universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Gene regulation by small RNAs at different altitudes in *Heliosperma pusillum* (Caryophyllaceae)

verfasst von / submitted by

Arezoo Fani

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

Master of Science (MSc)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 832
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Botanik
Betreut von / Supervisor:	Assoz. Prof. Dipl.-Ing. Dr. Ovidiu Paun

Preface

From the day that I, as a high schooler, decided to present the “plants” chapter of our biology book until now is quite some time. That was my first journey of professionally collecting different plants in the Alborz Mountain range around our home, as I already had this habit and interest of collecting flowers since I was small. Years passed and now here I am writing my master’s thesis about an alpine plant called *Heliosperma pusillum.*

It took me almost four years to finish my master’s degree. But it was not like that I drag my master study for no good reason, it was interesting for me that master students can study as long as they want here and participate in as many courses as they like to. I have been always so curious to know about different branches of a study and had as much experiences as possible, so I took this situation as an opportunity to explore, to attended different courses including additional ones and more importantly to feel that I am truly becoming a botanist.

Acknowledgments

I would like to express my appreciation to my supervisor Dr. Ovidiu Paun.

I would like to acknowledge the help provided by PhD student Aglaia Szukala.

I would also like to acknowledge Karen Koelzer for proofreading of „Zusammenfassung“.

I express my very profound gratitude to my parents and my brothers, Omidreza and Alireza, for their unconditional support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

Contents

Abstract

As a response to climate change, plants may need adaption to new ecological conditions. The plant species *Heliosperma pusillum* (Caryophyllaceae) forms two stable ecotypes in the Southeastern Alps as a result of adaptation to different elevations (alpine vs. montane belt). The two ecotypes show different phenotypic/ morphological characters, most importantly, glabrous vs. 'hairy' plants, also different leaf anatomical traits. Since the divergence happened between ancestral alpine populations and derived mountain populations that adapted to hotter and drier habitats, this system could offer a useful model to analyze alpine plants' responses to climate change. The aim of this project was to investigate the causes of this phenotypic and ecological divergence. More specifically, divergent regulations of gene expressions in alpine and montane populations of *Heliosperma pusillum* was investigated by analyzing of small RNAs from plants cultivated in a common garden. Based on this work, most importantly the differences in the light regimes received by alpine and montane ecotypes in their habitats caused the light related genes to be regulated (targeted) differently predominantly by 20-22 nt small RNAs (miRNA and tasiRNA) between these two ecotypes. Moreover, the number of overlapped differentially regulated genomic regions by tasiRNA and miRNA (20-22 nt) was higher in comparison to hetsiRNA (24 nt) among localities.

Introduction

As the environment is constantly changing its abiotic and/or biotic conditions, organisms must continuously adapt to the environmental conditions they are under. Particularly in mountains where habitats are diverse and there are various microclimates (Scherrer and Körner 2011, as cited in Bertel et al. 2016c); this may lead to formation of ecotypes (Bertel et al. 2016c). Ecotypes belong to the same species but show different phenotypic traits due to the local adaption to a specific microclimate (Hufford and Mazer 2003, as cited in Bertel et al. 2016c, Lowry 2012). The genetic variation between ecotypes, resulted from environmental contrast of landscape, is not accidental but at least in part well organized (Lowry 2012). As plants fitness depends on many biotic and abiotic factors at their growing sites, ecotypes are not static populations, but they can be even in the lane of speciation (Lowry 2012).

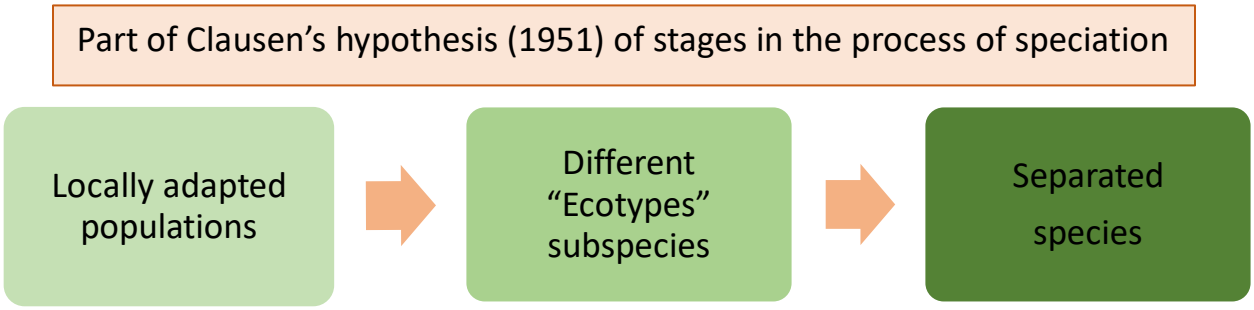


Figure. 1*: Stages of the speciation process. Adapted from Lowry 2012 © 2012 The Linnean Society of London*

Heliosperma pusillum

Alpine and montane ecotypes of *Helisperma pusillum* (in some references the montane ecotype of *H. pusillum* is referred to as *H. veselskyi*) grow at different elevations, in the alpine (1700–2300 m) vs. montane belt/ zone (500–1300 m) [Figure 2] (Bertel et al. 2016c, 2018). Where the alpine ecotype grows on moist screes in open areas, the montane ecotype grows underneath overhanging rocks and around shallow caves where it is shady with low irradiance (i.e., the montane ecotype does not experience direct sunlight during all the seasons), and water availability is limited in its habitat (Bertel et al. 2016a, 2016c). Additionally, daily temperature has greater fluctuation for the alpine ecotype, whereas it has a moderate amplitude for the montane ecotype (Bertel et al. 2016a). These two ecotypes even show different microbial communities (Trucchi et al. 2017).

The main morphological difference between these ecotypes is that the leaves of the alpine ecotype are glabrous, whereas the montane ecotype's leaves are slightly broader and covered with a thick layer of multicellular glandular hairs (Bertel et al. 2016c). Low water availability and/ or increased herbivory pressure, which is the case for the montane ecotype's natural growing site, can be a cause for more trichomes to form (Bertel et al. 2016c, 2018). The characteristic leaf anatomy differences are stable, independent if they grow in their natural growing sites or in a common garden, suggesting these traits are not representing phenotypic plasticity and they are heritably adapted to the ecological conditions of their habitats (Bertel et al. 2016c). As the alpine ecotype adapted to high irradiance and increased water availability, it has thicker palisade parenchyma and consequently thicker leaves (Bertel et al. 2016c). Palisade parenchyma which locates underneath upper epidermis in leaves contains most of chloroplasts and its higher thickness lead plants to have higher photosynthesis (Bertel et al. 2016c). In turn, the montane ecotype's leaf anatomy (in leaf thickness and stomatal pattern) reflects its adaption to the shady, dry habitat thus likely result in lower photosynthesis rate in the montane ecotype (Bertel et al. 2016c). Moreover, the montane ecotype's cell wall has more elasticity which may be a sign of adaption to its dry habitat (Bertel et al. 2016a). These morphological and phenotypic differences between the alpine and montane ecotypes are not due to phenotypic plasticity, as they show differences even when they grow for multiple generations under common garden conditions (Bertel et al. 2018).

There are no intrinsic barriers for gene flow between *Heliosperma pusillum* ecotypes (Bertel et al. 2016b), but hybrids are rare between alpine and montane ecotypes due to spatial barrier which separate their populations over altitude and cause isolation (Trucchi et al., 2017). In addition to spatial barrier, since the ecological conditions of ecotypes' habitats are distinct and the plants are well adapted to their habitats' conditions, the possible hybrid plants will likely confront a decreased fitness in the parental growing areas (Bertel et al. 2016b). Accumulation of differences between ecotypes resulting from both heritable and non-heritable traits may cause speciation over a period of time (Pfennig et al. 2010, Bonduriansky et al. 2012, Coyne and Orr 2004, as cited in Bertel et al. 2016c) and this makes *H. pusillum* an interesting model.

The alpine ecotype of *Heliosperma pusillum* is widespread in the central and southern mountains of Europe, but the montane ecotype is restricted to a few populations which are scattered in southeastern Alps and northernmost Balkan Peninsula (Bertel et al. 2016c). It is suggested that the montane ecotype of *H. pusillum* has diverged from its alpine relative after the last glaciation (Bertel et al. 2016c) several times independently (Trucchi et al. 2017). Similar selection pressures during independent divergence events may lead to parallel evolution (Barrett and Schluter 2008, as cited in Bertel et al. 2018) and it is suggested the independent evolution of montane ecotypes in different population which are under same environmental condition is due to natural selection (Bertel et al. 2018). There are several causes for a (possible) parallel adaption: various molecular adjustments which result in same phenotypes or different changes on the same gene /or molecular pathway (Stern 2013 as cited in Trucchi et al. 2017). It is also showed that differentiation among alpine populations is lower and suggested it is due to alpine ecotype's continual distribution, whereas montane populations indicates more differentiation as a result of their disjoint habitat (Poldini 2002, Wilhalm et al. 2014, Trucchi et al. 2017).

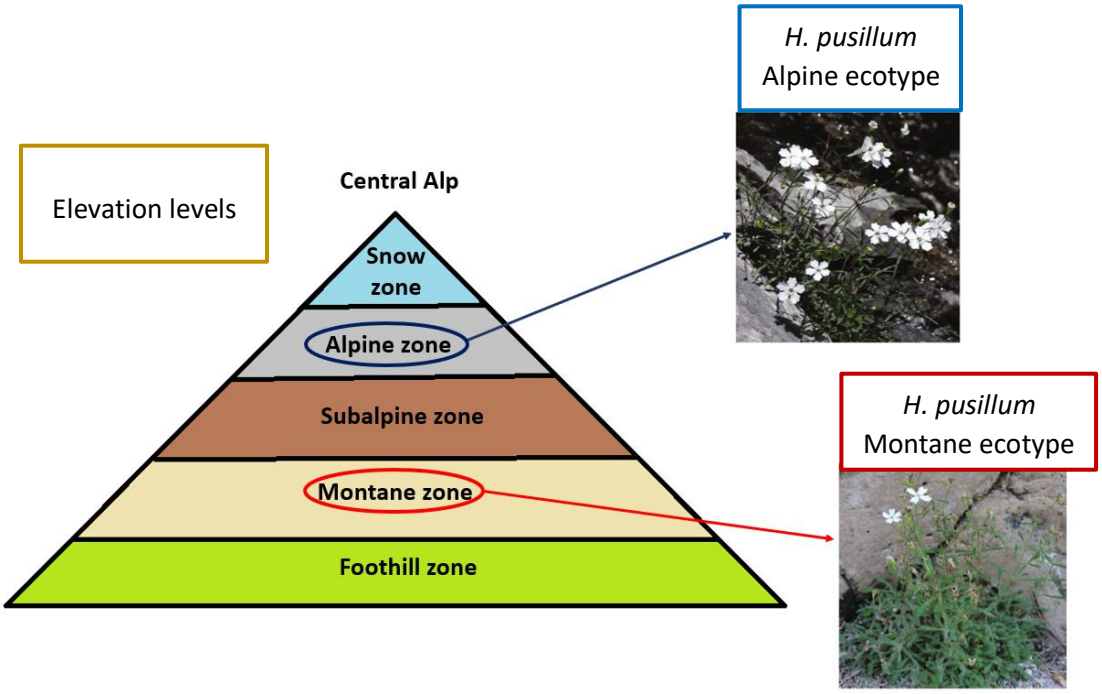


Figure 2. *Altitudinal zonation in central Alps and typical plants of the alpine and montane ecotypes of Helliosperma pusillum. Photos from Bertel et al. 2016a © 2016 The Linnean Society of London.*

Small RNAs in plants

Plants small RNAs are mostly 21 to 24 nucleotide long and have a role in shaping phenotypic differences through different processes like methylation of DNA, degradation of mRNAs and prevention of translation (Borges and Martienssen 2015).

There are different known classes of small RNAs including microRNAs (miRNAs), heterochromatic small interfering RNAs (hetsiRNAs), trans-acting siRNAs (tasiRNAs), hairpin derived siRNAs (hp-siRNAs), natural antisense siRNAs (natsiRNAs), phased siRNAs (phasiRNAs) and epigenetically activated siRNAs (easiRNAs) (Borges and Martienssen 2015).

While some small RNA groups including miRNAs are processed from single-strand hairpin-like structures, others like hetsiRNAs and tasiRNAs are derived from double-strand RNA precursors (Borges and Martienssen 2015). In general, dicer-like proteins are needed for small RNAs to be processed (Borges and Martienssen 2015). Another group of proteins associated with small RNAs is Argonaute proteins which bind to small RNAs and have an important role as a mediator for further recognition of small RNAs targets (Borges and Martienssen 2015). Before loading of small RNAs onto Argonaute proteins small RNAs are duplex and having or not having possible mismatches between their two strands cause them to load onto their specific Argonaute proteins and aim different targets (Borges and Martienssen 2015).

Although miRNAs mainly play a role in post-transcriptional gene silencing by degradation of mRNAs and prevention of translation, siRNAs such as tasiRNAs and hetsiRNAs are mostly involved in transcriptional gene silencing and RNA-directed DNA methylation (Borges and Martienssen 2015). Due to their fast evolution, miRNA genes are usually specific to a family or genus (Borges and Martienssen 2015). HetsiRNAs appear to be in general the most abundant small RNAs type (Borges and Martienssen 2015).

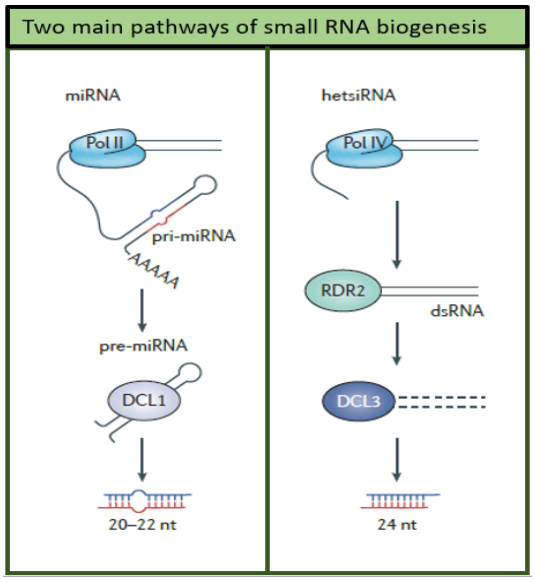


Figure 3. *Two main pathways of small RNA biogenesis in plants. Adapted from Borges and Martienssen 2015 © 2015 Macmillan Publishers Ltd*

Small RNAs 5'-end assists loading of small RNAs into their specific Argonaute proteins and establish their activities (Borges and Martienssen 2015). Additionally, modifications at 3'-end of small RNAs in plants is important as they influence not only the function and abundance of small RNAs, but also their target identification and the stability of small RNAs (Borges and Martienssen 2015). It is suggested that small RNAs modification happens in some tissues and cells that are under growth and leads to small RNAs diversity (Borges and Martienssen 2015). Some siRNAs are even produced through modification in miRNAs and overall miRNAs activate formation of some siRNAs (tasiRNAs/ easiRNAs/ phasiRNAs) from mRNAs, ncRNAs and transposable elements (Borges and Martienssen 2015).

