Distribution of predicted Retrotransposon LTRs in gene flanking regions of the *Pinus taeda* genome v.2.0.

Angelika Voronova[1]*, Martha Rendon[2], Par Ingvarsson[2], Dainis Ruņģis[1]

[1]Latvian State Forest research institute "Silava"; [2]Swedish University of Agricultural Sciences

*Corresponding author, E-mail: angelika.voronova@silava.lv

Introduction

Conifer genomes are large (*P.sylvestris* (2C) = 46,96 pg or 44 949 Mbp, Fuchs *et al*. 2008), are characterised by multiple gene families and pseudogenes, and contain large inter-gene regions and a high proportion of repetitive sequences. Up to 62% of the sequenced loblolly pine genome (*Pinus taeda*) consists of retrotransposon (RE) sequences and 70% of these are Long Terminal Repeat (LTR) REs (Neale *et al*. 2014). Transcription and transposition of REs is associated with stress conditions and/or meristematic tissues in various plant species. However, expression of the RE does not directly imply further transposition. In conifer genomes, it is possible to detect RE sequences co-expressed with stress associated genes (Voronova *et al*. 2013). It has been reported that transposable element (TE) composition varies considerably between individuals and can influence gene function by disruption of gene functional sequences, influencing of transcription, large insertions in introns could affect gene splicing, impact heterochromatin formation in the gene region, and play a part in functional non-coding RNA formation (Rebollo *et al.* 2012; Lisch 2013). TEs contribute to regulation of gene networks by embedding transcription factor binding sites (Feschotte 2008; Sundaram *et al.* 2014; Zhao *et al.* 2018). LTRs could contain transcription initiation and termination signals, cis-acting elements, polypurine tract (PPT), integrase binding signals, tRNA primer binding sites (Kumar, Bennetzen 1999). The aim of this study was the analysis of genes containing LTRs in flanking regions in the *Pinus taeda v.2.0.* genome.

Materials and methods

Reference genomes were downloaded from https://treegenesdb.org. The *UPPMAX* (*Uppsala Multidisciplinary Center for Advanced Computational Science*) resources with installed software were used for manipulation with genome sequences. Full-length elements were clustered with *CD-Hit* v.4.6.4 (Fu *et al.* 2012) utilising a sequence identity threshold of 0.8 resulting in 15622 TE representatives from 19700 originally found in the database. Additionally, all LTRs were extracted from the repeat database (24591 entries) and clustering performed, resulting in an additional database of LTR representatives (9659 entries).

PIER v.2.0. (*Pine Interspersed Element Resource*, Wegrzyn *et al.* 2013) was used for the LTR prediction. The *PIER* database contains all TEs and their constructs that were recognised by automatic genome annotation and a large portion of entries contain nested repeats with 2-4 pairs of direct LTRs, which negatively influence results. The database of LTR representatives resulted in sequences from 79 bp to 13295 bp in length, average length was 579 bp, median length 417 bp. In the literature, different LTR length ranges were considered for de novo LTR TE identification: 100-2000 bp (Rho *et al.* 2007), 100-3000bp (You, 2015). Several REs with LTRs that reach more than 5kB in length were reported (*tos17*, Hirochika 1992; *Sukkula*, Shirasu *et al*. 2000; *Grande*, Garcia-Martinez 2003; *Ogre*, Macas&Neumann, 2007). The presence of some longer predicted LTRs could be explained by the ubiquitous fraction of nested repeats that leads to the identification of false positive REs in the *PIER* database, resulting in false LTRs that represents the repeated bodies of REs. To partly overcome these problems, LTR sequences that were 0.1-2kb in length were used for further analyses, resulting in 9107 sequences (5.7 % reduction). BEDOPS v.2.4.35 was used for extraction of gene flanking regions (Neph *et al.* 2012). 5kB sequences from 5' and 3' flanking regions were extracted and were divided into 1kB regions. Each region had reference to its gene ID. Predicted conifer LTR sequences were BLASTed to the flanking sequences with following conditions: e-value >0.01; percent identity >80%, alignment length ≥100 bp, Query Coverage Per HSP ≥90%.

