universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

The Role of smRNA Regulation during Recurrent Ecotype Formation in *Heliosperma pusillum* (Caryophyllaceae)

verfasst von / submitted by

Arne Bielke, B.Sc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree program code as it appears on the student record sheet:	UA 066 220
Studienrichtung lt. Studienblatt / degree program as it appears on the student record sheet:	Joint-Masterstudium Evolutionary Systems Biology
Betreut von / Supervisor:	Assoz. Prof. Dipl.-Ing. Dr. Ovidiu Paun

Acknowledgments

First of all, I would like to thank Ovidiu Paun. Thank you for your warm welcome in your research group and for introducing me to such a great research field. And thank you for taking so much time to understand all the confusing abbreviations. I really appreciate that. Working with you has influenced my future academic path to a great extent. I would like to thank Aglaia in particular. During my internship, you made it easy for me to dive into this exciting topic. Even after you were done with your PhD, you have always been available to answer my questions. Special thanks for taking the extra time to make your comments not sound too harsh and notes like "your description of Heliosperma sounds like a fantasy epos". Thanks to Kathie for helping me to find all the answers to my small and big questions. Discussions with you have often helped me a lot, but just as often they really confused me (in English or German). I would like to thank rest of the Plant Ecological Genomics group. For the conversations during the smoking breaks, discussions about robot rights, for the souvenirs from the other side of the world or just for having a Glühwein. I really enjoyed our time together and will always have fond memories of it. I would like to thank Emily, Flo, and Stefan. I can't think about a better crew to start somewhere in a new city. Our time together really means a lot to me. Thank you, Markus, for supporting me whenever you can. You always kept my motivation high and introduced me to the deep secrets of computer science. I can't think of anyone I'd rather have by my side than you. Last but not least, I would like to thank my family and especially my parents and my brother. Without your help, I would not have been able to enjoy my studies as much as I ended up doing. I know that you know how grateful I am to have a family like you. I can always count on you.

For support during laboratory work and data acquisition we like to acknowledge Marie Huber, Daniela Paun and Carles Ortega. For help during setting up of the reciprocal transplantations we like to acknowledge Marie Huber, Daniela Paun, Carles Ortega and Pau Carnicero Campmany. For the germination of the plants, we like to acknowledge Martina Imhiavan and Daniel Schlorhaufer from the Botanical Garden of the University of Innsbruck. Computational resources were provided by the Vienna Life Science Compute Cluster (LiSC). Permission to carry out the research activities presented was granted by the Parco Naturale Dolomiti Friulane (no. 1943).

Abstract (English)

Due to the ongoing climate crisis, it is becoming increasingly important to understand how organisms adapt to different environmental challenges. It is still unclear to what extent molecular mechanisms such as small RNA (smRNA) regulation is involved in adaptation, but previous studies suggest that smRNAs can regulate phenotypic response to the environment. *Heliosperma pusillum* (Caryophyllaceae) exhibits alpine and montane ecotypes, which diverged independently at least three times. The presence (montane) or absence (alpine) of multicellular glandular trichomes represents the most striking phenotypic difference between the ecotypes. This phenotypic differentiation is stably inherited across multiple generations. Reciprocal transplantations have shown that *H. pusillum* ecotypes have a strong home-site advantage, as well as documented enhanced phenotypic plasticity in the derived montane ecotype. The present work aims to uncover patterns of smRNA regulation of gene expression likely linked to ecological and phenotypic ecotype differentiation. Analyses conducted showed that smRNA and transcriptomic data exhibit similar trends. By analyzing smRNA profiles of *H. pusillum* in a common garden, we uncovered a large proportion of differentially targeted regions (DTRs) linked to trichome development and responses to different biotic stresses. Comparing smRNA profiles of the same ecotypes during reciprocal transplantations showed a high proportion of unique DTRs, differentially-targeted genes (i.e., DTRs overlapping with genic regions; DTGs), but apparently converging gene functions. This further indicates polygenic control of the relevant traits. Overall, the results suggest a considerable link between smRNA activity and differential gene expression suggesting that smRNA regulatory divergence significantly shapes early stages of parallel evolution.

Abstract (German)

Aufgrund der derzeitigen Klimakrise wird es immer wichtiger zu verstehen, wie sich Organismen an verschiedene Umweltherausforderungen anpassen. Es ist noch unklar, inwieweit molekulare Mechanismen wie smallRNAs (smRNAs) an der Anpassung beteiligt sind. Frühere Studien deuten darauf hin, dass smRNAs die phänotypische Reaktionen auf die Umwelt regulieren können. *Heliosperma pusillum* (Caryophyllaceae) weist alpine und montane Ökotypen auf, die sich mindestens dreimal unabhängig voneinander weiterentwickelt haben. Das Vorhandensein (montan) oder Fehlen (alpin) von mehrzelligen Drüsentrichomen ist die auffälligste phänotypische Differenzierung zwischen den Ökotypen und wird über mehrere Generationen stabil vererbt. Reziproke Transplantationen haben gezeigt, dass die beiden Ökotypen einen starken Heimatstandortvorteil haben und dass der abgeleitete montane Ökotyp eine erhöhte phänotypische Plastizität aufweist. Ziel der vorliegenden Arbeit ist es, Muster der smRNA-Regulierung der Genexpression aufzudecken, die wahrscheinlich mit der ökologischen und phänotypischen Differenzierung der Ökotypen zusammenhängen. Die durchgeführten Analysen zeigen, dass smRNA- und transkriptomische Daten ähnliche Trends aufweisen. Durch Analysen von smRNA-Profilen von *H. pusillum* (aufgezogen unter identischen Bedingungen) wurden viele „differentially targeted regions“ (DTRs) aufgedeckt, die mit Trichom-Entwicklung und Reaktionen auf verschiedene biotische Stressfaktoren in Verbindung stehen. Der Vergleich von smRNA-Profilen bei wechselseitigen Transplantationen zeigt einen hohen Anteil DTRs und DTRs die sich mit genischen Regionen überschneiden aber auch konvergierende Genfunktionen. Dies unterstützt die Hypothese einer polygenen Kontrolle der relevanten Merkmale. Die Ergebnisse deuten auf einen Zusammenhang zwischen der Aktivität der smRNAs und Genexpression hin, was darauf schließen lässt, dass die regulatorische Divergenz der smRNAs eine führende Rolle während der frühen Stadien der parallelen Evolution spielt.

Content

ACKNOWLEDGMENTS **2**

ABSTRACT (ENGLISH) **3**

ABSTRACT (GERMAN) **4**

LIST OF TABLES **6**

LIST OF FIGURES **7**

ABBREVIATIONS ..... FEHLER! TEXTMARKE NICHT DEFINIERT.

INTRODUCTION **9**

2. MATERIALS AND METHODS **12**

2.1 MATERIAL AND SAMPLE PREPARATION 12

2.2 READS ALIGNMENTS AND DELIMITATION OF TARGET GENOMIC REGIONS 13

2.3 DIFFERENTIAL TARGETING ANALYSES 14

2.4 GO ENRICHMENT OF DIFFERENTIALLY TARGETED GENES 16

2.5 SHARED DTRS, DTGS, AND GO-TERMS 16

3. RESULTS **17**

3.1 PRINCIPAL COMPONENT ANALYSES 17

3.2 DIFFERENTIALLY TARGETED REGIONS 18

3.3 GENE ONTOLOGIES 22

3.4 SHARED DIFFERENTIALLY TARGETED REGIONS 24

4. DISCUSSION **29**

CONCLUSION **32**

REFERENCES **33**

APPENDIX **38**

CODE 38

SUPPLEMENTARY TABLES 41

List of Tables

Table 1 Samples used in the study, an identifier of the biological replicates and number of accessions per group and in total for samples of the common garden and the transplanted samples. * biological replicates with technical replicates. 13

Table 2 Information about the used contrast groups. The table contains used abbreviations and the respective details for samples of the common garden and the transplanted samples 15

Supp. Table 1 Information about sample grouping and details of the used abbreviation for samples of the common garden and the transplanted samples. 41

Supp. Table 2 Sample ID and the respective number of sequences before (Raw) and after trimming for all length classes together (20 - 24 nt) and separated for samples of the common garden and the transplanted samples. 41

Supp. Table 3 Total number of differential targeted annotated regions, including DTRs (differentially targeted regions) DTG (Differentially targeted genes) mRNAs, Exons, CDSs (coding regions), 3´UTRs, 5´UTRs and the total number of UTRs for the comparison groups of the Common Garden and Transplanted samples. 42

Supp. Table 4 Top five GO-terms related to leaf material for the Gene regulatory regions of the Common Garden samples. GO-terms of up-regulated DTRs are displayed on the left and GO-terms of down-regulated DTRs are displayed on the left. Abbreviations for the individual comparison groups are given in bold. Each GO-term has a p-value < 0.05. 43

Supp. Table 5 Top five GO-terms related to leaf material for the Gene regulatory regions of the Transplanted samples. GO-terms of up-regulated DTRs are displayed on the left and GO-terms of down-regulated DTRs are displayed on the left. Abbreviations for the individual comparison groups are given in bold. Each GO-term has a p-value < 0.05. 44

Supp. Table 6 Total number, number of unique (and unique in percentage) of up- or down-regulated differentially targeted regions (DTRs), differentially targeted genes (DTGs) and gene ontology (GO) terms for the comparison groups of the common garden data and the data of the transplanted samples. The mean of each group is given in bold..... 45

List of Figures

Figure 1 Alpine (a) and montane (b) ecotypes of Heliosperma pusillum. (photos modified after (a) R. Flatscher; (b) M. Sonnleitner). 9

Figure 2 Geographical map of the analyzed samples and experimental setup (Szukala et al., 2022a). The numbering of the ecotype pairs is consistent with previous work done by Bertel et al. (2018). Color coding is consistent across all three subfigures. Alpine individuals are represented by triangles and montane individuals by dots. Filled symbols indicate samples in their natural environment, unfilled symbols represent reciprocally transplanted individuals. (a) Geographical map of the analyzed population pairs. (b) Visualization of the experimental setup for the common garden data. (c) Visualization of the experimental setup for the reciprocal transplanted data. 13

Figure 3 Principal component analyses based on normalized read counts of smRNA data and RNAseq data for the common garden samples and the transplanted samples. Each dot represents one individual sample. Alpine individuals are represented by triangles and montane individuals by dots. Filled symbols indicate samples in their natural environment, unfilled symbols represent reciprocally transplanted individuals. Each color represents a different population pair. 18

Figure 4 MA-plots and the number of up- and down-regulated regions of the common garden and transplanted sample. Each MA-plot represents one comparison group, and each dot represents a targeted genomic region. Displayed is the log fold change (logFC) against the average log counts per million (logCPM). Sequences with an FDR smaller than 0.05 are displayed in red. 20

Figure 5 Bar chart in percentage of the differentially up- (left) and down- (right) expressed annotated regions for the common garden and the transplanted samples (with respect to the total number of up- or down-targeted regions). The abbreviation for each comparison group as well as the total number and the number of up- or down-targeted regions are shown in the middle. The bars refer to the number of coding sequences (CDSs), differentially targeted genes (DTGs), exons, and UTRs (untranslated regions) from up to down respectively. Total numbers are given in white. 22

Figure 6 Heatmap of significantly enriched GO-terms of the common garden and transplanted data. Color coding refers to the row-normalized number of differentially targeted genes exhibiting a specific GO-term. Total numbers and confidence intervals are also displayed within the heatmap. * indicates a p-value <0.05; ** indicates a p-value <0.01 and *** indicates a p-value <0.005.....24

Figure 7 Upset-plot of the common garden data. Displayed are the intersection sizes of the up- or down-regulated differentially targeted regions (DTR), differentially targeted genes (DTG), and GO-terms of the corresponding DTGs. The different colored bars below represent the different comparison groups. Each dot indicates respective comparison groups included in the number of intersections above. 25

Figure 8 Upset-plot of the transplanted data. Displayed are the intersection sizes of the up- or down-regulated differentially targeted regions (DTR), differentially targeted genes (DTG) and GO-terms of the corresponding DTGs. The different colored bars below representing the different comparison groups. Each dot indicates respective comparison groups included in the number of intersections above. 27

Figure 9 Percentage of up- or down-regulated unique differentially targeted regions (DTRs), differentially targeted genes (DTGs) and gene ontology (GO)-terms for the comparison groups of the common garden data and the data of the transplanted samples. Percentage of unique DTRs are given in green, percentage of unique DTGs in blue, and percentage of unique GO-terms in purple. 28

Introduction

It is well established that over evolutionary time horizons, the environment is changing constantly (Frakes, Francis & Syktus, 1992; Hubert et al., 2018). Especially in regard to the ongoing climate crisis (McCarty, 2001; Thuiller, 2007; Archer & Rahmstorf, 2010), it becomes more and more important to understand the mechanisms behind adaptation. During the last 100,000 years, the alpine biome experienced the greatest climate variations (Schär et al., 1998). After the Ice Age, most great glaciers melted and formed new and nearly unoccupied landscapes with various opportunities for species to colonize and adapt (Pielou, 2008; Prentice, Harrison & Bartlein, 2011). *Heliosperma pusillum* Waldst. & Kit. (figure 1) belongs to the family of Caryophyllaceae (Stevanović et al., 2014), which currently contains 15 different species (Frajman & Oxelman, 2007). This white flowering alpine plant results from a process of post-glacial colonization in the Central European mountain range and can be found from the Pyrenees to the Southern Carpathian and from upper Slovakia to Northern Greece (Strid & Tan, 1997; Shuka, Malo & Tan, 2011).