Small RNAs have an essential role in cell reprogramming, during gametogenesis, through DNA methylation (Borges and Martienssen 2015). In addition, small RNAs have a role in paramutation, a procedure which involves DNA methylation and causes epigenetic changes that can be passed to next generations (Borges and Martienssen 2015).

Small RNAs also contribute to maintenance of some crucial aspects like DNA damage repair, male fertility (by affecting pollen mother cells), chromosome segregation and some beneficial traits such as salt tolerance (Borges and Martienssen 2015).

- The main aim of this MSc thesis was to characterize small RNA classes (miRNA/ tasiRNA/ hetsiRNA) of the two ecotypes of *Heliosperma pusillum* when grown in a common garden. Further we aimed to identify differentially expressed small RNAs that may in turn affect differential gene expression (DE) and/or posttranscriptional regulation of genes, in particular those involved in phenotypic differences between the alpine and montane ecotypes.

Materials and methods

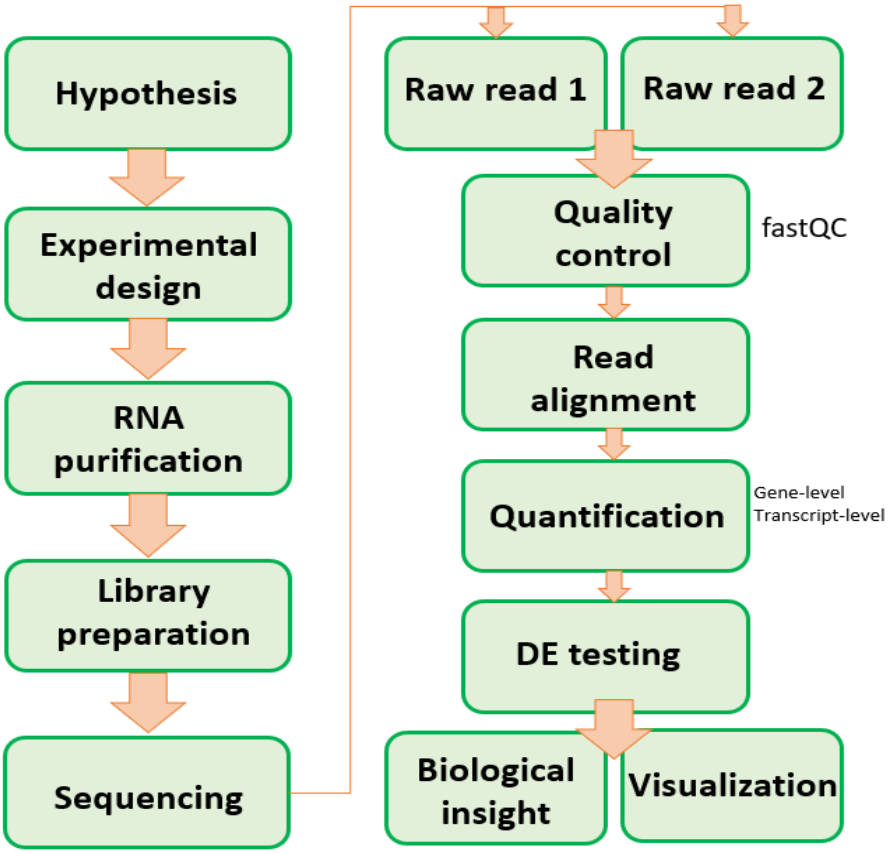


Figure 4. *General workflow for (small) RNA-sequence analysis. Diagram modified from http://databeauty.com/blog/tutorial/2016/09/13/RNA-seq-analysis.html (Wang 2019)*

Plant material

Plant material for the project was fixed from plants cultivated in a common garden in Innsbruck, Austria. Seeds of both ecotypes were collected from natural sites and cultivated under the same environmental conditions in a common garden. The first generation has been analyzed. The natural sites, labeled as locality 1, 3, 4 and 5 in this study, are the localities as described in Trucchi et al. (2017) and labeled there as A, C, D and E respectively. Previous analyses showed gene flow between montane populations of locality 3 and 4, therefore we pulled together the individuals from locality 3 and 4 in the analyses, as representing a single origin.

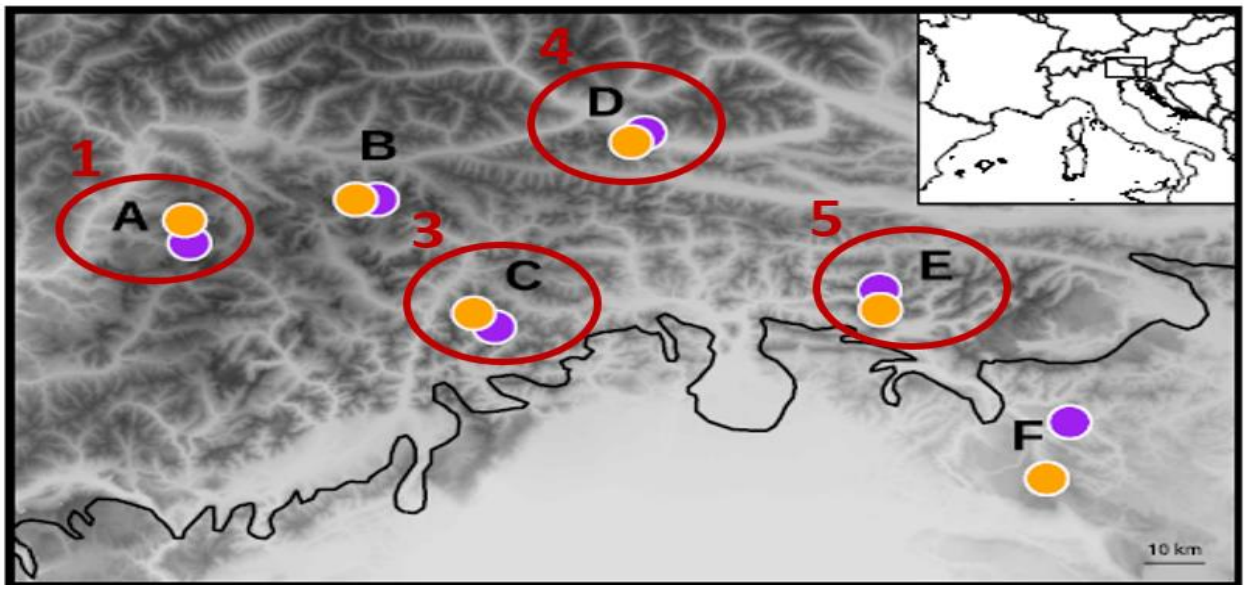


Figure 5. *Map of natural sites in southeastern Alps, where seeds of ecotype pairs of Heliosperma pusillum had been sampled (alpine ecotype in orange and montane ecotype in purple). Adapted from Trucchi et al. 2017.*

Initial workflow

The work started on already sequenced small RNA Illumina reads (small RNA extraction and library preparation for Illumina sequencing had been carried out by other members of the lab). Overall, small RNA sequencing was carried out for 12 alpine and 12 montane individuals (i.e. 3 biological replicates per ecotype and locality, for a total of 24 individuals).

TAG	DETAILS	TAG	DETAILS
P1A	Individual A, locality 1, alpine ecotype	V1A	Individual A, locality 1, montane ecotype
P1C	Individual C, locality 1, alpine ecotype	V1B	Individual B, locality 1, montane ecotype
P1D	Individual D, locality 1, alpine ecotype	V1C	Individual C, locality 1, montane ecotype
P3A	Individual A, locality 3, alpine ecotype	V1D	Individual D, locality 1, montane ecotype
P3B	Individual B, locality 3, alpine ecotype	V3A	Individual A, locality 3, montane ecotype
P3C	Individual C, locality 3, alpine ecotype	V3B	Individual B, locality 3, montane ecotype
P4A	Individual A, locality 4, alpine ecotype	V3C	Individual C, locality 3, montane ecotype
P4B	Individual B, locality 4, alpine ecotype	V4A	Individual A, locality 4, montane ecotype
P4C	Individual C, locality 4, alpine ecotype	V4B	Individual B, locality 4, montane ecotype
P5B	Individual B, locality 5, alpine ecotype	V4C	Individual C, locality 4, montane ecotype
P5C	Individual C, locality 5, alpine ecotype	V5C	Individual C, locality 5, montane ecotype
P5D	Individual D, locality 5, alpine ecotype	V5E	Individual E, locality 5, montane ecotype

Table 1. *List of alpine and montane individuals used in this study*

Analysis of small RNAs reads with CLC Genomic Workbench

The first step of the analyses included working with the program CLC Genomic Workbench Version 8.0 (available from https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/), which initially was used to trim off adaptors from the sequences (by other members of the lab). After extracting of small RNAs, counting, quality control and classification of small RNAs according to their size were carried out. MiRNAs, hetsiRNAs and tasiRNAs were identified based on their size (20-22 nt for microRNA and tasiRNA and 24 nt for hetsiRNA). Annotated miRNAs had been already separated from the other 20-22 nt small RNAs, by blasting against the miRBase Sequence Database, Release 22.1 (Kozomara and Griffiths-Jones 2010).

Figure 6. *Examples of original small RNA reads and their quality visualization in the program CLC Genomic Workbench Version 8.0*

Mapping of reads to the genome and transcriptome with STAR

Subsequently, the reads were mapped to a provided reference genome and, separately, transcriptome of *Heliosperma pusillum* using STAR Version 2.6 (Dobin et al. 2012). STAR aligned reads to the references and produced files with information regarding uniquely mapped reads, mismatches, multi-mapping reads and unmapped reads. The setting was as follow:

```
STAR --genomeDir ../reference --readFilesIn Heliosperma.fastq.gz --
readFilesCommand zcat --outSAMtype BAM SortedByCoordinate --
outFileNamePrefix mapping/Heliosperma.Starmap --outFilterMismatchNoverLmax
0.09 --outFilterMismatchNmax 2 --runThreadN 6
```

Maximum number of mismatches allowed (`--outFilterMismatchNmax`) was 2 and the ratio of mismatches to mapped length (`--outFilterMismatchNoverLmax`) was set for less /or equal to 0.09.

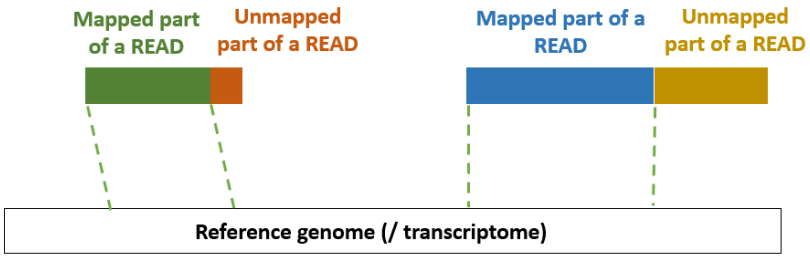


Figure 7. *Alignment of the reads to a reference (genome/ transcriptome) by STAR.*

```
                    Number of input reads |
                Average input read length |
                                UNIQUE READS:
            Uniquely mapped reads number |
                 Uniquely mapped reads % |
                   Average mapped length |
                Number of splices: Total |
     Number of splices: Annotated (sjdb) |
                Number of splices: GT/AG |
                Number of splices: GC/AG |
                Number of splices: AT/AC |
        Number of splices: Non-canonical |
               Mismatch rate per base, % |
                  Deletion rate per base |
                 Deletion average length |
                 Insertion rate per base |
                Insertion average length |
                      MULTI-MAPPING READS:
 Number of reads mapped to multiple loci |
      % of reads mapped to multiple loci |
 Number of reads mapped to too many loci |
      % of reads mapped to too many loci |
                           UNMAPPED READS:
% of reads unmapped: too many mismatches |
          % of reads unmapped: too short |
              % of reads unmapped: other |
                           CHIMERIC READS:
                Number of chimeric reads |
                     % of chimeric reads |
```

Figure 8. *Report produced by STAR*

Counting of reads with CORSET and featureCounts

After mapping, counting of the reads using both transcriptome and genome as reference was performed by using two different programs: CORSET Version 1.05 (Davidson and Oshlack 2014) which builds clusters using multi-mapping reads, and featureCounts (from subread-1.6.3) (Liao, Smyth, and Shi 2013) which counts exclusively uniquely mapping reads per genome feature (in our case, genic and intergenic regions) using an annotation file in GFF format, which was provided. When counting reads using featureCounts we allowed for counting multi-mappers, scaling the counts by the number of mapping positions. The CORSET and featureCounts (http://gensoft.pasteur.fr/docs/subread/1.5.3/SubreadUsersGuide.pdf) settings were as follow:

```
corset -g 1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2,2 -n
P1A,P1C,P1D,P3A,P3B,P3C,P4A,P4B,P4C,P5B,P5C,P5D,V1A,V1B,V1C,V1D,V3A,V3
B,V3C,V4A,V4B,V4C,V5C,V5E -m 10 -f true *.bam
```

In counting by CORSET, transcripts which had fewer than 10 reads aligning (across all individuals) were filtered out.

```
featureCounts -t gene -g ID -M --fraction -T 4 -
a ../genome.annotation.gff -o smallRNAcounts.txt *.bam
```

Counting of reads	
CORSET	**featureCounts**
Reads were clustered into genes and the counts per gene were calculated	Using an annotation file in GFF format, reads per genome feature (genic and intergenic regions) were counted

Figure 9. *Counting of reads by using CORSET and featureCounts*

Normalization (filter) of results

The tables of counts outputted by both CORSET Version 1.05 and featureCounts (from subread-1.6.3) were transformed using counts per million (cpm) and filtered (cpm > 0.5 in at least 3 samples) in Rstudio Version 1.1.463 (available from https://www.rstudio.com/). The setting used was:

```
cpm <- apply(data,2, function(x) (x/sum(x))*1000000)
keep=rowSums(cpm(data)>0.5)>=3
dataFilter =data[keep,]
write.table(dataFilter,    file="../Heliosperma.counts.txt",    quote=F,
sep="\t")
```

