De Novo Transcriptome Sequencing of *Ramonda serbica*: Identification of Late Embryogenesis Abundant Proteins

Ana Pantelić[1], Strahinja Stevanović[2], Nataša Kilibarda[3], Marija Vidović[2]

1) University of Belgrade, Faculty of Chemistry, Studentski trg 12-16, 11000 Belgrade, Serbia
2) Institute of Molecular Genetics and Genetic Engineering, Laboratory for Plant Molecular Biology, University of Belgrade, Vojvode Stepe 444a, Belgrade, Serbia
3) Singidunum University, Danijelova 32, 11000 Belgrade, Serbia

Introduction and aim

- Resurrection plant *Ramonda serbica* Panc. survives desiccation for a long period and fully recovers metabolic functions already within one day upon watering.
- Desiccation (extreme dehydration) induces protein unfolding and aggregation, destabilization or loss of cellular membrane integrity. Besides, desiccation provokes the accelerated generation of reactive oxygen species.

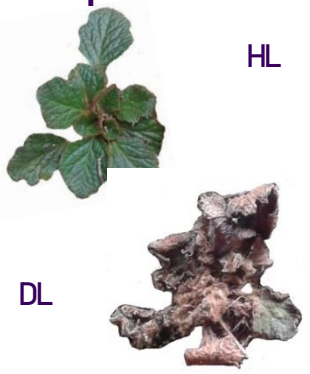
Aim:

To identify Late Embryogenesis Abundant Proteins (LEAPs) that contribute to desiccation tolerance in *R. serbica* by comparative transcriptomics of hydrated (HL) and desiccated leaves (DL).

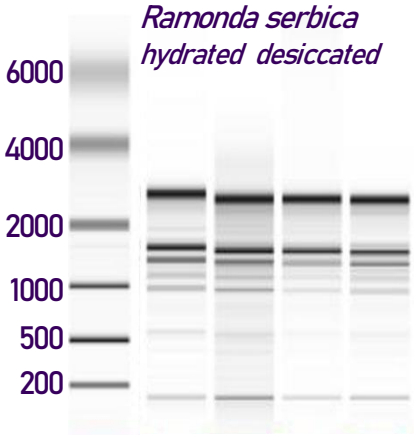
Workflow

Isolation of high-quality RNA from *Ramonda serbica* [1]

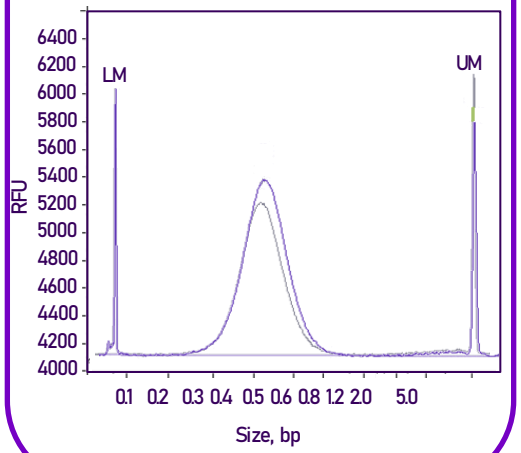
TRIzol-based mRNA extraction and purification[1]



Quality control and quantification



Synthesis of cDNA library



Transcriptome Reconstruction

Trinity (Inchworm, Chrysalis, Butterfly); v. r20140413p1. min_kmer_cov=2, min_glue=2; kmer=25

Annotation

NR, NT, Swiss-Prot, KOG/COG, KEGG, Pfam, GO

Quantification

Bowtie2-2.2.2, mismatch=0
RSEM-v1.3.0

Differential Expression Analysis

DESeq2 v. 1.10.1, p adjust < 0.05

Enrichment Analysis

[1] Vidović M, Ćuković K. 2020. 3 Biotech. 10(6):286.

