



NIBIO

NORWEGIAN INSTITUTE OF
BIOECONOMY RESEARCH

Complex regulation of repeat domain proteins during stress and development in Norway spruce

Igor Yakovlev

Tandem repeat domain (TRD) containing proteins are composed of repeated units of conserved stretches of 20–40 amino acids, that contain simple structural motifs, such as α -helices or β -sheets

Plants are enriched in tandem repeats containing proteins typically distributed into various families.

Several repeat protein gene families have been identified in plants. Such as Armadillo (ARM), Ankyrin (ANK), HEAT, Kelch-like repeats, Pentatricopeptide repeats (PPR), Tetratricopeptide (TPR), Leucine rich repeats (LRR) and WD-repeats.

The precipitous expansion of repeat proteins in plants is presumed to be through internal tandem duplications.

It is suggested that the ability of TRD proteins to bind to multiple ligands to establish interaction network in rapid response to external stimuli such as biotic and abiotic stress might play a role in their selection during evolution.

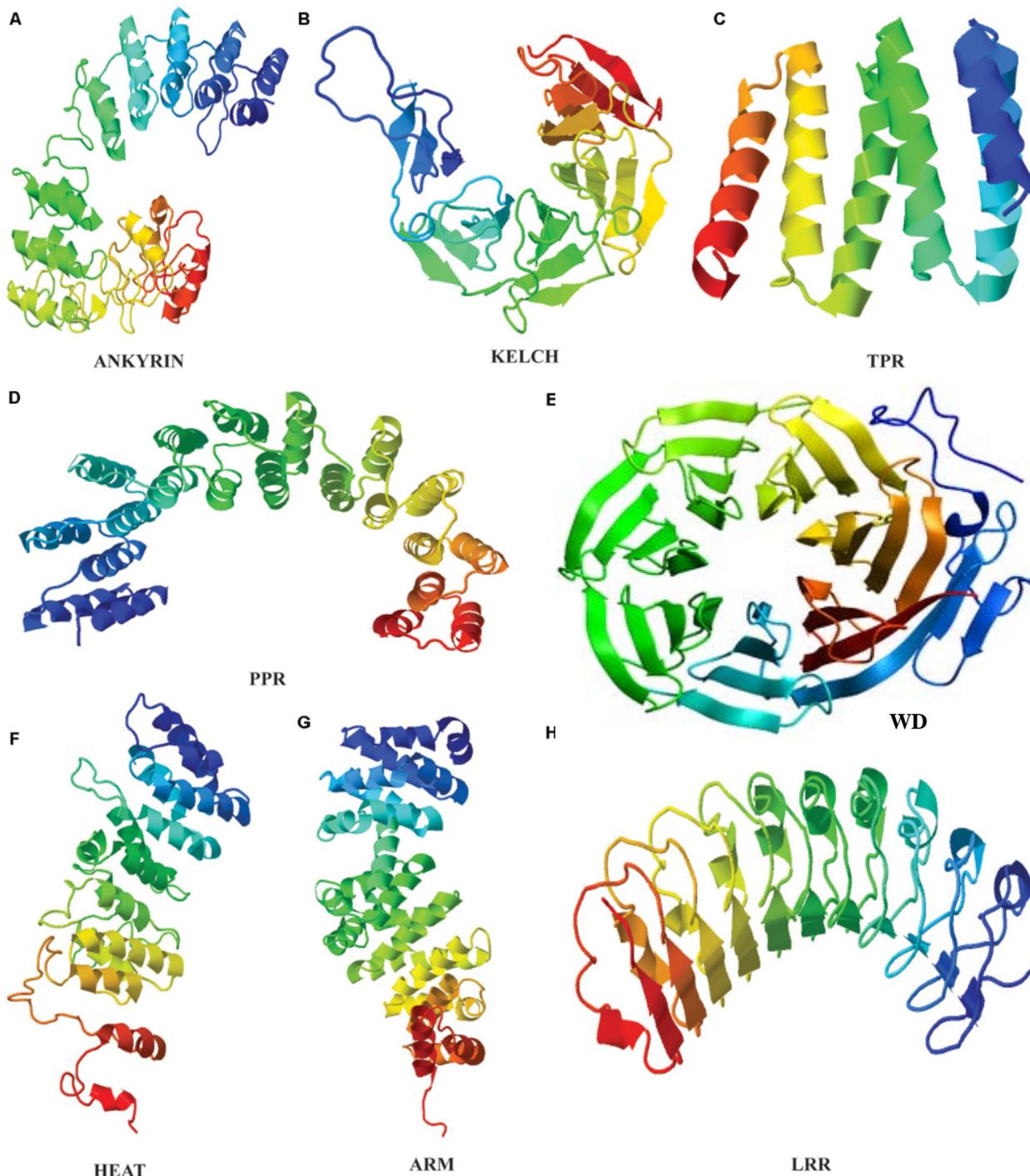
Putative number of repeats family proteins in various species