Heliosperma pusillum can be encounterede under several different conditions but most frequently prefers a humid environment along streams or on wet meadows above the treeline (Hegi, 1981). Interestingly, this plant adapted to two different environments, forming ecotypes. These two different ecotypes co-occur in the South-Eastern Alps (Trucchi et al., 2017). The alpine environment experiences great temperature fluctuations and is located approximately at a height between 1,700 and 2,300 m, mostly above the timberline. The ecotype which grows there adapted to moist soils and prefers direct sunlight in open areas. At altitudes between 500 to 1,300 m, the montane ecotype can be found, sheltered from direct sunlight, under overhanging rocks, or at shallow cave entrances. Due to the low water availability in this area, the

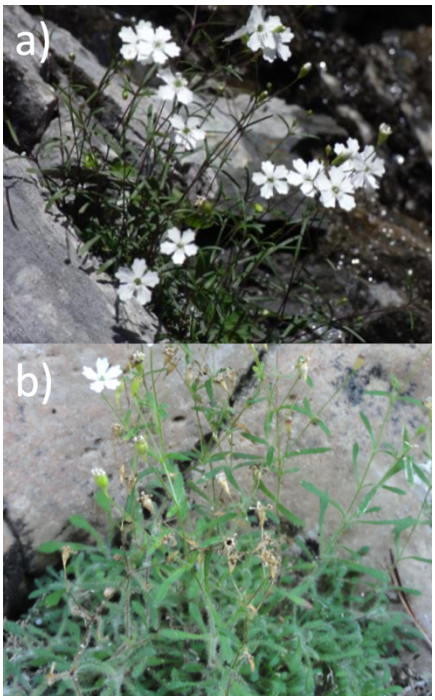

Figure 1 Alpine (a) and montane (b) ecotypes of Heliosperma pusillum. (photos modified after (a) R. Flatscher; (b) M. Sonnleitner).

montane ecotype has a higher drought tolerance (Bertel et al., 2016; Bertel et al., 2018). In comparison to the alpine area, the temperature at montane localities is described as higher on average and with a lower amplitude (Bertel et al., 2016). Reciprocal transplantation (RT) experiments done by Bertel et al. (2018) revealed that each ecotype has a strong home-side advantage, indicating that this divergence has an adaptive component. Furthermore, Bertel et al. (2018) and Szukala et al. (2022b) uncovered enhanced phenotypic plasticity in the derived montane ecotype.

As a result of the adaptation to different environments, the ecotypes also show several morphological differences. The thinner leaves and reduced amount of stomata of the montane ecotype could, for instance, help to adapt to different levels of irradiance and humidity. Furthermore, the leaves and stalks of the montane ecotype show a thick layer of multicellular trichomes, which could potentially indicate an adaptation to the higher herbivory pressure of the montane environment, and/or an adaptation to increased drought stress. Interestingly, these multicellular trichomes are absent on the alpine ecotype (Bertel et al., 2016). Planting individuals in a common garden (CG) showed that their morphological divergence stays stable for multiple generations (Bertel et al., 2018), which implies that morphological and physiological differentiation is, at least in part, genetically driven. Only little amounts of shared genetic differentiation were found to underlie functional convergence across different ecotype pairs, indicating a highly polygenic basis of trait differentiation (Szukala et al., 2022a). Analyses regarding their population structure revealed that neighboring alpine and montane populations tend to cluster more often with each other instead of the same ecotype of different populations (Trucchi et al., 2017; Szukala et al., 2022b.), indicating that alpine-montane populations diverged multiple times independently from each other. For all tested ecotype pairs, the time of divergence was inferred as postglacial, which further confirms the assumption that the ecotypes have polytopic origins and were therefore products of recurrent parallel evolution (Trucchi et al., 2017).

Other examples of parallel evolution have been documented for walking stick insects (Nosil et al., 2022), beach mice (Steiner et al., 2009), or saltmarsh beetles (Van Belleghem et al., 2018). Parallel evolution can occur when ancestral lineages split and common polymorphisms are

repeatedly selected in similar environmental niches. However, the mechanisms behind recurrent evolution are still not fully understood (Maeso, Roy & Irimia, 2012). Several molecular mechanisms are known which could be the driving force of recurrent phenotypic differentiation, like single nucleotide polymorphisms, epigenetic modification, and post-transcriptional targeting by smRNAs (Moss 2001; Wu and Zhang 2014; He et al., 2022). In plants, smRNAs have a length of 20 to 24 nucleotides and belong to the group of noncoding RNA chains. In plants, they can be sorted into two main classes of smRNAs: I) microRNAs (miRNAs) with a length of approximately 20–22 nucleotides, which are responsible for post-transcriptional gene silencing by mRNA cleavage and translational repression (Bartel, 2009; Li et al., 2018), and II) short-interfering RNAs (siRNAs) with a length of 21–24 nucleotides. SiRNAs are responsible for post-transcriptional gene silencing and transcriptional gene silencing by mRNA cleavage, translational repression, and DNA methylation (Yuan et al., 2004; Laganà et al., 2015). SmRNAs in eukaryotes are known as key regulators of gene expression networks (He & Hannon, 2004). Because of different functions like a regulatory and – among others – a silencing function, they can play an important role in the formation of different phenotypes (Bonnet, Van de Peer & Rouzé, 2006). Studies by Hewezi (2020) or Bonnet, Van de Peer and Rouzé (2006) have already shown that among other functions, smRNAs also play an important role in the enhancement of parasite resistance, or responses to abiotic and biotic stresses.

The following thesis investigates how smRNAs are involved in recurrent ecotype formation. By combining previous results with further analyses based on the smRNA profiles of the different *Heliosperma pusillum* individuals (transplanted and grown in a common garden), it might be possible to identify differentially targeted genes that are responsible for their specific phenotypes and associated with the already documented phenotypic plasticity in Heliosperma pusillum. This will help to gain a deeper understanding of adaptational processes and possibly aid more general predictions about how organisms adapt to new environments.

The aim of this thesis is to detect shared differentially targeted genes between evolutionary replicates grown in a common garden and to explore repeated altitudinal effects of smRNAs targeting within reciprocal transplantations. We expected the latter to also deliver information about the amount of plasticity in smRNA targeting depending on the

environment. The hypothesis is that smRNA targeting reflects the change in the altitudes as well as the demographic history, genetic variation, and gene expression patterns (Szukala et al., 2022). By comparing previously analyzed gene expression with smRNA signals, it might be possible to associate the different potential driving forces of phenotypic differentiation (i.e., smRNAs and genes). Because this project investigates ecotypes that diverged based on polygenic traits and benefit from naturally formed biological replicates with a polytopic origin (Szukala et al., 2022), we anticipated that we will observe differently targeted regions leading to similar phenotypic outcomes across population pairs.

2. Materials and Methods

2.1 Material and sample preparation

We sequenced smRNA that was extracted from leaves of four *Heliosperma pusillum* population pairs that have been previously analyzed with RNAseq (for more details on the experimental setup see Szukala et al., 2022a; Szukala et al., 2022b). The location of where the seeds of each sample set was collected is illustrated in figure 2a. The analysis of the resulting smRNA libraries was conducted with two sample sets (table 1). The first sample set contains smRNA data of four *H. pusillum* populations of each ecotype grown in a common garden (figure 2b). The number of biological replicates per population varies between two and four (thus, the experiment includes a total of 23 individuals). The second set contains smRNA data of individual plants that were reciprocally transplanted in Lienzer Dolomiten, Kärnten (Austria; alpine site: 46.762 N 12.877 E, 2,055 m; montane site: 46.774 N 12.901 E, 790 m) in 2017. Seeds of wild populations of both ecotypes were sampled in the Puez-Geisler region, Trentino-Südtirol/Alto Adige (Italy; alpine site: 46.601 N 11.768 E, 2,290 m; montane site: 46.564 N 11.77 E, 1,690 m). Each of both ecotypes of *H. pusillum* is represented by two populations in their original environment as controls as well as by transplanted individuals into the non-native environment (figure 2c). Each of the populations contains approximately 5 biological replicates, which lead to a total number of 37 accessions. Altogether, 60 accessions were analyzed including common garden and transplanted individuals. Further information

regarding the grouping and used abbreviations are displayed in table 2 and supp. table 1. The wet lab procedure to extract smRNAs and library preparation is described in Chapter 3 of Szukala 2022.

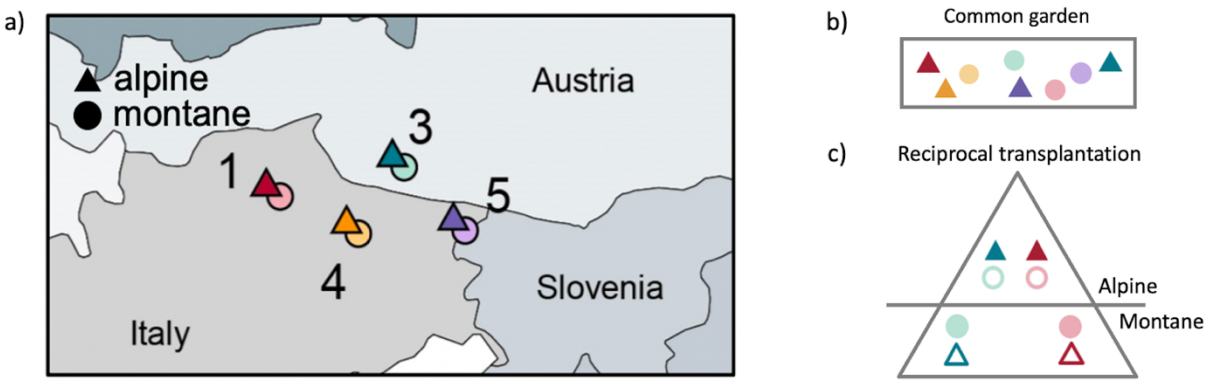

Figure 2 Geographical map of the analyzed samples and experimental setup (Szukala et al., 2022a). The numbering of the ecotype pairs is consistent with previous work done by Bertel et al. (2018). Color coding is consistent across all three subfigures. Alpine individuals are represented by triangles and montane individuals by dots. Filled symbols indicate samples in their natural environment, unfilled symbols represent reciprocally transplanted individuals. (a) Geographical map of the analyzed population pairs. (b) Visualization of the experimental setup for the common garden data. (c) Visualization of the experimental setup for the reciprocal transplanted data.

Table 1 Samples used in the study, an identifier of the biological replicates and number of accessions per group and in total for samples of the common garden and the transplanted samples. * biological replicates with technical replicates.

Group	Replicates	(n)	Group	Replicates	(n)
Common garden			**Transplanted**		
Total:		**23**	**Total:**		**37**
A1	A, C, D	3	A1_D	A, B, D, E, F	5
A3	A, B,	2	A1_U	A, B, C, D, E	5
A4	A, B, C	3	A3_D	B, C, D, G, J	5
A5	B, C, D	3	A3_U	A, B, C, D, F	5
M1	A, B, C, D	4	M1_D	A, B, C, E, F	5
M3	A, B, C	3	M1_U	A, B, C, E, F*	5
M4	A, B, C	3	M3_D	A, B, C, E, F	5
M5	C, E	2	M3_U	A*, B	2

2.2 Reads alignments and delimitation of target genomic regions

Trimgalore version v.0.6.4 (https://github.com/FelixKrueger/TrimGalore) (Krueger, 2012) was used for adapter trimming, trimming for length (20 nt – 24 nt), and quality trimming (--quality *20*). Regions shorter than 16 bp after quality trimming were considered not informative enough and therefore excluded. The bioinformatical pipeline used for this thesis is based on

Thornton (2022) and Chapter 3 of Szukala 2022. STAR v.2.7.9a (Dobin et al., 2013) was used to align the reads to the *Heliosperma pusillum* reference genome v.1.0 (NCBI accession number JAIUZE010000000; Szukala et al., 2022). Following manual recommendations, we did not allow to split reads around intronic regions and reduced the amount of allowed mismatches using the options --alignIntronMax 1 (i.e. allow only 1 nt intron per read) and --outFilterMismatchNoverLmax 0.05 (i.e. allow only 5 % mismatching base pairs per read). After sorting, the files were sorted by coordinates using samtools sort (Li et al., 2009; Danecek et al., 2021). Deeptools v.3.5.1 (Ramírez et al., 2016) was used to quantify the amount of smRNA reads mapped to 100 bp non overlapping genomic windows (*bamCoverage -b $file -o $file.bedgraph -of bedgraph --normalizeUsing None -bs 100 --numberOfProcessors* max/2). To account for uncertainties in the annotations of the reference *H. pusillum* genome, we chose an annotation-free approach for the further analytical steps. We therefore retained only those windows that were targeted by ten or more reads in that particular accession for each individual. Merging the regions of interest between accessions and combining adjacent features into a single bed file was done by using bedtools v.2.30.0 (Quinlan & Hall, 2010). By doing so, the information about peak regions of the smRNA alignments was stored in a bed file which was later converted into a .gff file, using the Bioconductor R packagertracklayer v.1.42.2 (Lawrence, Gentleman & Carey, 2009). The code used for the trimming, mapping and indexing is displayed in the appendix/code section.

2.3 Differential Targeting Analyses

A table of counts was produced by using FeatureCounts (Rsubread) v.2.0.3 (Liao, Smyth & Shi, 2014; Liao, Smyth & Shi 2019). Since smRNAs are known to also target multiple times in a genome like in members of transposable element (TE) families (Borges & Martienssen, 2015) or paralog genes/transcription factors (Jover-Gil et al., 2014), the *-M --fraction* option was used. This option counts multi-mapped reads as a fraction of the total mapping positions. If a smRNA is targeting three genomic regions it is counted as one-third each time. Analyses of the regions differentially targeted by smRNAs were done in R v.1.4.1106 (Team, 2013) using the Bioconductor package EdgeR v.3.24.3 (Robinson et al., 2010). In a further trimming step, every

differentially targeted region which had a mean count per million (cpm) above or equal to 0.5 was kept across individuals. This step was done to decrease the number of false positive differential targeted regions and therefore to maximize the power of the analyses. To account for library size, the table of counts was again normalized using the trimmed mean of the M-Value (*TMM*) method, which is explained in detail in Robinson & Oshlack (2010). To consider biological variance, a region-wise dispersion across all regions was estimated using the *estimateDisp* option. Afterwards, a quasi-likelihood negative binomial generalized log-linear model (EdgeR function *glmQLFit*) was fitted. Finally, the EdgeR function *glmQLFTest* was used to test for differentially targeted regions (DTRs) via a quasi-likelihood test. The different contrast groups used can be seen in table 2. The significance of the resulting DTRs was estimated by correcting p-values for multiple testing using Benjamini-Hochberg correction (Benjamini & Hochberg, 1995). Using the *makeContrast()* function, alpine smRNA counts were defined as base targeting level in a montane vs alpine comparison. For the reciprocal transplanted samples, smRNA counts of the native environment were used as base targeting level. Further information regarding the individual comparison groups can be found in table 2. The code used for differentially targeting analyses is also displayed in the appendix/code section.

Table 2 Information about the used contrast groups. The table contains used abbreviations and the respective details for samples of the common garden and the transplanted samples.

Abbreviation	Details
Common garden	
M1vsA1	Montane ecotype, location 1 vs Alpine ecotype, location 1
M3vsA3	Montane ecotype, location 3 vs Alpine ecotype, location 3
M4vsA4	Montane ecotype, location 4 vs Alpine ecotype, location 4
M5vsA5	Montane ecotype, location 5 vs Alpine ecotype, location 5
Transplanted	
MvsA_1D	Montane ecotype location 1 down transpl. vs Alpine ecotype, location 1 down transpl.
MvsA_3D	Montane ecotype location 3 down transpl. vs Alpine ecotype, location 3 down transpl.
MvsA_1U	Montane ecotype location 1 up transpl. vs Alpine ecotype, location 1 up transpl.
MvsA_3U	Montane ecotype location 1 up transpl. vs Alpine ecotype, location 1 up transpl.
AUvsAD_1	Alpine ecotype, location 1 down transpl. vs Alpine ecotype, location 1 up transpl.
AUvsAD_3	Alpine ecotype, location 3 down transpl. vs Alpine ecotype, location 3 up transpl.
MDvsMU_1	Montane ecotype, location 1 up transpl. vs Montane ecotype, location 1 down transpl.
MDvsMU_3	Montane ecotype, location 3 up transpl. vs Montane ecotype, location 3 down transpl.