Workflow

Analysis	Software	Version	Parameters	Remarks
Transcriptome Reconstruction	Trinity	r20140413p1	min_kmer_cov=2, min_glue=2, others are by default	
Annotation	Diamond	v0.8.22	R, Swiss-Prot: e value = 1e-5; KOG/COG: e value = 1e-3	
	KAAS	r140224	E-value = 1e-8	KEGG Annotation
	Blast	v2.2.28+	E-value = 1e-5	NT Annotation
	Hmmscan	HMMER 3.1b1	E-value = 0.01	Pfam Annotation
	Blast2go	b2g4pipe_v2.5	E-value = 1e-6	GO Annotation
Quantification	Bowtie2, RSEM	bowtie2-2.2.2, RSEM-v1.3.0	bowtie2: mismatch=0	Mapping to assembled transcriptome
Differential Expression Analysis	DESeq	1.12.0	|log2foldchang|>1 && padjust<0.005 TMM Poisson distribution BH	from 5 plant is take 3 leafs, and pull over
GO Enrichment	GOSeq, topGO	GOSeq-1.32.0, topGO-2.32.0	Corrected P-Value < 0.05	
KEGG Enrichment	KOBAS	v3.0	Corrected P-Value < 0.05	

Data quality

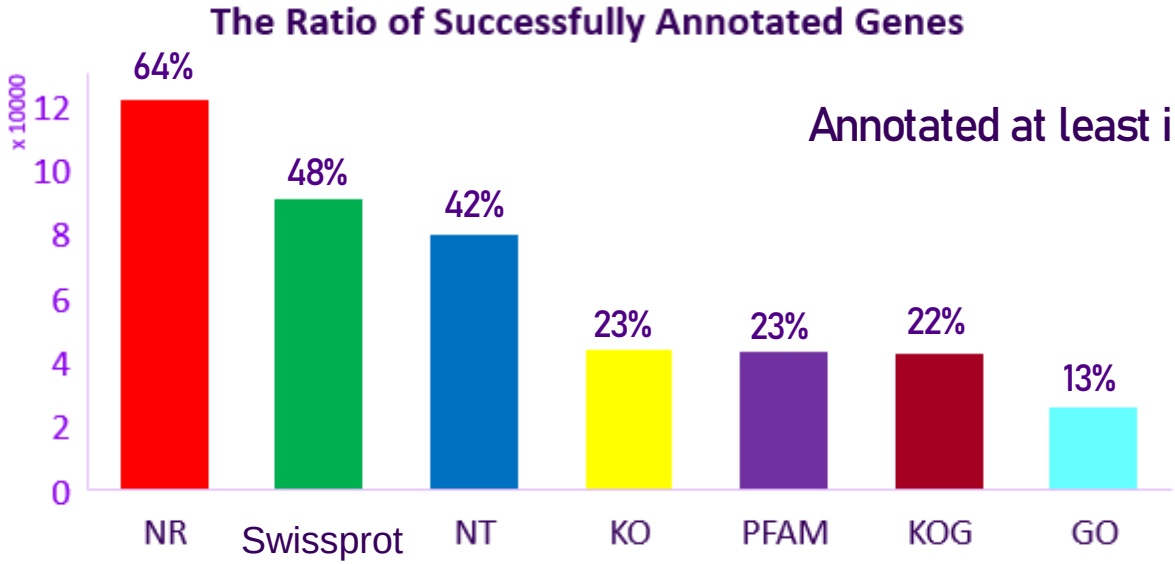
Overview of data production quality

sample	raw_reads	clean_reads	raw_data(G)	clean_data(G)	error_rate (%)	Q20 (%)	Q30(%)	GC_content (%)
H3	40137483	39608813	12.0	11.9	0.04	98.01	94.00	45.55
D4	38039070	37482969	11.4	11.2	0.04	98.02	94.10	46.01

Transcriptome Reconstruction

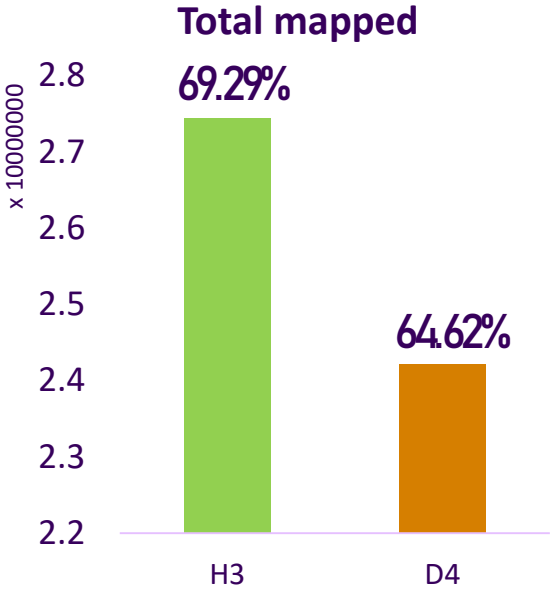
After *de novo* transcriptome analysis, 189 456 transcripts with 189 003 unigenes annotated with seven common databases.