	Armadillo	HEAT	Ankyrin	PPR	TPR	LRR	WD40	Kelch
<i>Oryza sativa</i>	144	17	411	985	259	947	507	75
<i>Arabidopsis thaliana</i>	124	17	308	564	194	794	449	120
<i>Picea abies</i>	233	302	127	1396	1198	1923	406	283

Number and functions of main groups of several repeat domain protein gene families in Norway spruce

RD gene family	Number of gene models	General functions (by Sharma et al., 2016)
Armadillo (ARM) repeats	233	Involved in the protein degradation mechanism (via ubiquitination) and regulate proteome plasticity to respond effectively under environmental stress conditions
Ankyrin (ANK) repeats	127	Transmembrane proteins, K ⁺ channel signal transduction and thermosensing mechanisms , essential for plant organogenesis and morphogenesis during developmental stages
HEAT repeats	302	Probably functions as a protein–protein interaction module and an integral component of plant defense signaling
Kelch-like repeats	283	Could play prominent roles in multiple physiological processes in plants, such as responses to various hormones, the circadian clock and photomorphogenesis; flower development, senescence, and induced defense responses
Tetratricopeptide (TPR)	1198	A number of TPRs are known to be involved in plant innate immune response, plant stress and hormone signaling pathways
Pentatricopeptide repeats (PPR)	1396	Mitochondria/chloroplast and nucleo-cytoplasmic localized PPR protein, are classified to perform diverse role in plant stress and development signaling
Leucine rich repeats (LRR)	1923	NB-LRR (nucleotide binding) proteins mainly correspond to R- (resistance) proteins and involved effector triggered immunity mechanisms . LRR kinase proteins have been characterized to mediate abiotic stress responses
WD repeats	406	The gene family of WD proteins are involved in a broad spectrum of crucial plant developmental processes. WD40 are epiregulators, involved into chromatin-based gene silencing and also in abiotic (temperature) stress regulation
Clathrin heavy chains/VPS repeat proteins	404	involved in protein sorting between vacuolar proteins (storage) and secreted proteins, intracellular protein transport, vesicle-mediated transport, mostly combined with TPR, PPR and LRR repeats
Total	4784	

