	AT4G24190	AtHsp90-7
	Solyc06g036290	AT5G52640	AtHsp90-1
	Solyc07g047790	AT3G07770	AtHsp90-6
	Solyc07g065840	AT5G56000|AT5G56010|AT5G56030	AtHsp90-4|AtHsp90-3|AtHsp90-2
	Solyc12g015880	AT5G56000|AT5G56010|AT5G56030	AtHsp90-4|AtHsp90-3|AtHsp90-2
Hsp100	Solyc02g088610	AT5G15450	ClpB3
	Solyc03g115230	AT1G74310	AtHsp101/ClpB1
	Solyc03g117950	AT5G51070	ClpD

	Solyc03g118340	AT5G50920	ClpC1
	Solyc06g011370	AT2G25140	ClpB4
	Solyc06g082560	AT1G74310	AtHsp101/ClpB1



Supplemental Figure 1: Overlap of down- and upregulated mRNAs between pollen stages
Venn diagrams illustrate the overlap of down- **(A)** and upregulated **(B)** mRNAs between tetrads (blue), post-meiotic (red) and mature pollen (yellow).

Supplemental Table 4: mRNAs representing the core set of the pollen HSR
Listed are the mRNAs (column 2) with upregulated levels in all three developmental stages. For each mRNA the assigned MapMan term (column 3) as well as the description (ITAG3.2; column 4) is shown. In addition, for mRNAs encoding Hsfs or Hsps, the respective family is indicated (column 1).

Hsf / Hsp family	SolycID	MapMan term(s)	description
Hsf	Solyc02g090820	stress.abiotic.heat	SolycHsfB1
	Solyc08g062960	stress.abiotic.heat;RNA.regulation of transcription.HSF,Heat-shock transcription factor family	SolycHsfA2
sHsp	Solyc03g082420	stress.abiotic.heat	heat shock protein
	Solyc03g123540	stress.abiotic.heat	heat-shock protein, putative
	Solyc04g014480	stress.abiotic.heat	heat-shock protein, putative
	Solyc06g076520	stress.abiotic.heat	class I heat shock protein
	Solyc06g076540	stress.abiotic.heat	class I heat shock protein
	Solyc06g076560	stress.abiotic.heat	17.6 kD class I small heat shock protein
	Solyc06g076570	stress.abiotic.heat	class I small heat shock protein
	Solyc08g062340	stress.abiotic.heat	heat-shock protein, putative
	Solyc08g062450	stress.abiotic.heat	heat-shock protein, putative
	Solyc08g078695	stress.abiotic.heat	heat-shock protein, putative

	Solyc09g015000	stress.abiotic.heat	heat-shock protein, putative
	Solyc09g015020	stress.abiotic.heat	heat-shock protein, putative
	Solyc11g020330	stress.abiotic.heat	leer-sHSP small heat shock protein
Hsp70	Solyc09g075950	stress.abiotic.heat	heat shock protein 70
Hsp90	Solyc06g036290	stress.abiotic.heat	heat shock protein 90
Hsp100	Solyc03g115230	stress.abiotic.heat	Solanum lycopersicum heat shock protein
	Solyc06g011370	stress.abiotic.heat	chaperone protein
not Hsf or Hsp	Solyc04g014600	development.unspecified	adenine nucleotide alpha hydrolases-like superfamily protein
	Solyc03g025650	misc.beta 1,3 glucan hydrolases.glucan endo-1,3-beta-glucosidase	beta-1,3-glucanase
	Solyc01g081250	misc.glutathione S transferases	glutathione s-transferase, putative
	Solyc10g084400	misc.glutathione S transferases	glutathione S-transferase family protein
	Solyc06g071750	not assigned.no ontology	octicosapeptide/Phox/Bem1p domain-containing protein kinase
	Solyc06g076670	not assigned.no ontology	RNA-binding (RRM/RBD/RNP motifs) family protein
	Solyc10g078930	not assigned.no ontology	activator of 90 kDa heat shock ATPase
	Solyc01g091260	not assigned.no ontology.glycine rich proteins	myeloid leukemia factor 1
	Solyc04g072160	not assigned.no ontology.glycine rich proteins	HSP20-like chaperones superfamily protein
	Solyc03g095700	not assigned.unknown	LOW QUALITY:Voltage-dependent T-type calcium channel subunit alpha-1I, putative
	Solyc03g118060	not assigned.unknown	NADH-ubiquinone reductase complex 1 MLRQ subunit
	Solyc03g123710	not assigned.unknown	purple acid phosphatase 17
	Solyc05g007900	not assigned.unknown	LOW QUALITY:RING/U-box superfamily protein, putative
	Solyc06g048633	not assigned.unknown	P-loop nucleoside triphosphate hydrolases superfamily protein with CH (Calponin Homology) domain-containing protein
	Solyc08g006155	not assigned.unknown	7-dehydrocholesterol reductase, putative
	Solyc09g008970	not assigned.unknown	LOW QUALITY:AT3g10020/T22K18_16
	Solyc09g059650	not assigned.unknown	vacuolar iron transporter family protein
	Solyc09g061310	not assigned.unknown	PPPDE putative thiol peptidase family protein
	Solyc09g075600	not assigned.unknown	GATA transcription factor-like protein
	Solyc10g084170	not assigned.unknown	BAG family molecular chaperone regulator 5
	Solyc08g016670	protein.degradation.ubiquitin.E3.SCF.cullin	calcyclin-binding protein, putative
	Solyc09g092690	protein.folding; cell.cycle.peptidylprolyl isomerase	peptidyl-prolyl cis-trans isomerase
	Solyc09g007270	redox.ascorbate and glutathione.ascorbate	ascorbate peroxidase
	Solyc05g055010	RNA.RNA binding	RNA-binding protein
	Solyc08g079170	stress	heat shock protein STI
	Solyc01g095320	stress.abiotic.heat	BAG family molecular chaperone regulator 6
	Solyc01g102960	stress.abiotic.heat	heat-shock protein, putative
	Solyc03g113930	stress.abiotic.heat	heat-shock protein, putative
	Solyc11g071830	stress.abiotic.heat	DnaJ protein homolog
	Solyc11g066100	stress.abiotic.heat; protein.folding	heat shock protein 70

Supplemental Table 5: Regulation of Hsf and Hsp mRNA levels after HS

Shown are members (column 2) of the Hsf and Hsp families (column 1) and the regulation of their mRNA levels in response to HS in tetrads (column 3), post-meiotic (column 4) and mature pollen (column 5). mRNA levels are either downregulated (↓), not regulated (→) or upregulated (↑). In addition, it is indicated whether an Hsf or Hsp is accumulated in a particular stage under non-stressed conditions (column 6), which was determined beforehand (Figure 5B).

family	SolycID	tetrads	post-meiotic	mature	stage-accumulated
Hsf	Solyc02g072000	→	→	→	mature
	Solyc02g090820	↑	↑	↑	---
	Solyc03g006000	→	→	---	tetrads
	Solyc03g026020	→	↑	---	tetrads
	Solyc03g097120	→	→	↑	---
	Solyc04g016000	→	→	---	tetrads
	Solyc04g078770	→	---	---	---
	Solyc06g072750	→	→	---	tetrads + post-meiotic
	Solyc07g040680	→	---	---	tetrads
	Solyc08g005170	→	→	→	---
	Solyc08g062960	↑	↑	↑	tetrads
	Solyc08g076590	→	→	---	tetrads
	Solyc08g080540	↑	→	↑	mature
	Solyc09g009100	→	→	→	---
	Solyc09g059520	→	→	→	tetrads
	Solyc09g065660	→	↑	↑	tetrads
	Solyc12g007070	→	↑	---	tetrads
	Solyc12g098520	→	→	→	mature
sHsp	Solyc02g080410	→	---	---	tetrads
	Solyc03g082420	↑	↑	↑	---
	Solyc03g123540	↑	↑	↑	tetrads
	Solyc04g014480	↑	↑	↑	---
	Solyc06g076520	↑	↑	↑	tetrads
	Solyc06g076540	↑	↑	↑	tetrads + post-meiotic
	Solyc06g076560	↑	↑	↑	---
	Solyc06g076570	↑	↑	↑	tetrads
	Solyc08g062340	↑	↑	↑	---
	Solyc08g062450	↑	↑	↑	tetrads
	Solyc08g078695	↑	↑	↑	tetrads
	Solyc08g078710	→	→	---	---
	Solyc08g078720	→	→	→	tetrads + post-meiotic
	Solyc09g015000	↑	↑	↑	---
	Solyc09g015020	↑	↑	↑	---
	Solyc11g020330	↑	↑	↑	---
	Solyc12g042830	→	↑	↑	---
Hsp10	Solyc01g088610	→	→	→	tetrads
	Solyc07g008800	→	→	→	post-meiotic
	Solyc07g042250	→	↑	↑	tetrads
	Solyc12g009250	→	→	→	tetrads + post-meiotic
Hsp40	Solyc01g006860	---	→	---	---

	Solyc01g044270	→	→	→	---
	Solyc01g065570	→	→	→	post-meiotic + mature
	Solyc01g066770	→	→	→	mature
	Solyc01g067780	---	↑	---	---
	Solyc01g079580	→	↑	---	tetrads
	Solyc01g079610	→	→	→	---
	Solyc01g086740	→	→	→	---
	Solyc01g088730	→	→	→	mature
	Solyc01g096700	→	→	→	tetrads
	Solyc01g098250	→	→	---	---
	Solyc01g099260	→	→	---	---
	Solyc01g100230	→	↑	↑	---
	Solyc01g105340	→	→	↑	---
	Solyc01g105780	→	↑	---	tetrads
	Solyc01g109890	→	→	→	---
	Solyc02g014860	→	→	---	---
	Solyc02g062160	→	→	→	tetrads
	Solyc02g064740	→	→	→	---
	Solyc02g077670	→	↑	↑	tetrads
	Solyc02g080130	→	→	→	mature
	Solyc02g083835	→	→	→	tetrads
	Solyc03g007280	→	→	→	---
	Solyc03g019640	→	→	---	tetrads + post-meiotic
	Solyc03g063350	→	→	→	---
	Solyc03g097240	→	→	→	mature
	Solyc03g115140	→	→	→	tetrads
	Solyc03g116710	→	↑	---	tetrads
	Solyc03g116790	→	→	---	tetrads
	Solyc03g117590	→	→	→	---
	Solyc03g118500	→	→	→	---
	Solyc03g123560	→	→	↑	tetrads + post-meiotic
	Solyc04g005820	→	→	→	post-meiotic + mature
	Solyc04g007370	→	→	→	---
	Solyc04g009770	→	→	→	---
	Solyc04g076080	→	→	→	tetrads
	Solyc04g077430	→	→	→	---
	Solyc04g080040	→	---	---	tetrads
	Solyc04g080600	→	→	→	---
	Solyc05g006240	→	→	→	mature
	Solyc05g006820	→	→	→	tetrads + post-meiotic
	Solyc05g008120	→	↑	→	tetrads
	Solyc05g009160	→	→	→	---
	Solyc05g011930	→	→	→	tetrads
	Solyc05g018610	→	→	→	---
	Solyc05g040050	→	→	---	tetrads
	Solyc05g053760	→	→	→	---
	Solyc05g055160	→	→	→	tetrads

	Solyc06g009670	→	→	→	tetrads
	Solyc06g068500	→	→	→	mature
	Solyc06g068820	→	→	↑	---
	Solyc06g071110	→	→	→	---
	Solyc06g073950	→	→	→	mature
	Solyc06g082840	→	→	→	---
	Solyc06g082970	→	→	---	tetrads
	Solyc07g026810	→	→	→	tetrads
	Solyc07g055260	→	→	→	tetrads
	Solyc07g065960	→	→	→	tetrads
	Solyc07g065970	→	→	---	tetrads
	Solyc07g066290	→	→	→	---
	Solyc08g006120	→	→	→	---
	Solyc08g021920	→	→	→	---
	Solyc08g029220	→	→	↑	---
	Solyc08g067795	→	---	→	mature
	Solyc08g068300	→	→	→	---
	Solyc08g078660	→	→	→	mature
	Solyc09g005120	→	→	↑	---
	Solyc09g005350	→	→	→	---
	Solyc09g007390	→	→	→	tetrads + post-meiotic
	Solyc09g007630	→	→	→	---
	Solyc09g009180	→	→	→	---
	Solyc09g015360	→	→	→	---
	Solyc09g092260	→	→	→	tetrads
	Solyc10g078560	→	→	→	---
	Solyc10g079090	→	→	→	---
	Solyc10g081220	→	→	→	mature
	Solyc10g086040	→	→	→	---
	Solyc11g005400	→	→	→	mature
	Solyc11g018670	→	→	→	tetrads
	Solyc11g065260	→	→	→	mature
	Solyc11g071930	→	→	→	---
	Solyc12g056850	→	→	→	---
	Solyc12g088890	→	→	→	post-meiotic + mature
	Solyc12g095800	→	→	→	mature
Hsp60	Solyc01g028810	→	→	→	tetrads + post-meiotic
	Solyc01g088080	→	→	→	---
	Solyc01g090750	→	→	→	---
	Solyc02g063090	→	→	→	tetrads + post-meiotic
	Solyc02g068730	→	→	---	tetrads + post-meiotic
	Solyc02g085790	→	→	→	tetrads + post-meiotic
	Solyc03g118910	→	→	→	tetrads + post-meiotic
	Solyc03g121640	→	→	→	---
	Solyc05g013990	→	→	→	---
	Solyc05g053470	→	↑	→	tetrads + post-meiotic
	Solyc05g056310	→	→	→	tetrads

	Solyc06g065520	→	→	→	---
	Solyc11g069000	→	→	→	tetrads + post-meiotic
	Solyc11g069790	→	↑	↑	tetrads + post-meiotic
Hsp70	Solyc01g103450	→	→	→	tetrads + post-meiotic
	Solyc01g106210	→	→	↑	---
	Solyc01g106260	→	→	→	---
	Solyc02g080470	→	→	→	mature
	Solyc03g082920	→	↑	↑	tetrads
	Solyc06g076020	→	→	↑	---
	Solyc07g043560	→	→	→	---
	Solyc08g082820	→	→	→	---
	Solyc09g010630	→	→	→	---
	Solyc09g075950	↑	↑	↑	---
	Solyc10g086410	→	→	↑	tetrads
	Solyc11g066060	↓	→	→	---
	Solyc12g043110	→	↑	↑	---
	Solyc12g043120	→	→	↑	---
Hsp90	Solyc03g007890	→	↑	↑	tetrads
	Solyc04g081570	→	→	→	---
	Solyc06g036290	↑	↑	↑	tetrads
	Solyc07g047790	→	→	→	tetrads + post-meiotic
	Solyc07g065840	→	→	→	---
	Solyc12g015880	→	→	→	---
Hsp100	Solyc02g088610	→	↑	↑	tetrads
	Solyc03g115230	↑	↑	↑	tetrads
	Solyc03g117950	→	↑	↑	tetrads
	Solyc03g118340	→	→	→	---
	Solyc06g011370	↑	↑	↑	---

Supplemental Table 6: HS regulation of mRNAs accumulated in tetrads

Listed are mRNAs accumulated in tetrads (column 2) that are either down- (↓) or upregulated (↑) in response to HS. For each mRNA the second hierarchy level MapMan term(s) (column 3) and the description (ITAG3.2; column 4) is shown.

HS regulation tetrads	SolycID	MapMan term(s)	description
↓	Solyc05g032680	amino acid metabolism.degradation	3-hydroxyisobutyryl-CoA hydrolase-like protein
↓	Solyc07g051820	cell wall.cellulose synthesis	cellulose synthase family protein
↓	Solyc01g087280	cell wall.degradation	pectin lyase-like superfamily protein
↓	Solyc08g077900	cell wall.modification	expansin-like protein
↓	Solyc09g092520	cell wall.modification	xyloglucan endotransglycosylase
↓	Solyc01g103950	Co-factor and vitamine metabolism.thiamine	thiamine biosynthesis bifunctional ThiED
↓	Solyc03g044710	development.storage proteins	patatin
↓	Solyc03g031910	DNA.synthesis/chromatin structure	P-loop containing nucleoside triphosphate hydrolases superfamily protein, putative
↓	Solyc06g069290	DNA.unspecified	leucine-rich repeat receptor-like protein kinase family protein, putative

↓	Solyc07g042015	DNA.unspecified	RecQ family ATP-dependent DNA helicase
↓	Solyc05g051200	hormone metabolism.ethylene	ethylene-responsive factor 1
↓	Solyc01g104185	lipid metabolism.lipid degradation	fatty acyl-CoA reductase
↓	Solyc03g006860	major CHO metabolism.degradation	fructokinase
↓	Solyc03g121430	misc.acid and other phosphatases	inorganic pyrophosphatase 2
↓	Solyc10g080840	misc.cytochrome P450	cytochrome P450
↓	Solyc01g108035	misc.dynamin	dynamin-related protein 3A
↓	Solyc11g006250	misc.GDSL-motif lipase	GDSL-motif lipase/hydrolase family protein 2
↓	Solyc04g007160	misc.gluco-, galacto- and mannosidases	alpha-glucosidase
↓	Solyc03g116130	misc.glutathione S transferases	glutathione s-transferase, putative
↓	Solyc07g054900	misc.nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases	methylketone synthase Id
↓	Solyc11g062350	misc.plastocyanin-like	blue copper protein
↓	Solyc08g014370	not assigned.no ontology	cysteine/Histidine-rich C1 domain family protein
↓	Solyc10g062220	not assigned.no ontology	ELMO/CED-12 domain-containing protein
↓	Solyc01g090160	not assigned.unknown	glycogen synthase kinase-3 homolog MsK-3
↓	Solyc01g097470	not assigned.unknown	neurogenic locus notch-like protein
↓	Solyc01g102630	not assigned.unknown	DNA mismatch repair protein MLH3-like protein
↓	Solyc01g103550	not assigned.unknown	LOW QUALITY:Transcription factor, MADS-box (M-type_MADS))
↓	Solyc01g104430	not assigned.unknown	nuclear pore complex protein Nup98b
↓	Solyc01g107675	not assigned.unknown	chloroplast envelope membrane protein
↓	Solyc01g111840	not assigned.unknown	major facilitator superfamily protein
↓	Solyc02g063180	not assigned.unknown	UDP-N-acetylenolpyruvoylglucosamine reductase
↓	Solyc02g072050	not assigned.unknown	calcium/calcium/calmodulin-dependent Serine/Threonine-kinase
↓	Solyc02g078240	not assigned.unknown	phosphoglycerate mutase, putative
↓	Solyc02g084500	not assigned.unknown	UPF0183 protein
↓	Solyc02g090500	not assigned.unknown	phenylalanine ammonia-lyase
↓	Solyc02g092810	not assigned.unknown	RING/U-box superfamily protein
↓	Solyc03g094110	not assigned.unknown	glycosyltransferase
↓	Solyc04g081165	not assigned.unknown	wall associated kinase 4
↓	Solyc05g009580	not assigned.unknown	aluminum activated malate transporter family protein
↓	Solyc05g012045	not assigned.unknown	Mov34/MPN/PAD-1 family protein
↓	Solyc06g009125	not assigned.unknown	with no lysine (K) kinase 1
↓	Solyc07g006310	not assigned.unknown	basic helix-loop-helix transcription factor
↓	Solyc07g015985	not assigned.unknown	purple acid phosphatase
↓	Solyc07g032260	not assigned.unknown	activating transcription factor 7-interacting 1
↓	Solyc08g007805	not assigned.unknown	NADH dehydrogenase 1 alpha subcomplex subunit 8
↓	Solyc09g092760	not assigned.unknown	RNA-binding (RRM/RBD/RNP motifs) family protein
↓	Solyc10g084745	not assigned.unknown	rRNA-processing protein UTP23, putative

↓	Solyc11g045140	not assigned.unknown	LOW QUALITY:Vacuolar protein sorting-associated protein, putative
↓	Solyc11g065300	not assigned.unknown	Beta-1,3-glucanase
↓	Solyc11g066000	not assigned.unknown	cyclin-dependent protein kinase inhibitor SMR14
↓	Solyc12g056750	not assigned.unknown	WRKY transcription factor 61 (WRKY)
↓	Solyc12g096310	not assigned.unknown	plant/protein (Protein of unknown function, DUF599)
↓	Solyc01g107690	protein.degradation	ATP-dependent Clp protease ATP-binding subunit ClpX
↓	Solyc01g009470	protein.postranslational modification	poly [ADP-ribose] polymerase
↓	Solyc05g013777	protein.postranslational modification	serine/threonine-protein kinase AFC2
↓	Solyc06g068450	protein.postranslational modification	non-specific serine/threonine protein kinase
↓	Solyc02g093330	protein.targeting	nuclear pore complex Nup98-Nup96
↓	Solyc01g100630	redox.dismutases and catalases	catalase
↓	Solyc02g088113	RNA.processing	polypyrimidine tract-binding protein-like
↓	Solyc01g057270	RNA.regulation of transcription	calmodulin-binding transcription activator (CAMTA)
↓	Solyc06g066340	RNA.regulation of transcription	Myb family transcription factor family protein (G2-like)
↓	Solyc09g074810	RNA.regulation of transcription	basic helix-loop-helix DNA-binding superfamily protein (bHLH)
↓	Solyc11g010940	RNA.regulation of transcription	Dof zinc finger protein (Dof)
↓	Solyc12g017370	RNA.regulation of transcription	Myb family transcription factor APL (G2-like)
↓	Solyc06g034140	RNA.regulation of transcription; DNA.unspecified	B3 domain-containing protein (B3)
↓	Solyc09g047870	RNA.RNA binding	Hemolysin A
↓	Solyc08g005890	secondary metabolism.phenylpropanoids	HXXXD-type acyl-transferase family protein, putative
↓	Solyc09g082620	stress.abiotic	EPIDERMAL PATTERNING FACTOR-like protein
↓	Solyc09g098130	stress.biotic	CC-NBS-LRR gene
↓	Solyc05g056470	transport.ABC transporters and multidrug resistance systems	ABC transporter family protein
↑	Solyc04g014600	development.unspecified	adenine nucleotide alpha hydrolases-like superfamily protein
↑	Solyc06g072430	development.unspecified	BAG family molecular chaperone regulator 5
↑	Solyc01g104740	hormone metabolism.ethylene	multiprotein-bridging factor, putative
↑	Solyc04g072160	not assigned.no ontology	HSP20-like chaperones superfamily protein
↑	Solyc06g071750	not assigned.no ontology	octicosapeptide/Phox/Bem1p domain-containing protein kinase
↑	Solyc00g014790	not assigned.unknown	LOW QUALITY:P-loop nucleoside triphosphate hydrolase superfamily protein
↑	Solyc00g142170	not assigned.unknown	Mitovirus RNA-dependent RNA polymerase
↑	Solyc03g118060	not assigned.unknown	NADH-ubiquinone reductase complex 1 MLRQ subunit
↑	Solyc05g046310	not assigned.unknown	N-acetyl-gamma-glutamyl-phosphate reductase
↑	Solyc09g059650	not assigned.unknown	vacuolar iron transporter family protein
↑	Solyc10g084170	not assigned.unknown	BAG family molecular chaperone regulator 5
↑	Solyc11g056270	not assigned.unknown	LOW QUALITY:Maturase
↑	Solyc09g092690	protein.folding; cell.cycle	peptidyl-prolyl cis-trans isomerase

↑	Solyc00g022107	protein.synthesis	ribosomal protein L2
↑	Solyc01g102960	stress.abiotic	heat-shock protein, putative
↑	Solyc03g113930	stress.abiotic	heat-shock protein, putative
↑	Solyc03g115230	stress.abiotic	Hsp100
↑	Solyc03g123540	stress.abiotic	sHsp
↑	Solyc06g036290	stress.abiotic	Hsp90
↑	Solyc06g076520	stress.abiotic	sHsp
↑	Solyc06g076570	stress.abiotic	sHsp
↑	Solyc08g062450	stress.abiotic	sHsp
↑	Solyc08g078695	stress.abiotic	sHsp
↑	Solyc11g071830	stress.abiotic	DnaJ protein homolog
↑	Solyc11g066100	stress.abiotic; protein.folding	heat shock protein 70
↑	Solyc08g062960	stress.abiotic; RNA.regulation of transcription	HsfA2

Supplemental Table 7: HS regulation of mRNAs accumulated in tetrads and post-meiotic pollen

Listed are mRNAs accumulated in tetrads and post-meiotic pollen (column 2) that are either down- (↓) or upregulated (↑) in at least one of the two stages in response to HS. For each mRNA the second hierarchy level MapMan term(s) (column 3) and the description (ITAG3.2; column 4) is shown.