Transcript length interval	200-500bp	500-1kbp	500-1kbp	500-1kbp	Total
No of transcripts	64728	61598	44999	18131	189456
No of unigens	64282	61591	44999	18131	189003

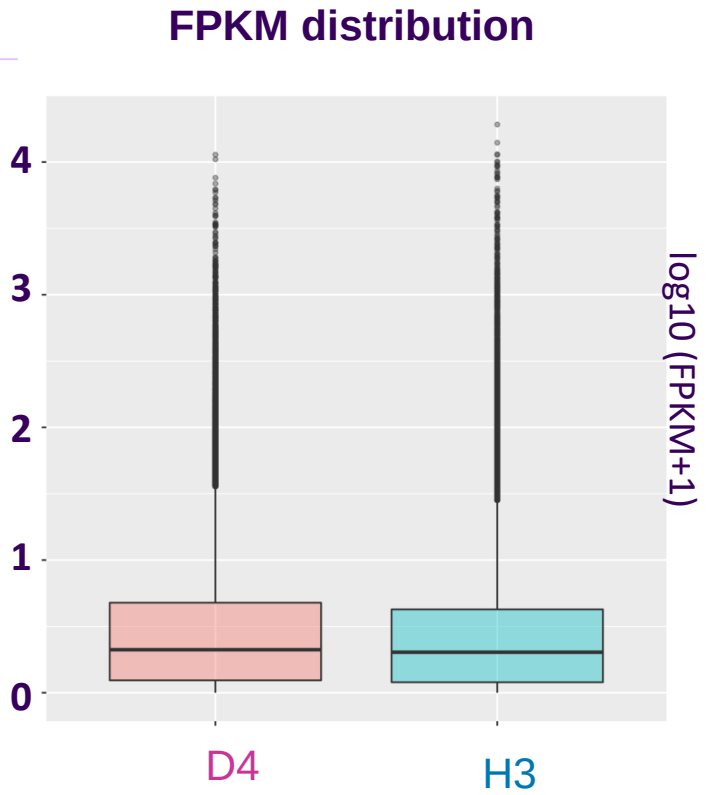


Annotated at least in one database: 127 176 (67%)

Quantification



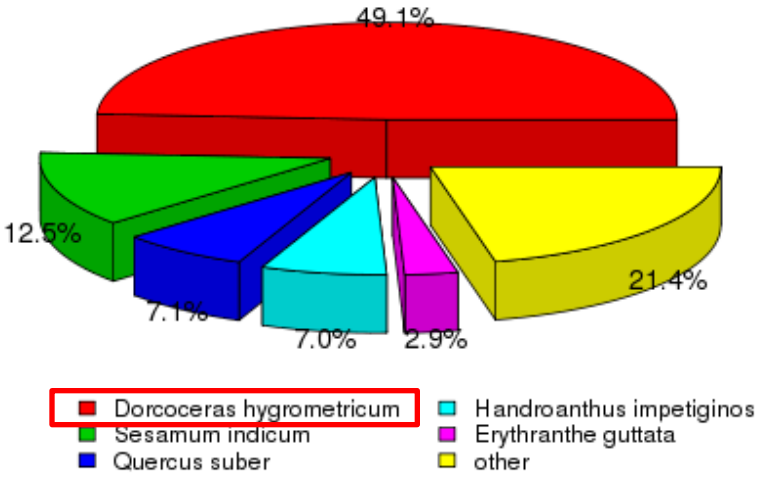
FPKM Interval	H3	D4
0~1	90100(47.67%)	93610(49.53%)
1~3	44332(23.46%)	45717(24.19%)
3~15	38949(20.61%)	35083(18.56%)
15~60	38949(20.61%)	11276(5.97%)
>60	38949(20.61%)	3317(1.75%)