PCA plots making

To visualize the clusters which are based on the connection of individuals within populations/ ecotypes and genetic distance among populations/ ecotypes, PCA (Principal Component Analysis) plots based on the filtered tables of counts were made in Rstudio Version 1.1.463. PCA is a meaningful way to flatten the data and capture the main directions where most of the variation belongs to, in order to identify individuals with similar patterns. The functions were as follow:

```
##Read the raw table of counts, output of corset
data    =    read.table("../Heliospermacounts.matrix",    header=TRUE,
row.names=1)

##Specify colors for biological replicates
BioReplicates = c(rep("green",3),rep("red",3),… )
##Make factor for ecotypes
Ecotype.fac <- as.factor(c(rep("A",12),rep("M",12)))

##RemoveUnwantedVariation
makeRUVset = function (dat){
  plotRLE(as.matrix(dat),outline=FALSE,ylim=c(-4,4),col=BioReplicates,
main = "Unwanted Variance not removed")
  plotPCA(as.matrix(dat), col=BioReplicates, cex=1.2)

##between-lane normalization for sequencing depth and possibly other
distributional differences between lanes


 uq = betweenLaneNormalization(as.matrix(dat), which = "full")
```

```
  set = newSeqExpressionSet(uq)
  plotRLE(set,outline=FALSE,ylim=c(-4,4),col=BioReplicates,    main    =
"Unwanted Variance removed")

 plotPCA(set,col=BioReplicates, k=2, cex = 1.7, pch = Ecotype, main =
"First two PCs'", labels = T)#pch = geoNum

  return(set)
}
makeRUVset(data)
```

Differential regulation analysis with edgeR

To detect the clusters (CORSET output) and genomic regions (featureCounts output) that are differentially targeted by small RNAs, edgeR version 3.24.3 (as implemented in Trinity) (Robinson, McCarthy, and Smyth 2009) with a false discovery rate (FDR) < 0.05 was used. The setting was:

```
$mv ./counts.txt ./counts.matrix
$/usr/local/Trinity/Analysis/DifferentialExpression/ run_DE_analysis.pl
--matrix ./counts.matrix --method edgeR --samples_file ./groups.txt
```

```
/usr/local/bin/trinityrnaseq_r2013_08_14/Analysis/DifferentialExpressi
on/analyze_diff_expr.pl --matrix ../counts.matrix -P 0.05 -C 1.5
```

Making Venn diagrams of DR (differentially regulated) regions

On the next step, based on edgeR output (FDR < 0.05), Venn diagrams of differentially regulated genomic regions by hetsiRNA, tasiRNA and miRNA were made to identify the number of overlapped regions among localities. After that number of the regions which were expected to overlap by chance only was calculated to see if the overlap exceeds stochastic expectations or not. The chance calculation was by using R package SuperExactTest version 1.0.7.1 (Wang, Zhao, and Zhang 2015) and the function was as follow:

```
install.packages("SuperExactTest")


library("SuperExactTest")

setwd("path/to/directory/")


list1 <- read.table("loc1_DElist.txt", header = FALSE, row.names =
NULL)
list34 <- read.table("loc34_DElist.txt", header = FALSE, row.names =
NULL)
list5 <- read.table("loc5_DElist.txt", header = FALSE, row.names =
NULL)

input  <- c(list1, list34, list5)
names(input)=c("Loc1","Loc34","Loc5")
str(input)
(length.input=sapply(input, length))
total=#total number of regions present in the table of count
(num.expcted.overlap=total*do.call(prod,as.list(length.input/total)))
```

Also, to find the overlapped regions among localities which were regulated in the same direction, Venn diagrams based on the number of regions which were targeted (downregulated) by small RNA classes in alpine ecotype and montane ecotype were made separately.

Visualization of small RNA targeting patterns with IGV

IGV version 2.3 (Robinson et al. 2011) was used to visualize the small RNAs and their target regions in the reference genome and transcriptome. Provided RNAseq data that originated from the same samples as the small RNA data, has been used to check the context of gene expression around the potential targets.

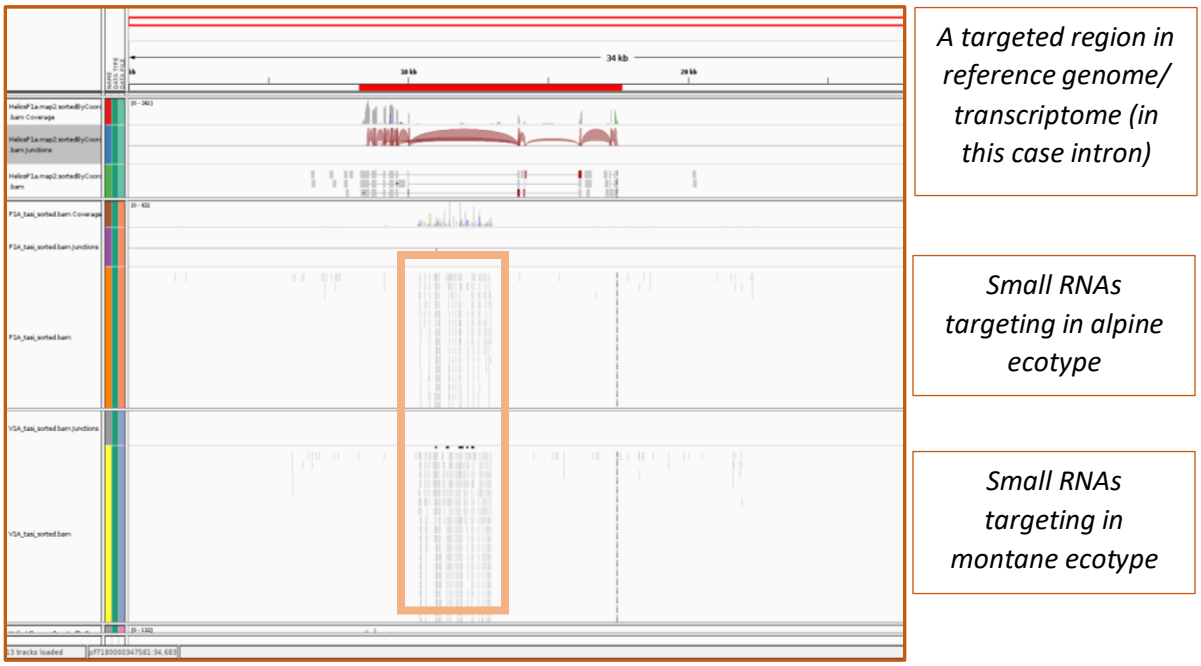


Figure 10. *IGV visualization of an example region including small RNAs and their target patterns in alpine and montane ecotypes*

Functional annotation with Blast2GO

Functional enrichment was performed using Blast2GO Version 5 (Gotz et al. 2008), which was used to annotate the genic regions which were targeted at the same direction at least between two localities by 20 to 22 nt small RNAs (tasiRNA and miRNA) and carry out Fisher's test for enrichment (filter value: FDR 0.1). Blast2GO provides Gene Ontology (GO) annotation where information regarding genes and gene products including biological processes, molecular functions and cellular components is gathered and the relationships among GO terms were visualized.

Summarizing GO terms with REVIGO

Finally, GO terms were then summarized based on FDR value and following settings in REVIGO (Supek et al. 2011) and scatter plots were made.

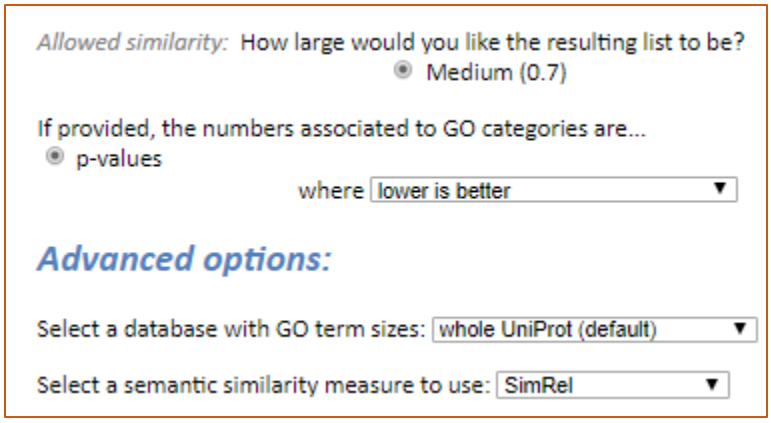


Figure 11. *REVIGO settings*

Results

The percentage of uniquely vs. multi mapped small RNA reads in *Heliosperma pusillum* using the genome and, respectively, the transcriptome as references (Table 2) shows except for the category of annotated miRNA (20-22 nt), more uniquely small RNA reads mapped to the genome than transcriptome. In addition, hetsiRNA had more uniquely mapped reads than multi mapped reads in genome mapping result but miRNA and tasiRNA were vice versa.

H. pusillum		**hetsiRNA** (24 nt)				**annotated miRNA** (20-22 nt)				**tasiRNA** and miRNA (20-22 nt)			
		Uniquely mapped reads %		Multi mapped reads %		Uniquely mapped reads %		Multi mapped reads %		Uniquely mapped reads %		Multi mapped reads %	
		Transcriptome	Genome	Transcriptome	Genome	Transcriptome	Genome	Transcripome	Genome	Transcriptome	Genome	Transcriptome	Genome
Average	All	20.55	36.90	35.25	29.57	51.19	14.84	43.90	80.54	22.40	24.05	55.89	40.57
	Alpine	20.51	36.92	35.91	29.73	52.79	13.61	43.60	81.84	22.25	24.16	57.53	40.83
	Montane	20.60	36.87	34.59	29.41	49.58	16.08	44.20	79.24	22.55	23.94	54.25	40.30
SD	All	1.00	2.58	1.53	1.50	7.96	3.22	6.63	4.45	1.71	1.97	2.85	4.07
	Alpine	1.31	2.86	1.81	1.84	8.98	3.11	7.38	4.85	1.89	1.97	2.39	5.04
	Montane	0.52	2.25	0.72	1.04	6.40	2.83	5.76	3.56	1.50	1.97	2.29	2.76

Table 2. *Heliosperma pusillum uniquely vs multi mapped small RNA reads. The percentages are based on final results of mapping with STAR (the remaining reads were mapped to too many loci, i.e. more than 10 mapping positions, or were unmapped). In this study, the group of 24 nt small RNAs were counted as hetsiRNA, annotated miRNAs are the group of miRNAs that were annotated by CLC Genomic Workbench Version 8.0 (not all the miRNAs can be annotated by this program) and finally the group of 20-22 nt small RNAs were considered tasiRNA and miRNA together. SD stands for standard deviation.*

These PCA plots shows how individuals within each locality were clustered together based on the reads counts performed by CORSET on different small RNA classes of *Helosperma pusillum* (figure 12). In general, the ecotypes differentiate along the first axis of the PCAs, whereas the second axis tends to differentiate more among the different mountain regions.

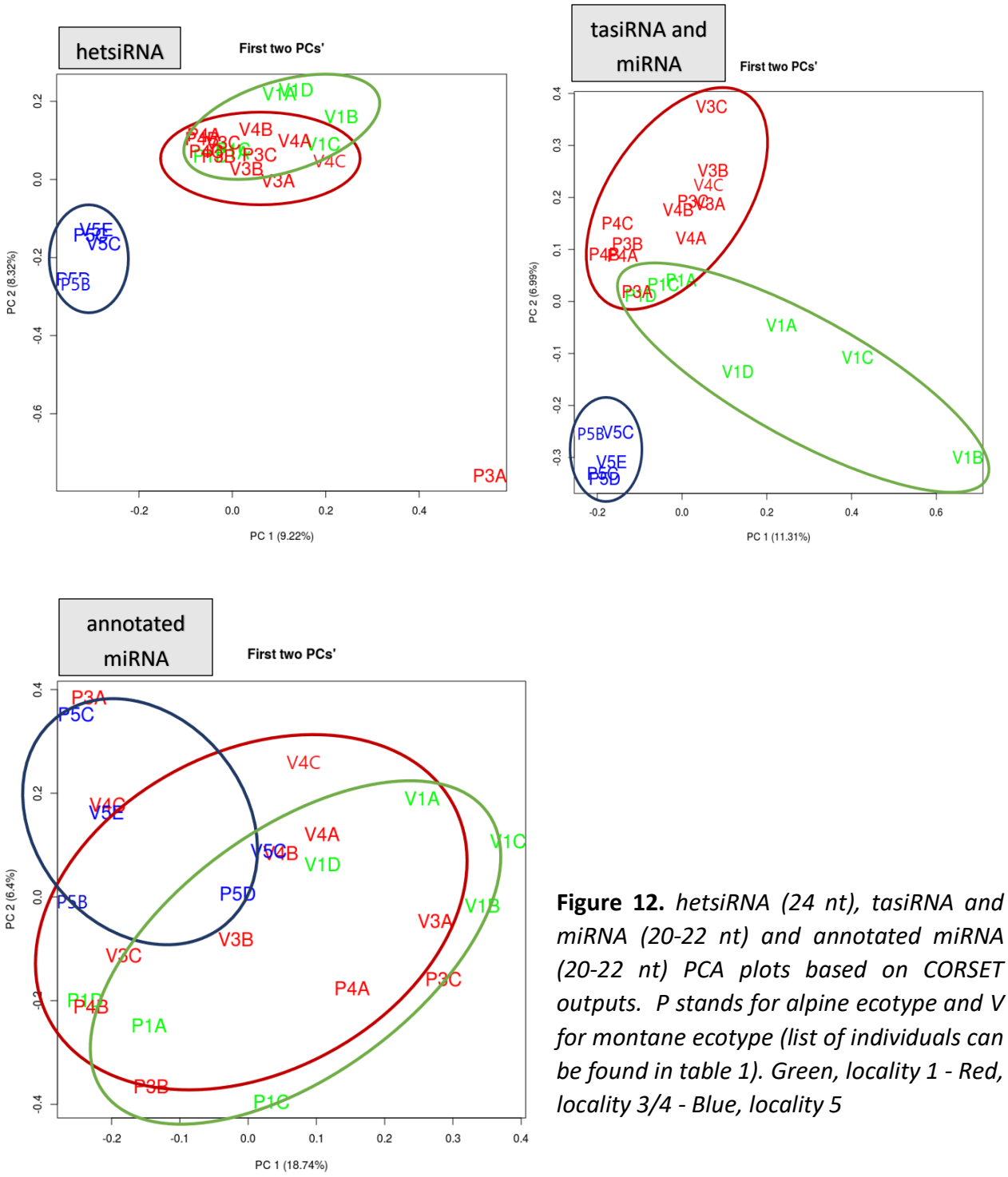


Figure 12. *hetsiRNA (24 nt), tasiRNA and miRNA (20-22 nt) and annotated miRNA (20-22 nt) PCA plots based on CORSET outputs. P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1). Green, locality 1 - Red, locality 3/4 - Blue, locality 5*

The PCA plots were also made according to the reads counts performed by featureCounts on different small RNA classes of *Helosperma pusillum* (figure 13). These PCA plots (figure 13) are clearer in comparison to PCA plots based on CORSET's output (figure 12) as the individuals within each locality cluster closer together and even the ecotypes within each locality cluster (mostly) together.