HS regulation				
tetrads	**post-meiotic**	**SolycID**	**MapMan term(s)**	**description**
↓	→	Solyc11g012600	not assigned.unknown	F-box/RNI-like superfamily protein
→	↓	Solyc01g103580	not assigned.unknown	transmembrane protein, putative (DUF679)
→	↓	Solyc02g067085	not assigned.unknown	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic
→	↓	Solyc03g025820	not assigned.unknown	shortage in chiasmata 1, putative
→	↓	Solyc10g074950	not assigned.unknown	transmembrane protein, putative (Protein of unknown function, DUF599)
→	↓	Solyc10g083580	not assigned.unknown	phytosulfokine 3, putative
↑	↑	Solyc01g091260	not assigned.no ontology	myeloid leukemia factor 1
↑	↑	Solyc08g016670	protein.degradation	calcyclin-binding protein, putative
↑	↑	Solyc05g055010	RNA.RNA binding	RNA-binding protein
↑	↑	Solyc06g076540	stress.abiotic	sHsp
↑	→	Solyc11g039980	mitochondrial electron transport / ATP synthesis.F1-ATPase	ATP synthase subunit alpha
↑	→	Solyc00g013140	mitochondrial electron transport / ATP synthesis.NADH-DH	NADH-ubiquinone oxidoreductase chain 5
↑	→	Solyc00g022090	not assigned.unknown	formyltetrahydrofolate deformylase 1, mitochondrial
↑	→	Solyc11g056320	not assigned.unknown	LOW QUALITY:Photosystem II CP47 reaction center protein
↑	→	Solyc05g052410	RNA.regulation of transcription	dehydration responsive element binding protein
→	↑	Solyc04g005310	not assigned.no ontology	translin-associated factor X
→	↑	Solyc09g011030	not assigned.no ontology	Hsp70-binding protein 1
→	↑	Solyc01g008900	not assigned.unknown	pyridoxal-5'-phosphate-dependent enzyme family protein
→	↑	Solyc04g016590	not assigned.unknown	global transcription factor C
→	↑	Solyc04g082720	not assigned.unknown	HSP20-like chaperones superfamily protein
→	↑	Solyc03g007000	protein.degradation	mitochondrial metalloendopeptidase OMA1
→	↑	Solyc05g053470	protein.folding	Hsp60l

→	↑	Solyc02g080790	protein.synthesis	deoxyhypusine synthase
→	↑	Solyc11g069790	PS.calvin cycle	Hsp60
→	↑	Solyc02g069120	RNA.regulation of transcription	zinc finger ZPR1 protein
→	↑	Solyc05g010670	stress.abiotic	heat shock family protein

Supplemental Table 8: HS regulation of mRNAs accumulated in post-meiotic pollen

Listed are mRNAs accumulated in post-meiotic pollen (column 2) that are either down- (↓) or upregulated (↑) in response to HS. For each mRNA the second hierarchy level MapMan term(s) (column 3) and the description (ITAG3.2; column 4) is shown.

HS regulation tetrads	SolycID	MapMan term(s)	description
↓	Solyc08g061772	hormone metabolism.cytokinin	adenylate isopentenyltransferase
↓	Solyc04g009850	hormone metabolism.ethylene; redox.ascorbate and glutathione	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
↓	Solyc12g009040	lipid metabolism.FA synthesis and FA elongation	long chain acyl-CoA synthetase
↓	Solyc06g011290	N-metabolism.ammonia metabolism	nodulin / glutamate-ammonia ligase-like protein
↓	Solyc03g063550	not assigned.unknown	LOW QUALITY:cotton fiber protein
↓	Solyc05g051300	not assigned.unknown	auxin-responsive protein
↓	Solyc07g041460	not assigned.unknown	LOW QUALITY:SNF2 domain-containing protein
↓	Solyc08g005015	not assigned.unknown	microspore-specific promoter 2
↓	Solyc09g082350	not assigned.unknown	vicilin-like protein
↓	Solyc11g061776	not assigned.unknown	adenylate isopentenyltransferase
↓	Solyc07g053450	RNA.regulation of transcription	bZIP family transcription factor (bZIP)
↓	Solyc10g078610	RNA.regulation of transcription	ethylene-responsive transcription factor (ERF)
↓	Solyc12g014130	signalling.calcium	SUN-like protein 33
↓	Solyc09g008420	TCA / org transformation.carbonic anhydrases	carbonic anhydrase, putative
↑	Solyc05g052530	misc.gluco-, galacto- and mannosidases	endoglucanase

Supplemental Table 9: HS regulation of mRNAs accumulated in post-meiotic and mature pollen

Listed are mRNAs accumulated in post-meiotic and mature pollen (column 2) that are either down- (↓) or upregulated (↑) in the two stages in response to HS. For each mRNA the second hierarchy level MapMan term(s) (column 3) and the description (ITAG3.2; column 4) is shown.

HS regulation				
post-meiotic	mature	SolycID	MapMan term(s)	description
↓	↓	Solyc02g021420	cell.organisation	villin 4
↓	↓	Solyc03g079970	not assigned.unknown	F-box SKIP31-like protein
↓	↓	Solyc06g069150	not assigned.unknown	LOW QUALITY:Zinc finger BED domain-containing protein DAYSLEEPER
↓	↓	Solyc10g080360	not assigned.unknown	LOW QUALITY:Transmembrane protein, putative
↓	↓	Solyc12g036900	not assigned.unknown	LOW QUALITY:evolutionarily conserved C-terminal region 9
↓	↓	Solyc01g057190	protein.degradation	LOW QUALITY:S-locus F-box protein type-12
↓	↓	Solyc03g111710	protein.degradation	BTB/POZ and TAZ domain protein

↓	↓	Solyc07g049130	RNA.regulation of transcription	LOW QUALITY:PLATZ transcription factor family protein
↓	↓	Solyc06g051000	stress.abiotic	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
↓	↓	Solyc02g093870	transport.amino acids	amino acid transporter, transmembrane
↓	↓	Solyc02g062870	transport.sugars	polyol monosaccharide transporter 7
↓	→	Solyc12g098500	amino acid metabolism.degradation	adenosylhomocysteinase
↓	→	Solyc03g116530	Biodegradation of Xenobiotics.lactoylglutathione lyase	lactoylglutathione lyase / glyoxalase I family protein
↓	→	Solyc04g008280	C1-metabolism	4-hydroxy-4-methyl-2-oxoglutarate aldolase
↓	→	Solyc05g051980	cell.organisation	annexin
↓	→	Solyc06g008500	cell.organisation	ankyrin repeat domain protein, putative
↓	→	Solyc09g092210	cell.organisation	nuclear migration protein nudC
↓	→	Solyc02g065100	development.storage proteins	patatin
↓	→	Solyc06g075260	development.storage proteins	vicilin-like antimicrobial peptides 2-2
↓	→	Solyc06g075280	development.storage proteins	vicilin-like antimicrobial peptides 2-2
↓	→	Solyc01g097960	development.unspecified	late embryogenesis abundant D-like protein
↓	→	Solyc03g097340	development.unspecified	WD40 repeat protein
↓	→	Solyc09g091600	development.unspecified	tetratricopeptide repeat-like superfamily protein
↓	→	Solyc06g082130	fermentation.PDC	pyruvate decarboxylase
↓	→	Solyc01g108460	hormone metabolism.brassinosteroid; protein.degradation	carboxypeptidase
↓	→	Solyc01g005360	hormone metabolism.salicylic acid; misc.O-methyl transferases	S-adenosyl-L-methionine-dependent methyltransferase superfamily protein
↓	→	Solyc03g114600	lipid metabolism.exotics(steroids, squalene etc)	alternaria stem canker resistance
↓	→	Solyc03g116620	lipid metabolism.lipid degradation	Phospholipase D
↓	→	Solyc09g075710	lipid metabolism.lipid degradation; Biodegradation of Xenobiotics	alpha/beta-Hydrolases superfamily protein
↓	→	Solyc00g007010	lipid metabolism.lipid degradation; signalling.phosphinositides	phosphoinositide phosphatase family protein
↓	→	Solyc09g090310	lipid metabolism.lipid transfer proteins etc	glycolipid transfer protein domain-containing protein
↓	→	Solyc03g058950	misc.acid and other phosphatases	inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase
↓	→	Solyc01g088590	misc.invertase/pectin methylesterase inhibitor family protein	invertase/pectin methylesterase inhibitor
↓	→	Solyc02g088380	misc.plastocyanin-like	early nodulin-like protein
↓	→	Solyc07g054670	mitochondrial electron transport / ATP synthesis.NADH-DH	external alternative NAD(P)H-ubiquinone oxidoreductase B2, mitochondrial
↓	→	Solyc01g091560	not assigned.no ontology	LOW QUALITY:Acyl-CoA thioesterase, putative
↓	→	Solyc05g052080	not assigned.no ontology	LOW QUALITY:transmembrane protein
↓	→	Solyc07g006040	not assigned.no ontology	DNA-directed RNA polymerase subunit beta
↓	→	Solyc09g011270	not assigned.no ontology	cysteine/Histidine-rich C1 domain family protein
↓	→	Solyc10g053910	not assigned.no ontology	B-cell receptor-associated 31-like protein
↓	→	Solyc10g085030	not assigned.no ontology	SOUL heme-binding family protein
↓	→	Solyc12g017310	not assigned.no ontology	defective in cullin neddylation protein
↓	→	Solyc12g035710	not assigned.no ontology	DUF21 domain-containing protein

↓	→	Solyc12g096950	not assigned.no ontology	TLD-domain containing nucleolar protein, putative isoform 1
↓	→	Solyc01g056360	not assigned.unknown	LOW QUALITY:Dehydrin family protein, expressed
↓	→	Solyc01g057852	not assigned.unknown	protein-S-isoprenylcysteine O-methyltransferase
↓	→	Solyc01g098600	not assigned.unknown	transmembrane protein, putative (DUF1068)
↓	→	Solyc01g100070	not assigned.unknown	LOW QUALITY:Pentatricopeptide repeat-containing protein
↓	→	Solyc01g106100	not assigned.unknown	LOW QUALITY:zein-binding protein (Protein of unknown function, DUF593)
↓	→	Solyc02g078280	not assigned.unknown	DNA ligase-like protein
↓	→	Solyc02g080780	not assigned.unknown	orotidine 5'-phosphate decarboxylase
↓	→	Solyc02g083440	not assigned.unknown	F-box
↓	→	Solyc02g089580	not assigned.unknown	late embryogenesis abundant protein family protein
↓	→	Solyc02g092000	not assigned.unknown	transmembrane protein
↓	→	Solyc02g092850	not assigned.unknown	LOW QUALITY:Sister chromatid cohesion protein PDS5 like B-B
↓	→	Solyc03g031530	not assigned.unknown	sulfite exporter TauE/SafE family protein
↓	→	Solyc03g046320	not assigned.unknown	late embryogenesis abundant protein family protein
↓	→	Solyc03g080000	not assigned.unknown	LOW QUALITY:transmembrane protein, putative (Protein of unknown function, DUF599)
↓	→	Solyc03g082350	not assigned.unknown	LOW QUALITY:Avr9/Cf-9 rapidly elicited protein 65
↓	→	Solyc03g082660	not assigned.unknown	major facilitator superfamily domain
↓	→	Solyc03g111797	not assigned.unknown	sieve element occlusion a
↓	→	Solyc03g113480	not assigned.unknown	LOW QUALITY:arginine N-methyltransferase, putative (DUF688)
↓	→	Solyc03g120180	not assigned.unknown	transmembrane protein, putative
↓	→	Solyc04g009840	not assigned.unknown	TRICHOME BIREFRINGENCE-LIKE 5
↓	→	Solyc04g015140	not assigned.unknown	low PSII Accumulation 3
↓	→	Solyc04g040095	not assigned.unknown	RNA-binding region RNP-1 protein
↓	→	Solyc04g080540	not assigned.unknown	DNA polymerase epsilon catalytic subunit A, putative
↓	→	Solyc05g005520	not assigned.unknown	LOW QUALITY:Peptidyl-prolyl cis-trans isomerase
↓	→	Solyc05g053555	not assigned.unknown	NADH-ubiquinone reductase complex 1 MLRQ subunit
↓	→	Solyc06g064610	not assigned.unknown	LOW QUALITY:Glucan endo-1,3-beta-glucosidase
↓	→	Solyc06g075710	not assigned.unknown	LOW QUALITY:MEI2-like 2
↓	→	Solyc06g082340	not assigned.unknown	F-box family protein
↓	→	Solyc07g065720	not assigned.unknown	complex 1 protein, LYR family protein
↓	→	Solyc08g081520	not assigned.unknown	LOW QUALITY:Plant/F1M20-13 protein
↓	→	Solyc09g009930	not assigned.unknown	YELLOW STRIPE like 3
↓	→	Solyc09g010730	not assigned.unknown	LOW QUALITY:Cyclin-dependent protein kinase inhibitor SMR3
↓	→	Solyc09g098120	not assigned.unknown	oil body-associated protein 1A
↓	→	Solyc10g018050	not assigned.unknown	elicitor-responsive protein 3
↓	→	Solyc11g012500	not assigned.unknown	zinc finger CCCH domain protein
↓	→	Solyc12g009180	not assigned.unknown	RmlC-like jelly roll fold protein
↓	→	Solyc12g010230	not assigned.unknown	calcineurin-like metallo-phosphoesterase superfamily protein, putative
↓	→	Solyc12g010850	not assigned.unknown	LOW QUALITY:BnaCnng66650D protein
↓	→	Solyc12g015810	not assigned.unknown	transmembrane protein C9orf5 protein
↓	→	Solyc12g017820	not assigned.unknown	golgi to ER traffic 4

↓	→	Solyc12g042850	not assigned.unknown	late embryogenesis abundant protein (LEA) family protein
↓	→	Solyc12g088910	not assigned.unknown	sulfotransferase
↓	→	Solyc08g083370	nucleotide metabolism.phosphotransfer and pyrophosphatases	soluble inorganic pyrophosphatase
↓	→	Solyc03g006200	protein.degradation	cysteine protease, putative
↓	→	Solyc06g007500	protein.degradation	ubiquitin-conjugating enzyme family protein
↓	→	Solyc08g021995	protein.degradation	26S protease regulatory subunit-like protein
↓	→	Solyc08g048550	protein.degradation	protease Do-like protein
↓	→	Solyc10g055560	protein.degradation	SBP (S-ribonuclease binding protein) family protein
↓	→	Solyc04g056560	protein.postranslational modification	protein phosphatase-2c, putative
↓	→	Solyc07g053180	protein.postranslational modification	leucine-rich repeat family protein
↓	→	Solyc10g086125	protein.synthesis; protein.aa activation; protein.degradation	ubiquitin family protein
↓	→	Solyc10g006530	PS.lightreaction	photosynthetic NDH subcomplex L 3
↓	→	Solyc01g098710	RNA.regulation of transcription	eukaryotic aspartyl protease family protein
↓	→	Solyc11g068940	RNA.regulation of transcription	U-box domain-containing family protein
↓	→	Solyc01g108310	RNA.RNA binding	LA-related protein 6 LA RNA-binding domain protein
↓	→	Solyc06g035610	RNA.RNA binding	RNA-binding family protein
↓	→	Solyc01g005900	secondary metabolism.phenylpropanoids	HXXXD-type acyl-transferase family protein
↓	→	Solyc05g031600	secondary metabolism.sulfur-containing	tryptophan aminotransferase related 2
↓	→	Solyc01g010360	signalling.14-3-3 proteins	14-3-3 protein, putative
↓	→	Solyc02g094000	signalling.calcium	calcium-binding protein
↓	→	Solyc01g105810	signalling.G-proteins	GDP dissociation inhibitor
↓	→	Solyc04g076140	signalling.G-proteins	Rho GTPase-activating protein
↓	→	Solyc02g072440	signalling.receptor kinases	leucine-rich repeat protein kinase family protein, putative
↓	→	Solyc02g062310	stress.abiotic	BURP domain protein RD22
↓	→	Solyc10g086670	stress.abiotic	adenine nucleotide alpha hydrolases-like superfamily protein
↓	→	Solyc01g059900	stress.biotic	dirigent protein
↓	→	Solyc07g006700	stress.biotic	pathogenesis-related family protein
↓	→	Solyc06g061260	transport.amino acids	LOW QUALITY:Amino acid transporter, putative
↓	→	Solyc06g060630	transport.metal	cation/H(+) antiporter 15
↓	→	Solyc05g005720	transport.potassium	potassium transporter family protein
↓	→	Solyc08g081230	transport.potassium	BTB/POZ domain-containing protein
↓	→	Solyc02g062860	transport.sugars	polyol monosaccharide transporter 6
↓	→	Solyc07g024030	transport.sugars	polyol monosaccharide transporter 2
↓	→	Solyc12g008320	transport.sugars	sugar transporter protein 18
↓	→	Solyc06g084155	transport.sulphate	sulfate transporter, putative
→	↓	Solyc03g093130	cell wall.degradation	xyloglucan endotransglucosylase-hydrolase 3
→	↓	Solyc07g063610	cell.organisation	dynein light chain family protein
→	↓	Solyc08g006820	DNA.unspecified	transmembrane 9 superfamily member
→	↓	Solyc05g008680	hormone metabolism.brassinosteroid	methyltransferase
→	↓	Solyc08g068160	misc.oxidases - copper, flavone etc	flavin-containing monooxygenase
→	↓	Solyc01g087020	not assigned.unknown	LOW QUALITY:Transmembrane protein, putative

→	↓	Solyc07g066350	not assigned.unknown	serine/arginine repetitive matrix-like protein
→	↓	Solyc07g019497	protein.postranslational modification	serine/threonine protein phosphatase 7 long form isogeny
→	↓	Solyc10g007240	redox.thioredoxin	TPR repeat thioredoxin TTL1-like protein
→	↓	Solyc06g062480	RNA.regulation of transcription	protein EARLY FLOWERING 3-like protein
→	↓	Solyc02g032860	S-assimilation.APR	5'-adenylylsulfate reductase
→	↓	Solyc03g031620	S-assimilation.APR	5'-adenylylsulfate reductase
→	↓	Solyc01g091460	signalling.G-proteins	brefeldin A-inhibited guanine nucleotide-exchange protein
→	↓	Solyc12g006950	transport.sulphate	sulfate transporter, putative
↑	↑	Solyc11g045040	lipid metabolism.lipid degradation	glycerophosphodiester phosphodiesterase family protein
↑	↑	Solyc03g025650	misc.beta 1,3 glucan hydrolases	beta-1,3-glucanase
↑	↑	Solyc03g034010	protein.degradation	U-box domain-containing family protein
→	↑	Solyc04g015200	glycolysis.cytosolic branch	ATP-dependent 6-phosphofructokinase
→	↑	Solyc02g071750	not assigned.no ontology	RNA ligase/cyclic nucleotide phosphodiesterase family protein
→	↑	Solyc04g077850	not assigned.no ontology	U-box domain-containing protein
→	↑	Solyc00g048515	not assigned.unknown	TBP-ASSOCIATED FACTOR 6B
→	↑	Solyc01g097825	not assigned.unknown	50S ribosomal protein L5, chloroplastic
→	↑	Solyc03g117160	not assigned.unknown	double-stranded RNA binding protein, putative
→	↑	Solyc03g121280	not assigned.unknown	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
→	↑	Solyc04g015660	not assigned.unknown	serine/threonine protein phosphatase 7 long form isogeny
→	↑	Solyc08g008440	not assigned.unknown	oxysterol-binding protein-related protein 3C
→	↑	Solyc03g006960	protein.postranslational modification	protein phosphatase 2C family protein
→	↑	Solyc07g062060	protein.postranslational modification	methionine sulfoxide reductase
→	↑	Solyc01g108860	redox.ascorbate and glutathione	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
→	↑	Solyc02g036370	RNA.regulation of transcription	Myb family transcription factor

Supplemental Table 10: HS regulation of mRNAs accumulated in mature pollen

Listed are mRNAs accumulated in mature pollen (column 2) that are either down- (↓) or upregulated (↑) in response to HS. For each mRNA the second hierarchy level MapMan term(s) (column 3) and the description (ITAG3.2; column 4) is shown.