Gene Ontology (GO) classification file was obtained from Gene Ontology Consortium (http://www.geneontology.org/). Gene functional annotation for *P.taeda* v.2.0. was generated considering gene homology. Less than 50% of the genes (15534 of 36730) were categorized to any GO term. BINGO v. 3.0.3 (Maere *et al.* 2005) was used for overrepresentation tests of GO categories using custom annotation available. Cytoscape v.3.3.0. (Smoot *et al.* 2011) was used for gene network visualization. Gene networks were built and overrepresentation tests performed. Each network was formed from the group of genes that were in proximity to one RE. Network edges represent connections of GO terms, but the node size depends on the gene count categorized to a particular term, significantly enriched nodes (p-value ≤0.05) are colored (**Figure 3**). A hypergeometric test with Bonferroni correction implemented in BINGO was used for the gene enrichment test.

Table 1. Number of totally extracted flanking regions and number of high quality hits.

Genome & gene set		Flanking region from the gene start/ end coordinates									
		5' 0-1Kb	3' 0-1kB	5' 1-2 kB	3' 1-2 kB	5' 2-3kB	3' 2-3kB	5' 3-4kB	3' 3-4kB	5' 4-5kB	3' 4-5kB
P.taeda v.2.0.	Nb of extr.seq.	36726	36728	34711	34063	33184	32310	31767	30838	30349	29479
	Nb of hqh to LTRs	5851	6450	4362	3901	3750	3628	3310	3069	3202	2924
	ratio	0,16	0,18	0,13	0,11	0,11	0,11	0,10	0,10	0,11	0,10
	>50	17	22	10	10	4	2	1	0	0	0
	>100	8	9	1	0	0	0	0	0	0	0
P.taeda v.1.0. HQ genes	Nb of extr.seq.	4298	4239	4177	4128	4130	4091	4081	4028	4023	3967
	Nb of hqh to LTRs	784	779	2258	1890	3151	2693	3593	3222	3816	3539
	ratio	0,18	0,18	0,54	0,46	0,76	0,66	0,88	0,80	0,95	0,89
	>50	1	1	1	0	0	0	0	0	0	0
	>100	0	0	0	0	0	0	0	0	0	0
P.taeda v.1.0. LQ genes	Nb of extr.seq.	75425	75459	72840	72797	71554	71470	70002	69836	68237	68017
	Nb of hqh to LTRs	2317	2540	4188	4243	4979	5070	5256	5387	5645	5382
	ratio	0,03	0,03	0,06	0,06	0,07	0,07	0,08	0,08	0,08	0,08
	>50	2	2	5	5	6	5	4	7	7	6
	>100	1	1	3	4	1	1	0	1	0	0
P.lambertiana v.1.0 HQ genes	Nb of extr.seq.	8779	8778	8746	8742	8719	8708	8692	8673	8660	8640
	Nb of hqh to LTRs	71	55	163	187	278	277	315	296	355	357
	ratio	0,01	0,01	0,02	0,02	0,03	0,03	0,04	0,03	0,04	0,04
	>50	0	0	0	0	0	0	0	0	0	0
	>100	0	0	0	0	0	0	0	0	0	0
P.lambertiana v.1.0 LQ genes	Nb of extr.seq.	71162	71157	70386	70475	69773	69909	69217	69344	68660	68836
	Nb of hqh to LTRs	470	466	1063	1011	1556	10508	1789	1368	2038	1999
	ratio	0,01	0,01	0,02	0,01	0,02	0,15	0,03	0,02	0,03	0,03
	>50	0	0	1	0	4	3	6	1	7	7
	>100	0	0	0	0	0	0	0	0	0	0

*Nb of extr.seq.(Number of extracted sequences); Nb of hqh to LTRs (Nb of high quality hits to LTR); >50 (Nb of LTRs that hit more than 50 regions); >100 (Nb of LTRs that hit more than 100 regions).