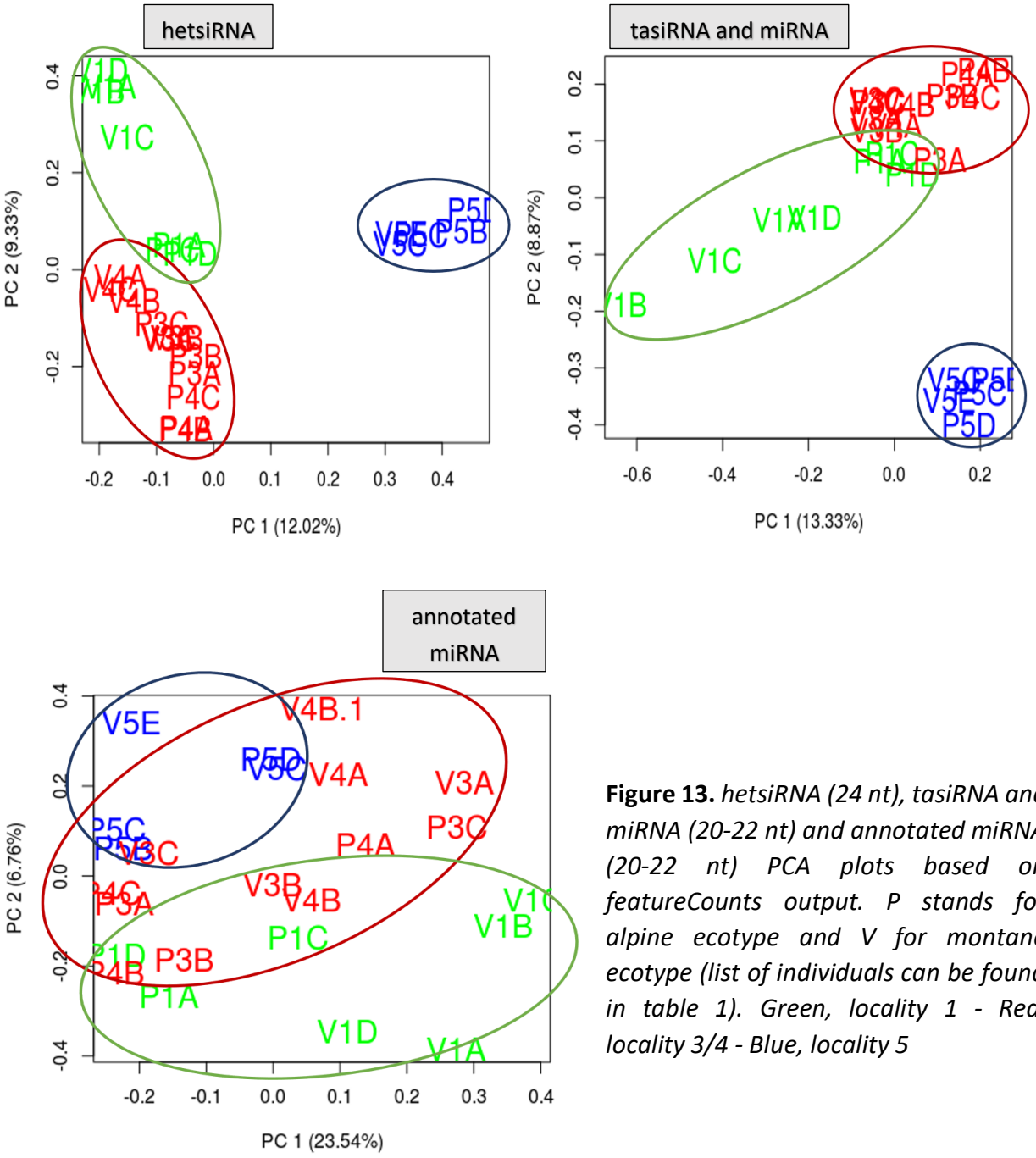


Figure 13. *hetsiRNA (24 nt), tasiRNA and miRNA (20-22 nt) and annotated miRNA (20-22 nt) PCA plots based on featureCounts output. P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1). Green, locality 1 - Red, locality 3/4 - Blue, locality 5*

The PCA plots of locality 1 indicates clearer clusters for alpine and montane ecotypes (supplementary figure S1) in comparison to locality 3/4 and 5 (supplementary figures S2 and S3).

Also, heatmaps of locality 1 shows stronger differentiation of ecotypes (supplementary figure S4) and the differentiation patterns for locality 3/4 and 5 (except for locality 5 hetsiRNA) are not as much clear (supplementary figure S5 and S6).

In order to have a better understanding of the (genes') functions that are differentially targeted by small RNAs and how different they are between *Heliosperma pusillum* ecotypes, edgeR outputs (FDR < 0.05) were investigated to find the differentially regulated genomic (genic and intergenic) regions among localities (table 3). Here alpine and montane plants show ecotype-related differential targeting by small RNAs at hundreds of genomic regions (i.e. excluding annotated miRNAs for which no DR in locality 3/4/5 was detected).

Heliosperma pusillum		GENOME								
		hetsiRNA			**tasiRNA and miRNA**			**annotated miRNA**		
No. of **differentially**		loc1	Loc3/4	loc5	loc1	Loc3/4	loc5	loc1	Loc3/4	loc5
regulated	**Montane** > alpine	**1495**	**103**	**38**	**1855**	**374**	**115**	**134**	**0**	**0**
genomic regions	**Alpine** > montane	**1368**	**175**	**15**	**975**	**226**	**31**	**12**	**0**	**0**
(cpm > 0.05) ≥ 3 FDR < 0.05	**All**	**2863**	**278**	**53**	**2830**	**600**	**146**	**146**	**0**	**0**

Table 3. *Number of differentially regulated genomic (genic + intergenic) regions for each locality based on edgeR outputs (FDR < 0.05). Table of counts were filtered by counts per million (cpm)> 0.5 in at least 3 samples. The regions that are more targeted in alpine ecotype (Alpine > Montane) and, respectively, in the montane ecotype (Montane > Alpine).*

The overlapped differentially regulated genomic regions by hetsiRNA, tasiRNA and miRNA among localities were identified (figure 14). After that also the number of regions which were expected to overlap by chance only was calculated (figure 15). In our case, overlap exceed the chance expectation.

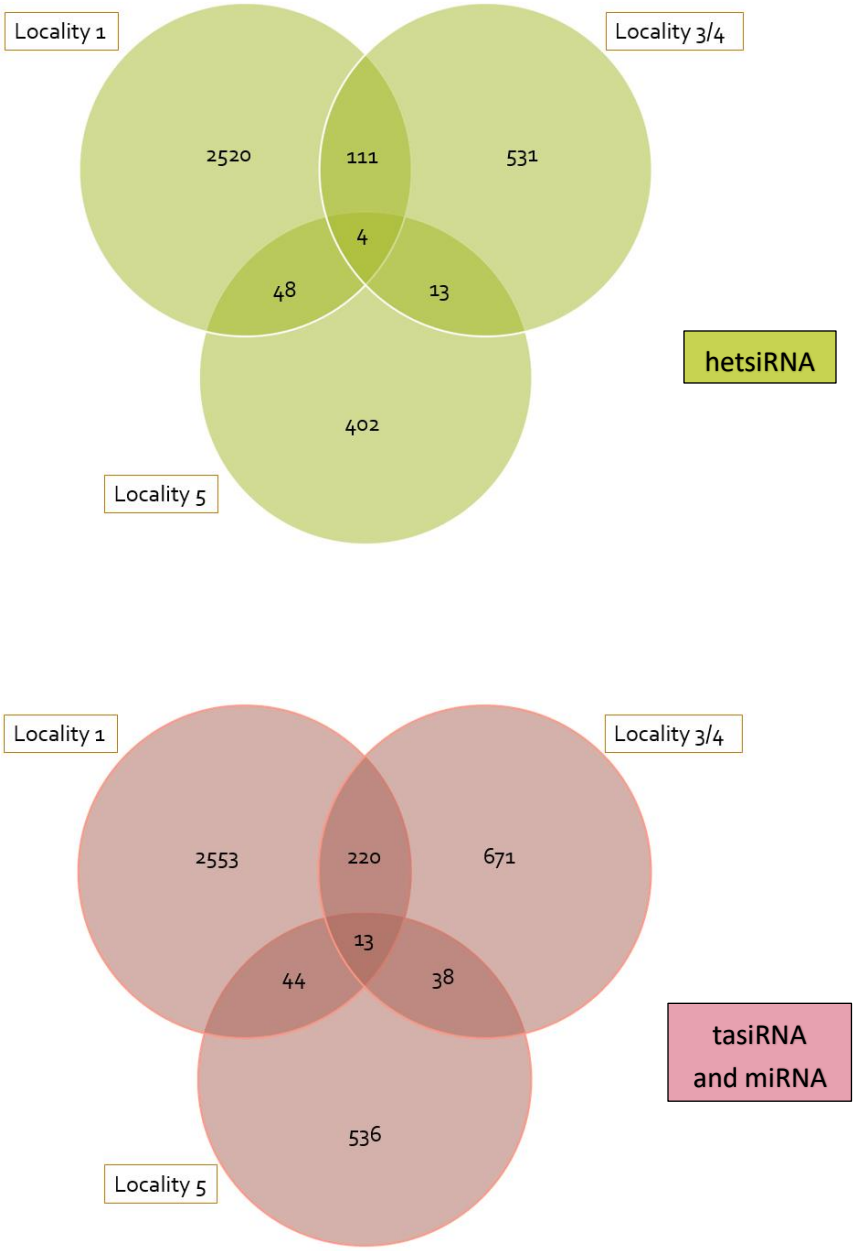


Figure 14. *Venn diagrams of differentially regulated genomic regions by hetsiRNA (24 nt) and tasiRNA and miRNA (20-22 nt) among localities, based on edgeR outputs (FDR < 0.05).*

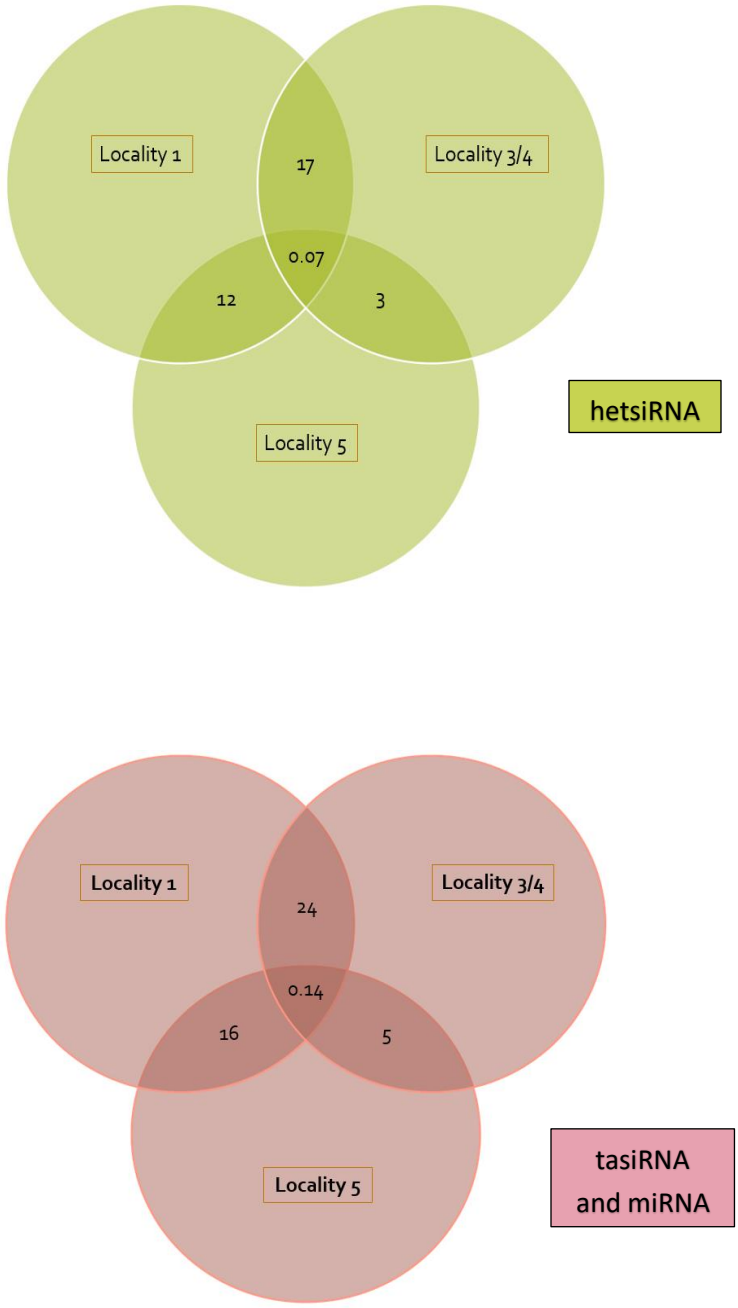


Figure 15. *Venn diagrams of differentially regulated genomic regions by hetsiRNA (24 nt) and tasiRNA and miRNA (20-22 nt) which may overlap by CHANCE only among localities.*

Previous Venn diagrams (figure 14) give us a whole perspective regarding number of overlap differentially regulated regions, which is mostly higher in tasiRNA and miRNA (20- 22 nt) in comparison to hetsiRNA (24 nt). To find the overlap regions among localities which were regulated in the same direction by different small RNA classes, the number of genomic regions which were targeted (downregulated) in alpine ecotype and respectively in montane ecotype were counted and Venn diagrams were made (figure 16).

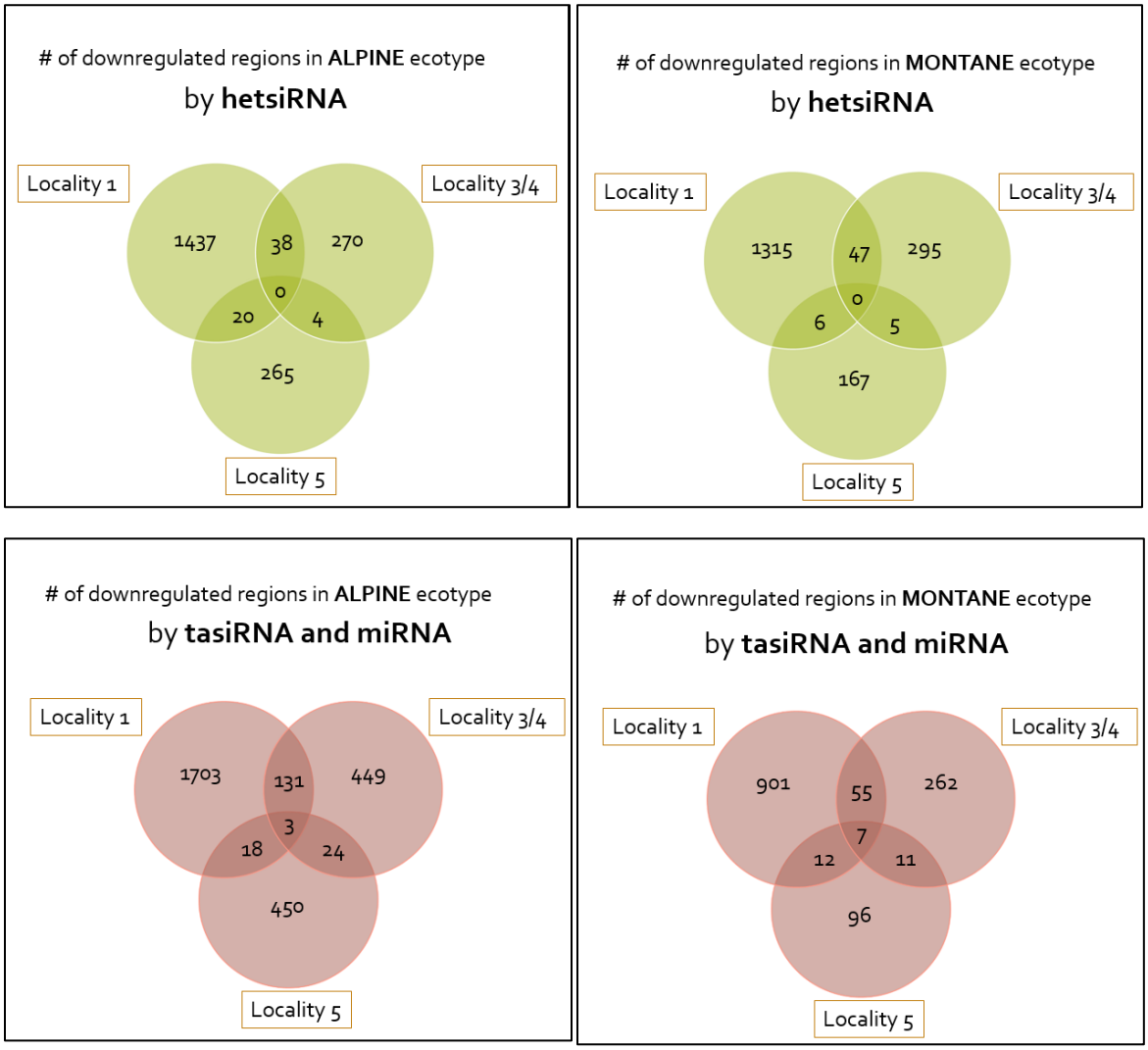


Figure 16. *Venn diagrams of targeted (downregulated) regions by hetsiRNA (24 nt) and tasiRNA and miRNA (20-22 nt) in the same direction (in alpine ecotype and montane ecotype), based on edgeR outputs (FDR < 0.05).*

In hetsiRNA case, there were no region which were regulated in the same direction among all (1,3/4 and 5) localities (figure 16). But there were ten (three in alpine ecotype + seven in montane ecotype) overlap regions which were targeted in the same direction by 20-22 nt small RNAs (tasiRNA and miRNA) (figure 16). Nine out of ten regions were intergenic and one was genic region. The targeted genic region was annotated by B2GO Version 5 as the "UDP-glycosyl transferase superfamily proteins" gene. In the next step, the targeted region(s) were visualized in IGV to see whether intron, exon, upstream or downstream of the gene(s) were targeted by small RNAs.

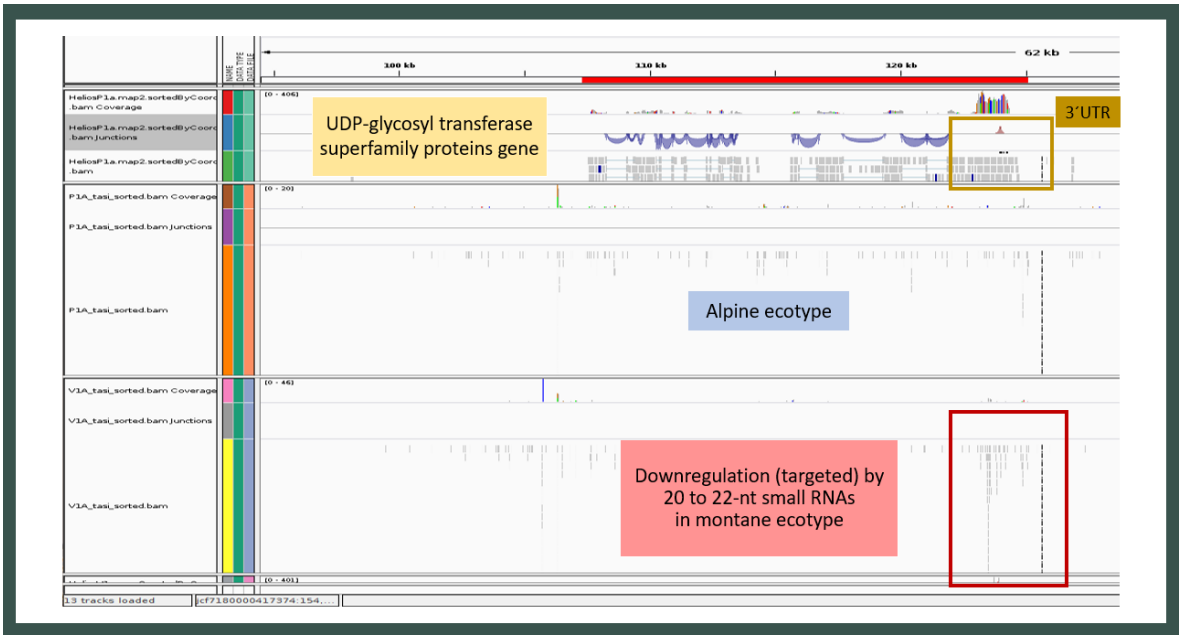


Figure 17. *Downregulation of "UDP-glycosyl transferase superfamily proteins" gene in montane ecotype of H. pusillum at downstream (3´UTR) by 20 -22 nt small RNAs (tasiRNA and miRNA). This figure is from visualization in IGV.*

After this step, Fisher's exact test (enrichment analysis) was performed on the genic regions which were targeted (downregulated) at least between two localities by 20-22 nt small RNAs (tasiRNA and miRNA) in alpine ecotype and respectively montane ecotype and summarized gene ontology scatter plots were made (supplementary figures S7 and S8). More significant and frequent GO cluster that is targeted in alpine ecotype belongs to monocarboxylic acid catabolism and after that to protein farnesylation and less significant term relates to prenylation (supplementary figure S7). In montane ecotype, all three GO clusters with similar frequency and high significance belongs to olefin (alkene) metabolism, ethylene biosynthesis and fruit ripening (supplementary figure S8).

Discussions

In our analyses the aim was to explore mapping percentages on both transcriptome and (particularly) newly available genome references as the characteristics of these two types of references are different. Where the transcriptome may include more isoforms of each gene, the genome may contain more TEs.

Except for the category of annotated miRNA (20-22 nt), more uniquely small RNA reads mapped to the genome than transcriptome, as expected since the target of many small RNAs falls outside the coding regions.

Uniquely mapped reads map to exactly one location within the reference genome (/transcriptome). Here (table 2), hetsiRNA had more uniquely mapped reads than multi mapped reads in genome mapping result but miRNA and tasiRNA were vice versa. High percentage of multi mapped reads for miRNA and tasiRNA in genome can be because they target similar UTR regions between some genes. Additionally, higher percentage of uniquely mapped reads of hetsiRNA in genome in comparison to transcriptome (table 2) may be as a result of hetsiRNA role in transcriptional silencing of TEs (Borges and Martienssen 2015) which found more in genome. Overall, the reason for high percentage of multi mapping and mapping to too many loci can be most likely because there is more than one target to which the read maps as small RNAs are normally acting in trans or because of very short sequence of small RNAs (RNA-seq tools are not designed for really short sequences and it is a challenge to confidently map short fragments with them).

Since counting step is the major step in our analysis, by using different programs it was possible to compare their performances and make different PCA plots to have a better understanding of the way individuals cluster. The PCA plots that were made using featureCounts' output was clearer (figure 12) in comparison to CORSET's output (figure 11) as the individuals within each locality cluster closer together and even the ecotypes within each locality cluster (mostly) together. As the separation of clusters was better using featureCounts, for the further steps featureCounts' outputs were used.

It was expected the same ecotypes within all the localities to behave similarly, but our study showed the ecotypes at each locality divergent independently, confirming previous results (Trucchi et al. 2017). As an example, individuals from locality 1 (supplementary figure S1)

clustered by ecotype but this separation is not as strong for locality 3/4 and 5 (supplementary figures S2 and S3). One reason for locality 1 ecotypes to separate completely can be that locality 1 is older than all the other localities. Some other reasons can be that the ecological condition of locality 3/4/5 were more similar to each other whereas locality 1 was under different condition. In addition, signs of adaptation to different locations may be weakened over time in presence of gene flow (Trucchi et al. 2017). These differences in the level of differentiation between alpine and montane ecotypes within each locality could be a consequence of the independent divergence. This pattern has been also observed in morphological and phenotypic difference between alpine and montane ecotypes even when they grow for multiple generations under common garden conditions (Bertel et al. 2018). The leaves of alpine ecotype are glabrous, whereas the montane ecotype's leaves are slightly broader and covered with a thick layer of multicellular glandular hairs (Bertel et al. 2016c). Low water availability and/ or increased herbivory pressure, which is the case for the montane ecotype's natural growing site, can be a cause for more trichomes to form (Bertel et al. 2016c, 2018). Alpine ecotype adapted to high irradiance and increased water availability and consequently has thicker palisade parenchyma, thicker leaves and higher photosynthesis (Bertel et al. 2016c). In turn, the montane ecotype's leaf anatomy (in leaf thickness and stomatal pattern) reflects its adaption to the shady, dry habitat thus likely result in lower photosynthesis rate (Bertel et al. 2016c) and the montane ecotype's cell wall has more elasticity which may be a sign of adaption to its dry habitat (Bertel et al. 2016a).

Although alpine and montane plants were cultivated under the same environmental conditions in a common garden, they show ecotype-related differential targeting by small RNAs at hundreds of genomic regions (table 3) (i.e. excluding annotated miRNAs for which no DE in locality 3/4/5 was detected). This result indicates that a significant portion of the ecotypic divergence in *H. pusillum* is not only plastic but rather stable. However, these could also be differential plastic responses to the artificial environment of the common garden. Reciprocal transplantations between the two natural habitats could offer more insights on the relative importance of plastic versus stable effects in the divergence at the level of small RNAs between the two ecotypes.

The calculation of expected overlap (by chance only) (figure 15) showed the (actual) overlap based on edgeR outputs (figure 14) exceeds stochastic expectations and therefore may represent the effect of natural selection with phenotypic consequences. Beside natural selection, introgression (gene flow due to hybridization and backcrossing) could also play a role in increasing the common patterns.

Based on annotation in Blast2GO and visualization in IGV (figure 17), downstream (3´UTR) of the gene which is responsible for "UDP-glucosyltransferase superfamily proteins" is targeted by 20-22 nt small RNAs (tasiRNA and miRNA) in montane ecotype of all localities. One of the transferase in this superfamily, plants flavonol O(3)-glucosyltransferase, is an enzyme which has the important role in anthocyanin pigment biosynthesis by transferring of glucose from UDP-glucose to a flavanol (Mitchell et al. 2018). It is suggested that anthocyanin is responsible for UV-B tolerance in leaves and this pigment protects chloroplasts against extreme light (Gould 2004). As montane ecotype of *H. pusillum* grows underneath overhanging rocks where is shady with low

irradiance and does not experience direct sunlight during all the seasons (Bertel et al. 2016c, Bertel et al. 2016a), this may cause the gene that is responsible for protection against strong light to be downregulated. The interesting fact is, although in our study the alpine and montane ecotypes grew under the same condition in a common garden, this gene was still downregulated in montane ecotype which may show that this regulation by small RNAs (and probably some other phenotypic differences between *H. pusillum* ecotypes) are most likely stable than a result of environmental plasticity.

Regarding the genic regions (GO clusters) which were targeted at least between two localities by 20-22 nt small RNAs (tasiRNA and miRNA) (supplementary figures S7 and S8); prenylation is post-translational addition of isoprenoid lipids to proteins and it was reported that it may cause phenotypic differences in some plants (Rodriguez-Concepcion 1999) and farnesylation is a type of prenylation. Protein farnesylation has a role in response to biotic stresses by initiating plant immunity and resistance against bacterial and oomycete pathogens (Goritschnig et al. 2008). As alpine and montane ecotypes of *H. pusillum* show different microbial communities (Trucchi et al. 2017) this may represent the differences in alpine and montane environments. It is also suggested that farnesylation can regulate the assembly of some of transcription factors and change their functions in flowering plants (Yalovsky et al. 2000).

In respect to alkene metabolism; there are some plants secondary metabolites (PSM) like terpenes which derive from alkene (Pavarini et al. 2012). Environmental effects such as herbivory have direct influence in PSM synthesis and accumulation which have a role in plant signaling and defense (Pavarini et al. 2012). This may also represent the possible effect of higher herbivory in montane ecotype of *H. pusillum* in comparison to alpine ecotype which also reflect in trichome formation in montane ecotype (Bertel et al. 2016c, 2018).

It was suggested that light quantity and ethylene biosynthesis are in contrast with each other and this response is rapid in the plants (Vandenbussche et al. 2003), as montane ecotype confront more intense light in the common garden in comparison to its natural growing site, it may cause downregulation of ethylene biosynthesis in montane ecotype.

This work was the first step to investigate the regulation roles of small RNAs in phenotypic differences between alpine and montane ecotypes of *Heliosperma pusillum*. To know more about the contribution of small RNAs, as well as to see directly if the parallel adaptation is caused by natural selection, and not by stochastic drift, comparing the plants grown in both native habitat and a common garden is needed (Kawecki and Ebert 2004, as cited in Bertel et al. 2018).

References

Bertel, Clara, Othmar Buchner, Peter Schönswetter, Božo Frajman, and Gilbert Neuner. 2016a. "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. doi:10.1111/boj.12467.

Bertel, Clara, Karl Hülber, Božo Frajman, and Peter Schönswetter. 2016b. "No Evidence Of Intrinsic Reproductive Isolation Between Two Reciprocally Non-Monophyletic, Ecologically Differentiated Mountain Plants At An Early Stage Of Speciation". *Evolutionary Ecology* 30 (6): 1031-1042. doi:10.1007/s10682-016-9867-y.

Bertel, Clara, Ivana Rešetnik, Božo Frajman, Brigitta Erschbamer, Karl Hülber, and Peter Schönswetter. 2018. "Natural Selection Drives Parallel Divergence In The Mountain Plant Heliosperma Pusillum S.L". *Oikos* 127 (9): 1355-1367. doi:10.1111/oik.05364.