Threshold for gene expression is FPMK <0.3

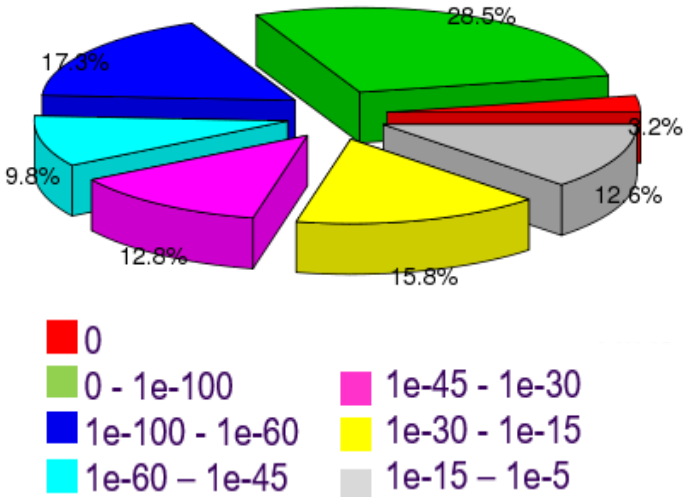
Gene Annotation Results

Species Classification

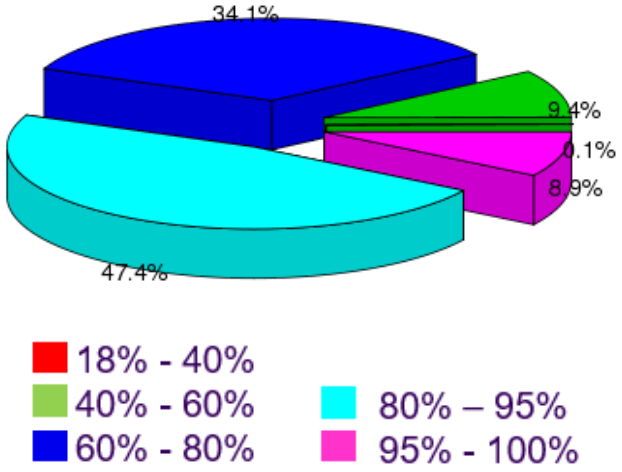


Nr databse, E-value 1E-5

E-value Distribution



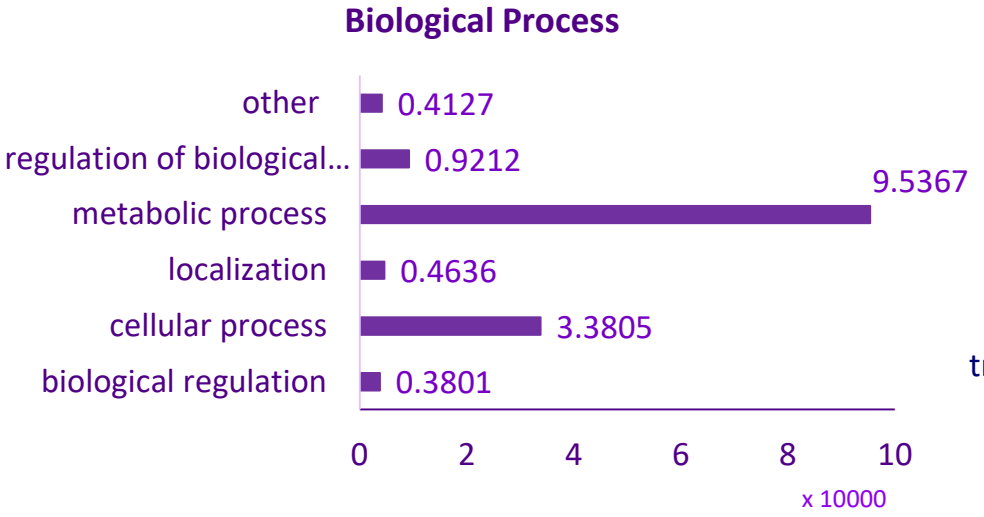
Similarity Distribution

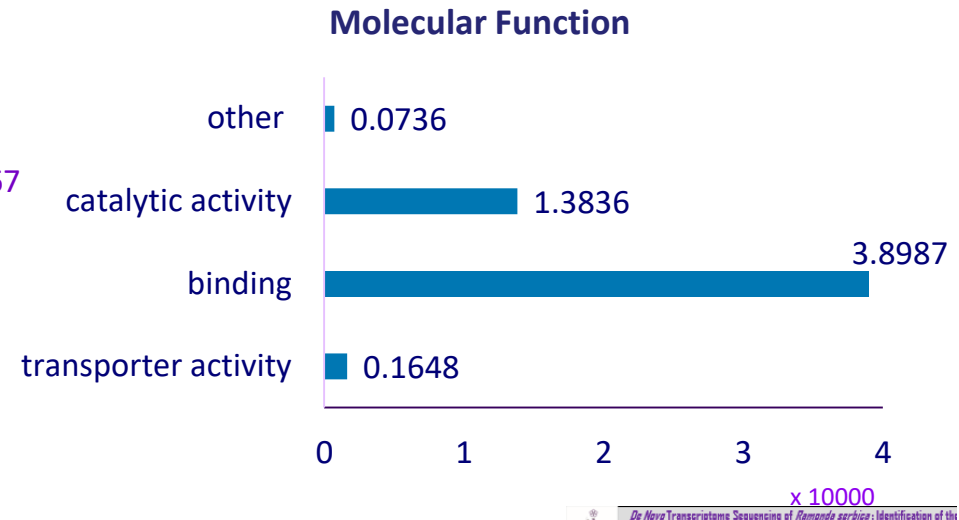