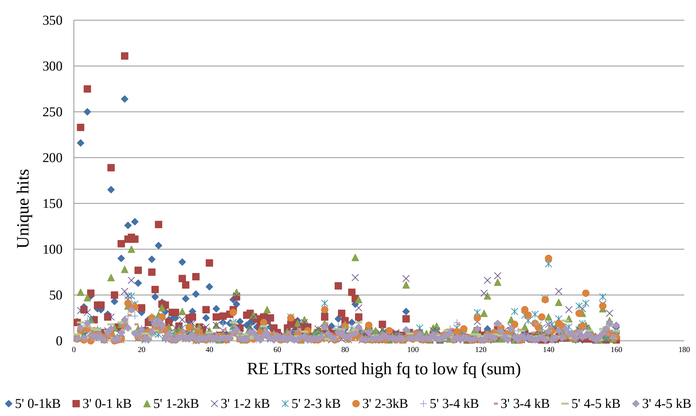

Figure 1. Distribution of all REs LTRs found in flanking gene regions of *P.taeda* v.2.0.

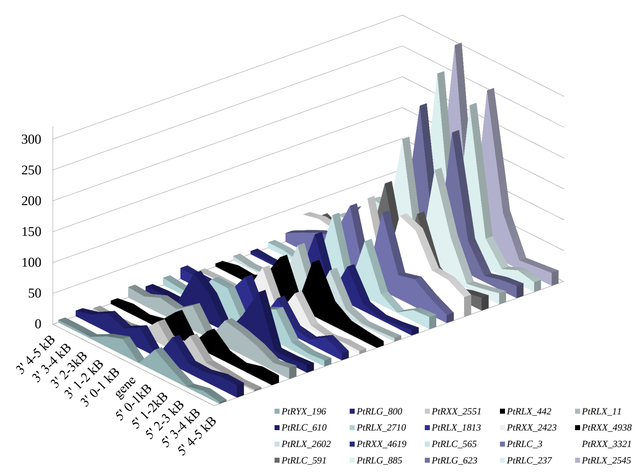

Figure 2. Distribution of 20 most frequent LTRs in gene flanking sequences of *P.taeda* v.2.0.

Results

Overview of extracted gene regions for each genome 5' and 3' flanking region is presented in **Table 1**. Ratio of hit number to the number of extracted flanking regions was similar for all regions (0.1-0.11), except for the 0-1 kB region (0.16-0.18), indicating that in proximal regions, the higher hit number is not dependent on the higher number of extracted regions. In all other analysed genomes this ratio is higher in more distant regions. In *P.taeda* v.1.0. and *P.lambertiana* v.1.0. genomes all gene sets (and especially high quality gene sets) contains significantly lower amount of total hits to LTRs in 0-1 kB regions. The number of repeats gradually increase with increasing distance from genes, which could reflect filtration results against repeats near genes. Furthermore, *P.taeda* v.1.0. gene flanking regions were enriched only with one RE family. The *P.lambertiana* high quality gene set lacks any RE family was more frequently present in any gene flanking region studied. These differences indicate that sequence scaffolding workflows and assigned quality frames greatly influence the results associated with highly repetitive sequences.

The genome of *P.taeda* v.2.0. was produced using *Single molecule Real time sequencing technology* (PacBio) by assembling long and short reads resulting in a significant increase in quality (Zimin *et al.* 2016). 160 predicted LTRs that were present in all ten regions were analyzed, each LTR hit from 19 to 820 unique gene flanking regions. 40 LTRs were significantly (p=0.001, t-test) more frequent in the 0-1 kB region. Frequencies of seven LTRs were significantly increased in the 0-2 kB region (p=0.001). Gene groups defined by presence of particular LTRs were further functionally analysed using Gene Ontology terms (**Table 2**). The largest molecular function group present in all identified gene networks was binding, catalytic activity; transferase activity. In biological process category largest group belongs to primary metabolic process involved genes. In many RE depentant networks genes were involved in defence response, methylation, photosynthesis; transmembrane transport; cell wall organization or biogenesis, cellular homeostasis and small molecule metabolism.