Bertel, Clara, Peter Schönswetter, Božo Frajman, Andreas Holzinger, and Gilbert Neuner. 2016c. "Leaf Anatomy Of Two Reciprocally Non-Monophyletic Mountain Plants (Heliosperma Spp.): Does Heritable Adaptation To Divergent Growing Sites Accompany The Onset Of Speciation?". *Protoplasma* 254 (3): 1411-1420. doi:10.1007/s00709-016-1032-5.

Borges, Filipe, and Robert A. Martienssen. 2015. "The Expanding World Of Small Rnas In Plants". *Nature Reviews Molecular Cell Biology* 16 (12): 727-741. doi:10.1038/nrm4085.

Davidson, Nadia M, and Alicia Oshlack. 2014. “Corset: Enabling Differential Gene Expression Analysis for de Novoassembled Transcriptomes.” Genome Biology 15 (7). https://doi.org/10.1186/s13059-014-0410-6.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2012. “STAR: Ultrafast Universal RNA-Seq Aligner.” Bioinformatics 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Goritschnig, Sandra, Tabea Weihmann, Yuelin Zhang, Pierre Fobert, Peter McCourt, and Xin Li. 2008. “A Novel Role for Protein Farnesylation in Plant Innate Immunity.” Plant Physiology 148 (1): 348–57. https://doi.org/10.1104/pp.108.117663.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talon, J. Dopazo, and A. Conesa. 2008. “High-Throughput Functional Annotation and Data Mining with the Blast2GO Suite.” Nucleic Acids Research 36 (10): 3420–35. https://doi.org/10.1093/nar/gkn176.

Gould, Kevin S. 2004. "Nature's Swiss Army Knife: The Diverse Protective Roles Of Anthocyanins In Leaves". Journal Of Biomedicine And Biotechnology 2004 (5): 314-320. doi:10.1155/s1110724304406147.

Kozomara, A., and S. Griffiths-Jones. 2010. “MiRBase: Integrating MicroRNA Annotation and Deep-Sequencing Data.” Nucleic Acids Research 39 (Database): D152–57. https://doi.org/10.1093/nar/gkq1027.

Liao, Y., G. K. Smyth, and W. Shi. 2013. “FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features.” Bioinformatics 30 (7): 923–30. https://doi.org/10.1093/bioinformatics/btt656.

Lowry, David B. 2012. "Ecotypes And The Controversy Over Stages In The Formation Of New Species". *Biological Journal Of The Linnean Society* 106 (2): 241-257. doi:10.1111/j.1095-8312.2012.01867.x.

Mitchell, Alex L, Teresa K Attwood, Patricia C Babbitt, Matthias Blum, Peer Bork, Alan Bridge, Shoshana D Brown, et al. 2018. “InterPro in 2019: Improving Coverage, Classification and Access to Protein Sequence Annotations.” Nucleic Acids Research 47 (D1): D351–60. https://doi.org/10.1093/nar/gky1100.

Pavarini, Daniel Petinatti, Saulo Petinatti Pavarini, Michael Niehues, and Norberto Peporine Lopes. 2012. “Exogenous Influences on Plant Secondary Metabolite Levels.” Animal Feed Science and Technology 176 (1–4): 5–16. https://doi.org/10.1016/j.anifeedsci.2012.07.002.

Robinson, James T, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S Lander, Gad Getz, and Jill P Mesirov. 2011. “Integrative Genomics Viewer.” Nature Biotechnology 29 (1): 24–26. https://doi.org/10.1038/nbt.1754.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2009. “EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data.” Bioinformatics 26 (1): 139–40. https://doi.org/10.1093/bioinformatics/btp616.

Rodriguez-Concepcion, M. 1999. “The Prenylation Status of a Novel Plant Calmodulin Directs Plasma Membrane or Nuclear Localization of the Protein.” The EMBO Journal 18 (7): 1996–2007. https://doi.org/10.1093/emboj/18.7.1996.

RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Supek, Fran, Matko Bošnjak, Nives Škunca, and Tomislav Šmuc. 2011a. “REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms.” Edited by Cynthia Gibas. PLoS ONE 6 (7): e21800. https://doi.org/10.1371/journal.pone.0021800.

Trucchi, Emiliano, Božo Frajman, Thomas H. A. Haverkamp, Peter Schönswetter, and Ovidiu Paun. 2017. "Genomic Analyses Suggest Parallel Ecological Divergence In Heliosperma Pusillum (Caryophyllaceae)". *New Phytologist* 216 (1): 267-278. doi:10.1111/nph.14722.

Vandenbussche, Filip, Willem H. Vriezen, Jan Smalle, Lucas J.J. Laarhoven, Frans J.M. Harren, and Dominique Van Der Straeten. 2003. “Ethylene and Auxin Control the Arabidopsis Response to Decreased Light Intensity.” Plant Physiology 133 (2): 517–27. https://doi.org/10.1104/pp.103.022665.

Wang, Minghui, Yongzhong Zhao, and Bin Zhang. 2015. “Efficient Test and Visualization of Multi-Set Intersections.” Scientific Reports 5 (1). https://doi.org/10.1038/srep16923.

Wang, Ting-You. 2019. "RNA-Seq Data Analysis | Http://Databeauty.Com". Databeauty.Com. http://databeauty.com/blog/tutorial/2016/09/13/RNA-seq-analysis.html.

Yalovsky, Shaul, Manuel Rodríguez-Concepción, Keren Bracha, Gabriela Toledo-Ortiz, and Wilhelm Gruissem. 2000. “Prenylation of the Floral Transcription Factor APETALA1 Modulates Its Function.” The Plant Cell 12 (8): 1257–66. https://doi.org/10.1105/tpc.12.8.1257.

Appendix

The first principal component (explaining between 32 and 43% of the data) indicates alpine vs montane differential regulation of individuals from locality 1 for all analyzed classes of small RNAs (Figure S1).

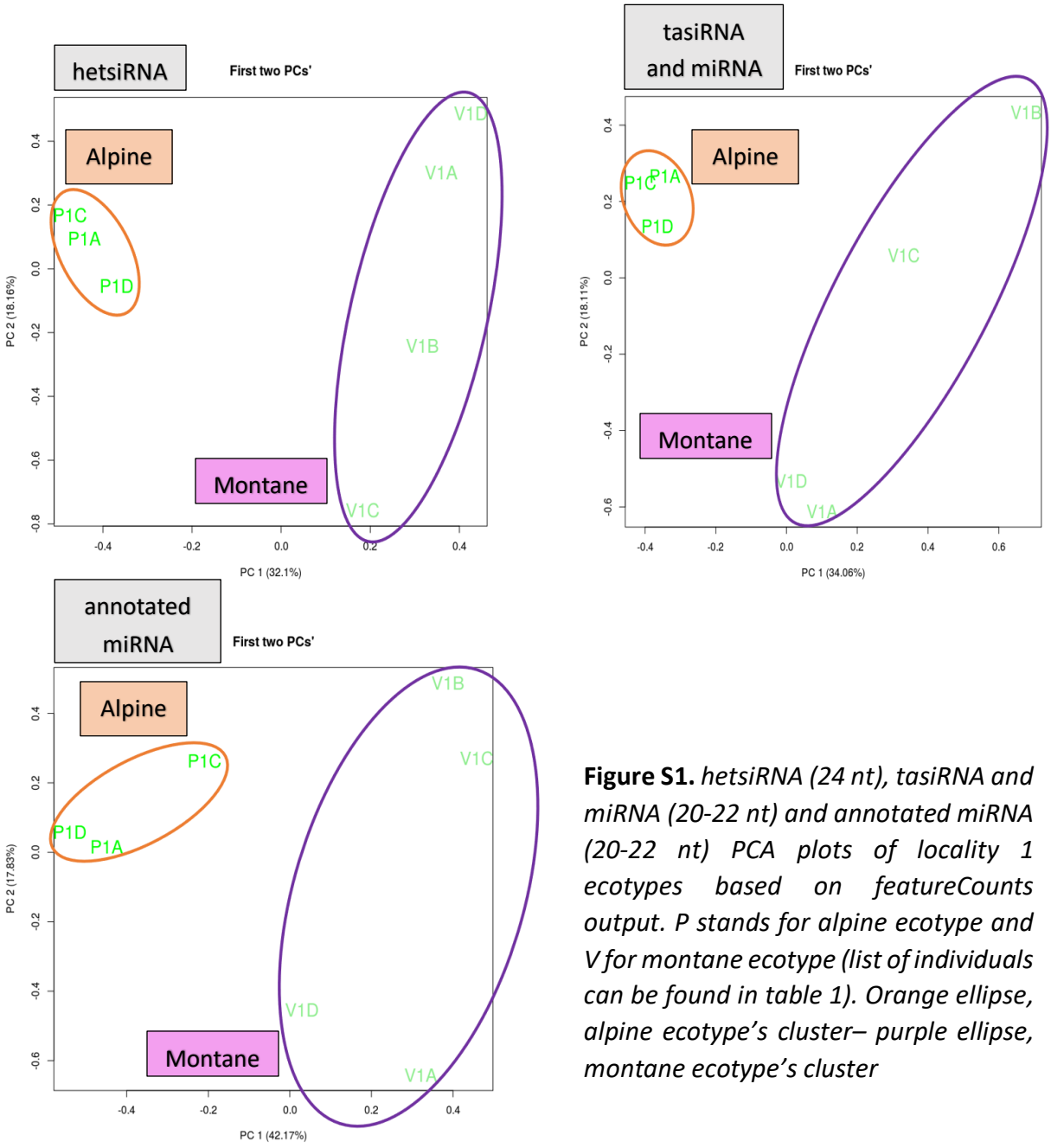


Figure S1. *hetsiRNA (24 nt), tasiRNA and miRNA (20-22 nt) and annotated miRNA (20-22 nt) PCA plots of locality 1 ecotypes based on featureCounts output. P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1). Orange ellipse, alpine ecotype's cluster– purple ellipse, montane ecotype's cluster*

The first principal component (explaining between 17 and 26% of the data) indicates alpine vs montane differential regulation of individuals from locality 3/4 for all analyzed classes of small RNAs (Figure S2).

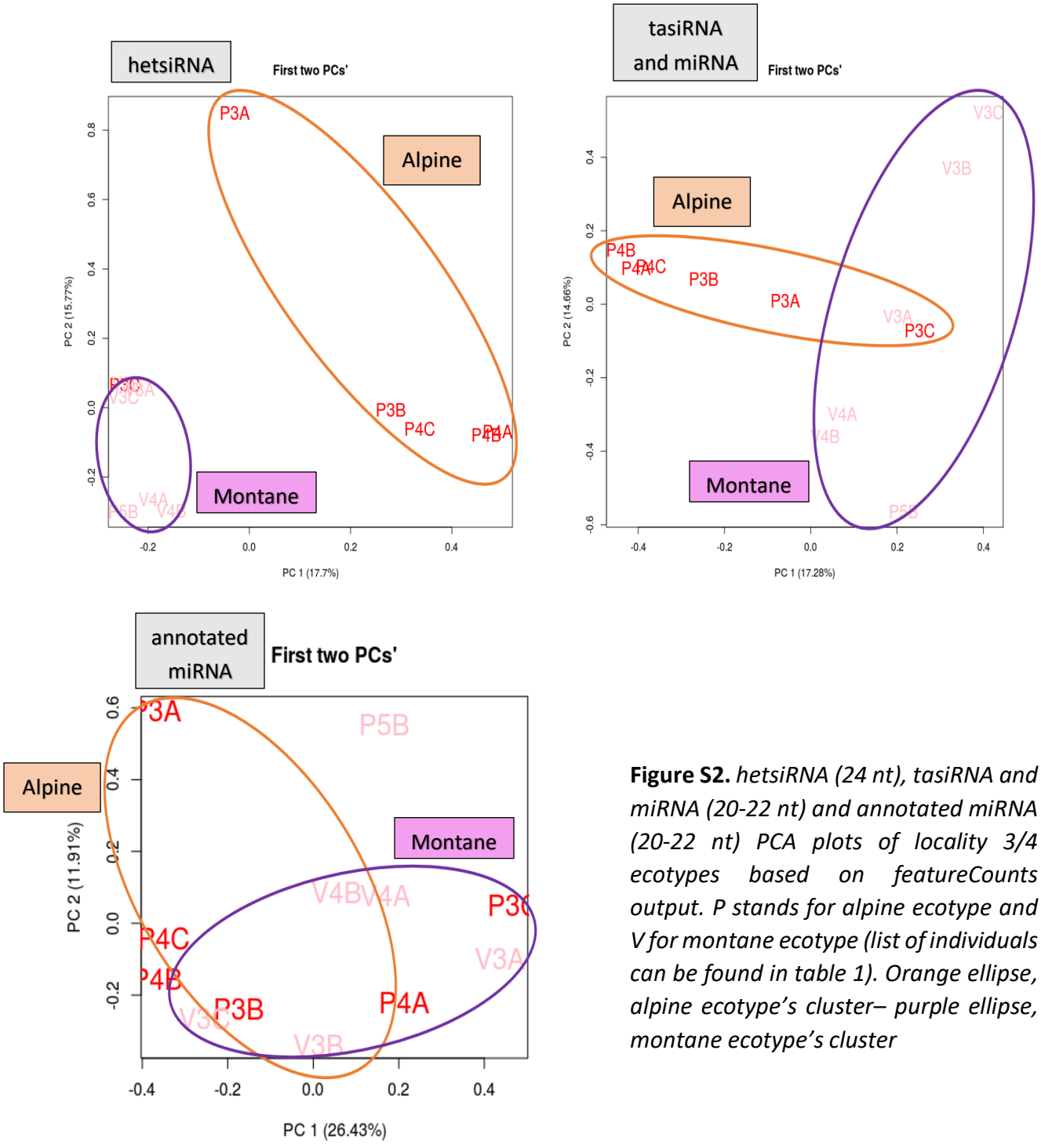


Figure S2. *hetsiRNA (24 nt), tasiRNA and miRNA (20-22 nt) and annotated miRNA (20-22 nt) PCA plots of locality 3/4 ecotypes based on featureCounts output. P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1). Orange ellipse, alpine ecotype's cluster– purple ellipse, montane ecotype's cluster*

The first principal component (explaining between 33 and 34% of the data) indicates alpine vs montane differential regulation of individuals from locality 5 for all analyzed classes of small RNAs (Figure S3).