Tertiary structures of plant repeat proteins



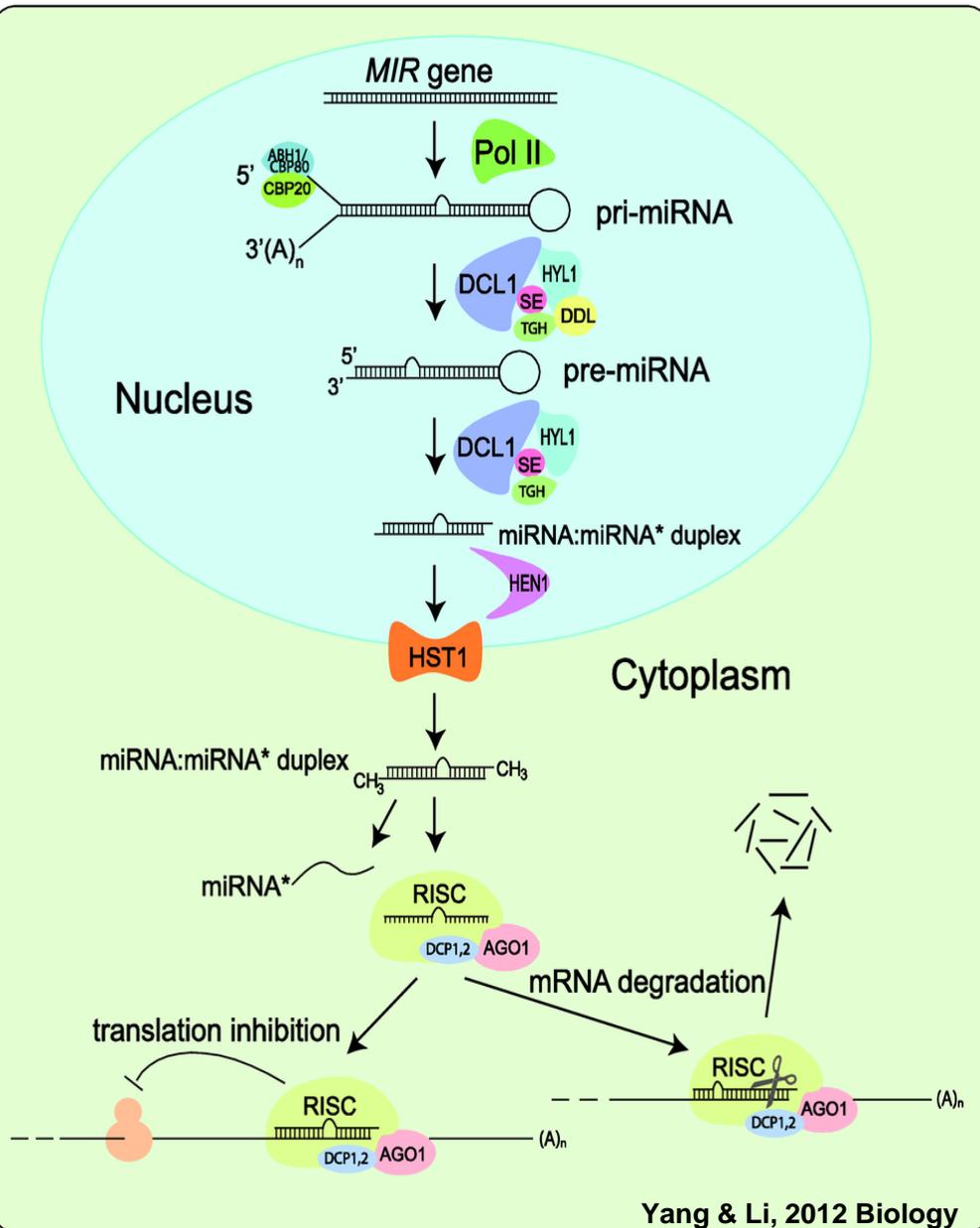
	Type of Helices
Armadillo	α -helix
HEAT	β -helix
Ankyrin	helix-turn-helix
PPR	helix-turn-helix
TPR	helix-turn-helix
LRR	β strand-turn- α helix
WD40	β -propeller
Kelch	β -propeller



NIBIO

NORWEGIAN INSTITUTE OF
BIOECONOMY RESEARCH

miRNA pathway



Plant microRNAs (miRNAs), a class of small non-coding regulatory RNAs, are canonically 20–24 nucleotides in length and bind to complementary target RNA sequences, guiding target attenuation via mRNA degradation or translation inhibition.

They should be traceable back to precursor with a hairpin structure





miRNA profiling in somatic embryos of Norway spruce under different growth temperatures

We defined about **2300** miRNA candidates with prevailing length 21 nt (41%) and 22 nt (34%).

Among them we selected 1115 highly expressed miRNAs (>100 reads).

Allowing up to 2 mismatches, we defined **522** conserved miRNAs belonged to 58 miRNA families and **593** novel miRNAs.

Differential expression analysis revealed 676 differential expressed miRNAs



Prediction of targets for highly expressed miRNAs

We defined 470 miRNAs targeting to 1139 gene models, incl.
930 annotated gene models from around 212 gene families with diverse
biological functions and
209 gene models without match at the Databases

We found high redundancy in miRNA – mRNA target pairs

We found that 317 DEGs could be regulated by several miRNAs and 290
miRNAs could regulate more than one gene



Characterization of numbers of differentially expressed miRNAs and their putative differentially expressed targets* preliminary annotated based on Pfam domains