2.4 GO Enrichment of differentially targeted genes

To identify which DTRs are overlapping with genomic regions of interest (genes, mRNAs, exons, coding sequences (CDS) and untransmitted regions (3'UTRs and 5'UTRs), bedtools intersect (Quinlan & Hall, 2010) was used. DTRs overlapping with genes (including potential regulatory regions and promoter sites – defined as 1,000 bp up- and down-stream of genes) were defined as differentially targeted genes (DTGs). The .gff file containing the annotated genome of *H. pusillum* was provided by Szukala et al. (2022a). The GO-enrichment was performed on basis of Fisher's exact test, which was done by using the Bioconductor package topGO v.2.34.0 (Alexa & Rahnenfuhrer, 2016). A biological process was considered as significantly overrepresented when the FDR adjusted p-value was lower than 0.05. By chance, multiple DTRs can overlap with one gene. Should this case arise, one DTR was randomly chosen and its differential targeting values were retained. This step is necessary because the GO-term enrichment analyses can only be performed if each gene of the DTG list has a unique value. To summarize the GO-terms and visualize them, Revigo (http://revigo.irb.hr) (Supek et al., 2011) was used. The resulting list of GO-terms was filtered manually for GO-terms of interest (i.e. related to biotic stress, circadian rhythm, drought, epigenetics, light response, ozone, pigmentation, stress, temperature, trichome, and triterpene). After row-wise normalization of the number of DTGs exhibiting a specific GO-term, a heatmap was produced by using pheatmap v. 1.0.12 (Kolde, 2019). The number of genes and significance was added manually. The code used for GO-Terms enrichment analyses is also displayed in the appendix/code section.

2.5 Shared DTRs, DTGs, and GO-Terms

The number of uniquely shared DTRs, DTGs, and GO-terms was inferred by using the R packages Upset v.1.4.0 (Jake Conway, 2019). For further visualization ComplexHeatmap v.2.13.1 (Gu, Eils & Schlesner, 2016) was used. The code used to calculate the intersection between DTRs, DTGs, and GO-terms is also displayed in the appendix/code section.

3. Results

3.1 Principal Component Analyses

For the common garden data, we obtained approximately 20 million reads per sample. After filtering, approximately 13 million reads remained. For the transplanted samples, we obtained approximately 11 million reads and after filtering, approximately 6 million reads remained. Number of reads for each sample before and after trimming can be found in supp. table 2. A principal component analysis was performed based on normalized read counts.

For the data of the common garden (figure 3 smRNA Data PCA – Common Garden), PC1 explains 16.7 % of the given variation. PC2 explains 12.4 % of the variation. The first PC separated ecotype pair 5 from the rest, while PC1 and 2 in combination tend to group the population pairs by shared ancestry (Trucchi et al., 2017; Szukala et al., 2021a), showing that more similarity is found among neighboring ecotypes of each pair than among different populations pairs of the same ecotype. Especially for pairs 3 and 4, the montane and the alpine ecotypes are aligned more with each other. The highest similarity between the two ecotypes of one group can be observed within pair 5 followed by pairs 3 and 4, whereas pair 1 shows the lowest similarity. For the smRNA data of the transplanted samples, PC1 explains 10.4 % of the variation in the data and groups the samples by ecotype pair (e.g., pair 1 vs pair 3). PC2 (8.4 % of variation explained) tends to group the samples by ecotype (i.e., alpine vs montane).

Similar patterns can be found for the RNAseq data of the common garden and transplanted samples (figure 3). For the common garden samples, PC1 (15.5 %) divides population pair 5 from the rest of the samples and PC2 (13.1 %) tends to divide the sample set by ecotypes. Compared to the smRNA data, the different population pairs show more similarities among each other but do not show a clear grouping. The PCA-plot of the RNAseq data for the transplanted samples shows more dispersed patterns. PCA1, which explains 16.9 % of the given variation, tends to group the different datasets by population pairs.

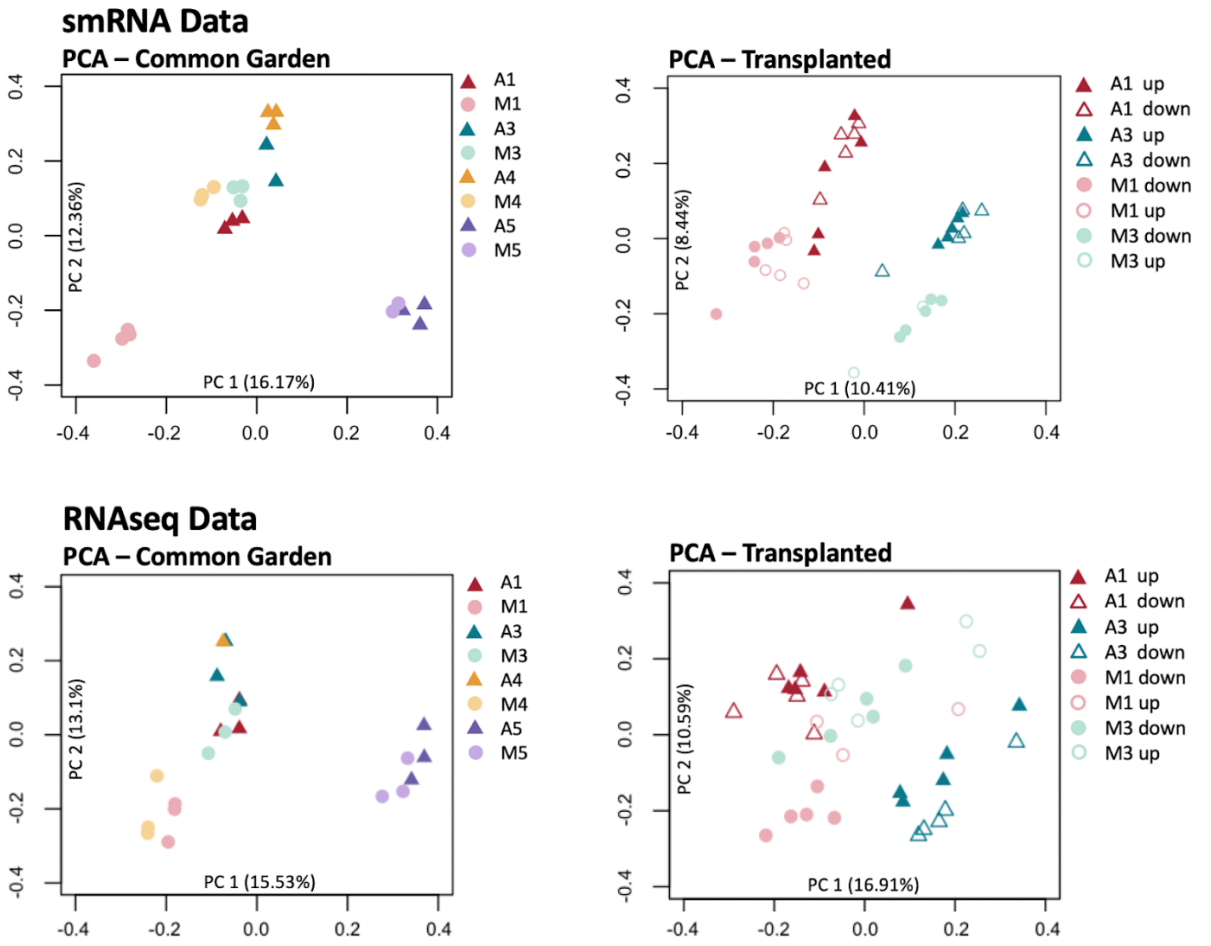

Figure 3 Principal component analyses based on normalized read counts of smRNA data and RNAseq data for the common garden samples and the transplanted samples. Each dot represents one individual sample. Alpine individuals are represented by triangles and montane individuals by dots. Filled symbols indicate samples in their natural environment, unfilled symbols represent reciprocally transplanted individuals. Each color represents a different population pair.

3.2 Differentially Targeted Regions

Figure 4 shows the MA-plots for the comparison groups of the common garden as well as for the comparison groups of the transplanted individuals. To compare the response of the population pairs and ecotypes of the common garden samples, four different MA-plots were generated (figure 4, upper four MA-plots). The highest number of DTRs can be observed for the comparison between montane and alpine population pair 1 (33,837 (i.e., 19,736 up and 14,101 down)). The lowest number of DTRs can be observed for the comparison between montane population pair 5 and alpine 5 (1,101 (i.e., 706 up and 395 down)). The number of up-targeted regions differs between 19,736 (montane 4 vs. alpine 4) and 706 (montane 5 vs.

alpine 5). The number of down-targeted regions differs between 15,045 (montane 4 vs. alpine 4) and 395 (montane 5 vs. alpine 5). Comparison group 5 shows the lowest amount of differentially targeted regions for both (up- and down-regulated).

For the ecotype comparison of the transplanted individuals (figure 4 transplanted, middle four MA-plots), the highest number of differentially targeted regions can be observed for the comparison between montane and alpine 1 in the montane environment (18,782 (i.e., 8,713 up and 10,069 down)). The lowest number of differentially targeted regions can be observed for the comparison between montane 1 and alpine 1 in the alpine environment (10,086 (i.e., 5,244 up and 4,842 down)). The number of up-targeted regions differs between 13,614 (montane vs. alpine 3 up) and 5,244 (montane vs. alpine 1 up). The number of down-targeted regions differs between 10,069 (montane vs. alpine 1 down) and 4,138 (montane vs. alpine 3 up).

To check for the environmental response, four additional plots were generated (figure 4 transplanted, lower four MA-plots). The highest number of differentially targeted regions can be observed for the comparison between montane 3 up and montane 3 down (9,902 (i.e., 9,082 up and 820 down)). The lowest number of differentially targeted regions can be observed for the comparison between alpine 3 down and alpine 3 up (387 (i.e., 285 up and 102 down)). The number of up-targeted regions differs between 9,082 (montane 3 up vs. montane 3 down) and 285 (alpine 3 down vs. alpine 3 up). The number of down-targeted regions differs between 1,622 (montane 1 up vs. montane 1 down) and 102 (alpine 3 down vs. alpine 3 up). Interestingly, the alpine comparisons (i.e. ADvsAU1 and ADvsAU3) showed notably fewer DTRs than the montane comparisons (i.e. MUvsMD1 and MUvsMD3).

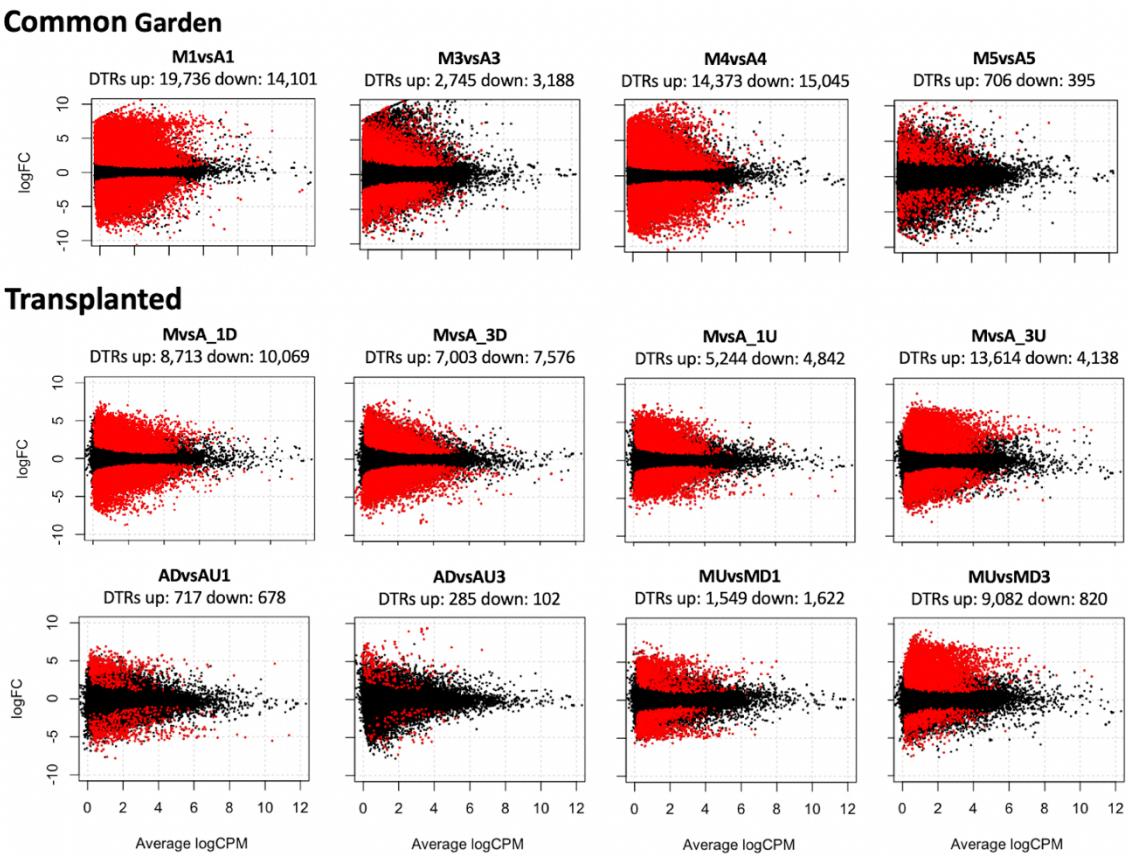

Figure 4 MA-plots and the number of up- and down-regulated regions of the common garden and transplanted sample. Each MA-plot represents one comparison group, and each dot represents a targeted genomic region. Displayed is the log fold change (logFC) against the average log counts per million (logCPM). Sequences with an FDR smaller than 0.05 are displayed in red.

After differential expression analyses, it was determined to which genomic components (coding sequences (CDSs), differentially targeted genes (DTGs), Exons, and UTRs (untranslated regions)) the DTRs belong. Figure 5 shows the number of DTRs for each comparison group. Raw numbers of the DTR composition are contained in supp. table 3. Comparing the counts of these differentially targeted annotated regions, significant differences between the comparison groups become clear: For nearly each comparison group the number of up-regulated CDSs, DTGs, exons, and UTRs is higher than the number of down-regulated CDSs, DTGs, exons, and UTRs for the same group.

For the up-regulated differentially targeted regions, the highest number of CDSs is given by comparison group M1vsA1 (2,638). The lowest number is given by M5vsA5 (164). DTGs are ranging between 5,045 (M1vsA1) and 255 (M5vsA5). Interestingly, A1vsM1 shows the lowest

(25.6 %) and A5vsM5 the highest (36.1 %) percentage of DTGs (in regard to the number of DTRs). For the exons, the total number ranges between 3,240 (M1vsA1) and 198 (M5vsA5). Untransmitted regions (UTRs) range between a total of 940 (M1vsA1) and 54 (M5vsA5). For the down-regulated common garden data, the total number of differentially targeted coding sequences (CDSs) ranges between 1,080 (M4vsV4) and 14 (M5vsA5). The total number of differentially targeted genes (DTGs) ranges between 4,300 (M1vsA1) and 78 (M5vsA5). For the exons, the total number ranges between 1,616 (M4vsA4) and 24 (M5vsA5). Untransmitted regions (UTRs) exhibit a total number between 789 (M4vsA4) and 12 (M5vsA5). In regard to the total number of DTRs for each comparison group, M1vsA1 is showing the highest percentage. Here 30.5 % of the DTRs are also DTGs. M5vsA5 is showing the lowest percentage. For this comparison group, 19.7 % of the DTRs are also DTGs. This is the lowest percentage of DTGs of the whole dataset.

For the up-regulated transplanted data, the total number of differentially targeted CDSs ranges between 3,032 (MvsA_3U) and 44 (ADvsAU3). The total number of DTGs ranges between 4,230 (MvsA_3U) and 82 (ADvsAU3). For the exons, the total number ranges between 3,518 (MvsA_3U) and 54 (ADvsAU3). UTRs show total numbers between 762 (MvsA_3U) and 20 (ADvsAU3). In regard to the total number of DTRs for each comparison group, ADvsAU1 is showing the highest percentage. Here 39.2 % of the DTRs are also DTGs. MvsA_U1 exhibits the lowest percentage. For this comparison group 27.4 % of the DTRs are also DTGs. For the down-regulated DTRs, the highest number of CDSs is given by comparison group MvsA_1D (1,548). The lowest number is given by ADvsAU3 (8). DTGs are ranging between 3,278 (MvsA_1D) and 32 (ADvsAU3). For the exons, the total number ranges between 1,896 (MvsA_1D) and 14 (ADvsAU3). For the UTRs, total numbers between 592 (MvsA_1D) and 8 (ADvsAU3) can be seen. Compared to the total number of DTRs, MUvsMD3 is showing the highest percentage of DTGs of the whole dataset. 44.4 % of the DTRs of this comparison group are also DTGs. With 24.5 %, the comparison group AvsM_3U shows the lowest percentage of DTGs for the down-regulated DTRs of the transplanted data. Interestingly, comparison groups of the same ecotype grown in different transplantation locations always show lower values of CDSs, DTGs, exons, and UTRs as comparison groups of different ecotypes grown in the same environment.

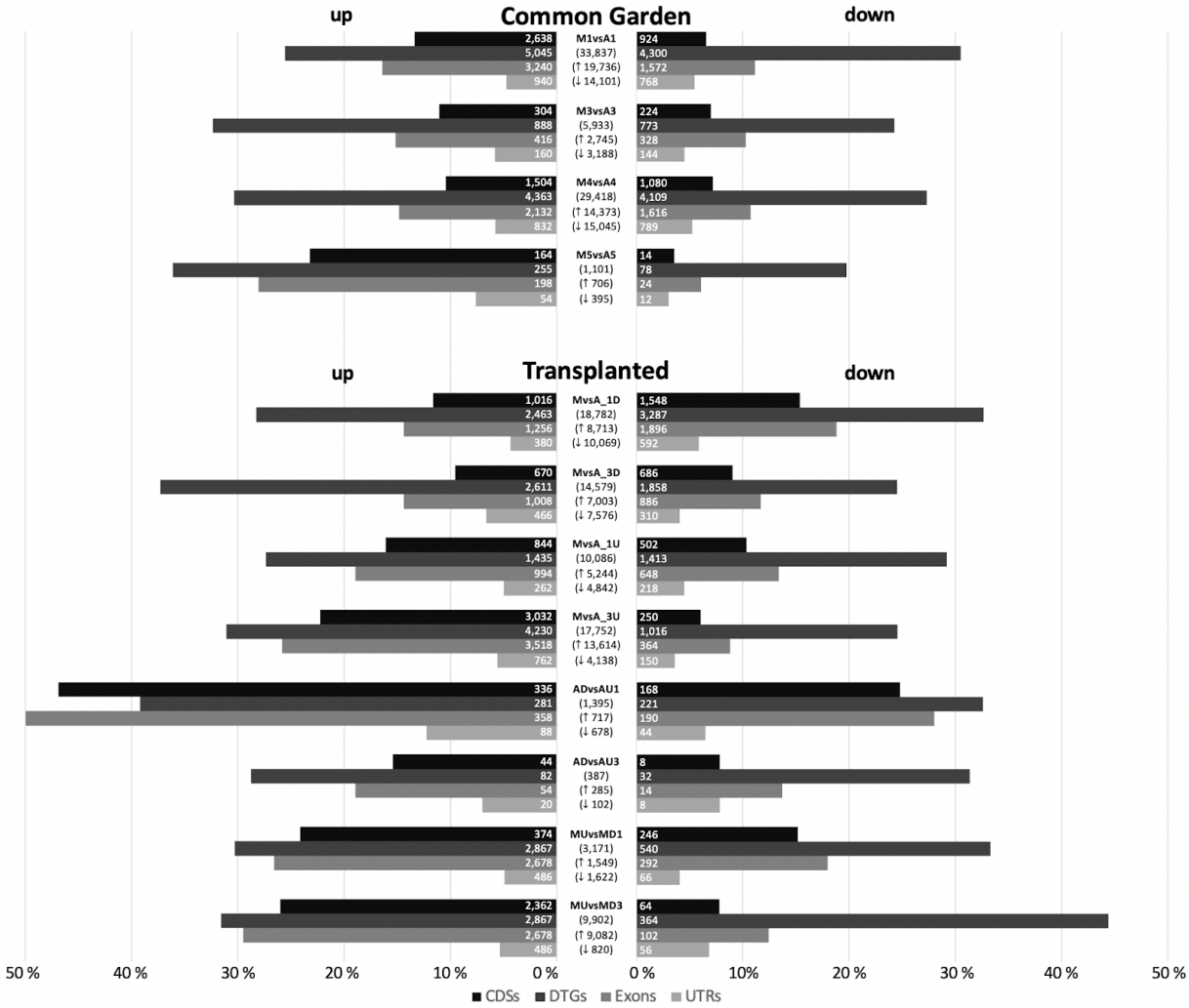

Figure 5 Bar chart in percentage of the differentially up- (left) and down- (right) expressed annotated regions for the common garden and the transplanted samples (with respect to the total number of up- or down-targeted regions). The abbreviation for each comparison group as well as the total number and the number of up- or down-targeted regions are shown in the middle. The bars refer to the number of coding sequences (CDSs), differentially targeted genes (DTGs), exons, and UTRs (untranslated regions) from up to down respectively. Total numbers are given in white.

3.3 Gene Ontologies

The total number of significant GO-terms for each comparison group varies between approximately 50 to 200. In total, the analysis led to 1,845 unique GO-terms that appear enriched in at least one analysis. Compared to the RNA-seq data, more enriched GO-terms in the smRNA analyses were observed. Supp. tables 4 and 5 contain the top five GO-terms (filtered by p-value) of the common garden and transplanted samples.

Comparing the number of DTRs exhibiting selected GO-terms associated to biotic and abiotic stresses and trichome development (figure 6), we can see that the common garden samples show a more dispersed pattern compared to the data of the transplanted samples. Interestingly, this group also shows the most significant values. While the comparison groups M3vsA3 and M5vsA5 show lower values, the comparison groups M1vsA1 and M4vsA4 show higher and more significant values for each up- and down-regulated GO-terms. Especially the montane ecotype of population pair 1 shows a high proportion of differentially targeted genes that are related to trichomes. The same comparison shows a down-regulation of DTGs related to biotic stress while two of them (GO:0009611 response to wounding and GO:0009607 response to biotic stimuli) are highly significant (p-value < 0.005). The same GO-terms are also significantly up-regulated in M4vsA4. Regardless of up- or down-regulation, the common garden data shows way more significant values for response to ozone (GO:001093) and response to reactive oxygen species (GO:0000302).

The GO-term analyses of the ecotype comparison of the transplanted data reveals a higher number of similar values. While the up-transplanted montane sample of population pair 3 shows the highest values, the same comparison group shows the lowest number of down-regulated GO-terms. The only high significant GO-term given by this group is related to demethylation (GO:0070988). Further focusing on GO-terms related to trichome development, we can see that the up-transplanted montane sample of population pair 3 exhibits highly significant signals for trichome branching (GO:0010091) and trichome morphogenesis (GO:0010090). Individuals of the same population pair and ecotype but down-transplanted exhibiting a significant down-regulation for the same GO-terms. The same comparison group's defense response to insects is highly significantly down-regulated (p-value < 0.005).

Having a look at the comparison groups of the environmental response, we can observe on average the lowest number of DTGs exhibiting a specific GO-term. Up-transplanted montane individuals compared to down-transplanted montane individuals of population pair 3 reveals the highest values for each GO-term within this subgroup. The only highly significant values here are given by up-transplanted montane individuals of population pair 1 (GO:1905156 and

GO:0043155 up-regulation of negative reaction of photosynthesis) and the same group of individuals but down-regulated (GO:1901000 regulation of response to salt stress and GO:0035065). Furthermore, transplanting alpine individuals of population pair 1 down produces a significant (p-value < 0.05) upregulation of different epigenetic-related GO-terms (histone H3-K27 trimethylation (GO:0098532), negative regulation of gene expression, epigenetic (GO:0045814) and regulation of histone acetylation (GO:0035065)).

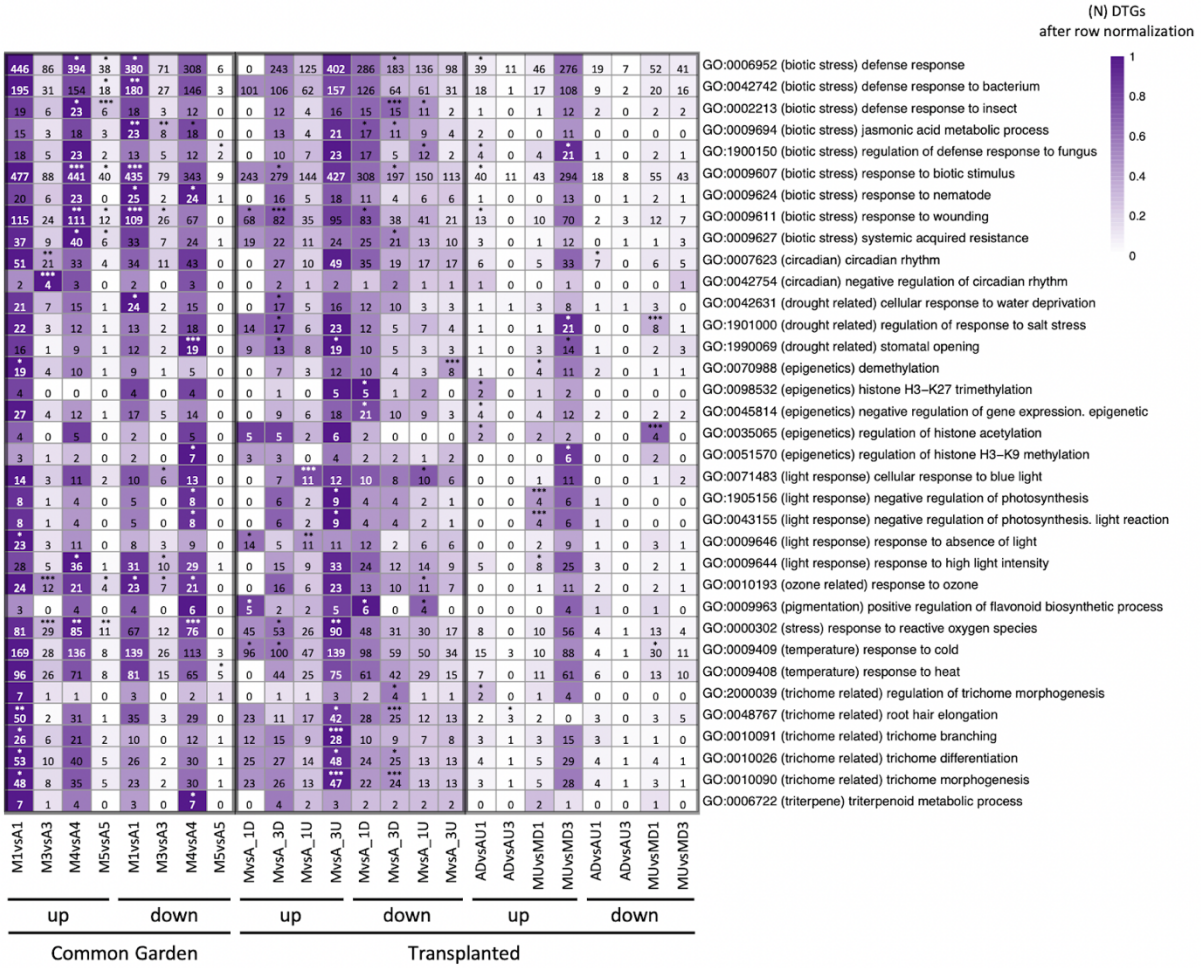

Figure 6 Heatmap of significantly enriched GO-terms of the common garden and transplanted data. Color coding refers to the row-normalized number of differentially targeted genes exhibiting a specific GO-term. Total numbers and confidence intervals are also displayed within the heatmap. * indicates a p-value <0.05; ** indicates a p-value <0.01 and *** indicates a p-value <0.005

3.4 Shared Differentially Targeted Regions

Differentially targeted genes and GO-term enrichment analyses overlaps were displayed in a combined Upset-plot (figure 7 and figure 8). Each Upset-plot contains the number of DTRs,

DTGs, and GO-terms shared between the different comparison groups and unique to each comparison group.