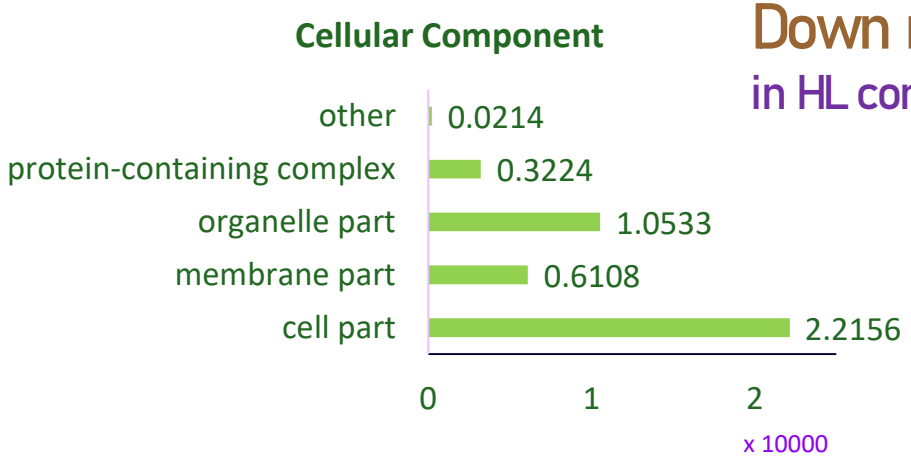
Xiao L et al., "The resurrection genome of Boea hygrometrica: A blueprint for survival of dehydration.", Proc Natl Acad Sci U S A, 2015 May 5;112(18):5833-7

Gene Ontology- Blast2go

Successfully annotated genes (HL & DL) were grouped into three main GO domains:



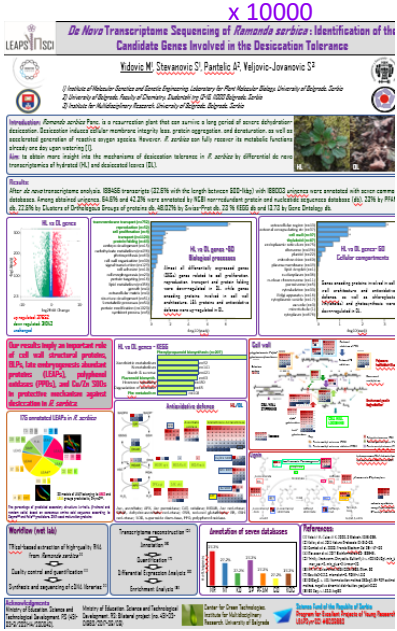




Up regulated in HL: 37652

Down regulated in HL: 31042

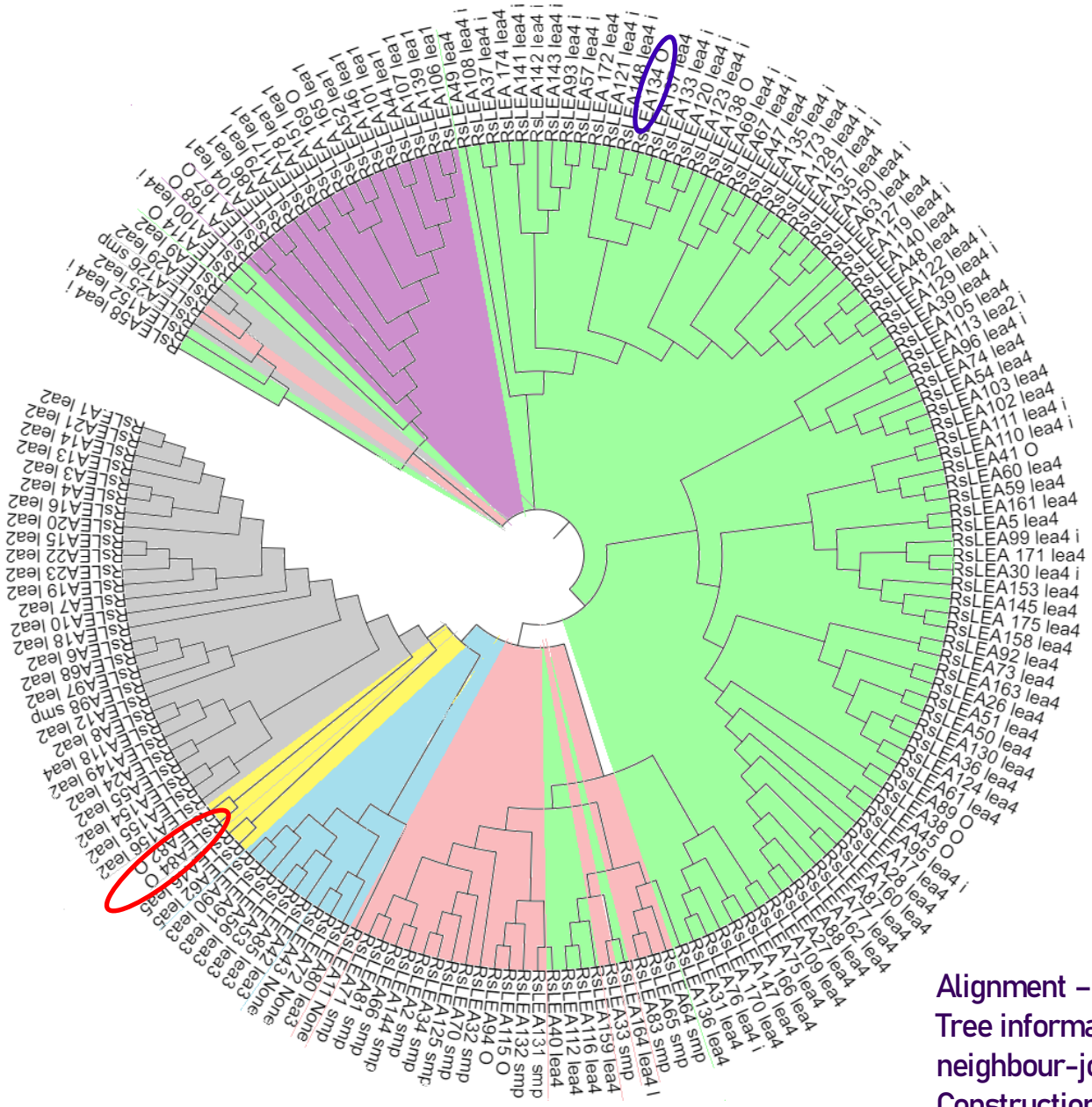
in HL compared with DL



Late embryogenesis abundant proteins

- Late embryogenesis abundant proteins (LEAPs) are intrinsically disordered proteins.
- The hallmark of desiccation tolerance is the accumulation of LEAPs in vegetative tissue of the resurrection plants.
- This heterogeneous group of anhydrobiosis-related intrinsically disordered proteins forms mostly random conformation when fully hydrated, turning into compact α-helices during desiccation
- Based on *in vitro* studies, LEAPs can be involved in water binding, ion sequestration, stabilization of both membrane and enzymes during freezing or drying, while by forming intracellular proteinaceous condensates they increase structural integrity and intracellular viscosity of cells during desiccation.