##	Pfam ID	Pfam description	Number of miRNAs	Number of gene models
1	PF00560	Leucine Rich Repeat (LRR)	274	163
2	PF00069	Protein kinase domain	139	97
3	PF00931	NB-ARC (nucleotide-binding adaptor R-gene shared) domain	169	90
4	PF01535	Pentatricopeptide (PPR) repeat	105	84
5	PF01582	Toll-Interleukin receptor (TIR) domain	138	52
6	PF00004	ATPase family associated with various cellular activities (AAA) – incl. ATP binding disease resistance protein (TIR-NBS-LRR class)	85	52
7	PF00515	Tetratricopeptide (TPR) repeat	57	44
8	PF00637	Clathrin heavy chain/ VPS (vacuolar protein sorting-associated)	40	30
9	PF00394	Multicopper oxidase	44	15
10	PF02536	mTERF (Mitochondrial transcription termination factor)	15	11
11	PF00646	F-box domain (WD repeats)	11	11
12	PF00249	Myb-like DNA-binding domain	12	10
13	PF00847	AP2 domain	8	8
14	PF01397	Terpene synthase, N-terminal domain	8	8
15	PF00418	Microtubule-associated protein (MAP) Tau, tubulin-binding repeat	7	5
16	PF06345	DRF (Diaphanous-related formins) autoregulatory domain	11	4
17	PF00566	Rab-GTPase-TBC domain	5	4
18	PF01715	tRNA Delta(2)-isopentenylpyrophosphate transferase (IPP transferase)	15	3
19	PF00046	Homeobox domain	11	3
20	PF00201	UDP-glucuronosyl and UDP-glucosyl transferase	5	3
21	PF08263	Leucine rich repeat N-terminal domain	4	3
22	PF11721	Di-glucose binding within endoplasmic reticulum	4	3
23	PF04937	Protein of unknown function (DUF 659)	3	3
24	PF08744	Plant transcription factor NOZZLE	3	3
25	PF03110	SBP (SQUAMOSA promoter binding protein-like) domain	40	2
26	PF00106	short chain dehydrogenase	7	2

* - evaluation based on variants with negative correlation (-0.6 ... -1.0) between DE miRNAs - target gene expression patterns. Targets were predicted by psRNATarget

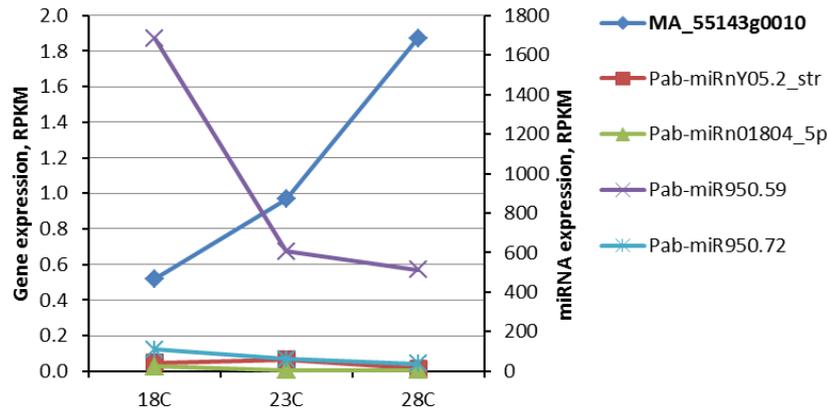


NIBIO

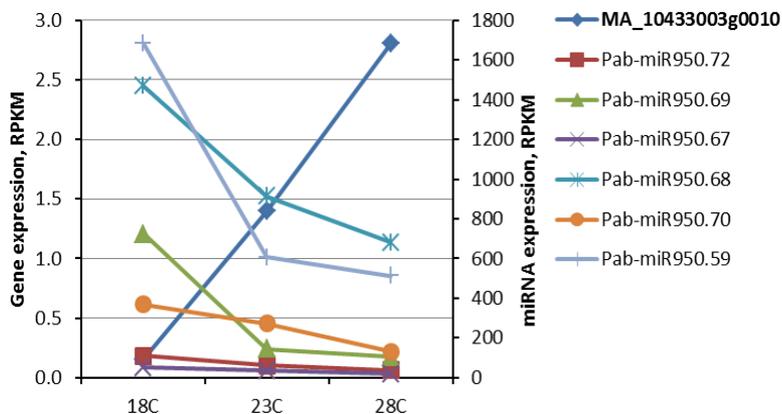
NORWEGIAN INSTITUTE OF BIOECONOMY RESEARCH

...