For the up-regulated common garden data (figure 6 left), no DTR or GO-term is uniquely shared between all four comparison groups, but four DTGs uniquely shared among all 4 comparison groups were found. While M1vsA1 shows the most unique DTRs (17,088) and DTGs (2,934), the most unique GO-terms can be observed in comparison group A4vsM4. The highest number of DTRs and DTGs shared between the two comparison groups is found between M1vsA1 and M4vsA4 (2,030 DTRs and 922 DTGs). Interestingly, the highest number of shared unique GO-terms can be found among comparison groups M1vsA1 and M3vsA3. Similar patterns were observed for the comparison of down-regulated DTRs, DTGs, and GO-terms (figure 6 right). Here, no uniquely shared DTRs or GO-terms between all 4 comparison groups are observed. The two comparison groups that shared the most unique DTRs are again M1vsA1 and M4vsA4 (1,878). These two groups also share the most DTGs (706) and GO-terms (14). M4vsA4 shows also the most unique DTRs within (12,311), while the most unique DTGs and GO-terms to one group are given by comparison group M1vsA1. Interestingly, even if no DTR is shared between all four comparison groups, 1 DTG is found uniquely shared among them. Furthermore, no uniquely shared GO-term was found between all four groups. This indicates that the GO-terms of the shared DTG can be also found among other differentially targeted genes.

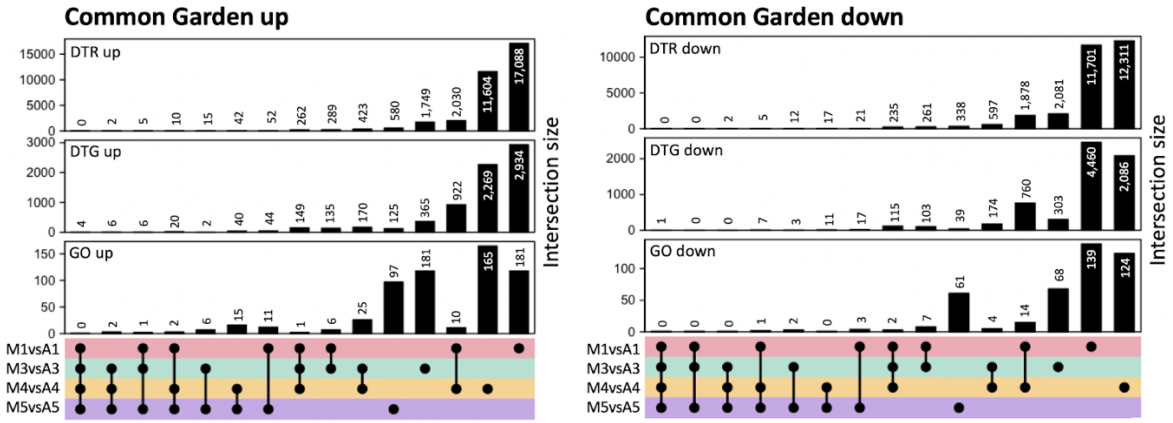

Figure 7 Upset-plot of the common garden data. Displayed are the intersection sizes of the up- or down-regulated differentially targeted regions (DTR), differentially targeted genes (DTG), and GO-terms of the corresponding DTGs. The different colored bars below represent the different comparison groups. Each dot indicates respective comparison groups included in the number of intersections above.

Figure 8 displays Upset-plots of up- and down-regulated DTRs, DTGs, and GO-terms for the ecotype comparison (upper two Upset-plots) and the environmental response (lower two plots). For the up-regulated data of the ecotype comparison, 154 DTRs shared between all 4 comparison groups can be observed, but only 60 DTGs and no GO-terms are uniquely shared between all four. The highest number of shared DTRs between the two comparison groups (1,668) can be found between MvsA_1D and MvsA_1U (same ecotype pair but different transplantation location). The highest number of DTGs between the two comparison groups (514) can be found between MvsA_3D and MvsA_3U (also the same ecotype pair but different transplantation locations). Interestingly, the highest number of shared GO-terms can be found between MvsA_1D and MvsA_3D (different ecotype pairs but the same transplantation location). The comparison group MvsA_3U shows the highest number of unique DTRs, DTGs, and GO-terms (11,302; 2,463 and 160 respectively). For the down-regulated data, similar patterns can be found. The number of DTRs, DTGs, and GO-terms shared among all four comparison groups is decreasing (112, 53, and 0 respectively). The highest number of DTRs and DTGs between two comparison groups is again shared between comparison groups of the same region but a different transplantation location (MvsA_1D and MvsA_1U as well as MvsA_3D/MvsA_3U). The highest number of GO-terms shared between two groups can be found between MvsA_1D and MvsA_3U. While MvsA_1D also shows the highest number of unique DTRs and DTGs, MvsA_3D shows the highest number of unique GO-terms.

For data on the environmental response, on average fewer DTRs, DTGs, and GO-terms were observed. For the up-regulated data, we cannot observe any uniquely shared DTRs between more than two comparison groups. The highest number of uniquely shared DTRs, DTGs, and GO-terms between two comparison groups can be found between MUvsMD1 and MUvsMD3 (same transplantation location but different region). The highest number of DTRs, DTGs, and GO-terms for one comparison group is given by MUvsMD3 (8,665; 2,224 and 206 respectively). Also, for the down-regulated data, we cannot observe any DTR, DTG, or GO-term that is uniquely shared between all 4 comparison groups. The only intersection of unique DTRs between two groups can be observed for ADvsAU1 and ADvsAU3 (7) as well as MUvsMD1 and MUvsMD3 (30). The highest number of DTRs and DTGs for one comparison group is given by

MUvsMD1 (1,592 and 487). The highest number of GO-terms unique to one comparison group is given by ADvsAU1 (181).

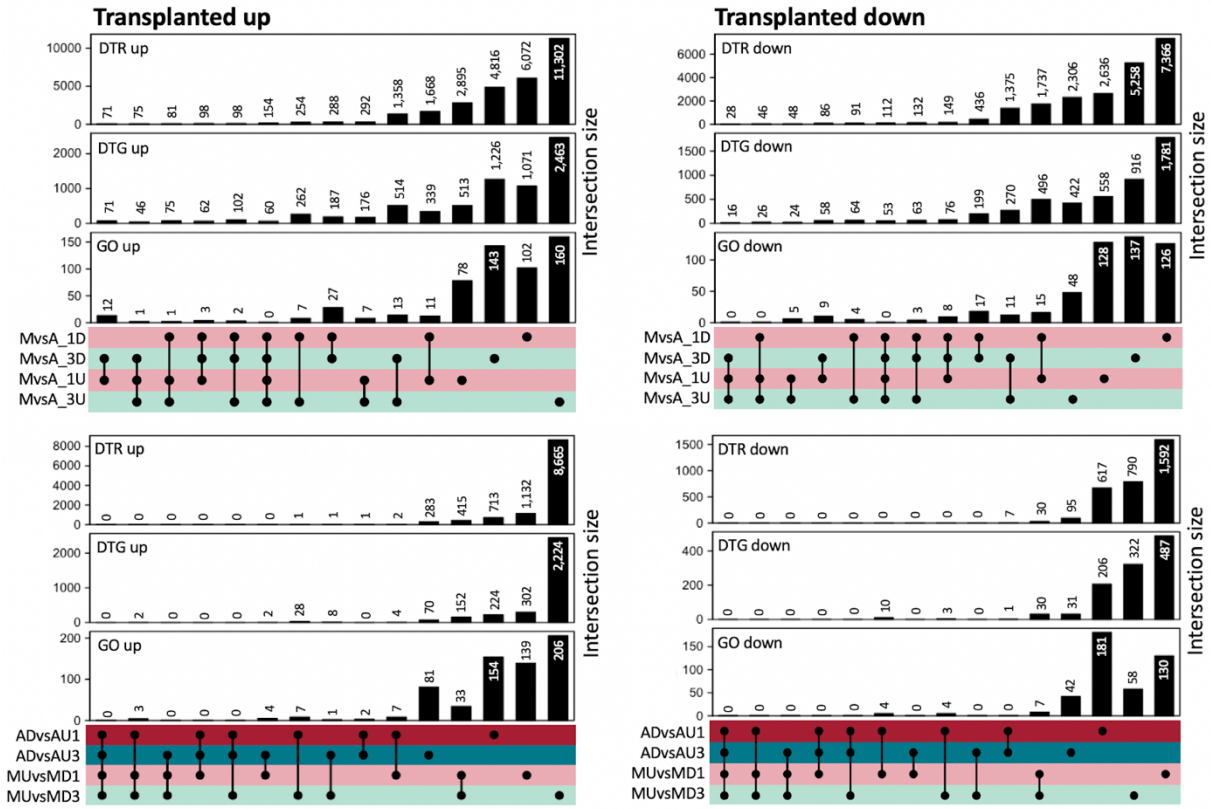

Figure 8 Upset-plot of the transplanted data. Displayed are the intersection sizes of the up- or down-regulated differentially targeted regions (DTR), differentially targeted genes (DTG) and GO-terms of the corresponding DTGs. The different colored bars below representing the different comparison groups. Each dot indicates respective comparison groups included in the number of intersections above.

The major proportion of DTRs, DTGs, and GO-terms turned out to be unique to each comparison group (displayed in figure 9). For the common garden data, the percentages of unique DTRs are between 86.6 % (M1vsA1 down) and 63.7 % (M3vsA3 down). The percentages of unique DTGs are between 71 % (M1vsA1 up) and 26.8 % (down M3vsA3). The highest percentage of unique GO-terms is exhibited by the up-regulated comparison group M5vsA5 (91.0 %), while the lowest percentage is given by the same comparison group but for the up-regulated GO-terms. For each comparison group, the percentage of unique DTGs is the lowest compared to DTRs and GO-terms. For the up-regulated data, the percentage of unique DTRs is always lower than the percentage of GO-terms. For the down-regulated data, the percentage of DTRs is always lower than the percentage of GO-terms (except for A3vsM3).

For the data of the ecotype comparisons of the transplanted samples (i.e., MvsA_1D, MvsA_3D, MvsA_1U, and MvsA_3U), the percentages of unique DTRs are between 83.0 % (MvsA_3U up) and 53.5 % (MvsA_1U up). The percentages of DTGs are between 73,2 % (MvsA_1D down) and 38.2 % (MvsA_1U up). The highest percentage of unique GO-terms is exhibited by the up-regulated comparison group MvsA_3U (83,8 %), while the lowest percentage is given by the up-regulated comparison group MvsA_1D (66,7 %). Most of the time, higher percentages of unique GO-terms than DTRs were observed. The exception is the comparison group MvsA_1D (up- and down-regulated data). For the comparison groups for the environmental response (i.e., MUvsMD1, MUvsMD3, ADvsAU1, and ADvsAU3), average higher percentages compared to the data of the ecotype comparisons can be observed. The mean for each comparison can be overserved in supp. table 6. Here, the highest percentage of differentially targeted regions is given by the up-regulated comparison group MUvsMD1 (99.4 %). The percentages of DTGs differ between 96.9 % (ADvsAU3 down) and 65,4 % (MUvsMD1 up). The percentages of GO-terms differ between 100 % (ADvsAU3 down) and 74,7 % (MUvsMD1 up).

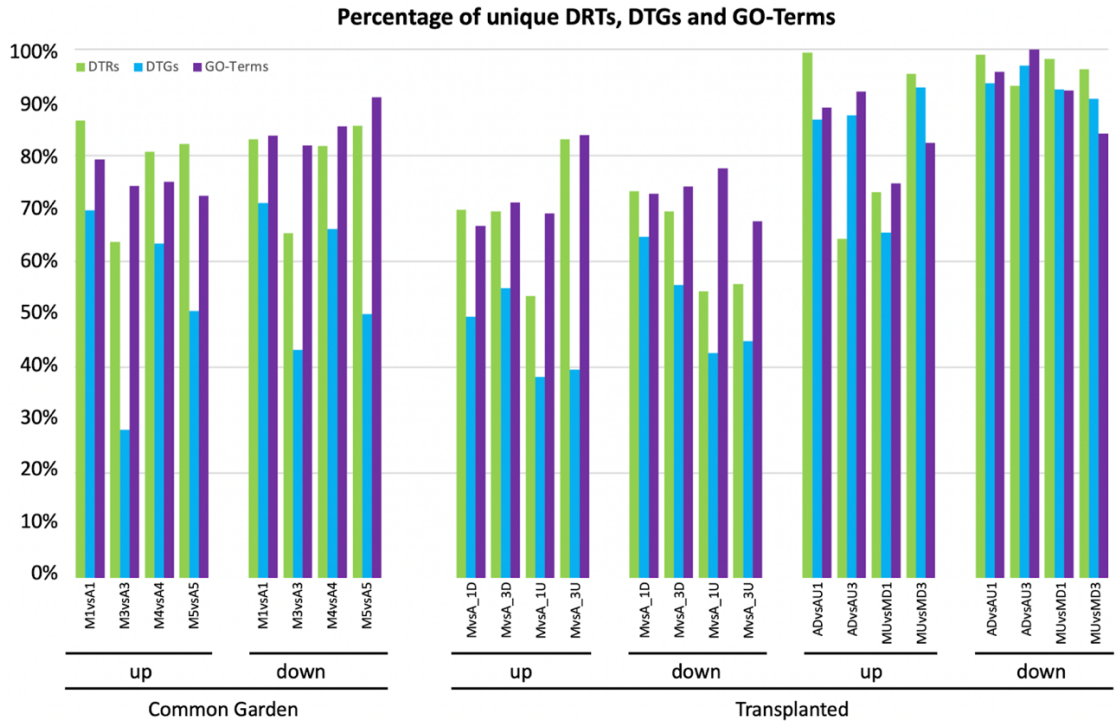

Figure 9 Percentage of up- or down-regulated unique differentially targeted regions (DTRs), differentially targeted genes (DTGs) and gene ontology (GO)-terms for the comparison groups of the common garden data and the data of the transplanted samples. Percentage of unique DTRs are given in green, percentage of unique DTGs in blue, and percentage of unique GO-terms in purple.

4. Discussion

For this analysis, the smRNA profiles were compared between ecotypes that differ in several traits that have been previously hypothesized to be highly polygenic (Szukala et al., 2022a). Naturally formed biological replicates of the two ecotypes exhibited by this plant were grown in a common garden and under a reciprocal transplantation setting. Afterward, it was investigated how the smRNA signals differ between the different ecotypes and their transplantation locations. Since it is known that the different ecotypes evolved in parallel (Szukala et al., 2022a), it was investigated how the smRNA composition changed during this early stage of ecological divergence and to what extent smRNAs are involved. It is already known that smRNAs play an important role in gene expression and phenotypic divergence (Hewezi, 2020; Bonnet, Van de Peer and Rouzé, 2006), but we still do not know much about how smRNAs are involved in ecological divergence (Silva et al., 2021). To change this circumstance, the biological implication of genes differentially targeted by smRNAs was compared with the biological implication of differentially expressed genes (Szukala et al., 2022a). Comparing transcriptional and post-transcriptional regulation helped to gain further insights into the role of smRNAs during recurrent evolution.

The cluster of normalized read counts of the smRNA shows similar patterns to the gene expression results of Szukala et al. (2022a). In the common garden data sets, one can see that the montane and alpine ecotypes of population pair 5 clearly separate from the rest of the dataset which is in line with genetic structure analyses done by Szukala et al. (2022a). Population pair 5 not only showed a lower degree of shared ancestry with the other population pairs but also appears to be the earliest diverging lineage (Trucchi et al., 2017). Interestingly, this fact seems to be also reflected in the smRNA data. For the transplanted dataset, the grouping of the gene expression data is not as clear as in the smRNA data but still shows similar trends. This could indicate that smRNAs and gene expression in *H. pusillum* are evolving in the same direction, but likely at different speeds. Since the PCA plot of smRNA data shows a clearer differentiation between the ecotypes of *Heliosperma pusillum*, it is likely that smRNAs expression shows a higher and faster response to a new environment than gene expression. This not only confirms earlier hypotheses that plants respond to a new

environment by accumulating mutations and epimutations (Jing et al., 2014) but also that epimutations in plants occur at a significantly higher rate than genetic mutations (Yao, Schmitz, & Johannes, 2021). The higher mutation rate of epigenetic features is also reflected in the circumstance that the smRNA data shows way more differentially targeted regions than the gene expression data (which also decreases the comparability between the two datasets).

It is known that differences regarding the smRNA activity accumulate over evolutionary time (Jing et al., 2014). This seems to be also reflected in the number of DTRs. As expected, more diverged ecotype pairs like population 1 and population 4 show more DTRs than less diverged ecotype pairs like populations 3 and 5. Having a look at the number of DTRs exhibited by comparing the same ecotype under different environmental conditions (i.e., up- and down-transplanted), significantly more differentially targeted regions for the montane ecotype than for the alpine ecotype are observable. In line with polygenic adaptation already hypothesized for the relevant adaptive traits in *Heliosperma pusillum* (Szukala et al., 2022a), this indicates that enhanced expression plasticity in the montane ecotype (Szukala et al., 2022b) is also reflected by the smRNA data.

Considering which specific genetic regions are targeted by smRNAs, it becomes obvious that not even half of smRNAs target genic regions (including promotors), and only a small portion of smRNAs is directly targeting exonic regions. This result suggests that a major portion of the smRNAs exert their regulatory effects at other genomic regions, likely most prominently affecting transposable elements. However, gene ontology analyses were performed on the basis of differentially targeted annotated genes. Interestingly, similarities to the gene ontology analyses based on differentially expressed genes by Szukala et al., 2022a were found. GO-terms related to trichome development, -morphogenesis, or -differentiation are also prominent within the smRNA data. Especially by comparing the smRNA data of population pair 3 of the transplanted samples, one can see that growing the montane ecotype in the alpine environment results in an increase of smRNA targeting of trichome-related genes (i.e., less transcription/translation). Growing the montane ecotype in the montane environment results in a decrease of smRNA targeting of trichome-related genes (i.e., more transcription/translation). This could indicate that trichrome development might not only be

regulated on the transcriptomic, but also on the post-transcriptomic level. Other biological processes can be observed to have different expression level for both smRNAs and genes. Examples are GO-terms like response to wounding or response to biotic stimuli which were mainly exhibited by comparing the two different ecotypes. This suggests that these traits may also be regulated on a post-transcriptomic level.

The number of shared DTRs, DTGs, and GO-terms differ widely between the comparison groups. The comparison groups that show more DTRs are likely to also show more shared DTRs. Interestingly, these groups also show notably fewer unique GO-terms. The reason for this could be that a major portion of these unique DTRs lead to more common GO-terms, indicating once more that different smRNAs can influence similar functions (Tsai et al., 2014). Comparisons between same population pairs grown in different environments (i.e., altitudes) showed a higher number of shared DTRs. Different population pairs grown in the same environment (i.e., altitude), on the other hand, show comparatively less DTRs, Indicating that the smRNAs profiles is mostly determined by their shared ancestry (Trucchi et al., 2017) and are stable across environment. This could indicate that the stability of morphological divergence (documented by Bertel et al. 2018) partly maintains by smRNAs.

Conclusion

SmRNA activity tends – compared to gene expression – to show a more prominent response to the environmental conditions. Since epimutations seem to accrue at a higher rate, the difference regarding the smRNA profiles increases notably over evolutionary time, resulting in different levels of differentiation between independently diverged populations. This makes smRNAs a potential candidate to be one of the important forces for ecological divergence. The montane ecotype showed a higher smRNA plasticity, which not only further confirms the key role of smRNAs during adaptational processes but also underpins the involvement of smRNAs during recurrent evolution. SmRNA signals regarding trichome development and response to biotic stress were detected, which implies that parts of the ongoing ecological divergence might be driven by post-transcriptional mechanisms. Further analysis should focus on smRNAs targeting shared regions between samples of different genotypes treated under the same conditions. This would help to identify smRNAs uniquely involved in specific adaptational processes.

References

Alexa, A., & Rahnenfuhrer, J. (2016). topGO: Enrichment analysis for Gene Ontology. R package version 2.28. 0. Cranio.

Archer, D., & Rahmstorf, S. (2010). The climate crisis: An introductory guide to climate change. Cambridge University Press.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. cell, 136(2), 215-233.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Bertel, C., Buchner, O., Schönswetter, P., Frajman, B., & Neuner, G. (2016). Environmentally induced and (epi-) genetically based physiological trait differentiation between Heliosperma pusillum and its polytopically evolved ecologically divergent descendent, H. veselskyi (Caryophyllaceae: Sileneae). Botanical Journal of the Linnean Society, 182(3), 658-669.

Bertel, C., Rešetnik, I., Frajman, B., Erschbamer, B., Hülber, K., & Schönswetter, P. (2018). Natural selection drives parallel divergence in the mountain plant Heliosperma pusillum sl. Oikos, 127(9), 1355-1367.

Bonnet, E., Van de Peer, Y., & Rouzé, P. (2006). The small RNA world of plants. New Phytologist, 171(3), 451-468.

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. Nature reviews Molecular cell biology, 16(12), 727-741.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy s. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. Gigascience, 10(2), giab008.

Dar, F. A., Mushtaq, N. U., Saleem, S., Rehman, R. U., Dar, T. U. H., & Hakeem, K. R. (2022). Role of epigenetics in modulating phenotypic plasticity against abiotic stresses in plants. International Journal of Genomics, 2022.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15-21.

Felix Krueger. Babraham bioinformatics—Trim Galore! (2012). http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ (Accessed 10 February 2022).

Frajman, B., & Oxelman, B. (2007). Reticulate phylogenetics and phytogeographical structure of Heliosperma (Sileneae, Caryophyllaceae) inferred from chloroplast and nuclear DNA sequences. Molecular Phylogenetics and Evolution, 43(1), 140-155.

Frakes, L. A., Francis, J. E., & Syktus, J. I. (1992). Climate modes of the Phanerozoic (p. 286).

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 32(18), 2847-2849.

He, L., & Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nature reviews genetics, 5(7), 522-531.

He, X., Wu, F., Qi, H., Meng, J., Wang, W., Liu, M., Li, L. & Zhang, G. (2022). Whole-genome resequencing reveals the single nucleotide polymorphisms associated with shell shape in Crassostrea gigas. Aquaculture, 547, 737502.

Hegi, G. (1981). Illustrierte Flora von Mitteleuropa: Pteridophyta, Spermatophyta, Band III Angiospermae Dicotyledones 1. Verlag Paul Parey.

Hewezi, T. (2020). Epigenetic mechanisms in nematode–plant interactions. Annual Review of Phytopathology, 58, 119-138.

Huber, B. T., MacLeod, K. G., Watkins, D. K., & Coffin, M. F. (2018). The rise and fall of the Cretaceous Hot Greenhouse climate. Global and Planetary Change, 167, 1-23.

Jake Conway, N. G. (2019). A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets.ISO 690

Jiang, C., Mithani, A., Belfield, E. J., Mott, R., Hurst, L. D., & Harberd, N. P. (2014). Environmentally responsive genome-wide accumulation of de novo Arabidopsis thaliana mutations and epimutations. Genome research, 24(11), 1821-1829.

Jover-Gil, S., Paz-Ares, J., Micol, J. L., & Ponce, M. R. (2014). Multi-gene silencing in Arabidopsis: a collection of artificial micro RNA s targeting groups of paralogs encoding transcription factors. The Plant Journal, 80(1), 149-160.

Kolde, R. (2019). pheatmap: Pretty Heatmaps. R package version 1.0. 12.

Laganà, A., Veneziano, D., Russo, F., Pulvirenti, A., Giugno, R., Croce, C. M., & Ferro, A. (2015). Computational design of artificial RNA molecules for gene regulation. In RNA Bioinformatics (pp. 393-412). Humana Press, New York, NY.

Lawrence, M., Gentleman, R., & Carey, V. (2009). rtracklayer: an R package for interfacing with genome browsers. Bioinformatics, 25(14), 1841-1842.

Li, A., Li, G., Zhao, Y., Meng, Z., Zhao, M., Li, C., Zhang, Y., Li, P., Ma, C., Xia H., Zhao, S., Hou, L., Zhao, P., & Wang, X. (2018). Combined small RNA and gene expression analysis revealed roles of miRNAs in maize response to rice black-streaked dwarf virus infection. Scientific reports, 8(1), 1-14.

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(16), 2078-2079

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general pur- pose program for assigning sequence reads to genomic features. Bioinformatics, 30(7):923-30.

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic acids research, 47(8), e47-e47

Maeso, I., Roy, S. W., & Irimia, M. (2012). Widespread recurrent evolution of genomic features. Genome biology and evolution, 4(4), 486-500.

McCarty, J. P., (2001). Ecological consequences of recent climate change. Conservation biology, 15(2), 320-331.

Moss, E. G., (2001). RNA interference: it's a small RNA world. Current Biology, 11(19), R772-R77

Nosil, P., Crespi, B. J., & Sandoval, C. P., (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature, 417(6887), 440-443.

Pielou, E. C., (2008). After the ice age. University of Chicago Press.

Prentice, I. C., Harrison, S. P., & Bartlein, P. J. (2011). Global vegetation and terrestrial carbon cycle changes after the last ice age. New Phytologist, 189(4), 988-998.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26(6), 841-842.

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic acids research, 44(W1), W160-W165.

Robinson, M. D., McCarthy, D.J., & Smyth, G. K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology, 11(3), 1-9.

Robinson, S. W., Afzal, A. M., & Leader, D. P. (2014). Bioinformatics: Concepts, methods, and data. Handbook of Pharmacogenomics and stratified medicine, 259-287.

Schär, C., Davies, T. D., Frei, C., Wanner, H., Widmann, M., Wild, M., & Davies, H. C. (1998). Current alpine climate. Views from the Alps: Regional perspectives on climate change, 21, 63.

Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson G. G., Owen-Hughes, Z., Blaxter, M., & Barton, G. J. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?. Rna, 22(6), 839-851.

Shuka, L., Malo, S., & Tan, K. (2011). New chorological data and floristic notes for Albania. Botanica Serbica, 35(2), 157-162.

Silva, W., Otto, S., & Immler, S. (2021). Evolution of plasticity in production and transgenerational inheritance of small RNAs under dynamic environmental conditions. PLOS Genetics, 17, e1009581.

Stevanović, V., Vladimirov, V., Niketić, M., Vukojičić, S., Jakovljević, K., Lubarda, B., & Tomović, G. (2014). Plant species and subspecies discovered by Dr. Josif Pančić 1-distribution and floristic importance. Botanica Serbica, 38(2), 251-268.

Steiner, C. C., Römpler, H., Boettger, L. M., Schöneberg, T., & Hoekstra, H. E. (2009). The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. Molecular Biology and Evolution, 26(1), 35-45.

Strid, A., & Tan, K. (1997). Flora hellenica. Koeltz Scientific Books.

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). "REVIGO summarizes and visualizes long lists of Gene Ontology terms" PLoS ONE 2011. doi:10.1371/journal.pone.0021800

Szukala, A. (2022). Parallel and adaptive processes during ecotype formation in "Heliosperma pusillum", Dissertation at the University of Vienna. 10.25365/thesis.71429

Szukala, A., Bertel, C., Frajman, B., Schönswetter, P., & Paun, O. (2022b). Parallel adaptation to lower altitudes is associated with enhanced plasticity in Heliosperma pusillum (Caryophyllaceae). bioRxiv, 2022-05.

Szukala, A., Lovegrove-Walsh, J., Luqman, H., Fior, S., Wolfe, T. M., Frajman, B., Schönwetter, P., & Paun, O. (2022a). Polygenic routes lead to parallel altitudinal adaptation in Heliosperma pusillum (Caryophyllaceae). Molecular Ecology.

Team, R. C. (2013). R: A language and environment for statistical computing.

Thornton, M. R. (2022). Ecological and evolutionary consequences of smRNA activity after allopolyploidization in marsh orchids (Dactylorhiza majalis s.l.), Master's thesis at the University of Vienna. 10.25365/thesis.71563

Thuiller, W. (2007). Climate change and the ecologist. Nature, 448(7153), 550-552.

Trucchi, E., Frajman, B., Haverkamp, T. H., Schönswetter, P., & Paun, O. (2017). Genomic analyses suggest parallel ecological divergence in Heliosperma pusillum (Caryophyllaceae). New Phytologist, 216(1), 267-278.

Tsai, C. -H., Liao, R., Chou, B., Palumbo, M., & Contreras, L. (2014). Genome-wide analyses in bacteria show small-RNA enrichment for long and conserved intergenic regions. Journal of Bacteriology, 197(1).

Van Belleghem, S. M., Vangestel, C., De Wolf, K., De Corte, Z., Möst, M., Rastas, P., De Meester, L., & Hendrickx, F. (2018). Evolution at two time frames: polymorphisms from an ancient singular divergence event fuel contemporary parallel evolution. PLoS genetics, 14(11), e1007796.

Wu, H., & Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell, 156(1-2), 45-68.

Yao, N., Schmitz, R. J., & Johannes, F. (2021). Epimutations define a fast-ticking molecular clock in plants. Trends in Genetics, 37(8), 699-710.

Yuan, B., Latek, R., Hossbach, M., Tuschl, T., & Lewitter, F. (2004). siRNA Selection Server: an automated siRNA oligonucleotide prediction server. Nucleic acids research, 32(suppl_2), W130-W134.