LEAPs phylogenetic tree

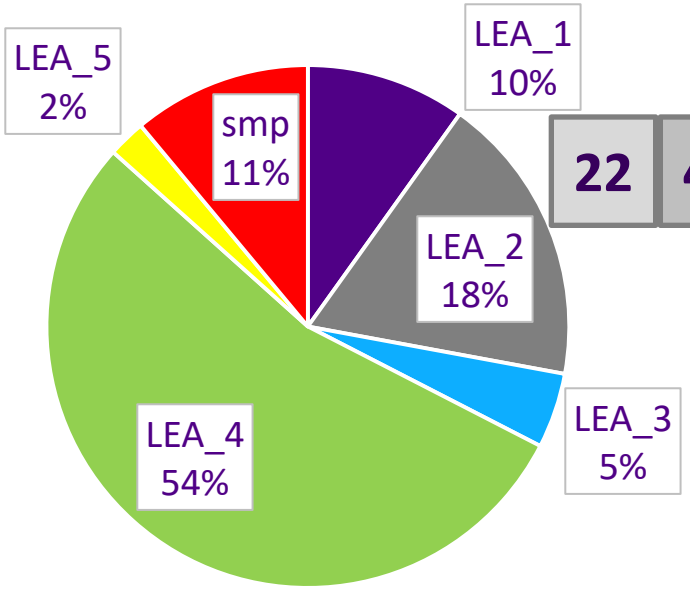


- LEA 1
- LEA 2
- LEA 3
- LEA 4
- LEA 5
- SMP

Alignment – MAFFT (G-INS-i 100 iteration)
Tree information – Simple phylogeny (clustering method neighbour-joinig)
Construction of the tree – iTOL.embl.de

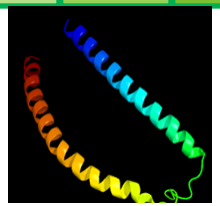
LEAPs classification

We identify and annotated 175 LEAPs in *Ramonda serbica*



22	45	30

84	12	22



3D models of LEAP belonging to LEA 2 and LEA 4 groups predicted by Phyre2[1].

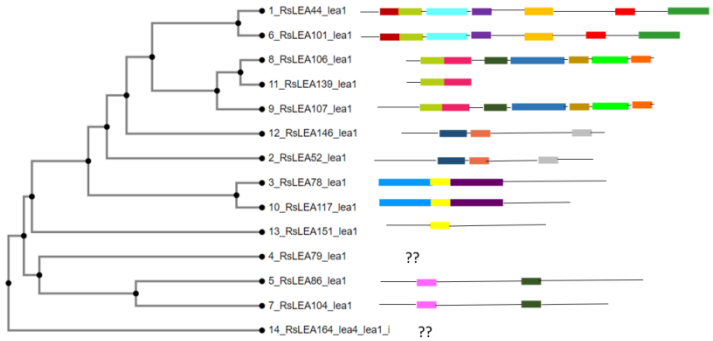
The percentage of predicted secondary structure (α-helix, β-strand and random coils) based on consensus amino acid sequence according to PSIPRED[2], Sopma[3], and Fells[4] predictors.
SMP – seed maturation proteins

α	β	coil

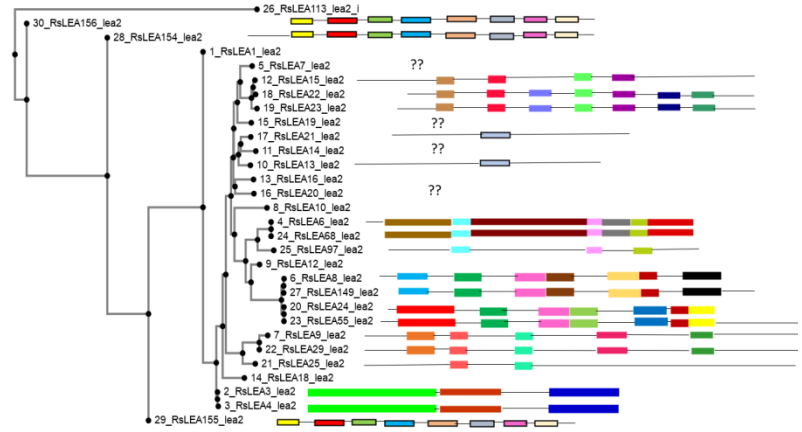
1) Kelley et al. 2015 Nature Protocols 10: 845-58.
2) Jones 1999 J. Mol. Biol. 292: 195-202.
3) Combet et al. 2000. Trends Biochem Sci 291: 147-150
4) Piovesan et al. 2017 Bioinformatics 33: 1889-91.