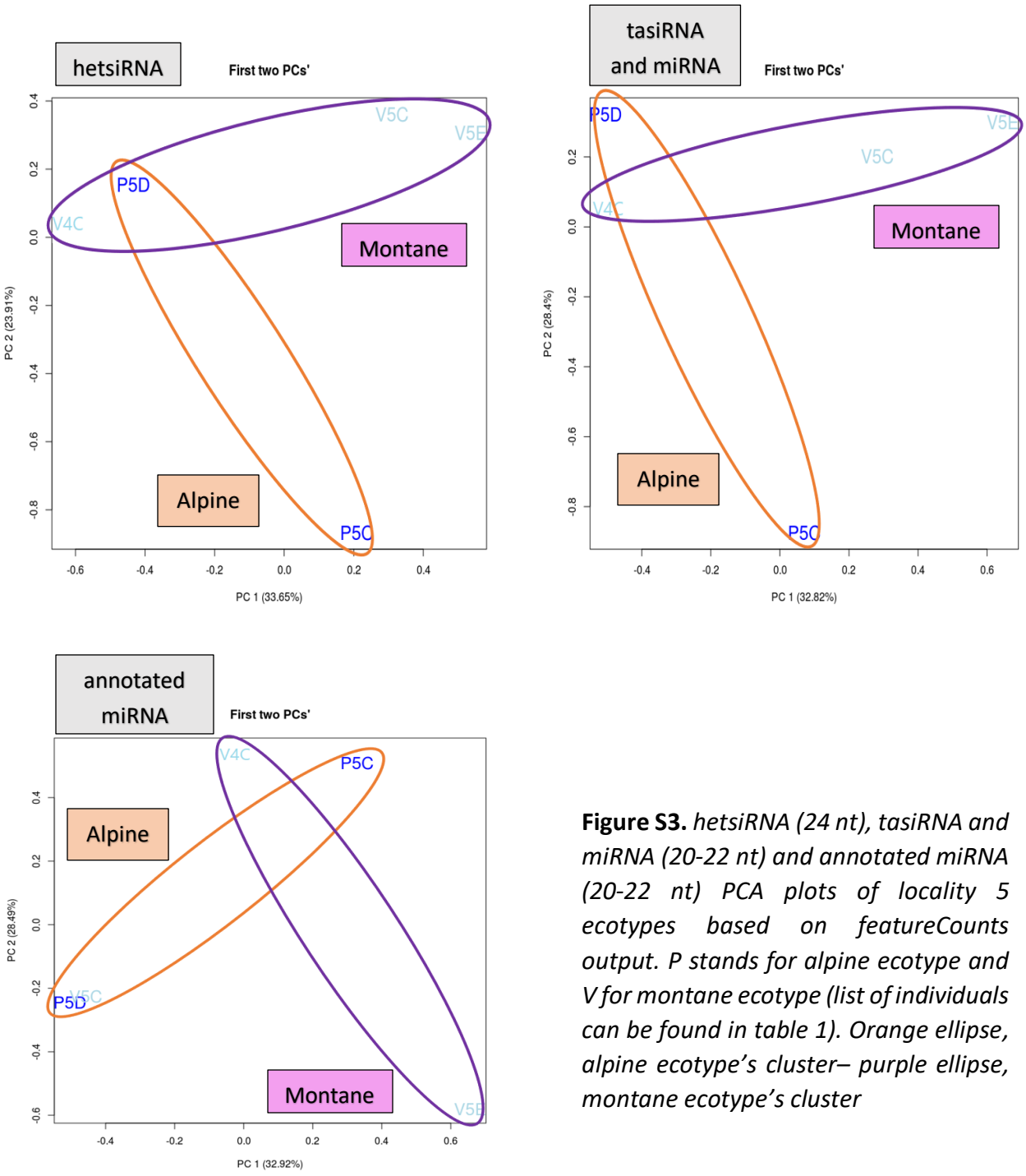
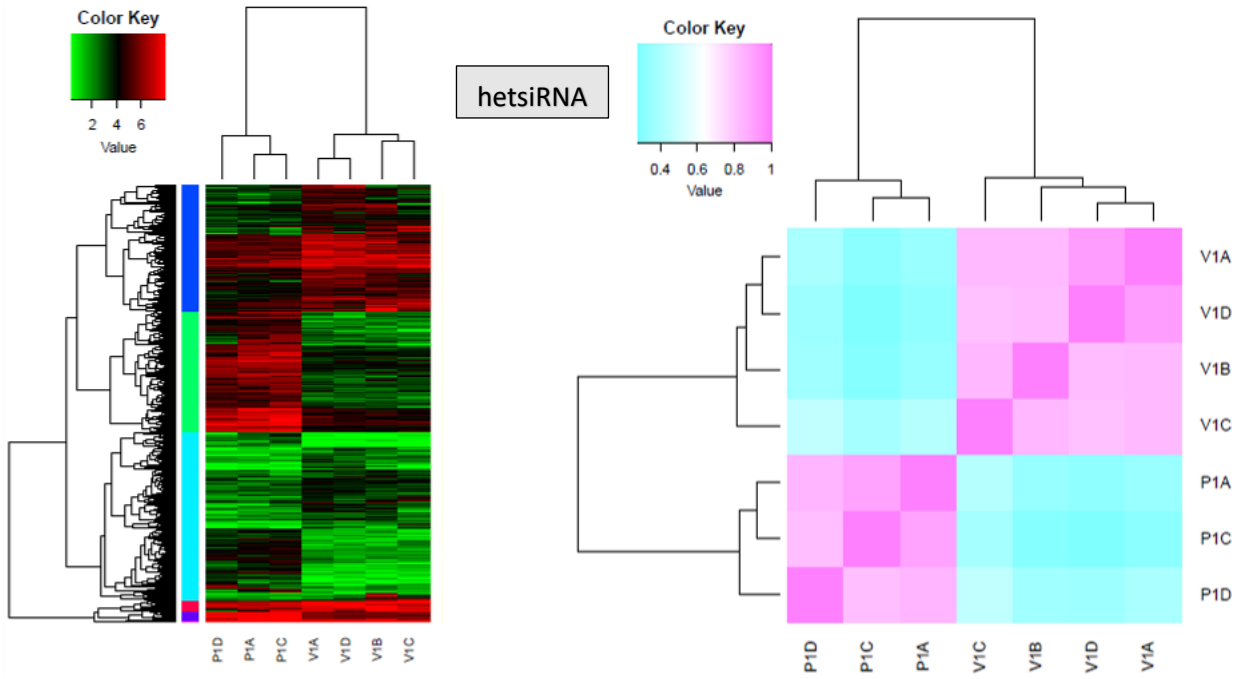


Figure S3. *hetsiRNA (24 nt), tasiRNA and miRNA (20-22 nt) and annotated miRNA (20-22 nt) PCA plots of locality 5 ecotypes based on featureCounts output. P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1). Orange ellipse, alpine ecotype's cluster– purple ellipse, montane ecotype's cluster*

Heatmaps of differentially regulated genomic (genic and intergenic) regions by hetsiRNA (24 nt) and tasiRNA and miRNA (20-22 nt) shows clearer differentiation in locality 1 (figure S4) compare to locality 3/4 and 5 (figure S5 and S6). Annotated miRNA did not show any differentiation in locality 3/4 and 5.



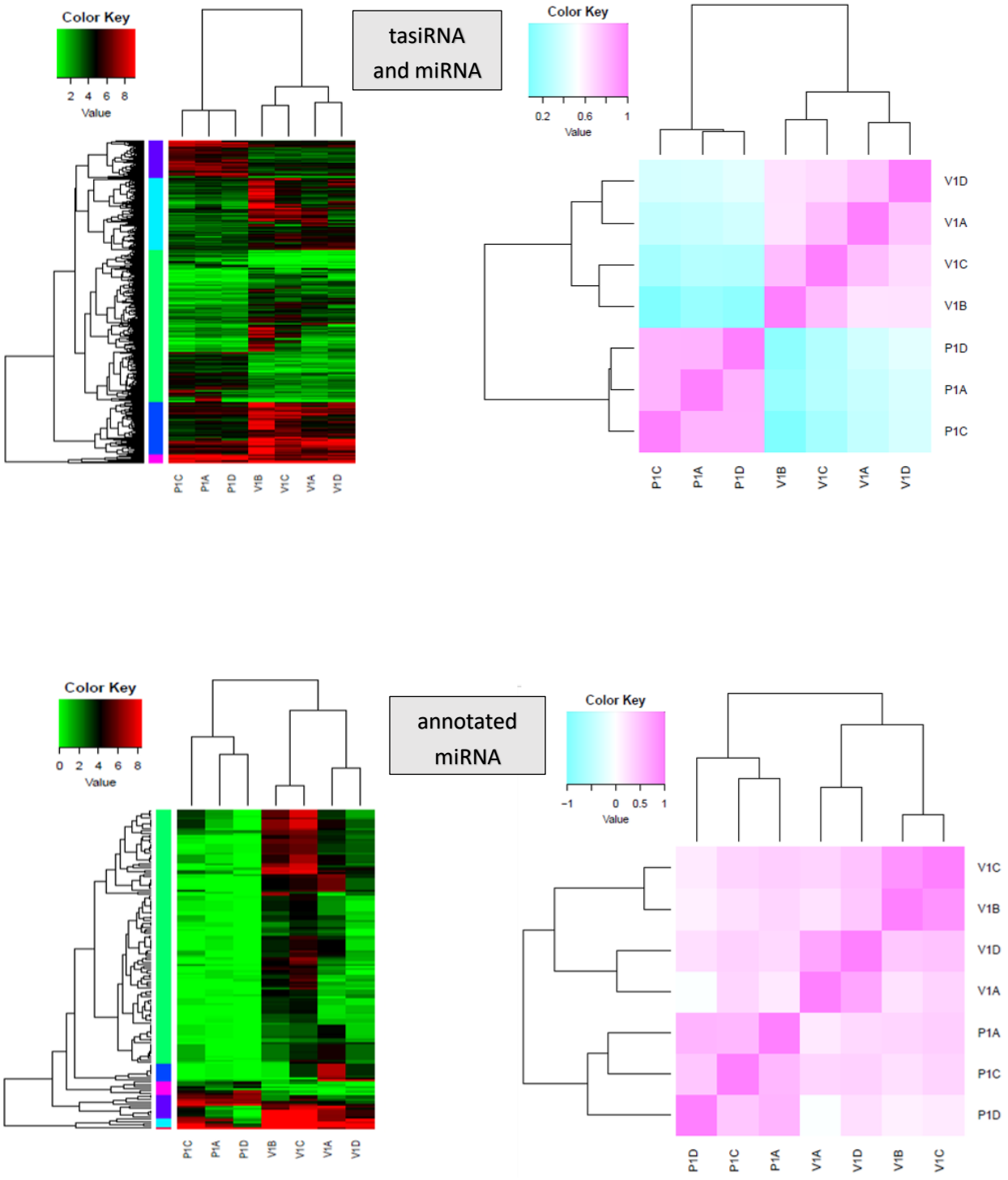


Figure S4. *Heatmaps of differentially regulated genomic regions by hetsiRNA, tasiRNA and miRNA and annotated miRNA in locality 1. These heatmaps were made based on edgeR output (FDR < 0.05). P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1).*

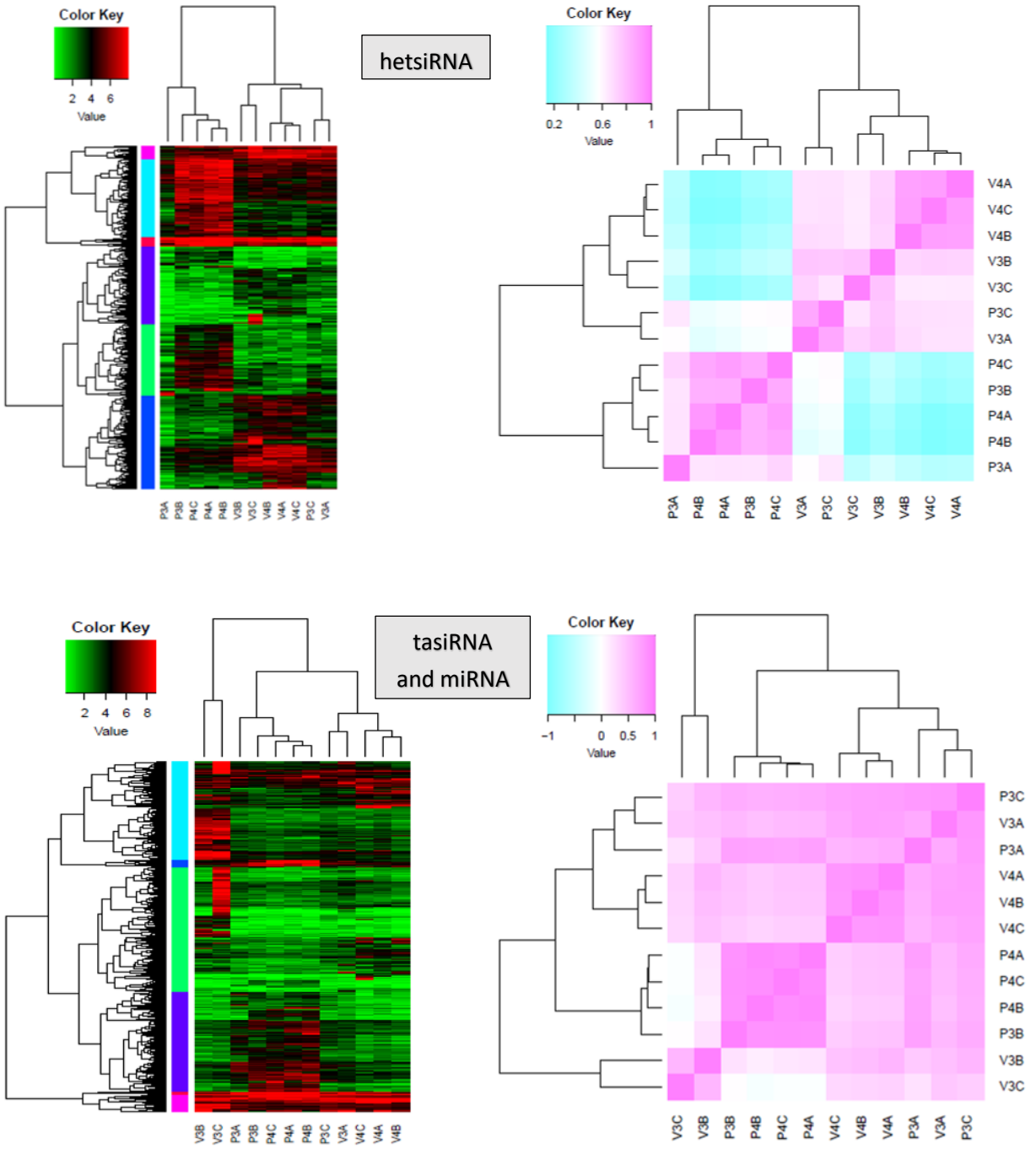


Figure S5. *Heatmaps of differentially regulated genomic regions by hetsiRNA, tasiRNA and miRNA in locality 3/4. These heatmaps were made based on edgeR output (FDR < 0.05). P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1).*

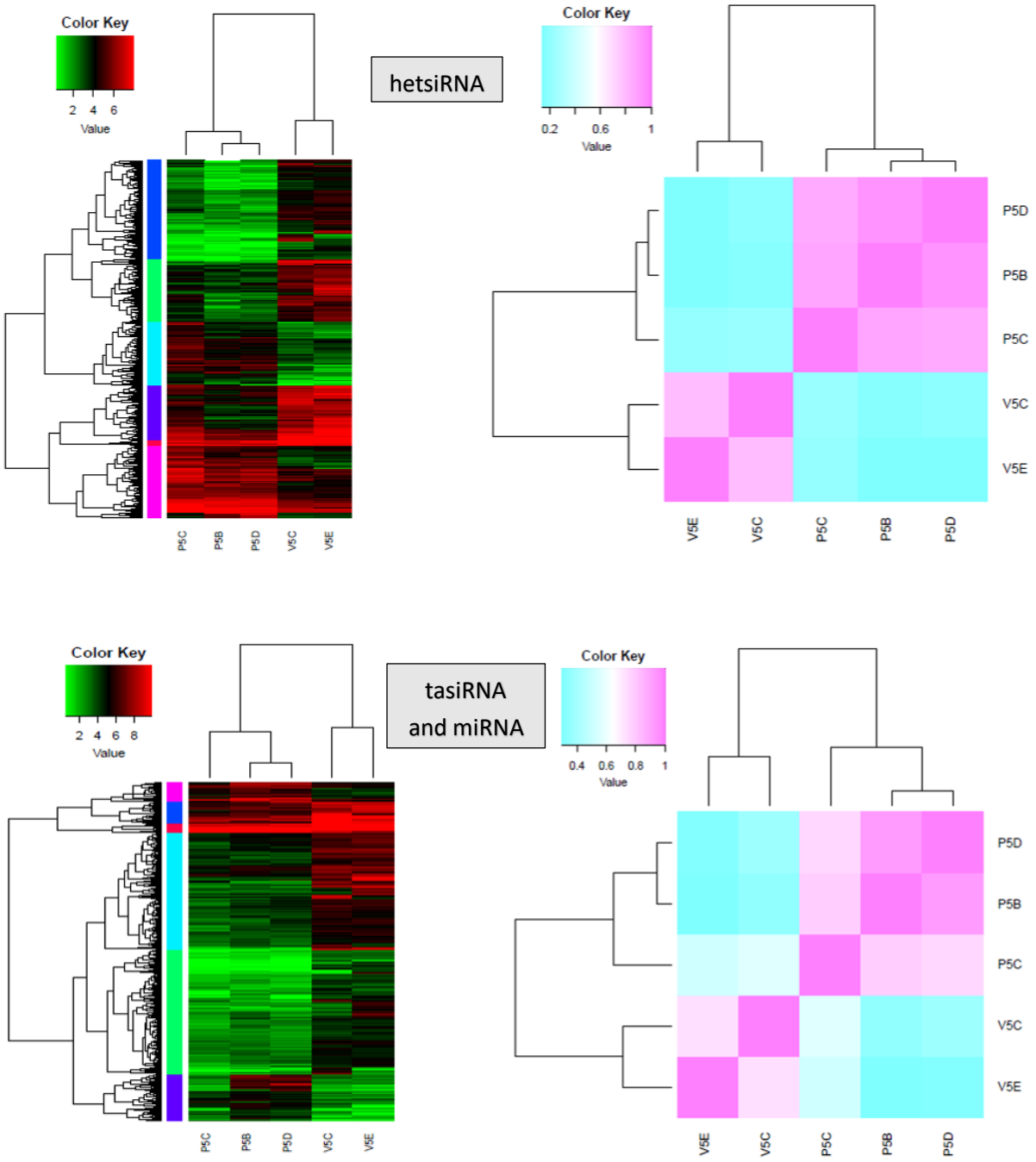


Figure S6. *Heatmaps of differentially regulated genomic regions by hetsiRNA, tasiRNA and miRNA in locality 5. These heatmaps were made based on edgeR output (FDR < 0.05). P stands for alpine ecotype and V for montane ecotype (list of individuals can be found in table 1).*

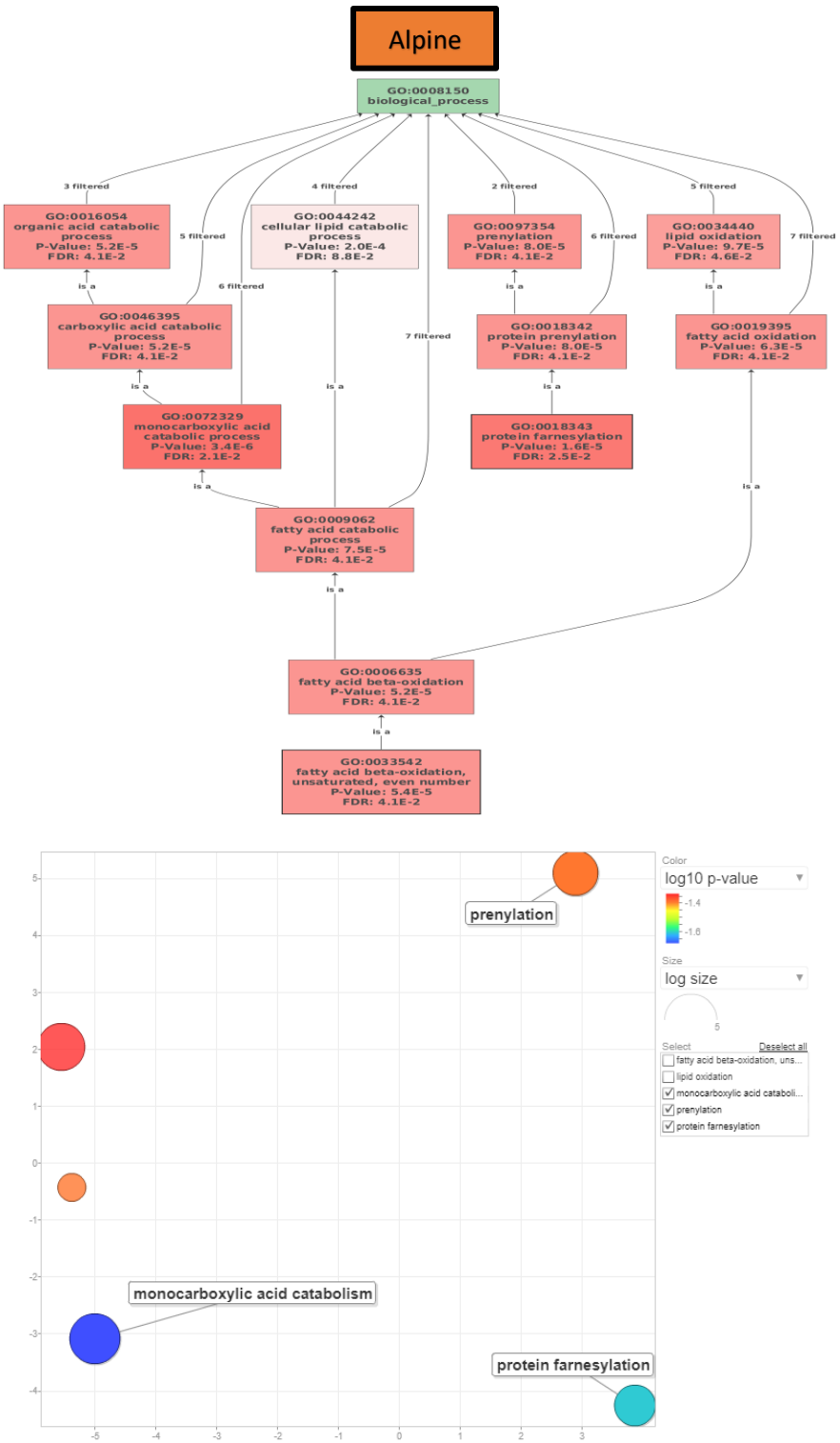


Figure S7. *Enrichment graph and Scatter plot of summarized GO terms (biological processes) for targeted (downregulated) genes by tasiRNA and miRNA in alpine ecotype. These genes are at least targeted in two out of three localities. Both enrichment graph and scatter plot show biological processes. Enrichment graph was made in Blast2GO Version 5 and was filtered (filter value for FDR and p-value were 0.1). Scatter plot was made by REVIGO and is based on FDR value. Blue bubbles are more significant GO terms than red and orange ones. Bubbles' size represents the frequency of the GO term.*

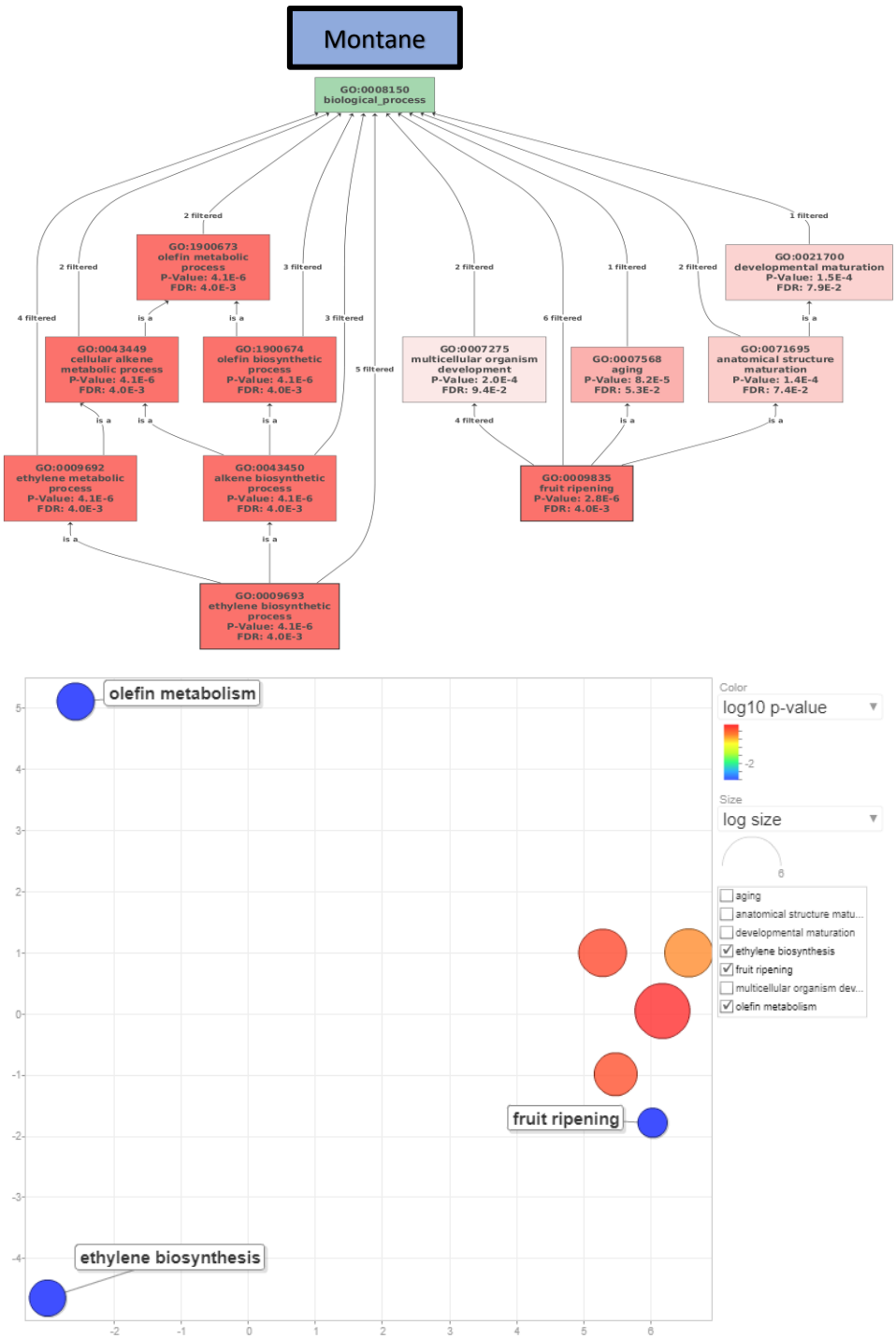


Figure S8. *Enrichment graph and Scatter plot of summarized GO terms (biological processes) for targeted (downregulated) genes by tasiRNA and miRNA in montane ecotype. These genes are at least targeted in two out of three localities. Both enrichment graph and scatter plot show biological processes. Enrichment graph was made in Blast2GO Version 5 and was filtered (filter value for FDR and p-value were 0.1). Scatter plot was made by REVIGO and is based on FDR value. Blue bubbles are more significant GO terms than red and orange ones. Bubbles' size represents the frequency of the GO term.*

Zusammenfassung

Als Folge des Klimawandels müssen sich Pflanzen an neue ökologische Bedingungen anpassen. Die Pflanzenart *Heliosperma pusillum* (Caryophyllaceae) bildet in den Südostalpen mehrfach zwei Ökotypen, die sich sowohl anhand der ökologischen als auch der morphologischen Merkmale voneinander unterscheiden. Dies erfolgte aufgrund der Anpassung an verschiedene Höhenlagen (alpine vs. Montane Vegetationsstufe). Vor allem weist der montane Ökotyp eine dichte Behaarung mit multizellulären Trichomen auf, was bei dem alpinen Ökotyp nicht vorhanden ist. Die Behaarung und weitere divergente Merkmale verleihen dem montanen Ökotyp einen selektiven Vorteil unter den heißeren und trockeneren Umweltbedingungen der montanen Stufe. Aus diesem Grund stellt dieses System ein nützliches Modell dar, das die Untersuchung der Anpassung von Alpenpflanzen an den Klimawandel ermöglicht. Ziel dieses Projektes ist, die möglichen Ursachen dieser phänotypischen und ökologischen Divergenz zu untersuchen. Es wurde die Regulation der Genexpression durch kleine RNAs (abgekürzt sRNAs) in alpinen und montanen Populationen von *Heliosperma pusillum* untersucht. Verschiedene Klassen von sRNA mit einer Länge von 20-22 bzw. 24 Nukleotiden wurden von Pflanzen extrahiert, die experimentell nebeneinander kultiviert wurden, um verschiedene Umweltfaktoren als Ursache für regulatorische Unterschiede auszuschließen. Unsere Ergebnisse zeigen, dass einige der für Licht zuständigen Gene, vorwiegend durch 20-22 nt sRNAs (miRNA und tasiRNA) in den verschiedenen Ökotypen unterschiedlich reguliert werden. Dieses Ergebnis ist auf die unterschiedlichen Lichtregimes zurückzuführen, die die alpinen und montanen Ökotypen in ihren natürlichen Habitaten vorfinden. Darüber hinaus ist die regulatorische Aktivität der 20-22 nt sRNAs (miRNA und tasiRNA), die zu einer unterschiedlichen Regulation der Genen zwischen Ökotypen führt, in den verschiedenen Lokalitäten viel einheitlicher im Vergleich zu den 24 nt sRNAs (hetsiRNA).