HS regulation tetrads	SolycID	MapMan term(s)	description
↓	Solyc06g048750	---	unknown protein
↓	Solyc02g075620	cell wall.pectin*esterases	pectinesterase
↓	Solyc12g044890	DNA.unspecified	transmembrane 9 superfamily member
↓	Solyc10g007290	glycolysis.cytosolic branch	phosphoenolpyruvate carboxylase
↓	Solyc12g005310	hormone metabolism.auxin	auxin-responsive GH3 family protein
↓	Solyc10g076200	lipid metabolism.lipid transfer proteins etc	non-specific lipid-transfer protein
↓	Solyc06g069700	not assigned.no ontology	pPentatricopeptide repeat-containing protein, putative
↓	Solyc01g081590	not assigned.unknown	non-specific lipid-transfer protein
↓	Solyc03g044480	not assigned.unknown	cyclin-dependent kinase inhibitor family protein

↓	Solyc06g053700	not assigned.unknown	ATBET12, putative
↓	Solyc08g066420	not assigned.unknown	LOW QUALITY:Lachrymatory-factor synthase, putative
↓	Solyc11g063500	not assigned.unknown	LOW QUALITY:nitrite reductase 1
↓	Solyc12g042870	not assigned.unknown	LOW QUALITY:Ulp1 protease family, C-terminal catalytic domain containing protein
↓	Solyc01g057010	protein.degradation	LOW QUALITY:S-locus F-box protein type-11
↓	Solyc05g010590	protein.targeting	WPP domain-interacting tail-anchored protein 1
↓	Solyc01g090830	RNA.regulation of transcription	LOW QUALITY:Zinc finger family protein
↓	Solyc06g069080	signalling.calcium	tetratricopeptide repeat protein 7A
↓	Solyc12g088190	transport.amino acids	amino acid permease
↑	Solyc04g025440	cell wall.degradation	pectin lyase-like superfamily protein
↑	Solyc01g068365	not assigned.unknown	beta-glucosidase, GBA2 type family protein
↑	Solyc06g076960	not assigned.unknown	ELF4-like 4
↑	Solyc09g075600	not assigned.unknown	GATA transcription factor-like protein
↑	Solyc06g054150	stress.abiotic	LOW QUALITY:17.6 kDa class II heat shock protein
↑	Solyc08g080540	stress.abiotic	HsfB2b
↑	Solyc12g056560	stress.abiotic	heat shock family protein

Supplemental Table 11: Predicted miRNAs for developing and heat-stressed pollen

Listed are all detected miRNAs (column 1 and 3) and their nucleotide sequence (column 2 and 4). miRNAs are named in ascending order based on their lexicographically sorted nucleotide sequences.

miRNA identifier	sequence (5' to 3')		miRNA identifier	sequence (5' to 3')
solyc-miR001	AAAAAAAGACUGACAGAAAUUUG		solyc-miR398	AUGACAAGUUGCGACUUUUAUGAA
solyc-miR002	AAAAAAAUGAAAUCGCUGCUCAAG		solyc-miR399	AUGACGUGACCGAAUUAUAAUUGG
solyc-miR003	AAAAAAGAUGCAGGACUAGAC		solyc-miR400	AUGACUGCCGCGCUCGCCUUGGAC
solyc-miR004	AAAAAAUGGGCUAGCCCGGCCUGA		solyc-miR401	AUGACUUGGUGAAUCAUGGCAG
solyc-miR005	AAAAACACAAUAUACUAGUGAUUU		solyc-miR402	AUGACUUUAAUGAAGAGUUGCGAC
solyc-miR006	AAAAAGUGACCUGCAUUUUACUGA		solyc-miR403	AUGACUUUUAUGAAAAGUUGUGAC
solyc-miR007	AAAAAUCAGAUAAUUUGUUGGUGA		solyc-miR404	AUGACUUUUAUGAAAGGUUGUGAC
solyc-miR008	AAAAAUGACAUGUAGGACACGUGU		solyc-miR405	AUGACUUUUAUGGAAAGUUGUGAC
solyc-miR009	AAAAAUGACAUGUAGGACAUAUGU		solyc-miR406	AUGAGUUAGGCUCAAGAACUAAUU
solyc-miR010	AAAAAUGGGCUAGCCCGACCUGAC		solyc-miR407	AUGAUAGUGGAACCCUAUUGGUGU
solyc-miR011	AAAAAUGGGCUAGCCCGGCCUGAC		solyc-miR408	AUGAUCUAGCUUUUGGGAUUGAGU
solyc-miR012	AAAACCGUUAAAUAAGUGAUCAAU		solyc-miR409	AUGCAACAUUUAUGAUGUCUACGU
solyc-miR013	AAAACGUGUCGAUCCAGUCUCACC		solyc-miR410	AUGCAAUAUUUAUGAUGUCCACGU
solyc-miR014	AAAAGAGCUGUUCGCAGAGAG		solyc-miR411	AUGCAAUAUUUAUGAUGUUCACGU
solyc-miR015	AAAAGUAAAACGUUACUUACAGUG		solyc-miR412	AUGCACUGAUGACUUGCAGUGACU
solyc-miR016	AAAAGUAGAAAUAUAUAUCCU		solyc-miR413	AUGCCAUGUAGGACACGUGUGUCU
solyc-miR017	AAAAGUGACCUGCAUUUUACUGAC		solyc-miR414	AUGCUCCUCGUUAUGAAAGUGACU
solyc-miR018	AAAAGUUGUGACUUUUAUGAAGAG		solyc-miR415	AUGCUCUAAUUUAUGGACGACAGG
solyc-miR019	AAAAUAAAGAAAAAAUUAUAAAUA		solyc-miR416	AUGCUCUAGUUUUUGAACAGCAGA
solyc-miR020	AAAAUGAGGCGUAUGCUCGCUCCG		solyc-miR417	AUGGAGUCACGGAGUGUGCCACGU
solyc-miR021	AAAAUUCAUUGACCAGUAAAGAGA		solyc-miR418	AUGGCACUAUUUUGUGAGCCCAAC
solyc-miR022	AAAAUUGUCAUGUAGAACGAAUGU		solyc-miR419	AUGGGCCUUGUUGAAUGGCGUAUG
solyc-miR023	AAAAUUGUCAUGUAGGACGAAUGU		solyc-miR420	AUGGGUCUGAAGAGGUCUGAAUGA
solyc-miR024	AAACCGAUAACCCAAUAAGACGAU		solyc-miR421	AUGGGUGGGCUAGACCAUAUUGAC
solyc-miR025	AAACCUUACUUGACAUAUACUUAU		solyc-miR422	AUGGUAUCGUGAGAUGGAAUCAGC
solyc-miR026	AAACGAUGUCUCAUCCACGUCAGA		solyc-miR423	AUGGUGUCUCAAAAUAUCAAG
solyc-miR027	AAACGCAUAAACGUAACUGUGAGU		solyc-miR424	AUGUAAGGAUAUGACACAUGAGCC
solyc-miR028	AAACGCUGAUUCCAUCUCACGAUA		solyc-miR425	AUGUAGGACAUGUGUGUCUACUUG
solyc-miR029	AAACGCUGAUUCUAUCUCAUGAUA		solyc-miR426	AUGUAGGCUCAGGCUAUGACUGCC

solyc-miR030	AAACUUAGACCAACCCUUUGUAAA
solyc-miR031	AAAGAACUAGACCUUUAGCGGCGA
solyc-miR032	AAAGACAAAAUAAAAUGGGUCAUA
solyc-miR033	AAAGACUGAUUUAUUCACUGGAAG
solyc-miR034	AAAGACUUUAGGGACAUAUACAAA
solyc-miR035	AAAGAGGGGCAUACUGUGAACACC
solyc-miR036	AAAGAUAUAUAUGUGUCCACGUGG
solyc-miR037	AAAGAUUGAGACUAUCCUAUA
solyc-miR038	AAAGCCAUGGACGGAAGUGGGA
solyc-miR039	AAAGCCUGUAUGGGUUAGCCCGGA
solyc-miR040	AAAGCGAGAGAAAAUUGUAUAUAU
solyc-miR041	AAAGCGCACAAGGGUCGCCUCGCU
solyc-miR042	AAAGCUUAUAUGGAUUAGCCCGGA
solyc-miR043	AAAGGAGGGUAUACCAACUCUAGA
solyc-miR044	AAAGGUAUAUAUGUCCUCACGUGG
solyc-miR045	AAAGGUAUAUAUGUGCCCACGUGG
solyc-miR046	AAAGGUAUAUAUGUGCUCACGUGG
solyc-miR047	AAAGGUAUAUAUGUGUCCACGUGG
solyc-miR048	AAAGGUAUAUAUGUGUUCACGUGG
solyc-miR049	AAAGGUGUAUAUGUGCCUACGUGA
solyc-miR050	AAAGUAGGGCAAACAAAUUGAACC
solyc-miR051	AAAGUAUCUCAGAAGACUCAGAUG
solyc-miR052	AAAGUGAAAGCGACUAAUGCUAAG
solyc-miR053	AAAGUGAAGUUGAAUUCUGCUA
solyc-miR054	AAAGUGAGACGAACAAAUUGAAUC
solyc-miR055	AAAGUUGAACAAGUAGAUACAUGU
solyc-miR056	AAAUAGGGACACGUGACACAAUCU
solyc-miR057	AAAUAGGGACACGUGAUACAAUCU
solyc-miR058	AAAUCUCGAUUUACUUGGGUCACC
solyc-miR059	AAAUGAACUUAAUGGGCUGAAUCU
solyc-miR060	AAAUGAGUUGAUUUGGAAAACAUU
solyc-miR061	AAAUGCUAGUUGCACGUACAAAGG
solyc-miR062	AAAUUCUUGUUGGUGAUUAUUUAU
solyc-miR063	AACAACAUACUUACUGAAAUGCCA
solyc-miR064	AACAACUGAUAGUUGAGGUGU
solyc-miR065	AACACUUGCUGUAGAUAAGUCAGU
solyc-miR066	AACCAAUGAAUUCAACCCCAACUU
solyc-miR067	AACCAGUAGGACGGACCGUCGUGG
solyc-miR068	AACCGGUCCUGUAUACUAAAAGAC
solyc-miR069	AACCUACGUGGCACUAUCUUGUGG
solyc-miR070	AACCUUUCAAAUGGUACUGACUGU
solyc-miR071	AACGAGAGAGAGCGGAAUAGCAAA
solyc-miR072	AACGUUGGACCCACAAGAUAGUGC
solyc-miR073	AACGUUGGACCCACAAGAUAGUGU
solyc-miR074	AACGUUGGACCUACAAGAUAGUGC
solyc-miR075	AACGUUGGGCCCACAAGAUAGUGC
solyc-miR076	AACGUUGGGCCCACAAGAUAGUGU
solyc-miR077	AACUGCGUCAUAUCUAACAAG
solyc-miR078	AACUGUCGGGAGACAUUAGCU
solyc-miR079	AACUUACGUGGCACUAUCUUGUGG
solyc-miR080	AACUUUAAGUGGCUAUCGAUGACU
solyc-miR081	AACUUUGAGAGUGCACAUCUAGAC
solyc-miR082	AACUUUUAUGAAAGGUUGUGAGCU
solyc-miR083	AAGAAAAUAAAGACUUCACAGACU
solyc-miR084	AAGAAUGAAUGACAUGAAAAAAGC
solyc-miR085	AAGACCCUGUGACGGUCCGUCGUG
solyc-miR086	AAGACGCUAAUAAGAAUCACGAAC
solyc-miR087	AAGACGGACAACGUCGCGACGA
solyc-miR088	AAGACUGGUCAACGGAACGGGACG
solyc-miR089	AAGAGUUGCGACUUUUAUGAAGAG
solyc-miR090	AAGAGUUGCGAUUUUUAUGAAGAG
solyc-miR091	AAGAUAGUGUCACGUAAGUCGAAA
solyc-miR427	AUGUAUCUAGAGUGAUUCACAUGU
solyc-miR428	AUGUAUCUAGUGUGAUUCACAUGU
solyc-miR429	AUGUCAUGUAGGACACGUGUGUCU
solyc-miR430	AUGUCAUGUAGGACAUGUGUGUCU
solyc-miR431	AUGUCAUGUAGGACUCAUGUGUUU
solyc-miR432	AUGUGACACUAGCUUGAAAAAAAA
solyc-miR433	AUGUGACACUAUCUUGUGGGCCCA
solyc-miR434	AUGUGAGCUCAUUUUGUAACGGCA
solyc-miR435	AUGUGGCAUUUAAAGAAAUGAAAA
solyc-miR436	AUGUGUCAUAAUCUUAUCCACCGA
solyc-miR437	AUGUUGUUCAGACUCUUCAAAAAU
solyc-miR438	AUGUUUUAGGAGUAAAAUGUU
solyc-miR439	AUUACAUGUAGGACAUGUGUGUCU
solyc-miR440	AUUACUAGAUCUUUGUUACGCAUG
solyc-miR441	AUUAGAGACUGCAUGAUGCGUAGA
solyc-miR442	AUUAGAGCUGUCAAUAUGGGCUAG
solyc-miR443	AUUAGGUUGAUAUACGUCACUGUA
solyc-miR444	AUUAGUAAUGUAAUGGGUUCUAAU
solyc-miR445	AUUAUGCCACGUAGAACGAGUGUG
solyc-miR446	AUUCAAAUUAUCUUGAGCUCAACC
solyc-miR447	AUUCAUGUAUACGAGAUGGUAAGA
solyc-miR448	AUUCCAUCUCACGAUACCAUAUCG
solyc-miR449	AUUCCAUCUCAUGAUUCCAUAUCA
solyc-miR450	AUUCCGACUAUUUAUUCGAUACUC
solyc-miR451	AUUCGAACCCACGACCUCCAAGUG
solyc-miR452	AUUCUAGAUAACGAAGUAUAAAAC
solyc-miR453	AUUCUCAAUAAACGUACUCCUGAG
solyc-miR454	AUUCUUUAUACGUAGAAGCCAAGA
solyc-miR455	AUUGAACAAGUAGAUACAUGUAUC
solyc-miR456	AUUGAACCCAAUUCAAAUUAUCUU
solyc-miR457	AUUGACAUUUUUGAAGAGUCUGAG
solyc-miR458	AUUGACUUUUUUUUCAAGCUAGUG
solyc-miR459	AUUGCCCAAGCCUUAUAAAGAGUC
solyc-miR460	AUUGUGCCAUGUAGGACUCAUGUG
solyc-miR461	AUUUCUUUCCCAAUAUAUGCA
solyc-miR462	AUUUGCCUCCUCUUGAGGCCGUUG
solyc-miR463	AUUUGGGAGUAUCGAACAAAUAGU
solyc-miR464	AUUUGUAUUUAUUCUAAAGGUAUG
solyc-miR465	AUUUGUGUCAUGCAGAACUCGUGU
solyc-miR466	AUUUGUUUCAAGCUAGUGCCACGU
solyc-miR467	AUUUUGCUUUUUUCGGUGGAGUAA
solyc-miR468	CAAAACGAAUGCUCCUUGCGA
solyc-miR469	CAAAAUAGUUUAGAUGUUUGG
solyc-miR470	CAAAAUAUCAAGCUACUGUUG
solyc-miR471	CAAACUAGUGUCACGUAGGCCGAA
solyc-miR472	CAACUUCACUUUGGAGUACAAG
solyc-miR473	CAAGAUAGUGUCACGUAGGCCGA
solyc-miR474	CAAGAUAGUGUCACGUAGGUUAAA
solyc-miR475	CAAGAUGAGGACUUUGAGAUCAAA
solyc-miR476	CAAUGAAACCUACUAUUUAAG
solyc-miR477	CACGUAGGCCAAAAAGGGGUAG
solyc-miR478	CAGCCAAGGAUGACUUGCCG
solyc-miR479	CAGCUGACGACUCGUUGAUUCU
solyc-miR480	CAGUCCACCGACAGCCGACGGG
solyc-miR481	CAGUUGCAAAAUACACUUUUU
solyc-miR482	CAUACAUUUAAUUAAGAUCACAAG
solyc-miR483	CAUAGUUUUGAACUGUGAUAAG
solyc-miR484	CAUAGUUUUGAACUGUGGUAAG
solyc-miR485	CAUAUAUGCUCUAGUUUUUGG
solyc-miR486	CAUGAACCUCCGGUGUCGCUGA
solyc-miR487	CAUGUGAUGAAGAUUACUACGU
solyc-miR488	CAUUUUACUGACCUACCGUUGAGG

solyc-miR092	AAGAUAGUGUCACGUAGAUCGAAA
solyc-miR093	AAGAUAGUGUCACGUAGGCCGAAA
solyc-miR094	AAGAUAGUGUCACGUAGGCUGAAA
solyc-miR095	AAGAUAUAUAUGUGUUCACGUGGA
solyc-miR096	AAGAUUAUGACACGUGUAUAUCCG
solyc-miR097	AAGAUUAUGACACGUGUAUGUCCG
solyc-miR098	AAGAUUCACUACAAAAGAACUUGA
solyc-miR099	AAGCAUGUCUUGUGGGCAUGGACU
solyc-miR100	AAGCCAUGGACGGAAGUGGGAG
solyc-miR101	AAGCCUAACAUGCAUGGAAGAGGC
solyc-miR102	AAGCUCAGGAGGGAUAGCGCC
solyc-miR103	AAGCUUAUAAGCUGGUCAAAUUGG
solyc-miR104	AAGGGUUGCACUACUAUCUGAUUU
solyc-miR105	AAGGUAUAUAUUUGUCCACGUGGA
solyc-miR106	AAGUAAACUUGGGCGGGUCAAGAC
solyc-miR107	AAGUACGUGGGCUACAGGCUUGCC
solyc-miR108	AAGUAGACAAACAUGUCCUACAUG
solyc-miR109	AAGUAGACACAUGAGUCAUACAUG
solyc-miR110	AAGUAGACACAUGAGUCCUACGUG
solyc-miR111	AAGUCAUCGAUAGCCACUUAAAGU
solyc-miR112	AAGUGAAUUGCAAGGUAAGACACC
solyc-miR113	AAGUGAGGUUUUACUCCUAGAACG
solyc-miR114	AAGUGUGAGGGAAAGCCUCAACCU
solyc-miR115	AAGUGUGUCUCUGAGAUUUCAGAC
solyc-miR116	AAGUUAGAACCAAGAGCUAAAAAG
solyc-miR117	AAGUUGUCGAACGACUAAAGAAGU
solyc-miR118	AAGUUGUGCCUUUAGAGUUACAUU
solyc-miR119	AAUAAGAUUAUGACACGUGUAUGU
solyc-miR120	AAUACAUGGCAGGAAGACAUGAGG
solyc-miR121	AAUACUUUGAGGGUAUUUUAGACC
solyc-miR122	AAUACUUUGAGGGUAUUUUAGGUC
solyc-miR123	AAUAGACACAUAUGUCCUACAUGU
solyc-miR124	AAUCAAUUUGUGUCCAUCUAGUGU
solyc-miR125	AAUCGUCAUGUAGGACGAAUGUGU
solyc-miR126	AAUCGUCGGGCAUGUUCAAACAUG
solyc-miR127	AAUCUUCAAAGUGUAAAAAAGCUU
solyc-miR128	AAUCUUGGAAUGUUUUGCUUUGGU
solyc-miR129	AAUGAAAGGUUGUGACUUUUAUGA
solyc-miR130	AAUGGCGUCAGAUUGAUAUACGUU
solyc-miR131	AAUGGUCGUGUUAAUUCUGAAAGG
solyc-miR132	AAUGUCGGAUCGAUAGAUAAGAUU
solyc-miR133	AAUGUGUACUUUGCAACUGUU
solyc-miR134	AAUGUUGGGCCCACAAGAUAGUGU
solyc-miR135	AAUUAGACUUUGUUCAUUUUAGAC
solyc-miR136	AAUUCAGUCUAUGGGAUAGUCUCA
solyc-miR137	AAUUCUGUGUAUAAAGCGAGAGAG
solyc-miR138	AAUUGAACAAAUAGACACAUUCGU
solyc-miR139	AAUUGUAGUUUCUGACAGAUU
solyc-miR140	AAUUGUCUACACAAAUAGGAGAAU
solyc-miR141	AAUUUUAUACAAAUUUAAGUGUCU
solyc-miR142	AAUUUUGACUGCGAAAACUGCUCU
solyc-miR143	ACAAAAGAACUUGAUCUUUAGUGG
solyc-miR144	ACAACACUCCAUACGUAGUGACUC
solyc-miR145	ACAACCUUGUUGAGCUUGAGGA
solyc-miR146	ACAAGACGGUAUCAUGUAGACCA
solyc-miR147	ACAAGAUAGUGCCACGUAGGCU
solyc-miR148	ACAAGAUAGUGCCACGUAGGUCA
solyc-miR149	ACAAGAUAGUGUCACGUAAGCUAA
solyc-miR150	ACAAGAUAGUGUCACGUAAGUCA
solyc-miR151	ACAAGAUAGUGUCACGUAGACAGA
solyc-miR152	ACAAGAUAGUGUCACGUAGAUCGA
solyc-miR153	ACAAGAUAGUGUCACGUAGGCCA
solyc-miR489	CAUUUUGCAUGAGAGUGUUUC
solyc-miR490	CAUUUUUCUUUUUCAUUCUCC
solyc-miR491	CCAAUGCACCAUACUCAUCUC
solyc-miR492	CCCUAUUCUCACAUCUCACUACCC
solyc-miR493	CCGAAGAGCUUGCGCGCCUCG
solyc-miR494	CCGGUGGAGCUUACAACAGCA
solyc-miR495	CCUGAAUUGUGUUCCAUUCUGG
solyc-miR496	CGAAGAGCUUGCGCGCCUCGG
solyc-miR497	CGAUUUUGCUUUUUCCGGGGAA
solyc-miR498	CGCCAAAGGAGAGCUGCCCUG
solyc-miR499	CGCUAUCCAUCCUGAGUUUCA
solyc-miR500	CGGCAGGUCAUCUUUGGCUAC
solyc-miR501	CGGUAAAAUCGCUGGUGAUGU
solyc-miR502	CGUAAAAUGUCAAAGCAUGUAUGG
solyc-miR503	CGUACGUGACAUUAUCUUGUGAAU
solyc-miR504	CGUAGUCCCUCUAAGAAGCUGG
solyc-miR505	CGUCAGUAUUCGCAGAGGGAG
solyc-miR506	CGUGAUAGAGUUUGAAUACAA
solyc-miR507	CGUGUAUCAUGUCAUGUAGGACA
solyc-miR508	CUAACUCUUCCCCUGCCCUCCAAA
solyc-miR509	CUAAUUUGGUUUGACUGAUUA
solyc-miR510	CUACAUGACAUAAUACACAUAGGA
solyc-miR511	CUACAUGACAUAAUACACGUAGGA
solyc-miR512	CUAGAUUCACGCACAAGCUCG
solyc-miR513	CUCUCCCUCAAAGGCUUCUGG
solyc-miR514	CUGAAGUGUUUGGGGGAACUC
solyc-miR515	CUGCAAUUCAUUAGUGUAUGA
solyc-miR516	CUGUGUAUAAAGUGAGAGAGGCGA
solyc-miR517	CUUACUGCACAGAGUGAGGGAC
solyc-miR518	CUUCACAUUCGACAUGCUAAAGGA
solyc-miR519	CUUCCUACCAGCUCAAGGAUGU
solyc-miR520	CUUCUGACAGAUUUCAUUGUGGAA
solyc-miR521	CUUCUGUGUAUAAAGCGAGAGAGA
solyc-miR522	GAAUUGACUGUUGGGCUUGAGUU
solyc-miR523	GAAUUGGCUGCUGGAUUUGAGU
solyc-miR524	GAAUUGUCUCCAGGAACUUUCAAG
solyc-miR525	GAAUUUCAUUGAGUAUGUUGUUGU
solyc-miR526	GACGCUCUCGCGUCCCACGCCUUC
solyc-miR527	GAUAAGAUUAUGACACGUGUAUGU
solyc-miR528	GCAGCACUAUUAAGAUUCUCA
solyc-miR529	GCGUAUGAGGAGCCAAGCAUA
solyc-miR530	GCUCACUGCUCUAUCUGUCACC
solyc-miR531	GGAGCAUCAUCAAGAUUCACA
solyc-miR532	GGAGGCAGCGGUUCAUCGAUC
solyc-miR533	GGGAUUGGUGAGUUGGAAAGC
solyc-miR534	GUAACUUUUUCAAAGAGUUGUGAC
solyc-miR535	GUACAUCCUACCCUCCCCAGA
solyc-miR536	GUCAAGAUGUUCAUUAAUAGGGAG
solyc-miR537	GUCAUACUUUUGGGACAUUGGUGC
solyc-miR538	GUCCACAAGAUAGUAUCAUGUAGA
solyc-miR539	GUGACUUUUAUGAAAUGUUGCGAC
solyc-miR540	GUGUAUUGUGUCAUGUAGGACUC
solyc-miR541	GUGUCACAUAGGCCGAAAGGAGU
solyc-miR542	GUUACUGUGAGUAACGUUUUAGGA
solyc-miR543	GUUUUAGGAGUAAAACGUUACUGU
solyc-miR544	UAAAAAGAAAAAAGGUAAAAG
solyc-miR545	UAAAAGUCUUGUGGGAUGACC
solyc-miR546	UAAAAUAAGUGGAACUACUAU
solyc-miR547	UAAACAAGGCUCUAUGACGGUC
solyc-miR548	UAAACCUCUGUUCUUGUAAUAA
solyc-miR549	UAAAGAAGAGAGAAUCUAGGA
solyc-miR550	UAAAGAUAUGACACAUGAGCCCA

solyc-miR154	ACAAGAUAGUGUCACGUAGGCUA
solyc-miR155	ACAAGAUAGUGUCACGUAGGCUG
solyc-miR156	ACAAGAUAGUGUCACGUAGGUCA
solyc-miR157	ACAAGAUAGUGUCAUGUAGGCCA
solyc-miR158	ACAAGAUAGUGUCAUGUAGGCCG
solyc-miR159	ACAAGCGUUGGACCCACGUAGGCC
solyc-miR160	ACAAGUACGUGGGCUACAGGCUUG
solyc-miR161	ACAAUUCCAUGUGGCAUUUAAAGA
solyc-miR162	ACACUUUUGUUGUAUAAAGCGAGA
solyc-miR163	ACAGCGAAUCGAAACUACGAG
solyc-miR164	ACAGGACUUCUGUCUAGGAUGAGU
solyc-miR165	ACAGUCUCCAACAGUGGCAGAAGU
solyc-miR166	ACAUAUGAGAGAUAGACUUAGA
solyc-miR167	ACAUCGAUGGUUAAUGAGAUGGGU
solyc-miR168	ACAUGACACAAUACAUGUAGGACA
solyc-miR169	ACAUGACAUAAUACAUGUAGGACG
solyc-miR170	ACAUGAUAGAGUUUGAAUACAAAU
solyc-miR171	ACAUUACAUGAUGGACACAACACU
solyc-miR172	ACAUUUUUGAAGAGUCUGAACAAC
solyc-miR173	ACCACGAAAAUCCACUUUGAAACU
solyc-miR174	ACCCAAAAGAUAGUGUCACGUAGA
solyc-miR175	ACCCACAAGAUAGUGCCACGUAGA
solyc-miR176	ACCCACAAGAUAGUGCCACGUAGG
solyc-miR177	ACCCACAAGAUAGUGUCACGUAAG
solyc-miR178	ACCCACAAGAUAGUGUCACGUAGA
solyc-miR179	ACCCACGAGAUAGUGUCACGUAGA
solyc-miR180	ACCCCUUUUCGGCCUACGUAGCAC
solyc-miR181	ACCCCUUUUUUGUCUACGUGGCAC
solyc-miR182	ACCGCCGUUUCAAUUUGUUUGAUC
solyc-miR183	ACCUACGUGGCACUAUCUUGUGGA
solyc-miR184	ACCUCUGGUUAAGGGUGAAACACU
solyc-miR185	ACCUUUUGUCUUACGUGACACACU
solyc-miR186	ACCUUUUGUCUUACGUGGCACACU
solyc-miR187	ACGACGGAUUCAUGCGACGGUC
solyc-miR188	ACGACUUUUAUGAAGAGUUGCGAC
solyc-miR189	ACGAGAGUCAUCUGUGACAGG
solyc-miR190	ACGAUAUGGUAUUGUGAGAUGGAA
solyc-miR191	ACGAUCAUUUGAAGUUUCACCAAA
solyc-miR192	ACGAUUGAUGUAAUGGCAAGA
solyc-miR193	ACGGGCCGUCAUGGCCUCCGUCGU
solyc-miR194	ACGUAAUUCUCAUGAGAUGAGACU
solyc-miR195	ACGUCUUACUCUUAAAACGUUACU
solyc-miR196	ACGUGACACUAGCUUGAAAAAAAA
solyc-miR197	ACGUGACACUAGCUUGAAAGAAAA
solyc-miR198	ACGUGACACUAUCUUGUGAGCCCA
solyc-miR199	ACGUGACACUAUCUUGUGGGACCA
solyc-miR200	ACGUGACACUAUCUUGUGGGCUCA
solyc-miR201	ACGUGACACUAUUUUGUGGGCCCA
solyc-miR202	ACGUGACAUUAUCUUGUGGGCCCA
solyc-miR203	ACGUGAUACUAUCUUGUAGACCCA
solyc-miR204	ACGUGGCACUAGUUCGAAAAAAAA
solyc-miR205	ACGUGGCACUAGUUUGAAAAAAAA
solyc-miR206	ACGUGGCACUAUCUUGUGGGACCA
solyc-miR207	ACGUGGCACUAUCUUGUGGGCUCA
solyc-miR208	ACGUGUAUUGUGUCAUGUAGAACA
solyc-miR209	ACGUGUAUUGUGUCAUGUAGGACU
solyc-miR210	ACGUGUAUUUUGUCAUGUAGGACU
solyc-miR211	ACGUGUCGUAAUCUUAUCCACCGA
solyc-miR212	ACGUUUUACUCCUAAAACGUUACU
solyc-miR213	ACUAAUAAAAAGUGUGCUCUC
solyc-miR214	ACUCACAAGAUAGUGACACGUAGG
solyc-miR215	ACUCACAAGAUAGUGCCACAUAGA
solyc-miR551	UAAAGAUAUGACACAUGAGCCCAA
solyc-miR552	UAAAUAGUAGGUUUCAUCCGU
solyc-miR553	UAAGAAGGGUACUUUUGAUCCAAU
solyc-miR554	UAAGAUUAUGACACGUGUAUGUCU
solyc-miR555	UAAGCGUUUAAAUUACCUCGAG
solyc-miR556	UAAGUUUUGGCGUUAGGUGUCAGU
solyc-miR557	UAAUUAUUUCUUGCUUGAUUG
solyc-miR558	UACAAGAGAAAUGACAACCUC
solyc-miR559	UACAUGACACUAUCUUGUGGAUCC
solyc-miR560	UACAUGACACUAUCUUGUGGGUCC
solyc-miR561	UACAUGUAUUGUGUCAUGUAGGAC
solyc-miR562	UACGUGACACUAGCUUGAAAGAAA
solyc-miR563	UACGUGACACUAGUUUGAAAGAAA
solyc-miR564	UACGUGACACUAUCCCGUGGGUCU
solyc-miR565	UACGUGACACUAUCUUGUGGAUCC
solyc-miR566	UACGUGACACUAUCUUGUGGGCU
solyc-miR567	UACGUGACACUAUCUUGUGGGUCC
solyc-miR568	UACGUGACACUAUCUUGUGGGUCU
solyc-miR569	UACGUGGCACUAUCUUGUAGGUCC
solyc-miR570	UACGUGGCACUAUCUUGUAGGUCU
solyc-miR571	UACGUGGCACUAUCUUGUCGAUCC
solyc-miR572	UACGUGGCACUAUCUUGUGGACCC
solyc-miR573	UACGUGGCACUAUCUUGUGGAUCC
solyc-miR574	UACGUGGCACUAUCUUGUGGGUC
solyc-miR575	UACGUGGCACUAUCUUGUGGGUCC
solyc-miR576	UACGUGGCACUAUCUUGUGGGUCU
solyc-miR577	UACGUGGCAGUAUCUUGUGGGUGC
solyc-miR578	UACGUGGCAUUAUCUUGACGGUCC
solyc-miR579	UACGUGGCAUUAUCUUGUGGGUC
solyc-miR580	UACGUGGCAUUAUCUUGUGGGUCC
solyc-miR581	UACGUGUAUUAUGUCAUGUAGGAC
solyc-miR582	UACUCCCUGACCGUCAUAGAGC
solyc-miR583	UACUCCUAAAACGUUACUGUGAGU
solyc-miR584	UAGAAUCUGUUCCUCUUGGGAG
solyc-miR585	UAGAUGGUGAUACUCCAGGUA
solyc-miR586	UAGAUGUGCAUACUCAAAGUU
solyc-miR587	UAGCCAAGGAUGACUUGCCUG
solyc-miR588	UAGCUUGACGUCUAACCCUAAAAC
solyc-miR589	UAGGAGUAAAACGUUACUAUGAGU
solyc-miR590	UAGGAUGUCAUGUAGAACAUGUGU
solyc-miR591	UAGUUAAAUUUUGGAAAAUUU
solyc-miR592	UAGUUAUUAUCUCAAAAAGUG
solyc-miR593	UAUGAAAAGUUGUGACUUUUAUA
solyc-miR594	UAUGAAGAGUUGUGACUUUUAUA
solyc-miR595	UAUGGAAAAUAUAUCUGCUAUUCC
solyc-miR596	UAUGGAAGAGGUGAUUGGAGA
solyc-miR597	UAUUAUGUCAUGUAGGACAUGUGU
solyc-miR598	UAUUAUGUCAUGUAGGACUCAUGU
solyc-miR599	UAUUGACACCAAAGAGAAGAU
solyc-miR600	UAUUGGAAUUUGGAAUAGUUUG
solyc-miR601	UAUUGGCCUGGUUCACUCAGA
solyc-miR602	UAUUGGUGCGGUUCAAUGAGA
solyc-miR603	UAUUGUGACAUGUAGGACUCAUGU
solyc-miR604	UAUUUCGAUCAUUGCAGUUGC
solyc-miR605	UAUUUCUGCAGCUUUGGAAUU
solyc-miR606	UAUUUGUUUUCUAGGAAAACC
solyc-miR607	UAUUUUGUCAUGUAGGACUCAUGU
solyc-miR608	UCAAACUAUACAGUUAUCAAG
solyc-miR609	UCAAAUCCAAUGUACAAGCAU
solyc-miR610	UCAAGAUUGUGAAUAUAAAUU
solyc-miR611	UCAAGCUCAACGAGGCCAUCUC
solyc-miR612	UCAAGUAGAUGCAUCACCUGCUUA

solyc-miR216	ACUCACAAGAUAGUGUCACGUAAG
solyc-miR217	ACUCACAGUAACGUUUUAGGAGUA
solyc-miR218	ACUCAGGAGUACGUUUAUUGAGAA
solyc-miR219	ACUCAUGACAAACUUAAGAAAAUG
solyc-miR220	ACUGACCGAACCACGUUAAAUCUU
solyc-miR221	ACUUACGUGACACUAUCUUGUGAG
solyc-miR222	ACUUCUGCCACUGUUGGAGACUGU
solyc-miR223	ACUUUGUCAUUUGAAGCUGAUAUA
solyc-miR224	ACUUUUAUGAAGAGUUGCGACUUU
solyc-miR225	AGAAACAACACUUGCUAAAGG
solyc-miR226	AGAAACAACACUUGCUAAAGGA
solyc-miR227	AGAAAGCUGACGAUUGGAACAUGU
solyc-miR228	AGAACAGAAUGGAUAUAGAGGAUU
solyc-miR229	AGAACGUGCAUAUGUUGCCACAUA
solyc-miR230	AGAACUCAAGUCAAAAGUUGAAUU
solyc-miR231	AGAAGUUGUGACUUUAAUAAAGAG
solyc-miR232	AGAAUAGUUGGCUCAUUAGGUUAA
solyc-miR233	AGAAUGAAAAAGAAAAAUGUA
solyc-miR234	AGAAUUUAAAAUACUAAUAUCAUG
solyc-miR235	AGAAUUUAGAAACGAUUUCACUGA
solyc-miR236	AGAAUUUGUGAAGUAUGGGACGAC
solyc-miR237	AGACAUGAUAAUAGAUGGUUA
solyc-miR238	AGACCUAUGUGGUACUAGCUUGAA
solyc-miR239	AGACGACGGAACCCACGACGGAUC
solyc-miR240	AGACGGCGUUAAGAAGCAUUACGU
solyc-miR241	AGACUAAAUAGGGACACGUGACAC
solyc-miR242	AGACUUGAUGAAACUCAAAGACGA
solyc-miR243	AGACUUGUUGACAACCUGGGCUCU
solyc-miR244	AGACUUGUUGACAACGAUACCAAU
solyc-miR245	AGACUUUUUUGAAGAGUUGUGACU
solyc-miR246	AGAGAAAAGACAUAAAGUGACACC
solyc-miR247	AGAGAAAUCAAGGAUUGUAUUGAU
solyc-miR248	AGAGAGUGUGAAUAGUAGUUCUCC
solyc-miR249	AGAGAUACAAUGAAUGUGCUC
solyc-miR250	AGAGAUAUUUGAAUGUCACGUAAG
solyc-miR251	AGAGCUGUCAAUAUGGGUUAGUCC
solyc-miR252	AGAGUGCAGAGUGUAGACUGCAGA
solyc-miR253	AGAGUUGCGAUUUUUAUGAAGAGU
solyc-miR254	AGAUAAGAUUAUGACACGUGUAUG
solyc-miR255	AGAUACUCAUUUGGGUAAGAAG
solyc-miR256	AGAUAGUGCCACGUAGGUCGAAAA
solyc-miR257	AGAUAGUGUCACGUAGGCCA
solyc-miR258	AGAUAUUGAUGCGGUUCAAUC
solyc-miR259	AGAUAUUUGGAUGUCACGUAAGUC
solyc-miR260	AGAUCGGUGGAUAAGAUUGUGACA
solyc-miR261	AGAUCGUCCGCACAGACACGUGGA
solyc-miR262	AGAUCUAGUCAUGCCUUGGGCC
solyc-miR263	AGAUGAUUAUGUGCUCUUGCA
solyc-miR264	AGAUUUAGGAGUAAAGUACAAUU
solyc-miR265	AGCAUAGUAAAGAGUAGAGACU
solyc-miR266	AGCAUUCCAUCAACAUUUUCUAGC
solyc-miR267	AGCCCACAAGAUAGUGUCACGUAG
solyc-miR268	AGCCUACGUGCAGUAUCUUGUGGG
solyc-miR269	AGCCUACGUGGCACUAGUUCGAAA
solyc-miR270	AGCCUACGUGGCACUAUCUUGUGA
solyc-miR271	AGCCUACGUGGCACUAUCUUGUGG
solyc-miR272	AGCGAUUUAACAUUAUUUUUUAAU
solyc-miR273	AGCGUUGGACCCACAAGAUAGUGU
solyc-miR274	AGCGUUGGAUCCACAAGAUAGUGU
solyc-miR275	AGCGUUGGGCCCACAAGAUAGUGC
solyc-miR276	AGCGUUGGGCCCACAAGAUAGUGU
solyc-miR277	AGCUAUCGAAUAACGUCGGAUGAC
solyc-miR613	UCAAUGAAUUUCCUGAUUUCUA
solyc-miR614	UCAAUGCUACAUACUCAUCCC
solyc-miR615	UCACAUUGCACAAAAAUGAAU
solyc-miR616	UCACAUUGUCAUAUUUGAAUC
solyc-miR617	UCACCUCAGCAUCAGUCAACCU
solyc-miR618	UCACUUAUUCCUUUUUUCUACC
solyc-miR619	UCAGGCGUCUAACUUCCAAGUUUA
solyc-miR620	UCAGGCUAUGACUGCCCAAGUC
solyc-miR621	UCAGUUAAAGGUACCUGAGAG
solyc-miR622	UCAUCAACCUUAAACAAGGCUCUA
solyc-miR623	UCAUCCCACAAGACUUUCACG
solyc-miR624	UCAUUAAGACUGAUAAUUUUACU
solyc-miR625	UCCAAAGGGAUCGCAUUGAUCC
solyc-miR626	UCCAGCGAUUUUGCCAUUUUG
solyc-miR627	UCCAGUCACCUCAUCCGUAUU
solyc-miR628	UCCAUGCAUGUUAGGCUUUCU
solyc-miR629	UCCCUUUUCGGCCUACGUGACACU
solyc-miR630	UCCGGAAUGUCGAGUUUGGUCU
solyc-miR631	UCCUCAAGAGUUUGUAGACGAG
solyc-miR632	UCGAUAAACCUCUGCAUCCAG
solyc-miR633	UCGAUCAAGGGUAUGGGUGUGC
solyc-miR634	UCGAUGGAAAUUGUACUUCAAG
solyc-miR635	UCGGACCAGGCUUCAUUCCCC
solyc-miR636	UCGUAGUUUCGAUUCGCUGUC
solyc-miR637	UCUACUCUUUGCUAUACUGAA
solyc-miR638	UCUAGGAACUUUUUUGGUAGCAUG
solyc-miR639	UCUAUCUCUCAUAUGCUACUA
solyc-miR640	UCUCUGACCAUGUGUCCACUCU
solyc-miR641	UCUGCAAUGAAGUGAAAACAG
solyc-miR642	UCUGUGGAUUCUUUAACUUCUU
solyc-miR643	UCUUACCAUAUCUCGUAUACAUG
solyc-miR644	UCUUGAGGACGCGAUGAGUUUC
solyc-miR645	UCUUGCCAAUACCGCCCAUUCC
solyc-miR646	UCUUGCCUACACCGCCCAUGCC
solyc-miR647	UCUUGGAUUAGCCUCUGGUGG
solyc-miR648	UCUUUAUUUAGUUGUCACCGU
solyc-miR649	UCUUUCCUACUCCUCCCAUACC
solyc-miR650	UGAAAUCGCUACCUAGGCAGC
solyc-miR651	UGAAAUCUGAAAACAAAUUGU
solyc-miR652	UGAACAAAAAGUUACCUGCAG
solyc-miR653	UGAACCCACAAGAUAGUGCCACGU
solyc-miR654	UGAACCCACAAGAUAGUGUCACGU
solyc-miR655	UGAACUCACAAGAUAGUGUCACGU
solyc-miR656	UGAAGAGUCUGAGCAACUUAG
solyc-miR657	UGAAGCUGCCAGCAUGAUCUA
solyc-miR658	UGAAGUGUUUGGGGGAACUC
solyc-miR659	UGAAUCUUCAAGUGAUGAGGCA
solyc-miR660	UGAAUCUUGAUGAUGCUGCAU
solyc-miR661	UGACAACUAAAUAAAGACGGA
solyc-miR662	UGACACUAUCUUGUGGGCUC
solyc-miR663	UGACAGAAGAGAGUGAGC
solyc-miR664	UGACAGAAGAGAGUGAGCAC
solyc-miR665	UGACCGUCUAUAAUCAUGUCU
solyc-miR666	UGACGGAUUAGCCUGUUUUGGC
solyc-miR667	UGACUAUUUUUUCAUAUCUGA
solyc-miR668	UGAGCCCACAAGAUAGUGCCACGU
solyc-miR669	UGAGCCCACAAGAUAGUGUCACGU
solyc-miR670	UGAGCCGAACCAAUAUCACUC
solyc-miR671	UGAGGUCUCUUUUUGUUUUACAAG
solyc-miR672	UGAUAUUUUGAGACACCAUUU
solyc-miR673	UGAUCUGUAUCUCUAUGACUG
solyc-miR674	UGAUGCAACUAUGAAACCAUC

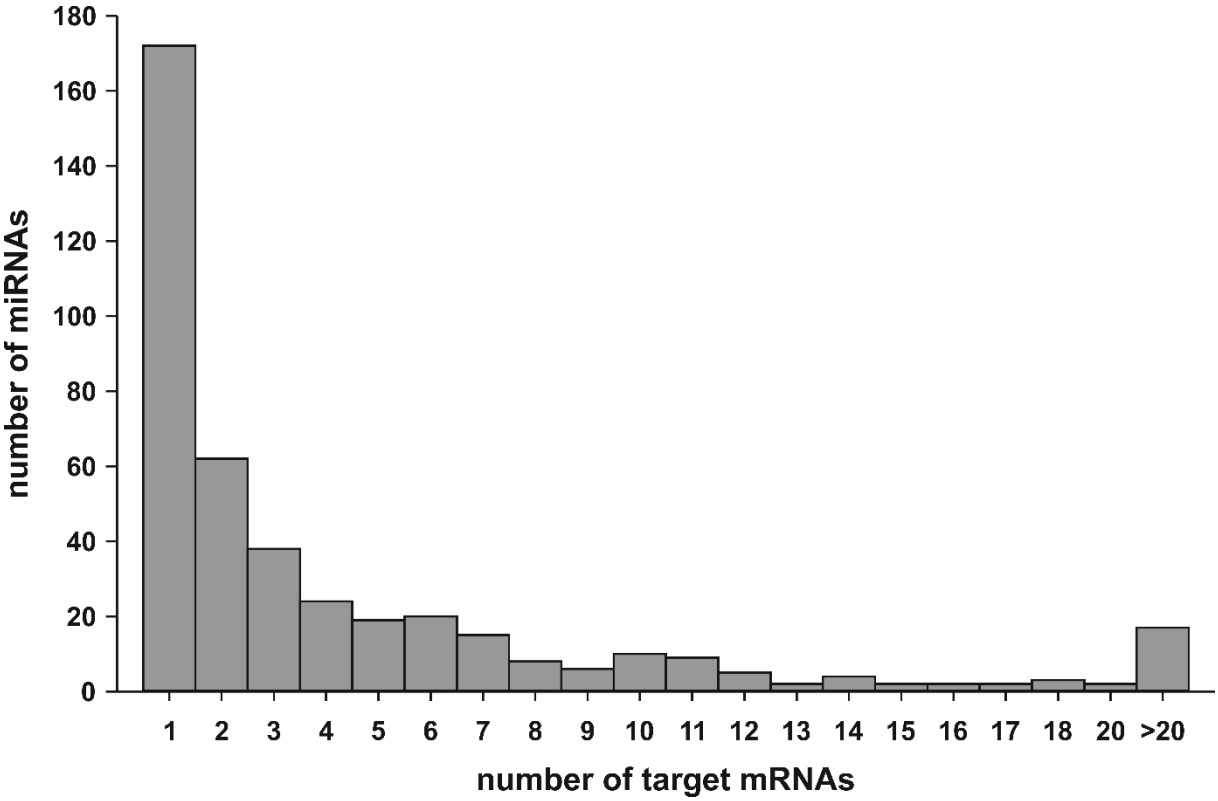
solyc-miR278	AGCUGCCUUCGAGAGGCUGAAG
solyc-miR279	AGCUGCUGACCUAUGGAUUCC
solyc-miR280	AGCUUACGUGACACUAUCUUAUGG
solyc-miR281	AGCUUACGUGGAACUAUCAUGUGG
solyc-miR282	AGCUUACGUGGCACUAUCUUGUAG
solyc-miR283	AGGAAAAGGAAACAACUGAUAGUU
solyc-miR284	AGGACCCAUAAGAUAGUGCCACGU
solyc-miR285	AGGACCUUAGUAUAGUAUAAGUAU
solyc-miR286	AGGACGUGUGUGUUUACUUGUUCA
solyc-miR287	AGGAGUGUCGGUCUGCAUGCUC
solyc-miR288	AGGAUCUGACAUGUAUCUGACGAU
solyc-miR289	AGGCCCAUGUGUCAUAUCUUUAC
solyc-miR290	AGGCCCAUGUGUCAUAUCUUUACA
solyc-miR291	AGGCGCAUGUGUCAUAUCUUUACA
solyc-miR292	AGGCUCUAUUUAUAUUGUGUCAUG
solyc-miR293	AGGGAUCUGUUUAUGUAUUUGACC
solyc-miR294	AGGGCAUAGAAUCUGUUCCUCU
solyc-miR295	AGGUAUGAAACUCACACUCUCAAU
solyc-miR296	AGGUCAACCUGUGGCUUAUAAGGC
solyc-miR297	AGGUUAUUCUUCUCUGGCAAG
solyc-miR298	AGUAACAUUUUACUCCUAAAACGU
solyc-miR299	AGUAACAUUUUAGGAGUAAAACGU
solyc-miR300	AGUAACAUUUUAGGUGUAAAACGU
solyc-miR301	AGUAACGUAUAUCAAUCUGAUGCC
solyc-miR302	AGUAACGUUUUACUCCUAAAACA
solyc-miR303	AGUAAUAUUUUAGGAGUAAAACGU
solyc-miR304	AGUAGGUGAUGCAACUACUUGUCU
solyc-miR305	AGUCACGUAGUGUGCACGUAAGAC
solyc-miR306	AGUCCGCUCGUUUUAUGGCCAGUC
solyc-miR307	AGUCCUACAUGGCAAAAUACACGU
solyc-miR308	AGUCUACGUGGCACUAUCUUGUGG
solyc-miR309	AGUGAAAGCGACUAAUGCUAAGGA
solyc-miR310	AGUGACGUAUAUCAACCUAAUGCU
solyc-miR311	AGUGACUUUCUGAAAUAAUAACUA
solyc-miR312	AGUGACUUUCUGAGGAAAGAGUGG
solyc-miR313	AGUGCCACCUAGGCUGAAAAAGGA
solyc-miR314	AGUGCCACGUAGGCCGAAAAGGGA
solyc-miR315	AGUGCCACGUAGGCUGAAAAGGGG
solyc-miR316	AGUGGACAAGUAAAGGUGGAUGGA
solyc-miR317	AGUGGAUAAGAUUAUGACACAUGU
solyc-miR318	AGUGGAUAAGAUUAUGACACGUGU
solyc-miR319	AGUGGCCGAACGCCGGUCUGAAGU
solyc-miR320	AGUGGGAGUGGCAUUGACUGUC
solyc-miR321	AGUGGGUGGUGUGGUAAGAUU
solyc-miR322	AGUGUCACGUAGGCCGAAAAGGGA
solyc-miR323	AGUGUCCCGUAUGUAAAAAGAACA
solyc-miR324	AGUGUGCCACGUAAGAUAAAAGGU
solyc-miR325	AGUGUGCCAUGUAAAACAAAAGAU
solyc-miR326	AGUGUGUCACGUAAGACAAAAGGU
solyc-miR327	AGUUAUGGUUAUAACGACUUAAG
solyc-miR328	AUAAAACGAGCGGACUACUCGAGU
solyc-miR329	AUAAAAUUGAACAAGUAGACACAU
solyc-miR330	AUAAACCUUACCUCCGAUCCUAUG
solyc-miR331	AUAAAUAGUAAUGUGUCCACGUGA
solyc-miR332	AUAAAUAGUAAUGUGUCCACGUGG
solyc-miR333	AUAACGUAGCUGCCAUGAUUCA
solyc-miR334	AUAAGAUAGUGCCACGUAGGACA
solyc-miR335	AUAAGAUUGUGCCAAGUGUCC
solyc-miR336	AUAAUAUCGUGAGAUGAAAUCAGC
solyc-miR337	AUACAAGAGAAAUGACAACCU
solyc-miR338	AUACACGUAUCAUAAUCUUAUCCA
solyc-miR339	AUACACGUGUCAUAAUCUUAUCCA
solyc-miR675	UGAUGUUGGAAUGGCUCAAUC
solyc-miR676	UGAUUGAGCCGUGCCAAUAUC
solyc-miR677	UGCACUGAUGACUUGCAGUGA
solyc-miR678	UGCACUGCCUCUUCCCUGGCU
solyc-miR679	UGCAGAUUAUUAAGGGUUUCAA
solyc-miR680	UGCAGGCUCAGGCUAUGACUGC
solyc-miR681	UGCCACGUAAGCUAAAAAGGGGUA
solyc-miR682	UGCCUAGGUAGCGAUUUCAUU
solyc-miR683	UGCCUGGCUCCCUGGAUGCCAA
solyc-miR684	UGCCUGGCUCCCUGUAUGCCA
solyc-miR685	UGCGGCGCGCGGCGUCGUUGAGGA
solyc-miR686	UGGAAGCUUCAUUGGUUGUCC
solyc-miR687	UGGACAGCGAAUCGAAACUAC
solyc-miR688	UGGACCAACCCUUUUUGACAACUC
solyc-miR689	UGGACCCACAAAAUAGUGUCACGU
solyc-miR690	UGGACCCACAAGAUAGUGUCACGU
solyc-miR691	UGGACCCAUAAGAUAGUGUCACGU
solyc-miR692	UGGAGAAUUAAGACUGUACCG
solyc-miR693	UGGAGGCAGCGGUUCAUCGAUC
solyc-miR694	UGGAGGUAGUAUGAAUUUUGU
solyc-miR695	UGGAUCCACAAGAUAGUGCCACGU
solyc-miR696	UGGGGACGGUCUUUCCAUAAU
solyc-miR697	UGGGGAGGGUGGUGUGUACGC
solyc-miR698	UGGGUUCUCCCUCUGAACAUCC
solyc-miR699	UGUAUUGUGUCAUGUAGGACUCAU
solyc-miR700	UGUCCACCCUUGAACUCAAGCU
solyc-miR701	UGUCGCAGAUGACUUUCGCCC
solyc-miR702	UGUCUCUCCCUCAAGGGCUCC
solyc-miR703	UGUCUUGCUGUUGUAAGCUCC
solyc-miR704	UGUGAAUGAUGCGGGAGAUAA
solyc-miR705	UGUGCAUCUUCUAAGUUGGAG
solyc-miR706	UGUGGGACAAAAAGAAGCCUA
solyc-miR707	UGUGGGUGGGGUGGAAAGAUU
solyc-miR708	UGUUAAAAGUUUGGCUCAUUU
solyc-miR709	UGUUAAUUCUGAAAGGGUUGCAAC
solyc-miR710	UGUUAGAUAUGACGGAGGUUU
solyc-miR711	UGUUCAAGUUGAUAAUGUUCC
solyc-miR712	UGUUGAUGGUAGAAUGUUUUUC
solyc-miR713	UUAAAGACAAAAUAAAAUGGGUCA
solyc-miR714	UUACAUGUAUUAUGUCAUGUAGGA
solyc-miR715	UUACGUGGCACACUCCGUGACUCC
solyc-miR716	UUACGUGGCACACUCUGUGACUCC
solyc-miR717	UUACGUGGCACUAGCUUGAAACAA
solyc-miR718	UUACUCAGAAUGUUGACUUAAG
solyc-miR719	UUACUCAUGCUACCAAAAAAG
solyc-miR720	UUAGAAUAGUAUAAGUGUGUCUCU
solyc-miR721	UUAGUAUAAUAUAAAUGUGUCUCU
solyc-miR722	UUAGUAUAGUAUAAGUGUCUCUGA
solyc-miR723	UUAUAUAUUUAAAUGUAGAGGA
solyc-miR724	UUAUAUGGACUUGGACAAACCUCC
solyc-miR725	UUAUGCUUCCUGUCAUCCAGA
solyc-miR726	UUAUGUCAUGUAGGACAUGUGUGU
solyc-miR727	UUAUGUGGCACAUUCCGUAACUCC
solyc-miR728	UUAUGUGUUAUGGGGCGGAUU
solyc-miR729	UUCAAACUCUAUCAUGUACAUCAU
solyc-miR730	UUCAAAUCUGACAACGUGAAC
solyc-miR731	UUCAAUUUGUUUGCCCUACUUUAC
solyc-miR732	UUCACAUUCGACAUGCUAAAGGAA
solyc-miR733	UUCAGAAACGUCGACGGGUGC
solyc-miR734	UUCCAGCGAUUUUGCCAUUUU
solyc-miR735	UUCCGAACUAGUGCCACGUAGGCU
solyc-miR736	UUCCGGUUCGCUGCCCAGAAC

solyc-miR340	AUACACGUGUCAUAAUCUUAUCUA
solyc-miR341	AUACACGUGUCAUAAUCUUAUUCA
solyc-miR342	AUACACGUGUCAUAAUUUUAUCCA
solyc-miR343	AUACCUAAACUAGUGAAAGUUUGA
solyc-miR344	AUACCUAAACUAUUGAGAGUGUGA
solyc-miR345	AUACCUGAACUAUUGAAAGUGUGA
solyc-miR346	AUACGUGGUACUAUCUUGUGGACU
solyc-miR347	AUAGACACAUUCGUUCUACAUGAC
solyc-miR348	AUAGAGAGGUGAAUAGUAGUUCUC
solyc-miR349	AUAGAGGUCUGAAUCUUAAUCAUU
solyc-miR350	AUAUACCCCUCGUUAUGAAAGUGA
solyc-miR351	AUAUACCCUCAAGAGUUAACACCC
solyc-miR352	AUAUACGUCACUGUAAAUAACGUU
solyc-miR353	AUAUACGUGUCAUAAUCUUAUCCA
solyc-miR354	AUAUAUAUUGCUACUUUUGAG
solyc-miR355	AUAUAUGCCCUUCACUCUAACGGA
solyc-miR356	AUAUAUGGAGAUAAAUAAUUAUGA
solyc-miR357	AUAUAUUAGGAAAGAAAUAAG
solyc-miR358	AUAUCAGCUUCAAAUGACAAAGUU
solyc-miR359	AUAUCAUGAUACCAUAUCGCAUGG
solyc-miR360	AUAUGCUCUAAUUUAUGAACGACA
solyc-miR361	AUAUGGACUUGGACAAACCUCCCC
solyc-miR362	AUAUUCGUGGUUAUAGGUUAGUAA
solyc-miR363	AUAUUCUAUCAAGUGGUAUCAGAG
solyc-miR364	AUAUUGGUGCGGUUCAAUUAG
solyc-miR365	AUCAAUUUUUCUUUUAAAUAAAGC
solyc-miR366	AUCAUGAUAUGGUAUCGUGAGAUG
solyc-miR367	AUCAUUGGACCCACAAGAUAGUGU
solyc-miR368	AUCCACAAGAUAGUACCACGUAGG
solyc-miR369	AUCCACAAGAUAGUGCCACGUAAG
solyc-miR370	AUCCACAAGAUAGUGCCACGUAGG
solyc-miR371	AUCCACAAGAUAGUGCCAUGUAGG
solyc-miR372	AUCCACAAGAUAGUGUCACGUAGA
solyc-miR373	AUCCACAAGAUAGUGUCACGUAGG
solyc-miR374	AUCCACAAGAUAGUGUCAUGUAGA
solyc-miR375	AUCCACAAGAUAGUGUCAUGUAGG
solyc-miR376	AUCCACAUUCAUUUCAUUACACGA
solyc-miR377	AUCCCUUUUCGGCCUACGUGACAC
solyc-miR378	AUCGAUGGAUAAGAUUAUGACACG
solyc-miR379	AUCGAUGGAUAAGAUUGUAUCACG
solyc-miR380	AUCGAUGGAUAAGAUUGUGCCAUG
solyc-miR381	AUCGGAUCGAUGGAUAAGAUUGUG
solyc-miR382	AUCGUUGGGCCCACAAGAUAGUGU
solyc-miR383	AUCGUUGGGCCUACAAGAUAGUGC
solyc-miR384	AUCGUUGGGCUCACAAAAUAGUGC
solyc-miR385	AUCGUUGGGCUCACAAGAUAGUGC
solyc-miR386	AUCUAGUGUGAUUCACAUGUAUCU
solyc-miR387	AUCUCCCGACUCAUUCAUUCA
solyc-miR388	AUCUGCAUACCAUUUACGAUAGUU
solyc-miR389	AUCUUACGUGGCACACUCGUGACU
solyc-miR390	AUCUUUAUCUCACCAAUCAGACGA
solyc-miR391	AUCUUUCUGACUUAUAUGACAUCU
solyc-miR392	AUCUUUUGUCUUAUGUGGCACACU
solyc-miR393	AUCUUUUUAGCUUACGUGGCAUCC
solyc-miR394	AUGAAAUAUUUAUGAUGUCCACGU
solyc-miR395	AUGAAAUGUUUCAAAUGACGUAUA
solyc-miR396	AUGAAGUCACAGAGUGUGUCACGG
solyc-miR397	AUGACAAGCUACCAUAAUGUGAUG
solyc-miR737	UUCCUGAUUGAUGAACAAAGA
solyc-miR738	UUCUGGGCAGCGAACCGGAAC
solyc-miR739	UUGACAGAAGAUAGAGAGCAC
solyc-miR740	UUGACGUUAUCUCAAUUGCUU
solyc-miR741	UUGAGCCGCGCCAAUAUCACG
solyc-miR742	UUGAGCCGUGCCAAUAUCACG
solyc-miR743	UUGAGGCUGCGAGCCAAUAAGA
solyc-miR744	UUGAGUGUUUUAGAGGCAUGCC
solyc-miR745	UUGAUGGAUAGAGUUACCUGAUA
solyc-miR746	UUGGCAUUCUGUCCACCUCC
solyc-miR747	UUGGUAGACUUGUUAGAUAUGACG
solyc-miR748	UUGGUCGCCCGAAGAGCUUGCG
solyc-miR749	UUGGUCUAGUCCUGCAUCUUU
solyc-miR750	UUGUAGCCCUUUGACAAAAGUG
solyc-miR751	UUGUAGUGAAAUGGAUGUUGAAUC
solyc-miR752	UUGUCCAAAUUGAGCUGCUAG
solyc-miR753	UUGUCCACUUUAGAAAUGACACAC
solyc-miR754	UUGUCUGCGACCUCUGAAAGGUCA
solyc-miR755	UUGUGAUCUUAAUUGAAUGUA
solyc-miR756	UUGUGUUCUCAGGUCACCCCU
solyc-miR757	UUUAAAUUACCUCGAGGGGAAG
solyc-miR758	UUUAAUUUUUUUUUUUGAAAAAUC
solyc-miR759	UUUACUUGUCGACAACAACUU
solyc-miR760	UUUAGCAAGAGUUGUUUUACC
solyc-miR761	UUUAGCUUUGUAUGCCUCUGA
solyc-miR762	UUUAGGUAUGAAACUCAAAAA
solyc-miR763	UUUAUGAAAAGUUGUGACUUCUAU
solyc-miR764	UUUAUUGUUUCGACAUUCUAG
solyc-miR765	UUUCAAAGUGGAUAUUCGUGG
solyc-miR766	UUUCCAAUUCCACCCAUUCCUA
solyc-miR767	UUUCGAACUAGUGCCACGUAGGCU
solyc-miR768	UUUCGAACUUGAGACCUUUAACAU
solyc-miR769	UUUCUCGCUUUAUACAAACACAGA
solyc-miR770	UUUCUGAUCCUACUAUCUAUA
solyc-miR771	UUUCUUUCUUUAUUUACUUGCG
solyc-miR772	UUUGACUAGUAGGCAGGAUCAG
solyc-miR773	UUUGAUCUCCAAGUUCUCAUC
solyc-miR774	UUUGAUUGUGCACAACUAAGC
solyc-miR775	UUUGCAUAUAUCAGGAGCUGC
solyc-miR776	UUUGCCAUGUCUGAAAUCAUC
solyc-miR777	UUUGUUGGUGCAACUGGCAUUCUC
solyc-miR778	UUUUAGGAGUAAAACGUUACUGAA
solyc-miR779	UUUUCGACCUACGUGACACUAUCU
solyc-miR780	UUUUCGAUCUACGUGAUACUAUC
solyc-miR781	UUUUCUACCCCUUUUUGGUCUA
solyc-miR782	UUUUCUGAGAAAACAAAUAGU
solyc-miR783	UUUUGGAAUCUCUGGAAGUUU
solyc-miR784	UUUUGGGCAGAUAGAAGCCGG
solyc-miR785	UUUUUAAUAUAUAACGGAACUUUU
solyc-miR786	UUUUUCAAGCUAGUGCCACGUAGG
solyc-miR787	UUUUUCAAGCUAGUGUCACGUAGG
solyc-miR788	UUUUUGGCCGACGUGGCACUAUCU
solyc-miR789	UUUUUGUGCAAUAUGAUUGAA
solyc-miR790	UUUUUGUUUGUAUAAAGCGAGAGA
solyc-miR791	UUUUUUACUUUUUGAAGAUUAACA
solyc-miR792	UUUUUUCAUGUCAUUCAUUCU
solyc-miR793	UUUUUUUUUAUAAAUCGCUGCACU

Supplemental Table 12: Known miRNAs detected in non- and heat-stressed pollen stages
Listed are predicted miRNAs of non- (CO) and heatstressed (HS) pollen developmental stages with a perfect hit in the miRBase. Based on their taxonomically closest miRBase hit, the miRNAs were classified (column 1) as having a hit in tomato, the Solanaceae family or the Viridiplantae kingdom. For each miRNA the identifier (column 2) and closest miRBase hits (column 3) in the corresponding category are shown. In addition, it is indicated in which samples the miRNA was identified (X, column 4 to 6).

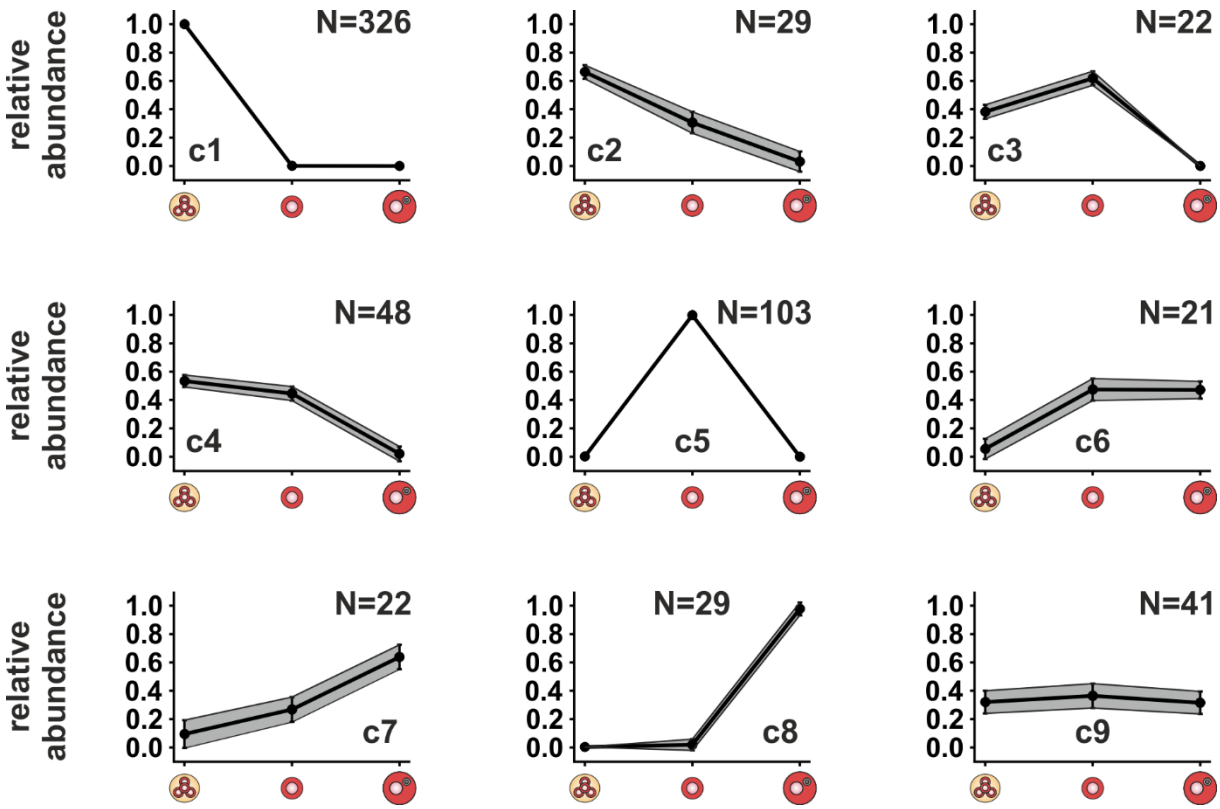
category	miRNA identifier	miRBase hits	tetrads		post-meiotic		mature	
			CO	HS	CO	HS	CO	HS
tomato	solyc-miR102	sly-miR390b-5p	X	X	X	X	X	X
	solyc-miR189	sly-miR1919a, sly-miR1919b, sly-miR1919c-3p	X	X	X	X	X	X
	solyc-miR512	sly-miR403-3p	X	X	X	X	X	X
	solyc-miR632	sly-miR162	X	X	X	X	X	X
	solyc-miR635	sly-miR166a, sly-miR166b	X	X	X	X	X	X
	solyc-miR645	sly-miR482c	X	X	X	X	X	X
	solyc-miR646	sly-miR482b	X	X	X	X	X	X
	solyc-miR649	sly-miR482e-3p	X	X	X	X	X	X
	solyc-miR657	sly-miR167a	X	X	X	X	X	X
	solyc-miR664	sly-miR156d-5p	X	X	X	X	X	X
	solyc-miR676	sly-miR171a	X	X	X	X	X	X
	solyc-miR684	sly-miR160a	X	X	X	X	X	X
	solyc-miR701	sly-miR1919c-5p	X	X	X	X	X	X
	solyc-miR702	sly-miR477-5p	X	X	X	X	X	X
	solyc-miR707	sly-miR482e-5p	X	X	X	X	X	X
	solyc-miR739	sly-miR156a, sly-miR156b, sly-miR156c	X	X	X	X	X	X
	solyc-miR746	sly-miR394-5p	X	X	X	X	X	X
	solyc-miR766	sly-miR482a	X	X	X	X	X	X
	solyc-miR614	sly-miR5304	X	X	X	X		X
	solyc-miR499	sly-miR390b-3p	X	X	X		X	X
	solyc-miR587	sly-miR169b	X	X		X	X	X
	solyc-miR742	sly-miR171b	X	X	X	X		
	solyc-miR530	sly-miR156d-3p			X	X	X	X
	solyc-miR602	sly-miR171c	X					
Solanaceae	solyc-miR531	stu-miR172d-5p	X	X	X	X	X	X
	solyc-miR601	stu-miR171a-5p, stu-miR171c-5p	X	X	X	X	X	X
	solyc-miR625	stu-miR393-5p	X	X	X	X	X	
	solyc-miR693	stu-miR162a-5p, stu-miR162b-5p		X	X	X	X	X
	solyc-miR514	nta-miR395a, nta-miR395b, nta-miR395c, stu-miR395a, stu-miR395b, stu-miR395c, stu-miR395d, stu-miR395e, stu-miR395f, stu-miR395g, stu-miR395h;stu-miR395i, stu-miR395j	X	X	X	X		
	solyc-miR321	stu-miR482e-5p	X		X		X	X
	solyc-miR678	nta-miR408, stu-miR408b-3p			X		X	

	solyc-miR529	stu-miR160a-3p					X	X
	solyc-miR756	stu-miR398b-3p			X			
	solyc-miR670	stu-miR479				X		
	solyc-miR498	nta-miR399a, nta-miR399b, nta-miR399c, nta-miR399d, nta-miR399e, nta-miR399f, nta-miR399g, stu-miR399j-3p, stu-miR399k-3p, stu-miR399l-3p, stu-miR399m-3p, stu-miR399n-3p, stu-miR399o-3p				X		
Viridiplantae	solyc-miR660	gma-miR172k	X		X	X	X	X
	solyc-miR658	tae-miR395b, bdi-miR395a, bdi-miR395b,bdi-miR395c-3p, bdi-miR395m, bdi-miR395e-3p, bdi-miR395f-3p, bdi-miR395g-3p, bdi-miR395h-3p, bdi-miR395j-3p, bdi-miR395k-3p, bdi-miR395l-3p, bdi-miR395n-3p, bdi-miR395o-3p, cpa-miR395a, cpa-miR395b, cpa-miR395c, cpa-miR395d, cpa-miR395e, bdi-miR395q, ata-miR395e-3p, ata-miR395a-3p, ata-miR395d-3p, ata-miR395f-3p, ata-miR395c-3p	X					
	solyc-miR693	ath-miR162a-5p, ath-miR162b-5p, csi-miR162-5p		X				



Supplemental Figure 2: Number of predicted target mRNAs per miRNA

Shown is the number of miRNAs with a certain number of predicted target mRNAs. miRNAs with more than 20 target transcripts were binned.



Supplemental Figure 3: Clustering of miRNA abundance profiles along the development of pollen
The relative abundance profiles of detected miRNAs were k-means clustered over the course of pollen development. Profiles represent the relative abundance of replicate averaged TPM values across the tetrad, post-meiotic and mature stage. Circles represent cluster averages and error bars as well as grey shadings standard deviations. In addition, the cluster number and number of miRNAs assigned to the clusters are given.

Supplemental Table 13: Functional annotation of developmentally relevant MTIs
Listed are MTIs that are likely to be developmentally relevant. For each MTI the assigned category (column 1), the miRNA (column 2), the target mRNA (column 3) as well as the second hierarchy level MapMan term(s) (column 4) and the description (ITAG3.2; column 5) of the mRNA is given. miRNAs of the same category targeting the same mRNA were merged. Further, it was possible that a miRNA had two binding sites on an mRNA, which is indicated by a (2x).

cat	miRNA(s)	SolycID	MapMan term(s)	description
A	solyc-miR663	Solyc10g086650	Biodegradation of Xenobiotics	LOW QUALITY:glyoxal oxidase-related protein
	solyc-miR472	Solyc12g095860	cell.cycle	LEY17226 cyclin-dependent protein kinase 2A-2
	solyc-miR663	Solyc06g008040	cell.division	CLIP-associating 1
	solyc-miR663	Solyc09g063030	cell.division	CLIP-associating family protein
	solyc-miR495	Solyc03g005620	cell.organisation	kinesin like protein
	solyc-miR608	Solyc02g068600	cell.organisation	ankyrin repeat-containing protein
	solyc-miR663	Solyc05g015510	development.squamosa promoter binding like (SPL)	squamosa promoter-binding protein 10
	solyc-miR663	Solyc05g015840	development.squamosa promoter binding like (SPL)	squamosa promoter binding protein 13
	solyc-miR663	Solyc07g062980	development.squamosa promoter binding like (SPL)	squamosa promoter-binding protein, putative
	solyc-miR663	Solyc10g009080	development.squamosa promoter binding like (SPL)	squamosa promoter binding protein 3
	solyc-miR663	Solyc12g011010	development.unspecified	protodermal factor 1
	solyc-miR704	Solyc02g082120	DNA.repair	DNA-3-methyladenine glycosylase, putative

	solyc-miR528	Solyc09g007330	DNA.synthesis/chromatin structure	minichromosome maintenance 8
	solyc-miR663	Solyc01g074000	DNA.synthesis/chromatin structure	histone H3
	solyc-miR663	Solyc12g098640	hormone metabolism.brassinosteroid; cell.vesicle transport	cycloeucalenol cycloisomerase
	solyc-miR606	Solyc12g087980	lipid metabolism.FA synthesis and FA elongation	HXXXD-type acyl-transferase family protein, putative
	solyc-miR536	Solyc04g016330	lipid metabolism.glyceral metabolism	glycerol-3-phosphate dehydrogenase [NAD(+)]
	solyc-miR663	Solyc12g010910	lipid metabolism.lipid degradation	alpha/beta-Hydrolases superfamily protein
	solyc-miR663	Solyc01g073640	misc.short chain dehydrogenase/reductase (SDR)	alcohol dehydrogenase-3,Pfam:PF13561
	solyc-miR064	Solyc09g019980	not assigned.no ontology	cytochrome b561/ferric reductase transmembrane with DOMON related domain, putative
	solyc-miR192	Solyc01g009160	not assigned.no ontology	harpin-induced1-like
	solyc-miR472	Solyc10g086210	not assigned.no ontology	translin family protein
	solyc-miR663	Solyc01g106020	not assigned.no ontology	RNA-binding KH domain protein
	solyc-miR663	Solyc02g084610	not assigned.no ontology	leucine-rich repeat receptor-like protein kinase
	solyc-miR663	Solyc08g005630	not assigned.no ontology	long-chain-alcohol oxidase
	solyc-miR704	Solyc08g008490	not assigned.no ontology	WD-repeat protein, putative
	solyc-miR776	Solyc07g056650	not assigned.no ontology	interactor of constitutive active ROPs protein, putative
	solyc-miR053	Solyc01g010520	not assigned.unknown	helicase with zinc finger protein
	solyc-miR192	Solyc06g073660	not assigned.unknown	Lipid transfer protein
	solyc-miR257	Solyc04g072320	not assigned.unknown	RNA polymerase I subunit 43
	solyc-miR380	Solyc05g054870	not assigned.unknown	U4/U6 small nuclear Prp3-like ribonucleoprotein
	solyc-miR528	Solyc11g071250	not assigned.unknown	protein EMBRYONIC FLOWER 1-like protein
	solyc-miR528	Solyc12g006480	not assigned.unknown	nuclear pore complex Nup205-like protein
	solyc-miR558	Solyc07g054610	not assigned.unknown	F-box/RNI-like superfamily protein
	solyc-miR606	Solyc02g077450	not assigned.unknown	vascular plant one zinc finger transcription factor protein
	solyc-miR663	Solyc01g079740	not assigned.unknown	LOW QUALITY:Fantastic four-like protein
	solyc-miR663	Solyc01g087190	not assigned.unknown	chaperonin-like RbcX protein
	solyc-miR663	Solyc03g120360	not assigned.unknown	LOW QUALITY:phospholipase-like protein (PEARLI 4) family protein
	solyc-miR663	Solyc04g051310	not assigned.unknown	transmembrane protein
	solyc-miR663	Solyc07g008670	not assigned.unknown	protein LONGIFOLIA 1
	solyc-miR663	Solyc12g038520	not assigned.unknown	squamosa promoter binding protein 6c
	solyc-miR690	Solyc02g069950	not assigned.unknown	KRR1 small subunit processome component
	solyc-miR734	Solyc05g032750	not assigned.unknown	EEIG1/EHBP1 protein amino-terminal domain protein
	solyc-miR734	Solyc09g064920	not assigned.unknown	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
	solyc-miR793	Solyc06g048610	protein.aa activation	nucleic acid-binding, OB-fold-like protein
	solyc-miR464	Solyc02g079400	protein.degradation	nitric oxide synthase-interacting protein homolog
	solyc-miR558	Solyc05g056140	protein.degradation	zinc finger (C3HC4-type RING finger) family protein
	solyc-miR608	Solyc01g103970	protein.degradation	RING finger protein
	solyc-miR663	Solyc07g005640	protein.degradation	26S proteasome non-ATPase regulatory subunit-like protein

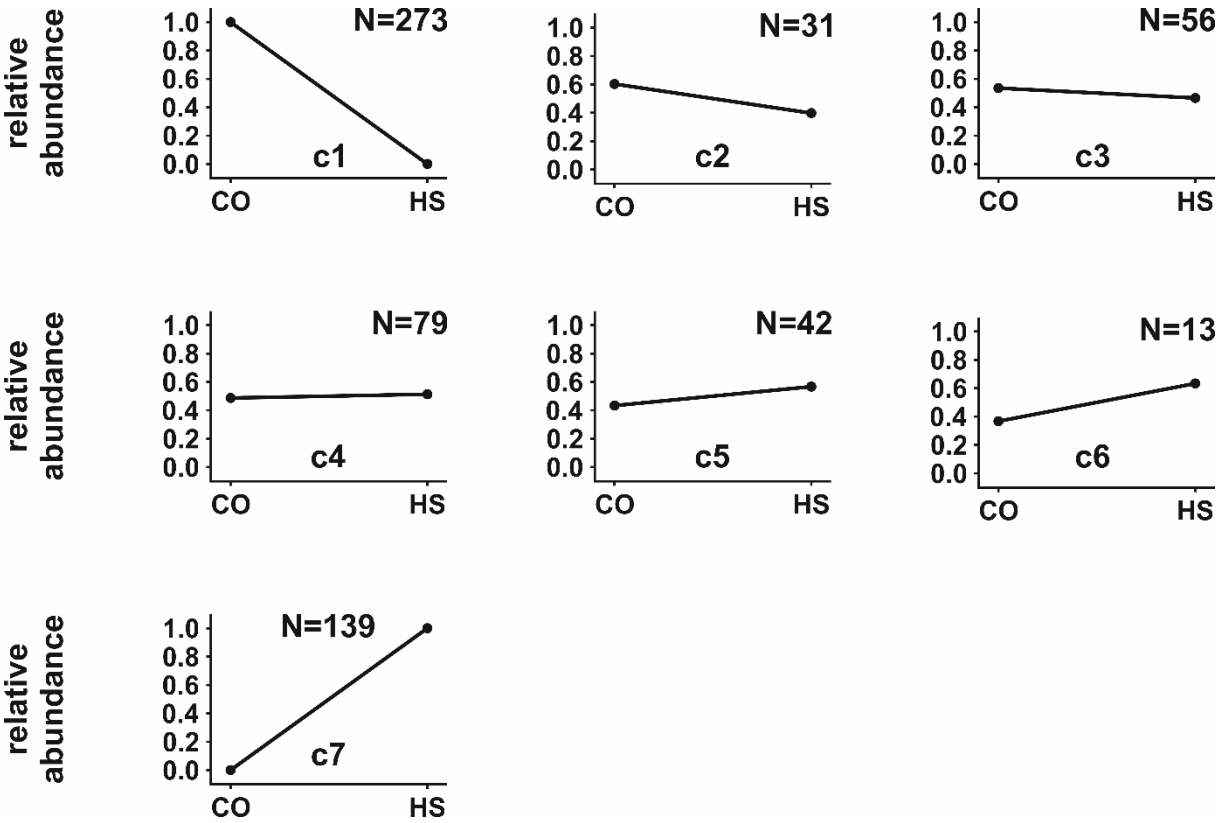
	solyc-miR663	Solyc07g049560	protein.postranslational modification	tyrosine specific protein phosphatase family protein
	solyc-miR663	Solyc09g011750	protein.postranslational modification	kinase family protein
	solyc-miR608	Solyc09g011330	protein.postranslational modification; signalling.receptor kinases; misc.myrosinases-lectin-jacalin	serine/threonine-protein kinase
	solyc-miR528	Solyc06g054400	protein.synthesis	translation initiation factor
	solyc-miR663	Solyc06g073090	protein.synthesis	chloroplast-specific ribosomal protein
	solyc-miR663	Solyc02g070210	protein.targeting	Sec14p-like phosphatidylinositol transfer family protein
	solyc-miR756	Solyc01g067740	redox.dismutases and catalases	superoxide dismutase [Cu-Zn] 1
	solyc-miR593	Solyc11g006490	RNA.regulation of transcription	basic-leucine zipper (bZIP) transcription factor family protein
	solyc-miR624	Solyc08g062210	RNA.regulation of transcription	nuclear transcription factor Y subunit
	solyc-miR662	Solyc01g094800	RNA.regulation of transcription	ATP-dependent helicase BRM
	solyc-miR709	Solyc02g062690	RNA.regulation of transcription	bHLH transcription factor 012
	solyc-miR756	Solyc06g071310	RNA.regulation of transcription	pollen-specific protein SF3, putative
	solyc-miR764	Solyc11g007580	RNA.regulation of transcription	DNA demethylase 3
	solyc-miR663	Solyc08g014440	RNA.regulation of transcription;protein.degradation	BTB/POZ and TAZ domain protein
	solyc-miR663	Solyc02g064540	RNA.RNA binding	RNA-binding (RRM/RBD/RNP motifs) family protein
	solyc-miR663	Solyc11g010470	signalling.14-3-3 proteins	14-3-3 protein 1
	solyc-miR663	Solyc04g079710	signalling.receptor kinases	kinase, putative
	solyc-miR663	Solyc12g010740	signalling.receptor kinases	leucine-rich repeat receptor-like protein kinase
	solyc-miR704	Solyc01g010030	signalling.receptor kinases	kinase family protein
	solyc-miR528	Solyc06g035690	stress.abiotic	post-GPI attachment-like factor-protein
	solyc-miR663	Solyc09g098130	stress.biotic	CC-NBS-LRR gene
	solyc-miR663	Solyc03g117547	transport.ABC transporters and multidrug resistance systems	ABC transporter family protein
	solyc-miR064	Solyc06g068560	transport.misc	major facilitator superfamily protein
	solyc-miR663	Solyc10g009180	transport.misc	secretory carrier membrane protein
	solyc-miR610	Solyc04g005780	Biodegradation of Xenobiotics	LOW QUALITY:glyoxal oxidase-related protein
	solyc-miR664	Solyc10g086650	Biodegradation of Xenobiotics	LOW QUALITY:glyoxal oxidase-related protein
	solyc-miR664	Solyc05g015510	development.squamosa promoter binding like (SPL)	squamosa promoter-binding protein 10
	solyc-miR664	Solyc05g015840	development.squamosa promoter binding like (SPL)	squamosa promoter binding protein 13
	solyc-miR664	Solyc07g062980	development.squamosa promoter binding like (SPL)	squamosa promoter-binding protein, putative
	solyc-miR664	Solyc10g009080	development.squamosa promoter binding like (SPL)	squamosa promoter binding protein 3
B	solyc-miR660	Solyc02g064960	development.unspecified	APETALA2b
	solyc-miR660	Solyc02g093150	development.unspecified	APETALA2c
	solyc-miR759	Solyc10g047670	hormone metabolism.abscisic acid	HVA22-like protein
	solyc-miR664	Solyc12g098640	hormone metabolism.brassinosteroid; cell.vesicle transport	cycloeucalenol cycloisomerase
	solyc-miR476	Solyc01g010810	not assigned.no ontology	Ran guanine nucleotide release factor-like protein
	solyc-miR610	Solyc07g054480	not assigned.no ontology	pentatricopeptide repeat-containing protein
	solyc-miR531	Solyc01g097580	not assigned.unknown	LOW QUALITY:Transcription factor GTE12

	solyc-miR610	Solyc06g073110	not assigned.unknown	DUF936 family protein
	solyc-miR660	Solyc09g059570	not assigned.unknown	B-cell receptor-associated protein 31-like
	solyc-miR660	Solyc09g075970	not assigned.unknown	Sec14p-like phosphatidylinositol transfer family protein
	solyc-miR664	Solyc12g038520	not assigned.unknown	squamosa promoter binding protein 6c
	solyc-miR737	Solyc02g083180	not assigned.unknown	UPF0548 protein
	solyc-miR664	Solyc07g005640	protein.degradation	26S proteasome non-ATPase regulatory subunit-like protein
	solyc-miR475	Solyc03g119290	protein.synthesis	elongation factor 1-alpha, putative
	solyc-miR609	Solyc12g035130	RNA.processing	RNA helicase DEAD36
	solyc-miR610	Solyc04g080130	RNA.regulation of transcription	zinc finger protein
	solyc-miR660	Solyc05g051060	RNA.regulation of transcription	homeodomain-like superfamily protein
	solyc-miR660	Solyc11g072600	RNA.regulation of transcription	APETALA2d
	solyc-miR759	Solyc08g005430	RNA.regulation of transcription	growth-regulating factor
	solyc-miR660	Solyc01g090460	RNA.regulation of transcription; development.unspecified	HD-ZIP
	solyc-miR660	Solyc09g007260	RNA.regulation of transcription; development.unspecified	AP2-like ethylene-responsive transcription factor
	solyc-miR660	Solyc10g084340	RNA.regulation of transcription; development.unspecified	AP2-like ethylene-responsive transcription factor
	solyc-miR438	Solyc10g078240	secondary metabolism.phenylpropanoids	cytochrome P450
	solyc-miR610	Solyc01g079340	signalling.receptor kinases	protein kinase superfamily protein
	solyc-miR609	Solyc11g006650	stress.abiotic	double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases superfamily protein, putative
	solyc-miR647	Solyc02g085930	stress.abiotic; misc.peroxidases	peroxidase
	solyc-miR664	Solyc09g098130	stress.biotic	CC-NBS-LRR gene
	solyc-miR664	Solyc03g117547	transport.ABC transporters and multidrug resistance systems	ABC transporter family protein
	solyc-miR660	Solyc06g075510	transport.peptides and oligopeptides	APETALA2e
C	solyc-miR233	Solyc04g081200	cell.division	regulator of chromosome condensation (RCC1) family protein
	solyc-miR490	Solyc07g065050	cell.division	cell division protein FtsZ-1-like protein
	solyc-miR263	Solyc02g069570	development.unspecified	pentatricopeptide repeat-containing protein, putative
	solyc-miR489	Solyc03g117370	development.unspecified	TOPLESS-related 1
	solyc-miR102	Solyc09g064800	major CHO metabolism.synthesis	isoamylase isoform 2
	solyc-miR225	Solyc07g053070	misc.acid and other phosphatases	purple acid phosphatase
	solyc-miR651	Solyc07g042220	misc.gluco-, galacto- and mannosidases	beta-galactosidase
	solyc-miR490	Solyc10g055700	not assigned.no ontology	pentatricopeptide repeat-containing protein
	solyc-miR651	Solyc04g049950	not assigned.no ontology	serine/threonine-protein kinase ATM
	solyc-miR233	Solyc04g081850	not assigned.unknown	single hybrid motif superfamily protein
	solyc-miR321; solyc-miR632	Solyc09g065170	not assigned.unknown	30S ribosomal protein S7, chloroplastic
	solyc-miR490	Solyc04g078830	not assigned.unknown	THO complex, subunit 5
	solyc-miR490	Solyc01g006800	protein.degradation	methionine aminopeptidase 2
	solyc-miR605	Solyc10g081240	protein.folding	GrpE protein homolog
	solyc-miR483	Solyc04g008790	protein.synthesis	ribosomal protein L1
	solyc-miR490	Solyc03g058350	protein.synthesis	translation initiation factor

	solyc-miR613	Solyc10g009070	protein.synthesis	RNA helicase DEAD32
	solyc-miR761	Solyc12g096150	protein.synthesis	60S ribosomal protein L13
	solyc-miR233	Solyc09g005340	RNA.processing	symplekin, putative
	solyc-miR605	Solyc10g050540	RNA.processing	cleavage stimulation factor subunit 3
	solyc-miR490	Solyc08g078300	RNA.regulation of transcription	homeobox-leucine zipper family protein
	solyc-miR490	Solyc12g010680	RNA.regulation of transcription	acyl-CoA N-acyltransferase
	solyc-miR683; solyc-miR684	Solyc06g075150	RNA.regulation of transcription	auxin Response Factor 10B
	solyc-miR683; solyc-miR684	Solyc11g013480	RNA.regulation of transcription	auxin Response Factor 17
	solyc-miR233	Solyc05g053090	RNA.regulation of transcription; development.unspecified	transcription factor GRAS
	solyc-miR233	Solyc09g083080	RNA.regulation of transcription; protein.synthesis	MAR-binding protein
	solyc-miR651	Solyc01g028860	signalling.calcium	evolutionarily conserved C-terminal region 2
	solyc-miR490	Solyc04g015970	transport.ABC transporters and multidrug resistance systems	ABC transporter family protein
	solyc-miR613	Solyc11g069260	transport.ABC transporters and multidrug resistance systems	ABC transporter family protein
D	solyc-miR656	Solyc04g005230	Biodegradation of Xenobiotics	alpha/beta-Hydrolases superfamily protein
	solyc-miR544	Solyc06g062430	minor CHO metabolism.myo-inositol	myo-inositol oxygenase
	solyc-miR354	Solyc12g056200	not assigned.no ontology	transferring glycosyl group transferase (DUF604)
	solyc-miR697	Solyc03g006560	not assigned.no ontology	alpha/beta-Hydrolases superfamily protein
	solyc-miR019	Solyc12g100120	not assigned.unknown	DUF1680 domain protein
	solyc-miR544	Solyc03g005420	not assigned.unknown	SPFH/Band 7/PHB domain-containing membrane-associated protein family
	solyc-miR659	Solyc01g098813	not assigned.unknown	K(+) efflux antiporter
	solyc-miR544	Solyc01g111880	protein.postranslational modification	MAP kinase kinase kinase 11
	solyc-miR019	Solyc07g053450	RNA.regulation of transcription	BZIP family transcription factor
	solyc-miR544	Solyc09g091260	signalling.receptor kinases	non-specific serine/threonine protein kinase
E	solyc-miR490	Solyc12g056160	transport.metal	cation/H(+) antiporter
F	solyc-miR544	Solyc02g080810	amino acid metabolism.degradation; PS.photorespiration	aminomethyltransferase
	solyc-miR544	Solyc09g005470	cell wall.precursor synthesis	UDP-glucuronic acid decarboxylase 1
	solyc-miR544	Solyc06g051310	cell.vesicle transport	Clathrin heavy chain
	solyc-miR544	Solyc01g097960	development.unspecified	late embryogenesis abundant D-like protein
	solyc-miR443	Solyc06g007140	lipid metabolism.FA desaturation	omega-3 fatty acid desaturase
	solyc-miR549	Solyc07g064960	lipid metabolism.Phospholipid synthesis	phosphatidate cytidylyltransferase
	solyc-miR658	Solyc10g081660	major CHO metabolism.synthesis	sucrose phosphatase AF493563
	solyc-miR096	Solyc03g005830	misc.acyl transferases; misc.O-methyl transferases	O-acetyltransferase family protein
	solyc-miR019	Solyc11g010550	not assigned.no ontology	GPI ethanolamine phosphate transferase
	solyc-miR334	Solyc04g015150	not assigned.no ontology	Phosphoglycolate phosphatase
	solyc-miR544	Solyc07g006040	not assigned.no ontology	DNA-directed RNA polymerase subunit beta
	solyc-miR627	Solyc09g008190	not assigned.no ontology	Acylphosphatase
	solyc-miR652	Solyc05g010150	not assigned.no ontology	HD domain-containing protein 2
	solyc-	Solyc05g008910	not assigned.unknown	class I glutamine amidotransferase-like

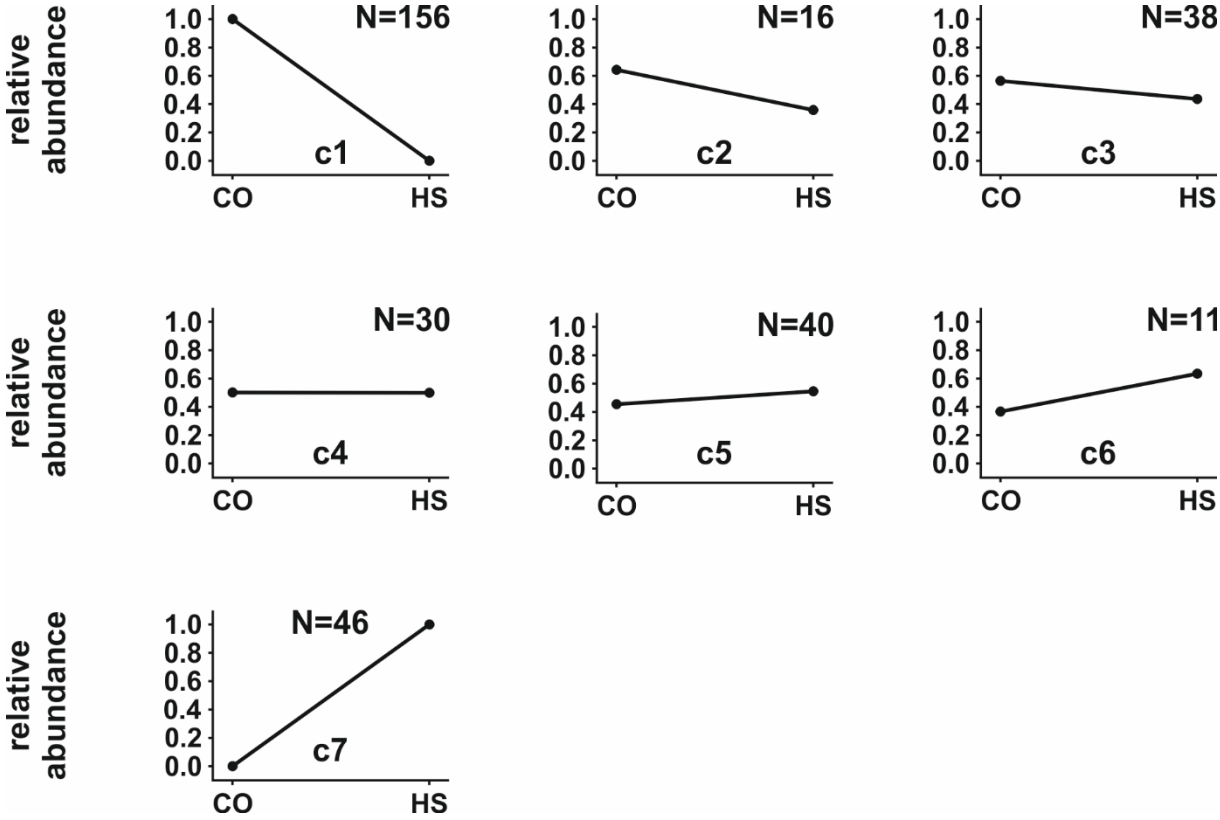
	miR172; solyc-miR457 (2x); solyc-miR656			superfamily protein
	solyc-miR354	Solyc10g048055	not assigned.unknown	major latex-like protein
	solyc-miR544	Solyc02g083440	not assigned.unknown	F-box
	solyc-miR544	Solyc06g067980	not assigned.unknown	late embryogenesis abundant protein
	solyc-miR544	Solyc11g005840	not assigned.unknown	cysteine desulfurase
	solyc-miR596	Solyc10g078420	not assigned.unknown	transmembrane protein
	solyc-miR708	Solyc12g042800	not assigned.unknown	dsRNA-binding domain-like superfamily protein
	solyc-miR758	Solyc11g072400	not assigned.unknown	defects in morphology protein 1
	solyc-miR549	Solyc03g111710	protein.degradation	BTB/POZ and TAZ domain protein
	solyc-miR019; solyc-miR337	Solyc03g006950	protein.postranslational modification	protein phosphatase 2C family protein
	solyc-miR055	Solyc08g082980	protein.postranslational modification; signalling.MAP kinases	MAP kinase kinase kinase 67
	solyc-miR652	Solyc03g117320	protein.synthesis	ribosomal protein L3
	solyc-miR337	Solyc05g015420	RNA.regulation of transcription	HCP family protein with MYND-type zinc finger protein
	solyc-miR658	Solyc03g005260	S-assimilation.ATPS	ATP-sulfurylase
	solyc-miR637	Solyc05g008000	signalling.phosphinositides	nudix hydrolase
	solyc-miR544	Solyc01g073740	TCA / org transformation.TCA	citrate synthase
	solyc-miR755	Solyc03g113060	transport.ABC transporters and multidrug resistance systems	ABC transporter A family protein
	solyc-miR019	Solyc07g007600	transport.H+ transporting pyrophosphatase	vacuolar-type H+-pyrophosphatase
	solyc-miR237	Solyc12g006950	transport.sulphate	sulfate transporter, putative
G	solyc-miR525	Solyc11g006220	cell.organisation	myosin XI, putative
	solyc-miR772	Solyc10g006820	cell.organisation	myosin VIII-2
	solyc-miR792	Solyc06g072320	lipid metabolism.exotics(steroids, squalene etc)	serine palmitoyltransferase
	solyc-miR364	Solyc12g032955	not assigned.no ontology	CBS domain-containing protein
	solyc-miR792	Solyc08g081360	not assigned.no ontology	PLAC8 family protein
	solyc-miR364	Solyc06g064660	protein.postranslational modification	protein kinase family protein
	solyc-miR723	Solyc06g071410	protein.postranslational modification	MAP kinase kinase kinase 40
	solyc-miR434	Solyc05g051830	RNA.regulation of transcription	MADS-box transcription factor family protein
	solyc-miR711	Solyc05g047640	RNA.regulation of transcription	S-acyltransferase
	solyc-miR731	Solyc12g005890	RNA.regulation of transcription	MYB-related transcription factor
	solyc-miR557	Solyc02g082540	signalling.calcium	calcium-binding EF-hand
	solyc-miR037	Solyc10g079030	signalling.G-proteins	Ras-related protein RABA4d
H	solyc-miR663	Solyc01g105100	not assigned.unknown	endoplasmic reticulum vesicle transporter protein
	solyc-miR663	Solyc03g093335	not assigned.unknown	tagatose-6-phosphate ketose/aldose isomerase, putative (Mog1/PsbP/DUF1795-like photosystem II reaction center PsbP family protein)
	solyc-miR663	Solyc08g067270	not assigned.unknown	hydroxyproline-rich glycoprotein family protein
	solyc-miR663	Solyc02g014350	protein.degradation	P-loop containing nucleoside triphosphate hydrolases superfamily protein

	solyc-miR192	Solyc05g055240	RNA.regulation of transcription	I-box binding factor
	solyc-miR608	Solyc10g077060	RNA.regulation of transcription	S-acyltransferase
	solyc-miR663	Solyc06g063150	signalling.receptor kinases	kinase
	solyc-miR704	Solyc04g006930	signalling.receptor kinases	kinase family protein



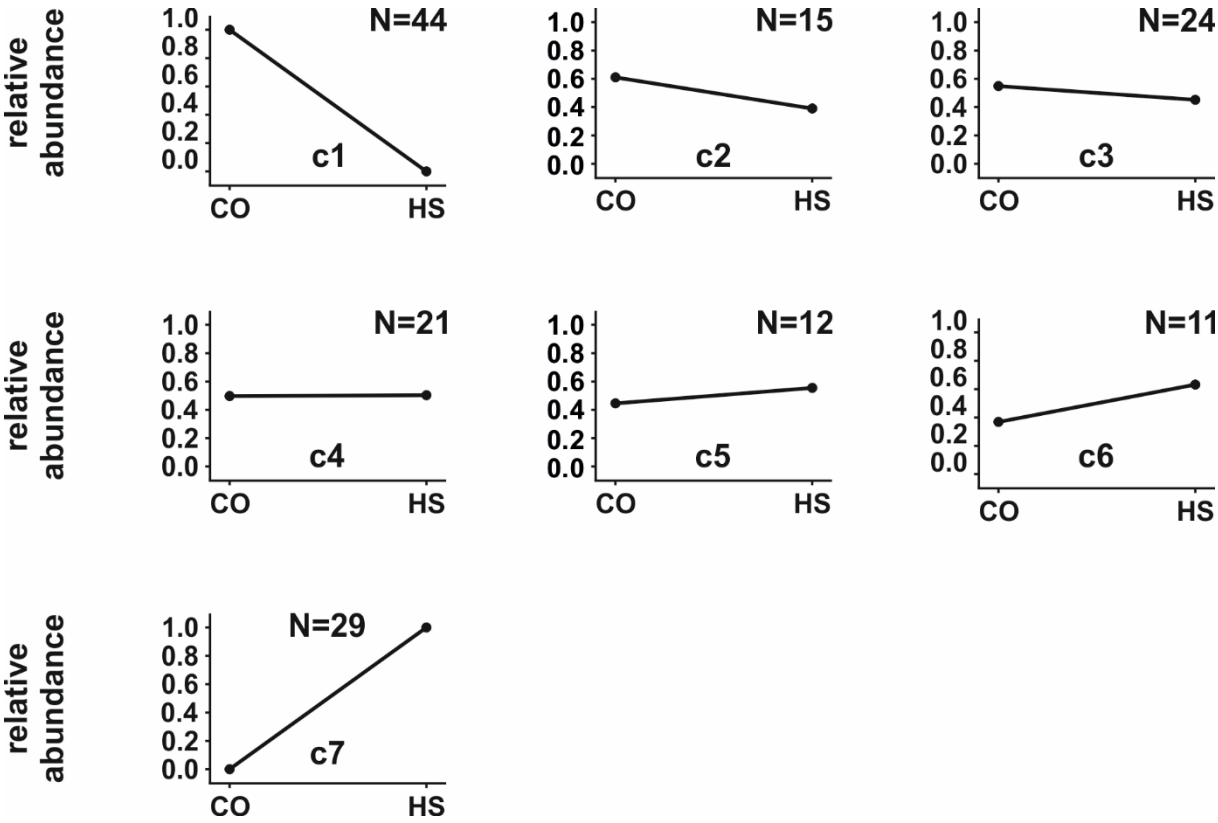
Supplemental Figure 4: miRNA clustering by abundance changes in tetrads after HS

The abundance profiles of miRNAs detected in non- (CO) and/or heat-stressed (HS) tetrads were k-means clustered. Profiles represent the relative abundance of replicate averaged TPM values across CO and HS. Circles represent cluster. In addition, the cluster number and number of miRNAs assigned to the clusters are given.

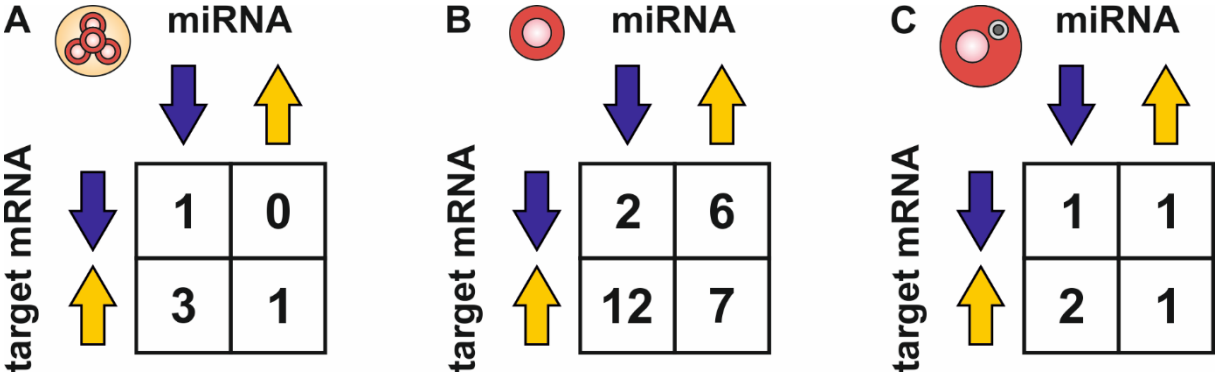


Supplemental Figure 5: miRNA clustering by abundance changes in post-meiotic pollen after HS

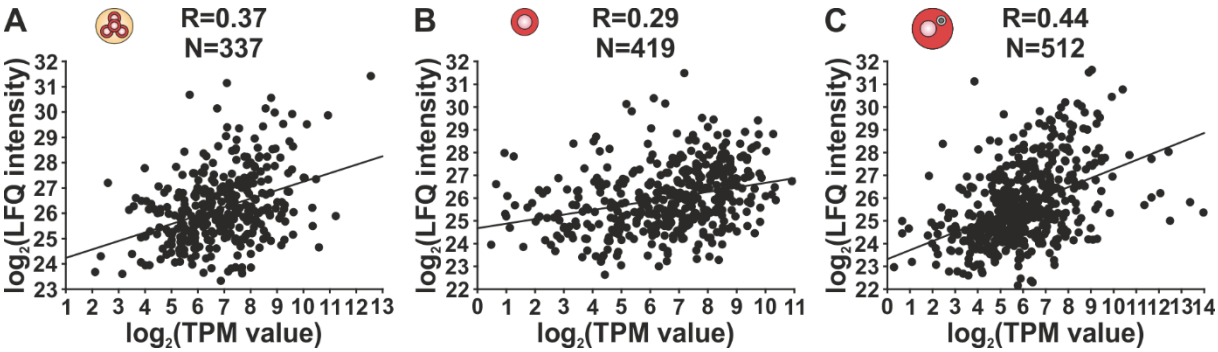
The abundance profiles of miRNAs detected in non- (CO) and/or heat-stressed (HS) post-meiotic pollen were k-means clustered. Profiles represent the relative abundance of replicate averaged TPM values across CO and HS. Circles represent cluster. In addition, the cluster number and number of miRNAs assigned to the clusters are given.



Supplemental Figure 6: miRNA clustering by abundance changes in mature pollen after HS
The abundance profiles of miRNAs detected in non- (CO) and/or heat-stressed (HS) mature pollen were k-means clustered. Profiles represent the relative abundance of replicate averaged TPM values across CO and HS. Circles represent cluster. In addition, the cluster number and number of miRNAs assigned to the clusters are given.



Supplemental Figure 7: HS-relevant MTIs in pollen developmental stages
Shown are MTIs detected in tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**, whose miRNAs and mRNAs are regulated in response to HS. miRNAs and target mRNAs had to be either down- (blue) or upregulated (yellow). Numerals indicate the number of MTIs showing a pattern of down- and upregulated miRNAs and target mRNAs.



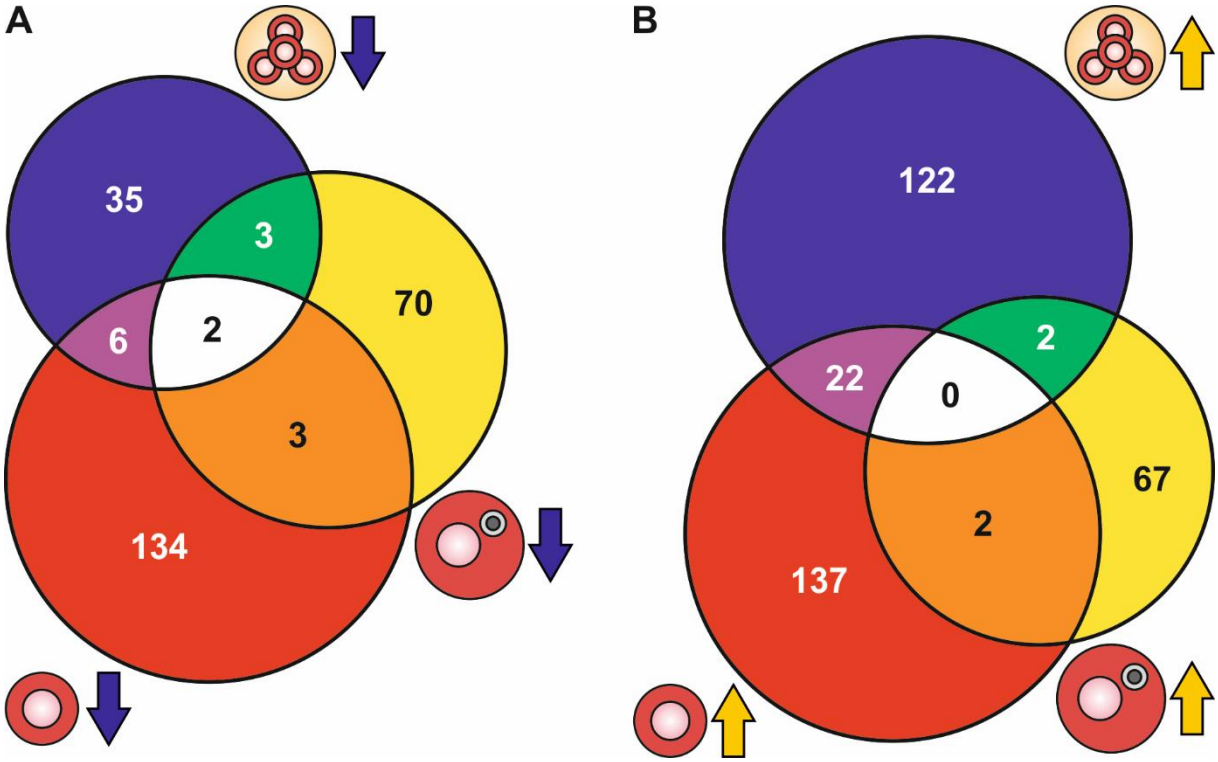
Supplemental Figure 8: Correlation of heat-stressed pollen transcriptomes and proteomes
(A, B and C) Shown are scatterplots indicating the correlation between transcript (TPM value) and protein levels (LFQ intensity) in tetrads **(A)**, post-meiotic **(B)** and mature pollen **(C)**. Only genes with measured transcript and measured protein levels were considered, whereby the proteins had to derive from a protein group containing only this protein. In addition, the pearson correlation coefficient (R) and the number of genes (N) used for the calculation are shown.

Supplemental Table 14: MapMan terms of mRNAs with direct and delayed translation
Listed are MapMan terms of the second hierarchy levels (column 1) and the number of mRNAs with direct and delayed translation assigned to them. The mRNAs include those with direct or delayed translation in tetrads (column 2 and 3), direct or delayed translation in post-meiotic pollen (column 4 and 5) and direct translation in mature pollen (column 6).

MapMan term	tetrads direct	tetrads delay	post-meiotic direct	post-meiotic delay	mature direct
amino acid metabolism.degradation	---	---	---	4	---
amino acid metabolism.synthesis	---	---	2	7	---
C1-metabolism.formate-tetrahydrofolate ligase	---	---	---	---	1
C1-metabolism.glycine hydroxymethyltransferase	---	---	---	1	---
C1-metabolism.methylenetetrahydrofolate reductase	---	---	---	1	1
cell wall.cell wall proteins	1	---	---	1	6
cell wall.degradation	2	---	---	---	4
cell wall.pectin*esterases	---	2	---	---	3
cell wall.precursor synthesis	---	---	---	3	7
cell.cycle	1	---	---	2	---
cell.division	---	---	---	1	4
cell.organisation	1	2	---	5	12
cell.unspecified	1	---	---	---	---
cell.vesicle transport	---	---	---	6	1
Co-factor and vitamine metabolism.pantothenate	---	---	---	1	---
development.late embryogenesis abundant	1	---	---	---	1
development.storage proteins	---	1	---	4	---
development.unspecified	---	2	---	4	7
DNA.repair	---	---	---	1	---
DNA.synthesis/chromatin structure	11	6	1	---	---
DNA.unspecified	---	---	1	1	---
fermentation.aldehyde dehydrogenase	1	---	1	---	---
fermentation.PDC	---	---	---	1	---
gluconeogenesis / glyoxylate cycle.citrate synthase	---	---	---	1	---
gluconeogenesis / glyoxylate cycle.Malate DH	---	---	---	1	---

gluconeogenesis / glyoxylate cycle.PEPCK	---	---	1	---	---
glycolysis.cytosolic branch	2	---	2	5	3
glycolysis.plastid branch	---	1	---	3	1
glycolysis.unclear/dually targeted	---	1	1	1	1
hormone metabolism.auxin	---	---	---	---	1
hormone metabolism.brassinosteroid	---	2	---	---	---
hormone metabolism.ethylene	---	1	---	---	---
lipid metabolism.exotics(steroids, squalene etc)	---	---	---	1	1
lipid metabolism.FA synthesis and FA elongation	---	---	1	7	2
lipid metabolism.glyceral metabolism	---	---	---	1	1
lipid metabolism.lipid degradation	---	---	---	6	---
lipid metabolism.Phospholipid synthesis	1	---	---	2	2
major CHO metabolism.degradation	---	1	2	2	1
major CHO metabolism.synthesis	---	---	1	3	---
metal handling.binding, chelation and storage	---	---	---	1	---
minor CHO metabolism.myo-inositol	---	---	---	---	2
minor CHO metabolism.others	---	1	1	1	---
minor CHO metabolism.sugar kinases	---	---	1	---	---
misc.acid and other phosphatases	1	1	---	2	---
misc.alcohol dehydrogenases	---	---	---	1	---
misc.beta 1,3 glucan hydrolases	1	2	---	---	2
misc.cytochrome P450	---	1	---	---	---
misc.dynamin	---	---	---	1	---
misc.GDSL-motif lipase	---	1	---	---	---
misc.gluco-, galacto- and mannosidases	1	4	---	---	1
misc.glutathione S transferases	---	---	---	---	1
misc.nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases	1	---	---	---	---
misc.O-methyl transferases	---	1	---	---	---
misc.oxidases - copper, flavone etc	1	3	---	1	2
misc.peroxidases	1	1	---	---	---
misc.short chain dehydrogenase/reductase (SDR)	---	1	---	2	1
misc.UDP glucosyl and glucoronyl transferases	---	---	---	---	1
mitochondrial electron transport / ATP synthesis.cytochrome c oxidase	---	---	---	1	---
mitochondrial electron transport / ATP synthesis.cytochrome c reductase	---	---	---	1	---
mitochondrial electron transport / ATP synthesis.F1-ATPase	---	---	1	3	---
mitochondrial electron transport / ATP synthesis.NADH-DH	1	---	1	5	---
N-metabolism.ammonia metabolism	---	---	---	2	---
N-metabolism.N-degradation	---	1	1	---	---
nucleotide metabolism.degradation	1	---	---	2	---
nucleotide metabolism.deoxynucleotide metabolism	1	---	---	---	---
nucleotide metabolism.phosphotransfer and pyrophosphatases	---	---	---	2	2
nucleotide metabolism.salvage	---	---	---	1	---
OPP.non-reductive PP	1	---	---	---	---
polyamine metabolism.synthesis	---	2	---	---	---

protein.aa activation	1	2	---	---	---
protein.degradation	9	6	1	3	2
protein.folding	4	3	---	---	---
protein.glycosylation	1	1	1	---	---
protein.postranslational modification	3	---	---	1	4
protein.synthesis	12	13	2	1	1
protein.targeting	2	1	1	2	2
PS.calvin cycle	2	1	1	---	---
PS.lightreaction	15	3	---	---	---
PS.photorespiration	---	---	---	2	1
redox.ascorbate and glutathione	1	1	1	1	---
redox.dismutases and catalases	---	---	---	1	---
redox.misc	---	1	---	---	---
redox.peroxiredoxin	1	1	---	1	---
redox.thioredoxin	4	---	1	2	---
RNA.processing	2	3	1	---	1
RNA.regulation of transcription	2	2	2	3	1
RNA.RNA binding	1	4	1	---	1
S-assimilation.sulfite redox	---	---	1	---	---
secondary metabolism.flavonoids	---	1	---	---	1
secondary metabolism.isoprenoids	---	---	---	1	1
secondary metabolism.N misc	2	---	---	1	---
secondary metabolism.phenylpropanoids	2	---	---	1	---
secondary metabolism.sulfur-containing	---	---	---	1	---
secondary metabolism.wax	1	---	---	---	---
signalling.14-3-3 proteins	---	---	---	3	1
signalling.calcium	1	1	---	1	3
signalling.G-proteins	---	2	---	3	8
signalling.in sugar and nutrient physiology	---	---	---	---	1
signalling.phosphinositides	---	---	---	2	1
signalling.receptor kinases	1	---	---	---	---
stress.abiotic	4	---	---	3	5
stress.biotic	---	---	---	---	1
TCA / org transformation.carbonic anhydrases	---	---	---	1	---
TCA / org transformation.other organic acid transformations	---	---	---	5	---
TCA / org transformation.TCA	---	---	3	4	2
transport.ABC transporters and multidrug resistance systems	---	---	---	2	---
transport.calcium	---	---	---	---	1
transport.H+ transporting pyrophosphatase	1	---	---	---	---
transport.Major Intrinsic Proteins	1	1	---	---	---
transport.metabolite transporters at the envelope membrane	---	---	---	1	---
transport.metabolite transporters at the mitochondrial membrane	---	---	1	1	1
transport.misc	---	---	1	---	---
transport.p- and v-ATPases	---	1	---	4	1
transport.peptides and oligopeptides	---	---	---	1	---
transport.porins	---	---	---	1	---
transport.unspecified cations	---	---	---	1	---



Supplemental Figure 9: Overlap of down- and upregulated protein groups between pollen stages
Venn diagrams illustrate the overlap of down- **(A)** and upregulated **(B)** protein groups between tetrads (blue), post-meiotic (red) and mature pollen (yellow).

Supplemental Table 15: Differentially regulated pairs of mRNAs and proteins
Listed are pairs of mRNAs and proteins that are both differentially regulated in response to HS in pollen developmental stages (column 1). For each pair the gene identifier (column 2), information about the mRNA and protein regulation (column 3 and 4; ↓ downregulation; ↑ upregulation) as well as the MapMan term(s) (column 5) and the functional description (ITAG3.2; column 6) are provided.

	SolycID	mRNA	protein	MapMan term(s)	description
tetrads	Solyc01g102960	↑	↑	stress.abiotic.heat	heat-shock protein, putative
	Solyc03g082420	↑	↑	stress.abiotic.heat	sHsp
	Solyc03g113930	↑	↑	stress.abiotic.heat	heat-shock protein, putative
	Solyc03g115230	↑	↑	stress.abiotic.heat	Hsp100
	Solyc03g123540	↑	↑	stress.abiotic.heat	sHsp
	Solyc06g076520	↑	↑	stress.abiotic.heat	sHsp
	Solyc08g062450	↑	↑	stress.abiotic.heat	sHsp
	Solyc11g020330	↑	↑	stress.abiotic.heat	sHsp
post-meiotic	Solyc12g098500	↓	↓	amino acid metabolism.degradation	adenosylhomocysteinase
	Solyc06g036290	↑	↓	stress.abiotic.heat	Hsp90
	Solyc08g079260	↑	↓	redox.thioredoxin	tetratricopeptide repeat-containing family protein
	Solyc01g105810	↓	↑	signalling.G-proteins	GDP dissociation inhibitor
	Solyc05g055760	↓	↑	secondary metabolism.isoprenoids	isopentenyl diphosphate isomerase
	Solyc01g102960	↑	↑	stress.abiotic.heat	heat-shock protein, putative
	Solyc02g077670	↑	↑	stress.abiotic.heat	Hsp40
	Solyc02g086830	↑	↑	protein.degradation	protease Do-like 1
	Solyc03g082420	↑	↑	stress.abiotic.heat	sHsp

	Solyc03g096850	↑	↑	protein.assembly and cofactor ligation	rubredoxin family protein
	Solyc03g113930	↑	↑	stress.abiotic.heat	heat-shock protein, putative
	Solyc03g123540	↑	↑	stress.abiotic.heat	sHsp
	Solyc04g012180	↑	↑	signalling.G-proteins	Ras-related protein Rab11D
	Solyc04g014480	↑	↑	stress.abiotic.heat	sHsp
	Solyc06g076520	↑	↑	stress.abiotic.heat	sHsp
	Solyc08g062340	↑	↑	stress.abiotic.heat	heat-shock protein, putative
	Solyc08g062450	↑	↑	stress.abiotic.heat	sHsp
	Solyc11g020330	↑	↑	stress.abiotic.heat	sHsp
mature	Solyc06g051000	↓	↓	stress.abiotic.drought/salt	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
mature	Solyc07g042250	↑	↓	protein.folding	Hsp10
mature	Solyc06g083190	↑	↑	protein.folding; cell.cycle.peptidylprolyl isomerase	Peptidyl-prolyl cis-trans isomerase
mature	Solyc10g081240	↑	↑	protein.folding	GrpE protein homolog

Supplemental Table 16: Analyzed RPs in this study

Shown are the RPs of the small (SSU) and lare subunit (LSU) discussed in Figure 25. For each RP the gene identifier (column 3), the protein family (column 1) and the name (column 2) are given.

SSU			LSU		
family	**name**	**SolycID**	**family**	**name**	**SolycID**
RPS3	RPS3_1	Solyc08g061850	RPL4	RPL4	Solyc10g084350
	RPS3_2	Solyc08g061960	RPL5	RPL5	Solyc09g020130
RPS4	RPS4	Solyc06g007360	RPL6	RPL6	Solyc11g012110
RPS5	RPS5	Solyc10g078620	RPL7A	RPL7A_1	Solyc06g050120
RPS6	RPS6	Solyc12g096300		RPL7A_2	Solyc06g064470
RPS7	RPS7	Solyc03g119360	RPL8	RPL8_1	Solyc07g063910
RPS8	RPS8_1	Solyc06g007570		RPL8_2	Solyc10g006580
	RPS8_2	Solyc09g057650	RPL10	RPL10_1	Solyc06g007520
	RPS8_3	Solyc10g006070		RPL10_2	Solyc06g082670
	RPS8_4	Solyc07g065170		RPL10_3	Solyc06g082650
RPS12	RPS12	Solyc12g042650	RPL10A	RPL10A	Solyc11g068420
RPS13	RPS13_1	Solyc01g091220	RPL12	RPL12_1	Solyc11g065670
	RPS13_2	Solyc03g083530		RPL12_2	Solyc11g069440
	RPS13_3	Solyc05g051000	RPL13	RPL13	Solyc12g096150
	RPS13_4	Solyc10g018250	RPL14	RPL14	Solyc09g066430
	RPS13_5	Solyc11g072260	RPL17	RPL17_1	Solyc06g071870
RPS15	RPS15	Solyc09g083370		RPL17_2	Solyc06g071880
RPS17	RPS17	Solyc10g083165	RPL18	RPL18	Solyc01g099900
RPS18	RPS18	Solyc03g078290	RPL19	RPL19_1	Solyc06g009210
RPS25	RPS25_1	Solyc02g069850		RPL19_2	Solyc09g075430
	RPS25_2	Solyc10g052660	RPL21	RPL21	Solyc10g078960
			RPL22	RPL22	Solyc01g099830
			RPL24	RPL24_1	Solyc09g008800
				RPL24_2	Solyc10g083740
				RPL24_3	Solyc10g085480
			RPL27	RPL27	Solyc06g073300
			RPL28	RPL28	Solyc12g044720

	RPL34	RPL34_1	Solyc02g087930
		RPL34_2	Solyc05g007560
	RPL35	RPL35_1	Solyc04g010240
		RPL35_2	Solyc05g054810
		RPL35_3	Solyc06g075620
	RPL35A	RPL35A_1	Solyc03g096360
		RPL35A_2	Solyc05g052800
		RPL35A_3	Solyc05g052810
	RPLP1	RPLP1	Solyc01g016470

Danksagung

Als Erstes möchte ich meinem Doktorvater Prof. Dr. Enrico Schleiff danken, dass er mir seit mittlerweile 6 Jahren die Möglichkeit gibt, an den verschieden Projekten seines Arbeitskreises mitzuwirken. Ich möchte ihm weiterhin dafür danken, dass er mich immer unterstützt hat und in vertrackten Situationen stets die zündenden Ideen hatte um Projekte voranzutreiben. Was mich jedoch immer am meisten an Enrico fasziniert hat, ist sein unermüdlicher Enthusiasmus und seine ansteckende Begeisterung für die Wissenschaft. Darüber hinaus möchte ich mich bei Jun. Prof. Dr. Michaela Müller-McNicoll dafür bedanken, dass sie sich dazu bereit erklärt hat, als Zweitgutachterin meiner Doktorarbeit zu fungieren.

Ein ganz besonderer Dank gilt meinem Betreuer Dr. Stefan Simm, der mich seit meiner Bachelorarbeit bei allen wichtigen Stationen meines wissenschaftlichen Werdegangs begleitet hat. Ich danke ihm für seine ständige Unterstützung, seinen Einfallsreichtum, die hitzigen Diskussionen und auch den Freiraum den er mir für die Realisierung meiner Projekte gegeben hat. Doch am meisten bin ich dafür dankbar, dass sich auch abseits der Wissenschaft eine Freundschaft entwickelt hat und hoffe, dass diese auch bestehen bleibt wenn sich unsere wissenschaftlichen Wege nach mehr als 6 Jahren trennen.

Ein weiterer Dank gilt Dr. Klaus-Dieter Scharf und Dr. Sotirios Fragkostefanakis, die mir bei Fragen rund um das Hitzestressverhalten von Tomatenpflanzen mit Rat und Tat zur Seite standen und die es mir ermöglicht haben auch die biologische Sichtweise besser zu verstehen.

Dem gesamten AK Schleiff möchte ich für die tolle Arbeitsatmosphäre, die diversen Feiern und die lustigen Abende während der verschiedenen Sommerschulen danken. Und natürlich auch dafür, dass Frieda so herzlich aufgenommen wurde. Besonders bedanken möchte ich mich bei Jannik und Niclas, die wie ich Teil der bioinformatischen Minderheit im AK Schleiff waren und mir insbesondere das Mittagessen und die anschließenden Kaffeepausen versüßt haben.

Meinen Eltern und Großeltern möchte ich für den Rückhalt und die Unterstützung während meines Studiums und meiner Promotion danken. Mein allergrößter Dank gilt aber meiner Frau Jenny, die in sämtlichen Lebenslagen zu mir gestanden hat und auch mal die ein oder andere schlechte Laune ausgehalten hat. Ich liebe dich und Blicke voller Vorfreude in unsere gemeinsame Zukunft.

Erklärung

Ich erkläre hiermit, dass ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den …….…….. ……..……...
Mario Keller

Versicherung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation über

„Molecular reprogramming in tomato pollen during development and heat stress”

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.

Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den …….…….. ……..……...
Mario Keller

Lebenslauf

Persönliche Daten

Name	Mario Keller
Geburtsdatum	31.05.1988
Geburtsort	Frankfurt am Main
Familienstand	verheiratet
Nationalität	deutsch

Beruflicher Werdegang

04/2015 - 08/2018	Wissenschaftlicher Mitarbeiter im Fachbereich Biowissenschaften (FB15) an der Goethe-Universität Frankfurt am Main bei Prof. Dr. Enrico Schleiff

Ausbildung

10/2018	Abgabe der Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften mit dem Titel „Molecular reprogramming in tomato pollen during development and heat stress" bei Prof. Dr. Enrico Schleiff
12/2014	Abgabe der Masterarbeit mit dem Titel „A High-throughput pipeline for variable NGS sample analysis" bei Prof. Dr. Enrico Schleiff
10/2012 – 03/2015	Studium im Fach Bioinformatik an der Goethe-Universität Frankfurt am Main mit Abschluss Master (Note 1,3)
10/2012	Abgabe der Bachelorarbeit mit dem Titel „Erstellung und Verwaltung eines cyanobakteriellen Pan-Genoms" bei Prof. Dr. Enrico Schleiff
10/2009 - 10/2012	Studium im Fach Bioinformatik an der Goethe-Universität Frankfurt am Main mit Abschluss Bachelor (Note 1,5)
09/1998 – 06/2008	Besuch der Altkönigschule in Kronberg mit Abschluss Abitur (Note 2,7)

Wissenschaftliche Konferenzen

09/2015	„German Conference on Bioinformatics" in Dortmund

Publikationen

Keller M, SPOT-ITN Consortium, Simm S (2018) The coupling of transcriptome and proteome adaptation during development and heat stress response of tomato pollen. BMC Genomics 19: 447

Lampe S, Kunze M, Scholz A, Brauß TF, Winslow S, Simm S, Keller M, Heidler J, Wittig I, Brüne B, et al (2018) Identification of the TXNIP IRES and characterization of the impact of regulatory IRES trans-acting factors. Biochim Biophys Acta - Gene Regul Mech 1861: 147–157

Keller M, Hu Y, Mesihovic A, Fragkostefanakis S, Schleiff E, Simm S (2017) Alternative splicing in tomato pollen in response to heat stress. DNA Res 24: 205–217

Simm S, Keller M, Selymesi M, Schleiff E (2015) The composition of the global and feature specific cyanobacterial core-genomes. Front Microbiol 6: 219

Simm S, Fragkostefanakis S, Paul P, Keller M, Einloft J, Scharf KD, Schleiff E (2015) Identification and expression analysis of ribosome biogenesis factor co-orthologs in solanum lycopersicum. Bioinform Biol Insights 9: 1–17