Appendix

Code

```
###TRIMMING
FILES=files
for f in $FILES
do
#adapter
    trim_galore --phred33 --quality 0 --illumina --length 1 --dont_gzip --fastqc
#length
    trim_galore --quality 0 --consider_already_trimmed 100 --length 20 --
    max_length 20 --dont_gzip --output_dir 20 --fastqc $f
#quality
    trim_galore --consider_already_trimmed 100 --phred33 --quality 20 --fastqc --
    length 16 $f
done


###MAPPING
STAR --runMode genomeGenerate --genomeDir genome/ --genomeFastaFiles
genome/helio.genome.fasta --runThreadN 24 --genomeChrBinNbits 15 --sjdbGTFfile
genome/helio.genome.gff --sjdbGTFtagExonParentTranscript Parent --
sjdbGTFfeatureExon CDS --sjdbOverhang 1

for file in *fq; do echo $file; STAR --genomeDir genome/ --readFilesIn $file --
outSAMtype BAM Unsorted --alignIntronMax 1 --outFilterMismatchNoverLmax 0.05 --
outFileNamePrefix ${file/.fq/_mapped} --runThreadN 24; done


###SORTING and INDEXING
for file in *.bam; do samtools sort -o ${file/_mappedAligned.out.bam/sorted.bam}
$file; done
for file in *sorted.bam; do samtools index $file; done
for file in *sorted.bam; do bamCoverage -b $file -o $file.bedgraph -of bedgraph -
-normalizeUsing None -bs 100 --numberOfProcessors max/2; done
#fitering by 10 (time targeted)
for file in *bedgraph; do awk '$4>=10' $file > $file.filtered10x; done
#sorting by scaffold and position
for file in *.filtered10x; do sortBed -i $file > $file.sorted.bedgraph; done
#concatenating and sorting
cat *_all_trimmed_finalsorted.bam.bedgraph.filtered10x.sorted.bedgraph | sortBed
-i - > all_sorted_all_Samples.bedgraph

bedtools merge -i all_sorted_all_Samples.bedgraph > all_merged.bedgraph


###FEATURECOUNTS
featureCounts -a all_reference_peaks.edit.gff -o all_counts.txt -t
sequence_feature -g ID -M --fraction -T 8 *all_trimmed_finalsorted.bam


###SAMPLE PREPARATION
##commongarden
groupGLM<-factor(c(rep("A1",3),rep("A3",2),rep("A4",3),rep("A5",3),
                   rep("M1",4),rep("M3",3),rep("M4",3),rep("M5",2)))

#keep rows with a mean higher then 0.5
y=x[apply(cpm(x),1,function(x){!(mean(x)<0.5)}),]
y<-DGEList(counts=y,group=groupGLM) #generating a DGEList object
ModelDesign=model.matrix(~0+groupGLM)
#generating contrast groups
my.contrasts=makeContrasts(A1vsM1=groupGLMM1-groupGLMA1,
                           A3vsM3=groupGLMM3-groupGLMA3,
                           A4vsM4=groupGLMM4-groupGLMA4,
                           A5vsM5=groupGLMM5-groupGLMA5,
                           levels=ModelDesign)

##transplanted
groupGLM<-factor(c(rep("A1_D",5),rep("A1_U",5),rep("A3_D",5),rep("A3_U",5),
                   rep("M1_D",5),rep("M1_U",5),rep("M3_D",5),rep("M3_U",2)))
```

```
#keep rows with a mean higher then 0.5
y=x[apply(cpm(x),1,function(x){!(mean(x)<0.5)}),]
y<-DGEList(counts=y,group=groupGLM) #generating a DGEList object
ModelDesign=model.matrix(~0+groupGLM)
#generating contrast groups
my.contrasts=makeContrasts(AvsM_1D=groupGLMM1_D-groupGLMA1_D,
                           AvsM_3D=groupGLMM3_D-groupGLMA3_D,
                           AvsM_1U=groupGLMM1_U-groupGLMA1_U,
                           AvsM_3U=groupGLMM3_U-groupGLMA3_U,
                           MUvsMD1=groupGLMM1_U-groupGLMM1_D,
                           MUvsMD3=groupGLMM3_U-groupGLMM3_D,
                           ADvsAU1=groupGLMA1_D-groupGLMA1_U,
                           ADvsAU3=groupGLMA3_D-groupGLMA3_U,
                           levels=ModelDesign)

###DTRs
run_DTR=function(dat,con){
  dat<-calcNormFactors(dat,method="TMM")
  DGE=estimateDisp(dat,design=ModelDesign,robust=T)
  GLM=glmQLFit(DGE,design=ModelDesign)
  lrt<-glmQLFTest(GLM,contrast=my.contrasts[,con])
  DTR=topTags(lrt,p.value=0.05,adjust.method="BH",n=dim(dat)[1])
  return(DTR)
}

###PCA-PLOT
run_PCA=function(dat,main){
  uq=betweenLaneNormalization(as.matrix(dat),which="full")
  set=newSeqExpressionSet(uq)
  plotPCA(set,col=colors,
          k=2,cex=1,pch=geoNum,cex.axis=1,
          main=paste("PCA - ", main),labels=FALSE,
          xlim=c(-0.4,0.4),ylim=c(-
0.4,0.4),cex.main=1.3,cex.axis=0.7,cex.lab=0.7)
  return(set)
}

###MA-PLOT
run_MA=function(dat,con,main){
  dat<-calcNormFactors(dat,method="TMM")
  DGE=estimateDisp(dat,design=ModelDesign,robust=T)
  GLM=glmQLFit(DGE,design=ModelDesign)
  lrt<-glmQLFTest(GLM,contrast=my.contrasts[,con])
  DTR=topTags(lrt,p.value=0.05,adjust.method="BH",n=dim(dat)[1])

  nUp<-length(rownames(DTR$table[DTR$table$logFC>0,]))
  nDown<-length(rownames(DTR$table[DTR$table$logFC<0,]))
  plotSmear(lrt,de.tags=rownames(DTR$table),
            xlab="Average logCPM ",ylab="logFC",xlim=c(0,12),ylim=c(-10,10),
            main=paste("MA-Plot - ", main,"\nDTRs up:", nUp, "down:",
nDown),cex.sub=0.7)
}

###FROM DTRs TO DTGs -> done in the cluster
#extracting genes and adding 1000 bp in front and behind
awk '$3 == "gene"' *.DTG |\
awk -F'\t' -vOFS='\t' '{ $3="DTG" }1' |\
awk -F'\t' -vOFS='\t' '{$4-=1000}1' |\
awk -F'\t' -vOFS='\t' '{$5+=1000}1' |\
awk -F'\t' -vOFS='\t' '$4+0<0{$4=1}1' > helio.DTG.gff

#adding DTG to input file
cat helio.DTG.gff 03_DTG/helio.genome.gff > new_helio.genome.gff

#one big file containing regions of intrest
awk '$3 == "DTG"' *.gff > helio.DTG.gff
awk '$3 == "gene"' *.gff > helio.gene.gff
awk '$3 == "mRNA"' *.gff > helio.mRNA.gff
awk '$3 == "exon"' *.gff > helio.exon.gff
awk '$3 == "CDS"' *.gff > helio.CDS.gff
awk '$3 == "three_prime_UTR"' *.gff > helio.three_prime_UTR.gff
awk '$3 == "five_prime_UTR"' *.gff > helio.five_prime_UTR.gff

#DTR to different anotated regions
module load bedtools/2.29.2
```

```
for file in *.bed; do bedtools intersect -wa -a helio.DTR.gff -b $file >
${file/_ID.bed/_DTG}; done

###GO-TERMS
run_topgo<-function(DTG){
  DTG_all_ID<-data.frame(data.frame(str_split_fixed(DTG$V9,";",3))$X1)
  names(DTG_all_ID)[1]<-'genes'
  DTG_all_ID$genes<-gsub("ID=","",as.character(DTG_all_ID$genes))
  gene_list_all<-DTG_all_ID
  geneSel_all<-(gene_list_all$genes)
  geneSel_all<-factor(as.integer(names(geneID2GO) %in% geneSel_all))
  names(geneSel_all)<-names(geneID2GO)

  sampleGOdata_all<-new("topGOdata",ontology="BP",
                        allGenes=geneSel_all,nodeSize=10,
                        annot=annFUN.gene2GO,gene2GO=geneID2GO)

  allGO=usedGO(sampleGOdata_all)
  resultFisher_all<-
runTest(sampleGOdata_all,algorithm="classic",statistic="fisher")
  resultKS_all<-runTest(sampleGOdata_all,algorithm="classic",statistic="ks")
  allRes_all<-GenTable(sampleGOdata_all,classicFisher=resultFisher_all,
                       classicKS=resultKS_all,orderBy="classicFisher",

ranksOf="classicFisher",topNodes=length(allGO),numChar=1000)

  topGo_final<-allRes_all[allRes_all$classicFisher<0.05,]
  return(topGo_final)
}

###UPSET-PLOT
Lt1=list(DTR_Group1=DTR_Group1
DTR=make_comb_mat(lt1)
Lt2=list(DTG_Group1=DTG_Group1
DTG=make_comb_mat(lt2)
lt3=list(GO_Group1=GO_Group1
GO=make_comb_mat(lt3)

top_ha=HeatmapAnnotation(
"DTRs_up"=anno_barplot(comb_size(DTR),gp=gpar(fill="black"),height=unit(2,"cm"),a
dd_numbers=TRUE),
"DTGs_up"=anno_barplot(comb_size(DTG),gp=gpar(fill="black"),height=unit(2,"cm"),a
dd_numbers=TRUE),
"GOs_up"=anno_barplot(comb_size(GO),gp=gpar(fill="black"),height=unit(2,"cm")
,add_numbers=TRUE),
gap=unit(2,"mm"),annotation_name_side="left",annotation_name_rot=0)

UpSet(DTR,set_order=c("Group1","Group2","Group3","Group4"),
comb_order=order(comb_size(DTR)),
bg_col=used_colour_code,
bg_pt_col=used_colour_code
top_annotation=top_ha,
right_annotation=upset_right_annotation(DTR,add_numbers=TRUE))

###HEATMAP
run_heatmap=function(dat,file){
  heatmap_input_matrix<-dat
  heatmap_input_matrix$GO_ID<-
paste(heatmap_input_matrix$GO_ID,heatmap_input_matrix$GO_term, sep="
  heatmap_input_matrix$GO_term<-NULL
  heatmap_input_matrix_raw<-heatmap_input_matrix[,-1]
  rownames(heatmap_input_matrix_raw)<-heatmap_input_matrix[,1]
  input <-t(scale(t(heatmap_input_matrix_raw),center=FALSE, scale = FALSE))

  pheatmap(input, kmeans_k = NA, breaks = NA, scale = "none", cluster_rows =
FALSE,
           cluster_cols = FALSE,
           show_rownames = TRUE, show_colnames = TRUE,
           filename="file",
           fontsize = 9, cellwidth = 20,cellheight = 15,
           colorRampPalette(c("white","purple4"))( 48 ))
  return(input)
}
```

Supplementary tables

Supp. Table 1 Information about sample grouping and details of the used abbreviation for samples of the common garden and the transplanted samples.

Abbreviation	Details	Abbreviation	Details
Common garden		**Transplanted**	
A1D	Ecotype: A = Alpine / M = Montane	**A**1_DA	Ecotype: A = Alpine / M = Montane
A**1**D	Location	A**1**_DA	Location
A1**D**	Individual sample	A1_**D**A	Transplantation location: U = Up / D = Down
		A1_D**A**	Individual sample

Supp. Table 2 Sample ID and the respective number of sequences before (Raw) and after trimming for all length classes together (20 - 24 nt) and separated for samples of the common garden and the transplanted samples.

ID	Raw	20-24_nt	20_nt	21_nt	22_nt	23_nt	24_nt
Common garden							
Mean	**20,941,102**	**13,222,420**	**1,145,319**	**1,434,851**	**2,124,471**	**3,762,228**	**4,755,552**
A1A	23,812,201	16,767,954	1,770,881	1,480,730	2,423,355	4,030,981	7,062,007
A1C	24,904,489	17,593,223	1,219,074	1,735,722	2,613,879	4,247,224	7,777,324
A1D	24,017,478	15,243,535	1,902,880	1,541,300	2,172,607	3,925,443	5,701,305
A3A	21,698,472	12,409,239	1,784,491	1,955,981	2,570,098	4,381,477	1,717,192
A3B	17,945,180	13,736,123	897,013	1,243,563	2,110,819	4,021,204	5,463,524
A4A	19,690,611	13,861,380	1,125,230	1,469,373	2,245,181	4,032,744	4,988,852
A4B	21,470,921	15,788,842	1,181,748	1,735,683	2,594,232	4,504,046	5,773,133
A4C	15,088,909	11,103,589	1,033,686	1,277,314	1,594,425	3,093,091	4,105,073
A5B	19,539,881	11,543,711	1,170,487	1,391,121	1,866,850	3,317,504	3,797,749
A5C	18,074,291	13,402,985	987,529	1,173,523	1,813,598	3,686,066	5,742,269
A5D	17,146,743	12,489,231	1,391,895	1,452,700	1,996,648	4,358,433	3,289,555
M1A	19,508,536	11,069,798	572,821	1,038,593	1,802,664	2,795,454	4,860,266
M1B	23,682,669	13,436,935	1,753,497	1,999,059	2,299,947	2,829,496	4,554,936
M1C	18,619,192	10,162,830	819,813	1,546,965	1,654,629	2,500,791	3,640,632
M1D	17,748,725	11,965,366	967,062	1,242,953	2,124,332	3,140,613	4,490,406
M3A	18,671,425	11,029,810	1,089,709	1,133,057	1,798,734	3,264,322	3,743,988
M3B	47,160,520	10,466,394	794,109	1,046,153	1,696,827	2,921,516	4,007,789
M3C	18,400,172	14,191,588	841,054	1,530,409	2,197,190	3,386,482	6,236,453
M4A	17,286,119	12,087,323	740,103	1,131,497	2,026,701	3,611,525	4,577,497
M4B	16,852,979	10,814,698	720,392	1,014,084	1,423,340	2,878,316	4,778,566
M4C	18,653,126	15,545,673	1,044,115	1,581,566	2,837,405	7,001,781	3,080,806
M5C	21,328,521	14,763,171	1,313,424	1,726,963	2,576,519	4,757,562	4,388,703
M5E	20,344,196	14,642,272	1,221,331	1,553,263	2,422,855	3,845,163	5,599,660
Transplanted							
Mean	**10,816,488**	**6,769,362**	**702,278**	**1,063,674**	**1,105,178**	**1,530,770**	**2,367,462**
A1_DA	7,616,736	3,418,835	147,379	377,377	501,815	980,076	1,412,188
A1_DB	9,666,001	2,037,160	86,557	196,377	292,396	590,946	870,884
A1_DD	7,118,891	3,496,596	209,888	470,292	496,976	834,662	1,484,778
A1_DE	20,722,277	9,882,656	2,105,110	1,713,591	1,589,385	2,071,635	2,402,935
A1_DF	8,468,386	6,120,493	521,866	948,076	1,036,841	1,431,631	2,182,079
A1_UA	11,732,014	9,453,229	654,271	1,258,705	1,608,010	2,098,121	3,834,122
A1_UB	17,577,945	11,589,001	1,119,519	1,884,859	2,081,097	2,433,723	4,069,803
A1_UC	6,360,498	1,884,736	359,524	539,258	290,288	292,391	403,275
A1_UD	14,400,118	11,270,087	1,013,704	1,369,151	1,640,328	2,507,147	4,739,757
A1_UE	4,357,274	2,434,092	294,449	320,484	387,941	510,076	921,142