Analysis of predicted LTRs reveal that some of them still represents internal RE sequences and more careful analysis of flanking sequences should be performed in order to evaluate respective LTRs. It was reported that the proportion of single LTRs in gymnosperms genomes is reduced compared to angiosperm genomes. Therefore full-length and chimeric elements are more widely distributed and could be expected to present more often in gene flanking regions. Furthermore, REs in gymnosperms are older and more diverged and therefore a smaller number of genes flanked by similar REs with large insertions/deletions may be identified if full-length elements are analysed. The shorter predicted LTR sequences used in this study reveal a potential gene set that could be further analysed in greater detail. 4 from 10 predicted LTRs presented in **Table 2** contained a PPT. *PtRLX_1813* contains a PPT and two AG-tracts in the reverse strand and one promoter on the direct strand. The $(AG)_4$A motif was found to be one of the most common TFBS for plant promoters (Liu *et al.* 2013), including in WRKY gene promoters, that are known regulators of plant responses to biotic and abiotic stresses (Zhang&Wang, 2005), and participate in light-responsive phototransduction processes in plants (Parida *et al.* 2009). *PtRLX_1813* contains two AG-rich tracts*: $(AGNN)_3(AG)_3(NNAG)_2$ and $(AGNN)_2(AGN)_4$*. For *PtRLX_11* all characteristic features of LTRs were detected (integrase binding signals; sense and antisense promoter (TATA-box); PPT, as well as several important cis-acting elements and TFBS (TCP, W-box*). PtRLC_3* also contain integrase binding signals, PPT, sense and antisense promoters, as well as G-box and ABRE motifs on both strands. G-box, ABRE motifs and GA-box were found frequently in plant gene promoters that are involved in metabolic processes (*Liu et al.* 2013). In this study *PtRLC_3* LTR was present in the vicinity of 337 genes, some of them are involved in kynurenine metabolism, L-phenylalanine biosynthesis, cellular metal ion homeostasis, response to water, flavonoid biosynthesis, xyloglucan metabolism, protein ubiqitination (**Table 2**, **Figure 3**).

Discussion

Automated TE detection relies on several strategies such as search for sequences with homology to known elements, de novo detection of repeated elements, detection of specific structural features and combined techniques (Bergman, Quesneville 2007). In the genome assembly process, repeated and highly similar sequences like TEs are causative of errors and gaps. There are 261 alternative reference genomes for human genome and in spite of that there are still unresolved hypervariable loci and different type of errors present (Chaisson *et al.* 2015). The six times larger pine genome was sequenced relying on NGS technologies and the first version of the *Pinus taeda* genome was announced in 2014 (Zimin *et al.*). The *P.taeda* v.2.0. genome is currently of the best quality regarding scaffold length. Version 2.0. was improved basically by merging very small contings (shorter than 500 bp) and while the first version of *P.taeda* genome contained 16.5 million contigs, the second version contains 2.9 million contigs (Zimin *et al.* 2016). The average *PacBio* read length used to reconstruct genome v.2.0. was 9665 bp, which is smaller than many REs. Mean repeat length in *PIER* database was 6273, median - 5383, but 3046 repeats are longer than 10 Kbp, and the longest construct was 35042 bp. Therefore ambiguous reconstruction of LTR versus internal repeat sequence is still possible. Some repeats were on opposite strands and strand-specific sequencing approaches should be used in order to evaluate correct positions. Separated flanking regions (1Kbp) were analysed in order to investigate location of the repeats. RE enrichment in 0-1 Kbp and 1-2 Kbp at 5' and 3' flanking gene regions was statistically significant. Internal sequences of five RE families were found frequently in vicinity to genes from 5' and 3' flanks, while five other repeats contains characteristic features of LTRs (TG-CA; polipurine tract, TATA-box). Detailed analysis of flanking regions revealed REs without a second LTR, and it is possible that the second LTR is lost during sequence reconstruction or was deleated via non-homologous recombination for some reasons. Presence of identical repeat sequnces in flanking region of multiple genes could indicate possible coordinated regulation, but the detailed structure of flanking repeats will be investigated in further analyses. We identified 10 highly represented RE families, which could be expected considering the larger genome size of gymnosperms. Smaller gene networks not presented here were identified and analysed. About one half of all *P.taeda* genes lack functional annotation, despite this fact stress responsive gene networks were identified.

Figure 3. Part of gene network containing *PtRLC_3* LTR in 0-1 Kbp 5' and 3' flanking region.

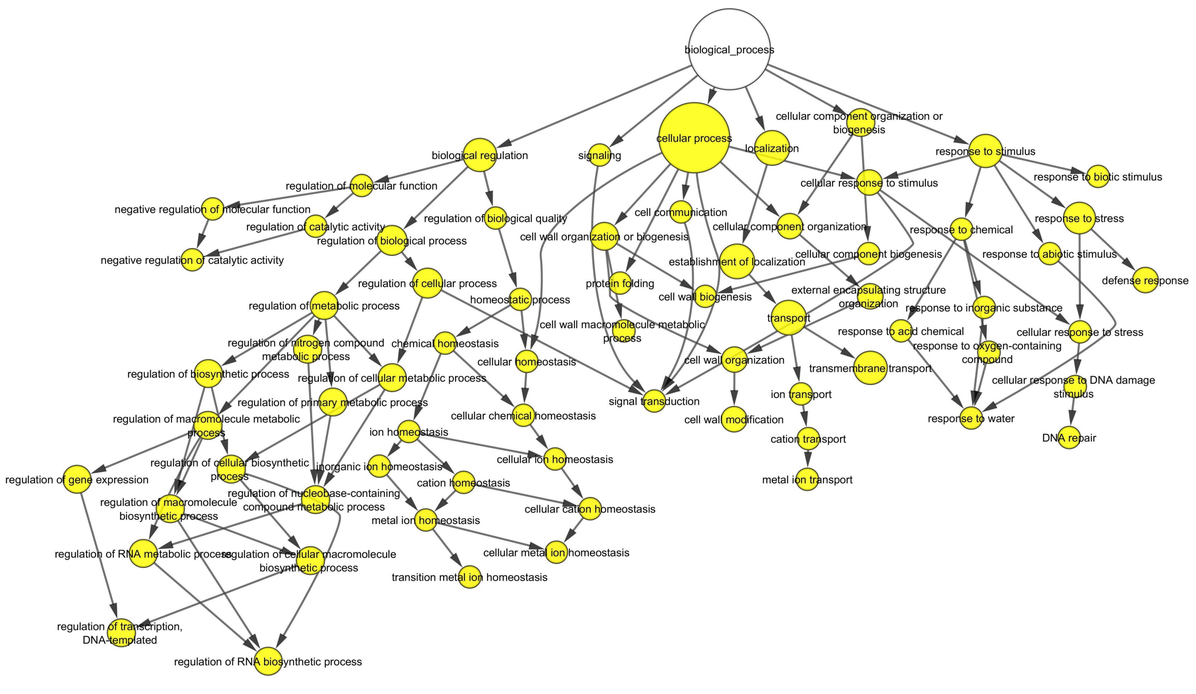

Table 2. Distribution of 10 predicted repeats that were most frequent (hit >200 unique genes) and statistically enriched in 0-1 Kbp flanking regions. Flanking region gene functional description (GO terms) included.

LTR ID	Biological process	Molecular function	Cellular compartment	Count of unique matching gene flanking regions										Total	t-test: t>5,04(p=0,001); t>2,31 (p=0,05)		
				5' 0-1 kB	3' 0-1 kB	5' 1-2 kB	3' 1-2 kB	5' 2-3 kB	3' 2-3 kB	5' 3-4 kB	3' 3-4 kB	5' 4-5 kB	3' 4-5 kB		0-1Kbp vs 2-5 Kbp	0-2 Kbp vs 3-5 Kbp	3' vs 5' 0-1Kbp
PtRLG_623	telomere capping; rRNA methylation; ornithine metabolic process; apoptotic process; vacuole fusion; glucose homeostasis; production of siRNA involved in RNA interference; deoxyribonucleoside diphosphate metabolic process; negative regulation of flower development; brassinosteroid biosynthetic process; glutamate metabolic process; hexose metabolic process; cytokinin metabolic process; lignin biosynthetic process; response to water; coenzyme A metabolic process; microtubule-based movement; protein glycosylation;	rRNA (adenine-N6,N6-)-dimethyltransferase activity; glycerol-3-phosphate O-acyltransferase activity; ornithine carbamoyltransferase activity; UDP-glucose 4-epimerase activity; chaperone binding; serine-type peptidase activity; damaged DNA binding; tubulin binding; pectinesterase activity; ubiquitin protein ligase activity; cytokinin dehydrogenase activity; cellulose synthase activity;	nuclear chromosome, telomeric region; RISC complex (RNAi effector); HOPS complex (tethering); Cdc73/Paf1 complex (transcription elongation); ER membrane; kinesin complex(microtubule); photosystem II oxygen evolving complex.	216	233	53	33	10	11	17	14	20	14	621	21,52	2,27	0,65
PtRLG_885 (0-1kb)	lignin catabolic process; lipid phosphorylation; strigolactone biosynthetic process; regulation of hormone levels; response to auxin; cellular response stimulus (DNA damage, stress) ; pollen-pistil interaction; cell communication;pectin biosynthetic process; nucleosome assembly;peptide transport; multicellular organism development;nucleobase-containing small molecule metabolic process; peptidyl-tyrosine phosphorylation;response to chemical; DNA recombination; protein folding; protein ubiquitination; plant-type cell wall organization; ubiquitin-dependent protein catabolic process;	hydroquinone:oxygen oxidoreductase activity; GDP-D-glucose phosphorylase activity; phosphatidylinositol phosphate kinase activity; calcium-dependent phospholipid binding; ADP binding; lyase activity; transferase activity, transferring glycosyl groups; channel activity;	-	165	189	69	26	7	7	12	14	18	13	520	11,76	2,59	1,16
PtRLG_885 (1-2 kb)	defense response; response to biotic stimulus; aromatic amino acid family biosynthetic process; nucleocytoplasmic transport; RNA phosphodiester bond hydrolisis, endonucleolytic; proteolisis; protein phosphorilation; ox-red process; nucleocytoplasmic transport; chitin catabolic process; carbohydrate metabolic process; RNA-dependant DNA biosynthetic process; aromatic amino acid family biosynthetic process; negative regulation of catalitic activity; cell wall modification; signal transduction	alpha-L-fucosidase activity; 3-deoxy-7-phosphoheptulonate synthase activity; Ran GTPase binding; chitin binding; ADP binding; ATP binding; heme binding; zinc ion binding; protein kinase activity; carbohydrate binding; RNA-DNA hybrid ribonuclease activity; terpene synthetase activity; pectinesterase activity;	-														
PtRLC_565	(-)-pinoresinol metabolic process; cell wall organization or biogenesis; methylation; chitin metabolic process; plant-type cell wall organization; gene expression; defense response; oxoacid metabolic process; hormone catabolic process; cell wall organisation	pinoresinol reductase activity; aminoacylase activity; terpene synthase activity; NAD binding; DNA-directed 5'-3' RNA polymerase activity; coenzyme binding; chitinase activity; zinc ion binding; terpene synthetase activity; heme binding; ADP binding; GTP binding; xyloglucan:xyloglucosyl transferase activity;	external encapsulating structure; Golgi apparatus; plasma membrane; mitochondrion; endoplasmic reticulum; apoplast; nucleus	90	106	18	8	1	2	18	9	20	9	281	10,58	1,83	1,41
PtRLX_2545 (0-1 Kb)	ornithine metabolic process; cellular glucose homeostasis; brassinosteroid biosynthetic process; cytokinin metabolic process; L-phenylalanine biosynthetic process; rRNA modification; apoptotic process; alpha-amino acid biosynthetic process; indolalkylamine biosynthetic process; vacuole fusion; symbiont process; production of siRNA involved in RNA interference; deoxyribonucleoside diphosphate metabolic process; nucleotide metabolic process; negative regulation of reproductive process; negative regulation of molecular function; protein ubiquitination; sulfate transmembrane transport; regulation of cell size;	rRNA (adenine) methyltransferase activity; dihydrofolate reductase activity; glycerol-3-phosphate O-acyltransferase activity; ornithine carbamoyltransferase activity; prephenate dehydratase activity; tryptophan synthase activity; hexokinase activity; UDP-glucose 4-epimerase activity; oxidoreductase activity, acting on the CH-NH group of donors; chaperone binding; enzyme inhibitor activity; phosphoglycerate mutase activity; ATPase activity; ubiquitin protein ligase activity;	RISC complex; HOPS complex; Cdc73/Paf1 complex; photosystem II oxygen evolving complex; intrinsic component of membrane; vacuole; protein (kinesin) complex; cell-cell junction; nucleosome, apoplast;	264	311	78	54	15	16	18	20	21	23	820	10,55	2,44	1,41
PtRLX_2545 (1-2 Kb)	phosphoenolpyruvate-dependent sugar phosphotransferase system; tRNA aminoacylation for protein translation; phosphorelay signal transduction system; methylation; response to chemical; drug transport; vesicle-mediated transport; protein folding; regulation of gene expression; proteolysis involved in cellular protein catabolic process; hydrogen peroxide metabolic process; protein phosphorylation; carbohydrate metabolic process;	active transmembrane transporter activity; glutathione peroxidase activity; aminoacyl-tRNA ligase activity; methyltransferase activity; cysteine-type endopeptidase activity; DNA binding transcription factor activity; drug transmembrane transporter activity; unfolded protein binding; ATPase activity, coupled to transmembrane movement of substances; phosphatase activity; passive transmembrane transporter activity;	extracellular region part; plastid; integral component of membrane; endoplasmic reticulum; lysosome;														
PtRLC_3 (0-1 Kb)	kynurenine metabolic process; L-phenylalanine biosynthetic process; cellular metal ion homeostasis; response to water; flavonoid biosynthetic process; response to biotic stimulus; cell wall modification; xyloglucan metabolic process; protein folding; protein ubiquitination; proteolysis; DNA repair;	dioxygenase activity; ATPase activity; catechol oxidase activity; prephenate dehydratase activity; exonuclease activity; protein heterodimerization activity; nutrient reservoir activity; protein dimerization activity; ubiquitin-like protein transferase activity;	endoplasmic reticulum; endoplasmic reticulum subcompartment; external encapsulating structure; nucleosome; apoplast; mitochondrion;	126	111	38	49	45	41	27	28	13	14	492	10,30	2,38	-1,29
PtRXX_3321 (0-1Kb)	regulation of seed germination; glutathione catabolic process; leaf morphogenesis; plant-type cell wall organization; lipid catabolic process; dUTP catabolic process; pollen tube growth; cellular response to chemical stimulus; methionyl-tRNA aminoacylation; regulation of defense response to fungus; ammonium transmembrane transport; protein folding; polysaccharide catabolic process; oxoacid metabolic process; chitin metabolic process; ubiquitin-dependent protein catabolic process; DNA replication; transmembrane transport; RNA phosphodiester bond hydrolysis, endonucleolytic; microtubule-based process; DNA repair;	gamma-glutamylcyclotransferase activity; raffinose alpha-galactosidase activity; methionine-tRNA ligase activity; NAD-dependent histone deacetylase activity (H3-K14 specific); dUTP diphosphatase activity; 4-hydroxyphenylpyruvate dioxygenase activity; glycopeptide alpha-N-acetylgalactosaminidase activity; peroxidase activity; ammonium transmembrane transporter activity; unfolded protein binding;beta-amylase activity; polygalacturonate 4-alpha-galacturonosyltransferase activity; structural constituent of cytoskeleton; chitin binding; ATPase activity, coupled to transmembrane movement of substances; endoribonuclease activity, producing 5'-phosphomonoesters;	apoplast; lytic vacuole; trans-Golgi network; chromosome; Golgi membrane; symplast;	111	113	100	66	49	35	47	40	32	34	627	9,51	5,26	0,20
PtRXX_3321 (1-2 Kb)			plant-type cell wall; extracellular region; pollen tube; microtubule; apoplast; mitochondrion;														
PtRLX_1813	transcription, DNA-templated; peptide transport; phosphate-containing compound metabolic process; cellulose biosynthetic process; protein dephosphorylation; response to stress; small GTPase mediated signal transduction; response to oxidative stress; cytoskeleton organization; regulation of transcription, DNA-templated; ion transport;	GTP binding; polygalacturonase activity; polysaccharide binding; calmodulin binding; protein serine/threonine phosphatase activity; cellulose synthase activity; DNA-directed 5-3' RNA polymerase activity; DNA binding; heme binding;	microtubule; nucleus; integral component of membrane; plasma membrane	48	56	13	18	8	16	24	16	11	18	228	8,63	1,69	1,66
PtRLC_591 (0-1 Kb)	molybdate ion transport; protein dephosphorylation; megagametogenesis; fucosylation; macromolecule metabolic process; RNA splicing; cellular water homeostasis; pectin biosynthetic process; nucleosome organization; response to oxidative stress; regulation of transcription, DNA-templated; multi-organism reproductive process; photosynthesis; translation; transport;	protein serine/threonine phosphatase activity; alpha-(1,2)-fucosyltransferase activity; molybdate ion transmembrane transporter activity; water channel activity; cooper ion binding; UDP-glycosyltransferase activity; RNA-directed DNA polymerase activity; DNA binding, RNA binding, calcium ion binding; structural constituent of ribosome; ATP binding; transmembrane transporter activity;	spliceosomal complex; photosystem II oxygen evolving complex; mitochondrion; ribosomal subunit; Golgi membrane; nucleosome; plasmodesma;	104	127	28	13	7	14	16	9	21	22	361	8,51	1,90	1,80
PtRLX_11	L-phenylalanine biosynthetic process; ubiquitin-dependent protein catabolic process; methylation; transmembrane transport; regulation of transcription, DNA-templated; oxidation-reduction process;	prephenate dehydratase activity; manganese ion binding; ubiquitin protein ligase binding; protein binding	-	42	40	36	21	27	26	13	15	18	17	255	8,22	3,09	-0,63
PtRXX_4938	lipid transport; megagametogenesis; protein folding; vacuolar transport; transport; nucleoside metabolic process; plant-type cell wall organization; cell wall organization; ubiquitin-dependent protein catabolic process;	-	integral component of membrane; cell wall; plant-type cell wall;	86	68	32	23	14	9	10	11	7	5	265	6,75	2,88	-2,13

This research was supported by The State Education Development Agency 1.1.1.2. "Post-doctoral Research Aid". Nb. *1.1.1.2/VIAA/1/16/094.*

PostDoc Latvia

NACIONĀLAIS ATTĪSTĪBAS PLĀNS 2020

EIROPAS SAVIENĪBA
Eiropas Reģionālās attīsitības fonds

IEGULDĪJUMS TAVĀ NĀKOTNĒ