Comparison of transcription profiles of selected differentially expressed miRNA and their RD targets



MA_55143g0010 PF00560-Leucine Rich Repeat; PF00931-NB-ARC domain; PF01582-TIR domain; PF08937-MTH538 TIR-like domain (DUF1863); PF12799-Leucine Rich repeats (2 copies); PF12826-Helix-hairpin-helix motif; PF13306-Leucine rich repeats (6 copies); PF13504-Leucine rich repeat; PF13676-TIR domain; PF13855-Leucine rich repeat



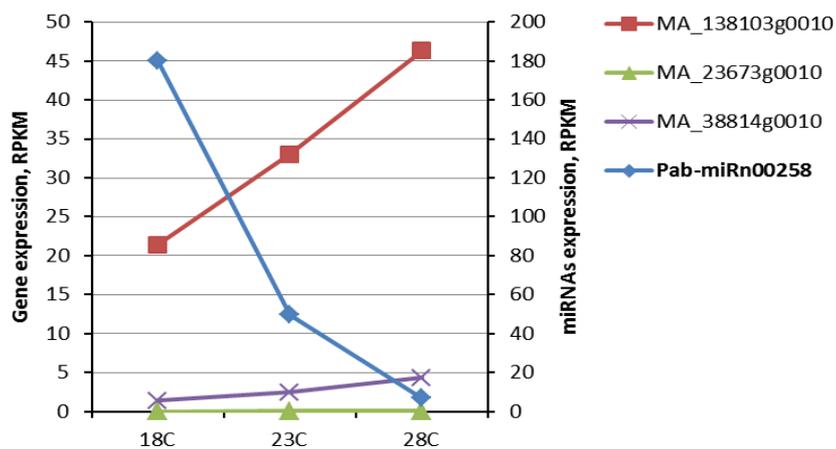
MA_10433003g0010
Disease resistance protein (TIR-NBS-LRR class) family
 PF00931 - NB-ARC domain; PF01582; PF13676 - TIR domain; PF01637 - Archaeal ATPase; PF01656 - CobQ/CobB/MinD/ParA nucleotide binding domain; PF03205-Molybdopterin guanine dinucleotide synthesis protein B; PF03308-ArgK protein; PF04665-Poxvirus A32 protein; PF05729-NACHT domain; PF06745-KaiC; PF08477-Miro-like protein; PF13191-AAA ATPase domain; PF13173; PF13207; PF13238; PF13401-AAA domain



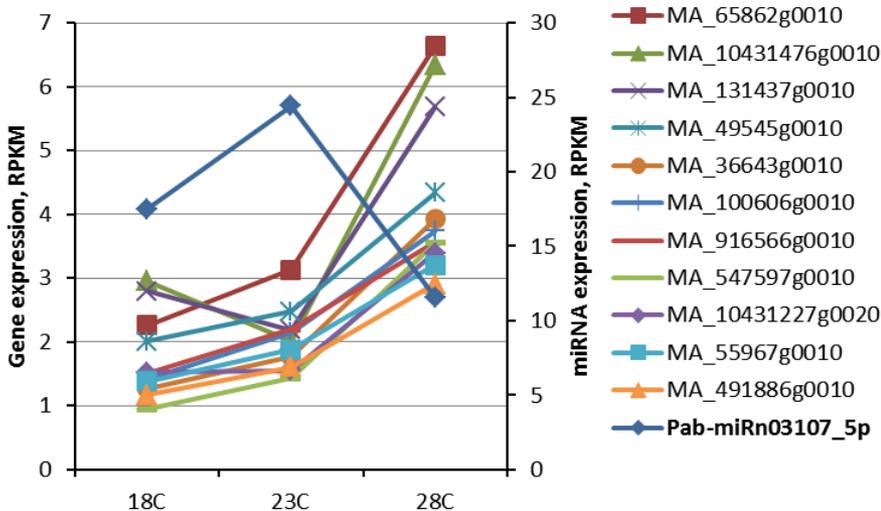
NIBIO

NORWEGIAN INSTITUTE OF BIOECONOMY RESEARCH

Comparison of transcription profiles of selected differentially expressed miRNA and their RD targets



MA_138103g0010	PF00646-F-box domain; PF12937-F-box-like
MA_23673g0010	PF00847-AP2 domain; PF02362-B3 DNA binding domain
MA_38814g0010	PF00646-F-box domain; PF01344-Kelch motif; PF07646-Kelch motif; PF12937-F-box-like; PF13964-Kelch motif



MA_65862g0010	PF00637-Region in Clathrin and VPS; PF01535-PPR repeat; PF04733-Coatomer epsilon subunit
MA_10431476g0010 MA_131437g0010	PF01535-PPR repeat; PF12854-PPR repeat; PF13041-PPR repeat family ; PF13431-Tetratricopeptide repeat
MA_49545g0010	PF01535-PPR repeat; PF12854-PPR repeat; PF13041-PPR repeat family ; PF13812-Pentatricopeptide repeat domain
MA_36643g0010	PF01535-PPR repeat; PF03704-Bacterial transcriptional activator domain; PF10602-26S proteasome subunit RPN7; PF12854-PPR repeat
MA_100606g0010	PF00418-Tau and MAP protein, tubulin-binding repeat; PF00515-Tetratricopeptide repeat; PF00637-Region in Clathrin and VPS; PF01535-PPR repeat
MA_916566g0010	PF00515-Tetratricopeptide repeat; PF01535-PPR repeat; PF02758-PAAD/DAPIN/Pyrin domain; PF04733-Coatomer epsilon subunit
MA_547597g0010	PF01535-PPR repeat; PF03704-Bacterial transcriptional activator domain; PF07443-HepA-related protein (HARP); PF07719-Tetratricopeptide repeat

Conclusions

- We showed that Norway spruce poses a variety of miRNAs and their isoforms with distinct temperature dependent expression patterns
- We defined 2050 miRNAs, from which were 1115 highly expressed in embryos, including 522 conserved and 593 novel
- They could target up to 6058 annotated gene models from around 1414 gene families with diverse biological functions and 4701 gene models without matches at the Databases.
- Annotated target genes are mostly represented by transcripts of multiple repeats proteins like TIR, NBS-LRR protein genes, PPR and TPR repeat, Clathrin heavy chains/VPS proteins, etc.



Conclusions

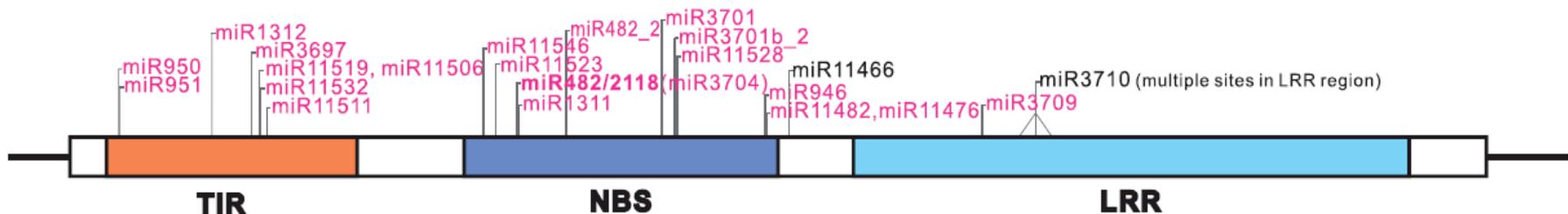
- We consider that TIR, NBS and LRR domain containing proteins could fulfill more general functions for signals transduction from external environment and conversion them into molecular response of any type.
- Unknown mechanisms provide fine-tuning of miRNAs pool content participating in developmental regulation and epigenetic memory formation.
- TRD gene models are also the source of different types of sRNAs (including miRNAs, phasi-sRNAs, etc.), which could be transported throughout the plant
- TRDs genes are highly mutable, where RD unit gain/loss may facilitate resistance to emerging pathogens.
- Taking together, our data confirm significant involvement and importance of miRNAs into post-transcriptional gene regulation and signal transduction regulation



Conclusions

1. TRD protein genes could be regulated by TF in regular manner
2. TRD protein genes could be targeted (i.e. regulated) by multiple miRNA families

Target site distribution of miRNAs identified in spruce targeting NB-LRRs. A prototypical NB-LRR gene encodes three conserved domains



3. Nearly 41 miRNA families in spruce were found to trigger phasiRNA (phased, secondary siRNAs or ta-siRs) production from diverse TRD loci which could target wide range of genes





People involved:

NIBIO: Igor Yakovlev, Vivian-Smith A. , Carneros E., and Carl G. Fossdal

NMBU (Norwegian University of life Sciences) :

Jorunn E. Olsen, Lee Y.K., Marcos Viejo



Funding: FRIPRO (NFR), EU FP7, TOPPFORSK (NFR)