A3_DB	12,027,306	8,031,505	1,069,774	1,524,704	1,211,672	1,705,490	2,519,865
A3_DC	5,132,869	2,823,583	240,999	599,979	474,051	578,498	930,056
A3_DD	5,445,292	3,916,807	191,532	757,719	523,682	877,307	1,566,567
A3_DG	5,134,853	3,133,174	206,579	430,516	480,944	860,878	1,154,257
A3_DJ	2,447,356	1,175,774	114,219	210,607	188,897	256,237	405,814
A3_UA	16,527,176	10,704,295	1,279,124	2,481,568	1,674,612	1,954,510	3,314,481
A3_UB	12,233,916	9,192,905	963,119	1,381,648	1,376,352	1,945,459	3,526,327
A3_UC	12,499,179	9,219,083	646,142	1,603,813	1,404,310	1,671,636	3,893,182
A3_UD	12,553,439	9,925,867	742,117	2,128,490	1,516,419	1,847,171	3,691,670
A3_UF	5,723,875	3,944,332	319,793	694,279	645,724	799,06	1,485,476
M1_DA	13,191,996	9,313,696	641,474	1,024,126	1,653,130	2,279,141	3,715,825
M1_DB	11,309,322	9,050,740	703,891	1,057,546	1,555,607	2,765,550	2,968,146
M1_DC	17,186,135	9,200,018	1,412,483	1,404,255	1,491,683	2,235,375	2,656,222
M1_DE	13,467,419	6,868,071	881,914	1,258,085	1,304,829	1,639,531	1,783,712
M1_DF	7,143,543	3,165,191	239,451	430,468	490,086	805,914	1,199,272
M1_UA	10,421,530	6,429,417	585,875	907,965	1,087,139	1,350,352	2,498,086
M1_UB	5,363,613	3,414,321	351,46	580,895	608,251	758,698	1,115,017
M1_UC	13,836,673	10,220,253	726,446	1,212,186	1,636,862	2,390,586	4,254,173
M1_UE	2,937,149	1,365,741	114,149	187,776	236,842	309,595	517,379
M1_UFII	18,614,706	12,607,291	1,487,883	1,980,483	2,111,834	2,895,906	4,131,185
M1_UFI	19,493,560	11,338,071	1,349,977	1,832,022	1,911,257	2,617,485	3,627,330
M3_DA	10,788,141	6,420,645	489,669	1,043,175	1,014,763	1,496,407	2,376,631
M3_DB	11,371,488	7,757,232	950,936	920,791	1,077,665	2,001,326	2,806,514
M3_DC	15,107,173	10,241,895	1,556,736	1,526,145	1,669,910	2,539,959	2,949,145
M3_DE	3,746,278	2,348,106	218,106	363,675	380,532	585,177	800,616
M3_DF	6,262,612	3,987,487	352,567	690,882	662,604	1,080,873	1,200,561
M3_UAII	15,055,545	7,452,406	1,026,127	1,320,581	1,417,776	1,571,410	2,116,512
M3_UAI	16,192,143	9,262,795	1,265,362	1,626,468	1,755,398	1,953,604	2,661,963
M3_UB	12,608,592	9,907,530	748,684	1,224,254	1,614,051	2,176,477	4,144,064

Supp. Table 3 Total number of differential targeted annotated regions, including DTRs (differentially targeted regions) DTG (Differentially targeted genes) mRNAs, Exons, CDSs (coding regions), 3´UTRs, 5´UTRs and the total number of UTRs for the comparison groups of the Common Garden and Transplanted samples.

	DTRs	DTGs	mRNAs	Exons	CDSs	3'UTRs	5'UTRs	UTRs (total)
Common Garden								
All								
A1vsM1	33,837	9,345	10,098	4,812	3,562	836	872	1708
A3vsM3	5,933	1,661	1,628	744	528	148	156	304
A4vsM4	29,418	8,472	8,424	3,748	2,584	842	788	1630
A5vsM5	1,101	333	364	222	178	28	38	66
Up								
A1vsM1	14,101	4,3	3,95	1,572	924	336	432	768
A3vsM3	3,188	773	674	328	224	60	84	144
A4vsM4	15,045	4,109	3,278	1,616	1,08	372	426	798
A5vsM5	395	78	56	24	14	8	4	12
Down								
A1vsM1	19,736	5,045	6,148	324	2,638	500	440	940
A3vsM3	2,745	888	954	416	304	88	72	160
A4vsM4	14,373	4,363	5,146	2,132	1,504	470	362	832
A5vsM5	706	255	308	198	164	20	34	54
Transplanted								
All								
AvsM_1D	18,782	575	63	3,152	2,564	568	404	972

AvsM_3D	14,579	4,469	4,414	1,894	1,356	464	312	776
AvsM_1U	10,086	2,848	3,258	1,642	1,346	296	184	480
AvsM_3U	17,752	5,246	7,046	3,882	3,282	538	374	912
ADvsAU1	1,395	502	780	548	504	88	44	132
ADvsAU3	387	114	118	68	52	16	12	28
MUvsMD1	3,171	1,009	1,404	704	620	82	60	142
MUvsMD3	9,902	3,231	4,732	2,78	2,426	336	206	542
Up								
AvsM_1D	8,713	2,463	278	1,256	1,016	216	164	380
AvsM_3D	7,003	2,611	2,672	1,008	670	298	168	466
AvsM_1U	5,244	1,435	1,802	994	844	148	114	262
AvsM_3U	13,614	423	6,16	3,518	3,032	454	308	762
ADvsAU1	717	281	454	358	336	64	24	88
ADvsAU3	285	82	94	54	44	10	10	20
MUvsMD1	1,549	469	702	412	374	36	40	76
MUvsMD3	9,082	2,867	4,414	2,678	2,362	292	194	486
Down								
AvsM_1D	10,069	3,287	352	1,896	1,548	352	240	592
AvsM_1U	4,842	1,413	1,456	648	502	148	70	218
AvsM_3D	7,576	1,858	1,742	886	686	166	144	310
AvsM_3U	4,138	1,016	886	364	250	84	66	150
ADvsAU1	678	221	326	190	168	24	20	44
ADvsAU3	102	32	24	14	8	6	2	8
MUvsMD1	1,622	540	702	292	246	46	20	66
MUvsMD3	820	364	318	102	64	44	12	56

Supp. Table 4 Top five GO-terms related to leaf material for the Gene regulatory regions of the Common Garden samples. GO-terms of up-regulated DTRs are displayed on the left and GO-terms of down-regulated DTRs are displayed on the left. Abbreviations for the individual comparison groups are given in bold. Each GO-term has a p-value < 0.05.

Term (up-regulated)	Term (down-regulated)
M1vsA1	
hydrogen peroxide transmembrane transport	manganese ion homeostasis
regulation of vesicle-mediated transport	response to bacterium
establishment of protein localization to organelle	myo-inositol hexakisphosphate biosynthetic process
response to arsenic-containing substance	myo-inositol hexakisphosphate metabolic process
regulation of cellular component size	response to biotic stimulus
M3vsA3	
regulation of secondary metabolic process	protein tetramerization
response to ozone	protein lipidation
mitochondrial fission	lipoprotein biosynthetic process
response to reactive oxygen species	fatty acid elongation
negative regulation of circadian rhythm	response to fungus
M4vsA4	
xylan metabolic process	trehalose metabolic process
xyloglucan biosynthetic process	stomatal opening
C4-dicarboxylate transport	cellular modified amino acid biosynthetic process
cell wall organization or biogenesis	disaccharide biosynthetic process
transmembrane transport	regulation of stomatal opening
M5vsA5	
response to xenobiotic stimulus	protein localization involved in auxin polar transport
reactive oxygen species metabolic process	polysaccharide localization

response to oxidative stress	negative regulation of small molecule metabolic process
pentose-phosphate shunt, oxidative branch	cell growth
defense response to insect	cellular polysaccharide catabolic process

Supp. Table 5 Top five GO-terms related to leaf material for the Gene regulatory regions of the Transplanted samples. GO-terms of up-regulated DTRs are displayed on the left and GO-terms of down-regulated DTRs are displayed on the left. Abbreviations for the individual comparison groups are given in bold. Each GO-term has a p-value < 0.05.

Term (up-regulated)	Term (down-regulated)
MvsA_1D	
cellular response to reactive nitrogen species	response to extracellular stimulus
response to arsenic-containing substance	homeostatic process
indole-containing compound metabolic process	cellular response to starvation
malate metabolic process	cellular response to nitric oxide
indole glucosinolate metabolic process	cellular response to nutrient levels
MvsA_3D	
hormone transport	response to herbivore
response to wounding	cellular process
negative regulation of multicellular organismal process	detoxification of cadmium ion
nitrate transport	monocarboxylic acid metabolic process
basipetal auxin transport	organic acid metabolic process
MvsA_1U	
protein insertion into mitochondrial membrane	cellular response to nitric oxide
fatty acid beta-oxidation	carboxylic acid biosynthetic process
mitochondrial membrane organization	protein-containing complex localization
establishment of protein localization to mitochondrial membrane	endonucleolytic cleavage of tricistronic rRNA transcript
negative regulation of secondary cell wall biogenesis	organic acid biosynthetic process
MvsA_3U	
cell development	DNA demethylation
biological_process	DNA dealkylation
COPII-coated vesicle cargo loading	demethylation
retrograde transport, endosome to Golgi	post-embryonic root morphogenesis
cytosolic transport	plant organ development
ADvsAU1	
histone H3-K4 trimethylation	diterpenoid biosynthetic process
response to nutrient levels	acetyl-CoA biosynthetic process
peptidyl-lysine trimethylation	diterpenoid metabolic process
cellular response to nitrogen starvation	monocarboxylic acid biosynthetic process
response to fungus	thioester biosynthetic process
ADvsAU3	
response to toxic substance	ribosome biogenesis
cadmium ion homeostasis	rRNA processing
metal ion homeostasis	rRNA metabolic process
positive regulation of response to biotic stimulus	ribonucleoprotein complex biogenesis
cation homeostasis	ribosomal large subunit biogenesis
MUvsMD1	
meiotic mismatch repair involved in recombination	sucrose transport
protein retention in ER lumen	cellular response to nitrate

maintenance of protein localization in endoplasmic reticulum	disaccharide transport
regulation of cellular protein metabolic process	oligosaccharide transport
chiasma assembly	positive regulation of response to salt stress
MUvsMD3	
organelle organization	regulation of phosphorus metabolic process
chromosome organization	regulation of phosphate metabolic process
amino acid activation	regulation of cellular protein metabolic process
tRNA aminoacylation	negative regulation of mitotic nuclear division
tRNA aminoacylation for protein translation	establishment of tissue polarity

Supp. Table 6 Total number, number of unique (and unique in percentage) of up- or down-regulated differentially targeted regions (DTRs), differentially targeted genes (DTGs) and gene ontology (GO) terms for the comparison groups of the common garden data and the data of the transplanted samples. The mean of each group is given in bold.

	DTR			**DTG**			**GO-terms**		
	Total	**Unique**	**Unique(%)**	**Total**	**Unique**	**Unique(%)**	**Total**	**Unique**	**Unique(%)**
Up									
Mean	8,182	6,608	80	1,849	1,222	66	115	98	85
A1vsM1	14,101	11,701	83	3,463	2,46	71	166	139	83
A3vsM3	3,188	2,081	65	699	303	43	83	68	81
A4vsM4	15,045	12,311	81	3,154	2,086	66	145	124	85
A5vsM5	395	338	85	78	39	50	67	61	91
Mean	8,644	626	72	2,377	1,078	45	165	121	73
AvsM_1D	8,713	6,072	69	2,158	1,071	49	153	102	66
AvsM_3D	7,003	4,861	69	2,308	1,266	54	201	143	71
AvsM_1U	5,244	2,805	53	1,342	513	38	113	78	69
AvsM_3U	13,614	11,302	83	3,698	1,463	39	191	160	83
Mean	2,908	2673	91	858	760	88	174	145	83
ADvsAU1	717	713	99	258	224	86	173	154	89
ADvsAU3	285	183	64	80	70	87	88	81	92
MUvsMD1	1,549	1132	73	462	302	65	186	139	74
MUvsMD3	9,082	8665	95	2,632	2,442	92	250	206	82
Down									
Mean	939	7,755	82	222	1,391	62	166	125	75
A1vsM1	19,736	17,088	86	4,214	2,934	69	149	118	79
A3vsM3	2,745	1,749	63	837	236	28	159	118	74
A4vsM4	14,373	11,604	80	358	2,269	63	220	165	75
A5vsM5	706	580	82	247	125	50	134	97	72
Mean	6,656	4,392	66	1,664	919	55	149	110	73
AvsM_1D	10,069	7,366	73	2,758	1,781	64	173	126	72
AvsM_3D	7,576	5,258	69	1,651	916	55	185	137	74
AvsM_1U	4842	2,636	54	1,307	558	42	165	128	77
AvsM_3U	4,138	2,306	55	938	422	45	71	48	67
Mean	806	787	97	284	262	92	110	103	93
ADvsAU1	678	671	99	220	206	93	189	181	95
ADvsAU3	102	95	93	32	31	96	42	42	100
MUvsMD1	1,622	1,592	98	527	487	92	141	130	92
MUvsMD3	820	790	96	355	322	90	69	58	84

Declaration

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Ausführungen, die aus fremden Quellen sinngemäß oder wörtlich entnommen wurden, versichere ich kenntlich gemacht zu haben. Ich habe mich bemüht, sämtliche Inhaber*innen der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

I hereby declare that I have written this master's thesis independently and have not used any sources or aids other than those indicated. I declare that I have marked all statements taken verbatim or analogously from external sources. I have made every effort to locate all owners of the image rights and have obtained their permission to use the images in this work. Should a copyright infringement nevertheless become known, I request that you notify me.

Vienna, March 14, 2023

Place, Date

Signature