Architecture of conserved protein motifs in LEAPs

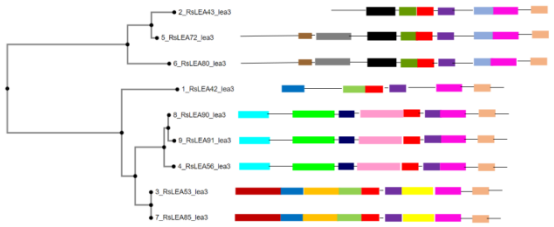
LEA 1



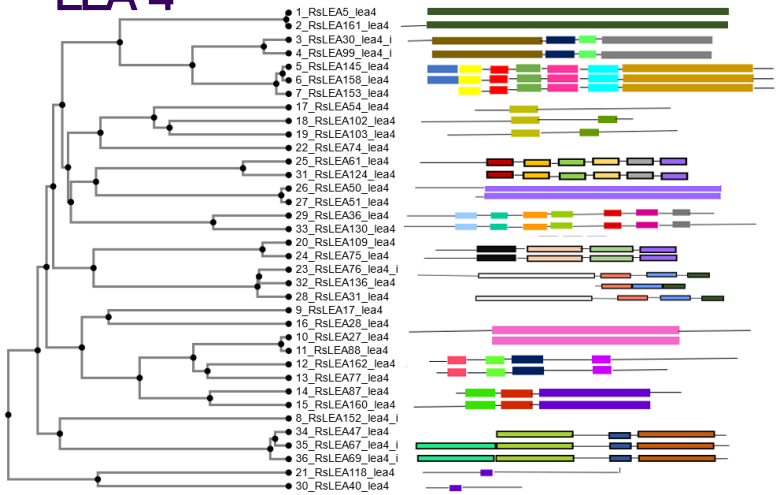
LEA 2



LEA 3

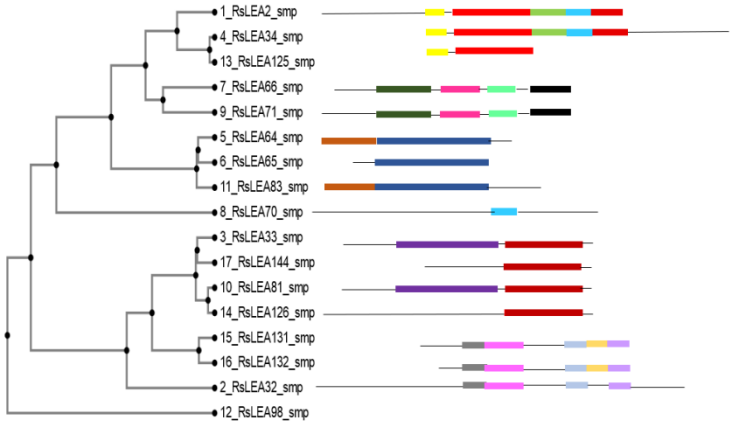


LEA 4



LEA 5

SMPs



Alingment and tree – MAFFT (NJ method, JTT substitution model, Ignore heterogeneity among sites)

Conclusion

1. Differential transcriptomics of HL and DL was performed for the first time.
2. 175 LEA proteins from *R.serbica* are identify and classify
3. The majority belongs to LEA 4 family group and they show high propensity for α helix formation (>80 %)
4. Getting more insights of structural properties of upaccumulated LEAPs under water lack will help us to reveal their role in desiccation tolerance mechanism (eg. protein stabilization role acting as molecular shields)

Thank you!

Acknowledgments
Ministry of Education, Science and Technological Development, RS (451-03-9/2021-14/200042).

Science Fund of the Republic of Serbia
Program for Excellent Projects of Young Researchers
LEAPSyn-SCI #6039663

NGP-net
Non-globular proteins in molecular physiopathology

EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY