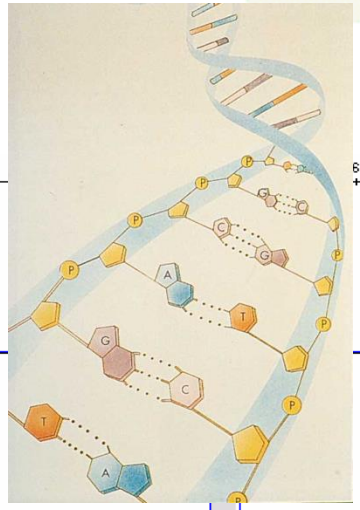
Genetic and genomic research in Latvia

Dainis Ruņģis

Genetic Resources Centre

dainis.rungis@silava.lv

www.silava.lv



Genetic Resource Centre

- This is a facility set up in 2006 at LSFRI Silava with funding from the Ministry of Agriculture
- This centre incorporates the Latvian Gene Bank, the Gene Bank database, and a genetic analysis laboratory

LATVIJAS REPUBLIKAS ZEMKOPĪBAS MINISTRIJA

Molecular tools

- Neutral DNA markers
- Expressed sequence tags
- Genome sequences
- Functional determination and validation
- Epigenetic effects (‘missing heritability’)

Application of molecular tools

- Genetic diversity
- Provenance determination
- QTL mapping
- Markers linked to traits
- Genomic selection

Background

- High genetic diversity within breeding programs
- Tree breeding programs differ from crop breeding
- Fst vs Qst
- Functional studies (gene expression) often rely on one individual
- Difficult to transfer functional studies to (breeding) populations
- Survey of ‘alternative’ genetic diversity in natural populations and breeding material

Latvian perspective

- Pine, (spruce, birch)
- Emphasis on endemic species and populations
- Proportionally large areas reforested with selected material

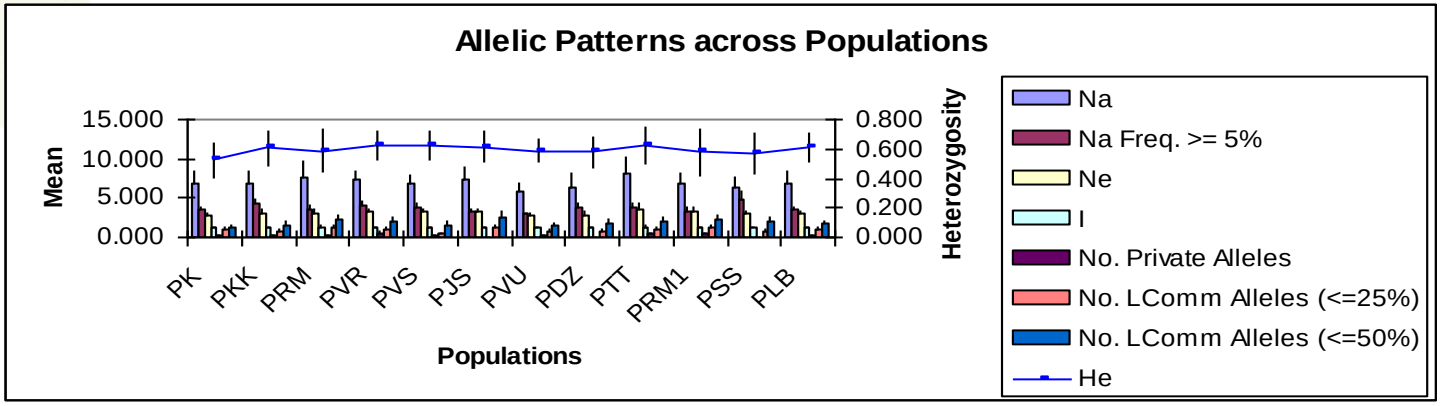
Darmstadt pines (Germany) vs a local provenance

Use of neutral DNA markers

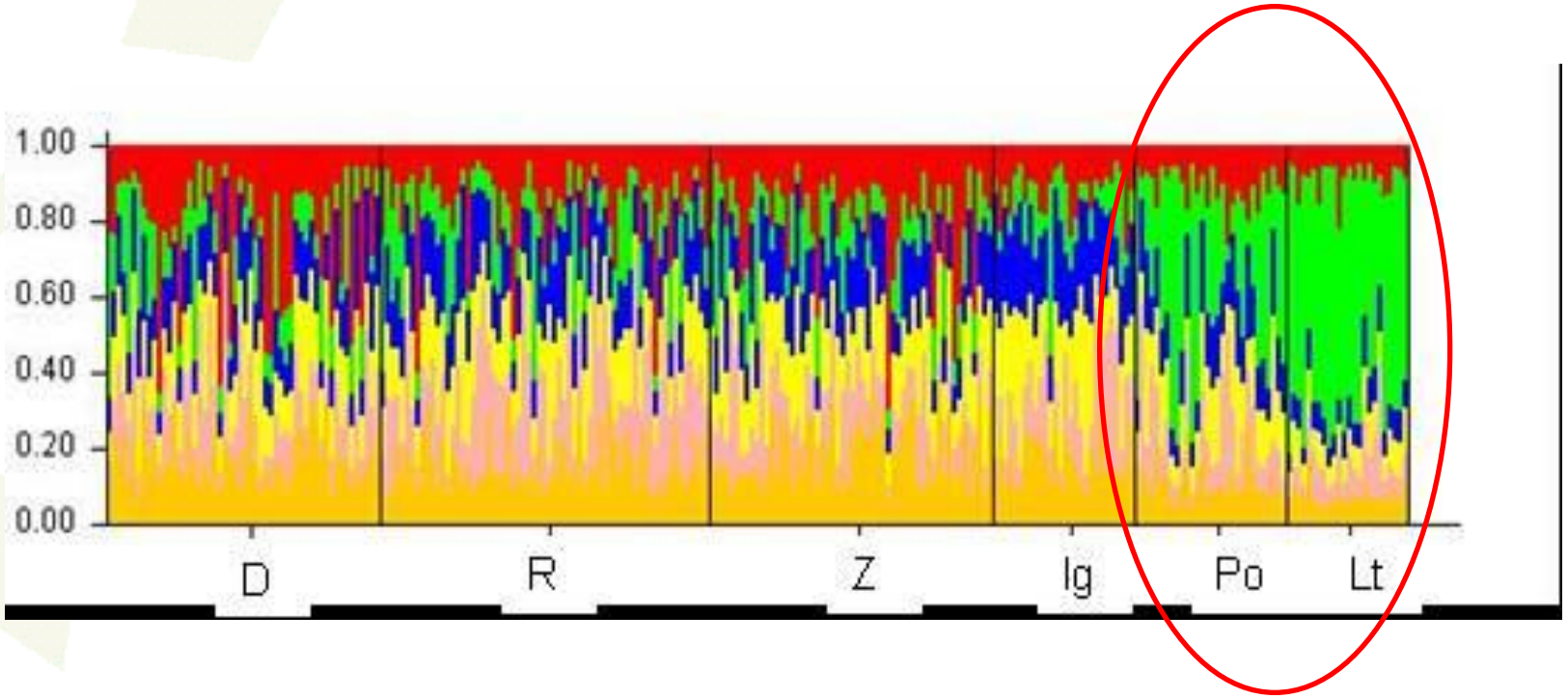
- Population studies of pine, spruce, birch, oak
- Nuclear SSR markers, chloroplast and mitochondrial markers
- Identification of hybrids (*Alnus* sp.)
- Fingerprinting of clones (pine, spruce, hybrid aspen)

Nuclear SSR markers

For pine and spruce:
-No differentiation found between Latvian populations and also neighbouring regions
- Genetic diversity comparisons between natural populations, seed plantations and breeding material
- Clone identification in seed plantations

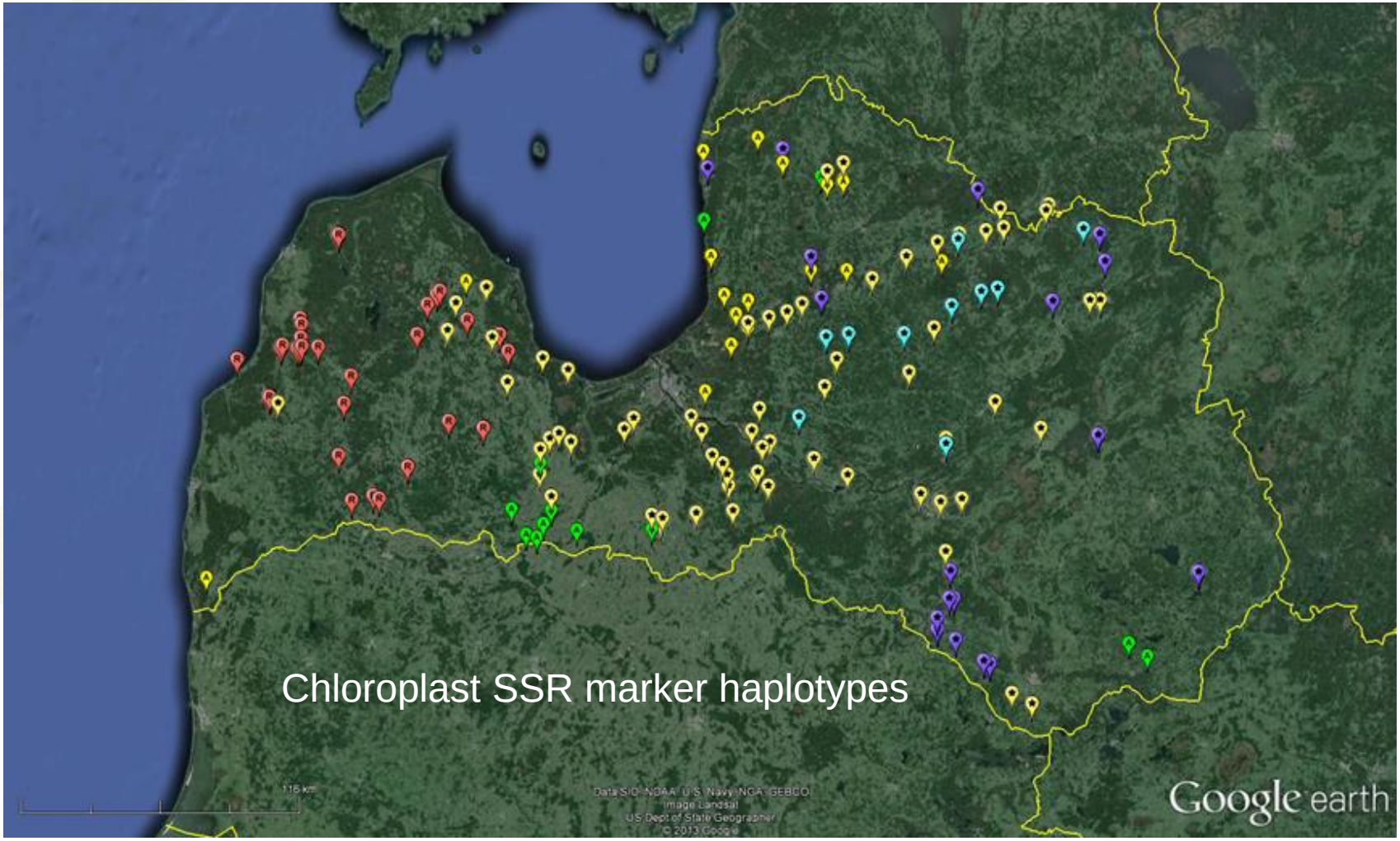


Birch nuclear SSR markers



Latvian birch breeding zones – south, west, north; Estonian; Polish; Lithuanian

Oak (*Quercus robur L.*)



Sources of ‘alternative’ genetic variation

- Conifer genomes are large
- Different to angiosperm genomes (no polyploidy)
- Phenotypic plasticity

- Epigenetic effects and ‘missing heritability’

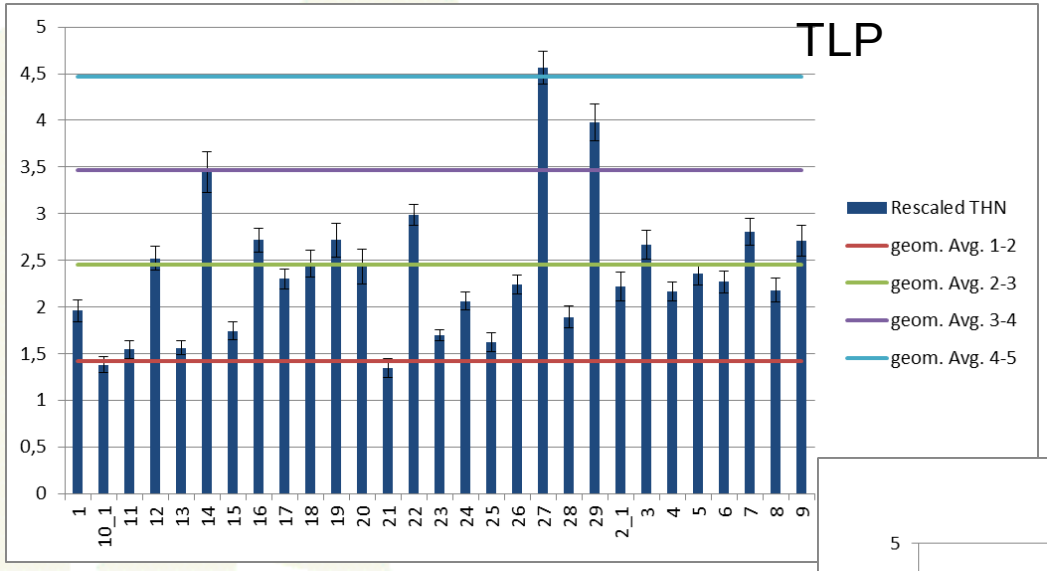
‘alternative’ genetic variation

- Research in Scots pine
  - Gene copy number variation (CNV)
  - microRNAs
  - retrotransposons
- Results from two projects
  - *Heterobasidion annosum* (Vilnis Šķipars)
  - Retrotransposon activation and distribution (Angelika Voronova)

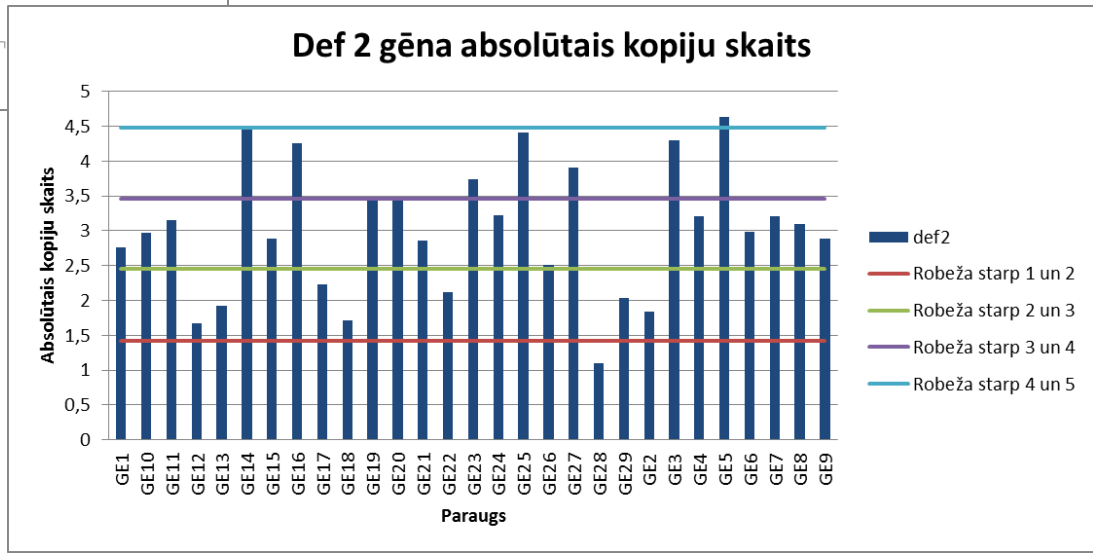
Heterobasidion resistance

- Candidate gene analysis
  - Thaumatin-like protein (TLP) gene
  - Defensin 2 gene (Def2)
  - *Pinus sylvestris* resistance gene (PsACRE)
  - Pinosylvin synthase gene (PsBBs)

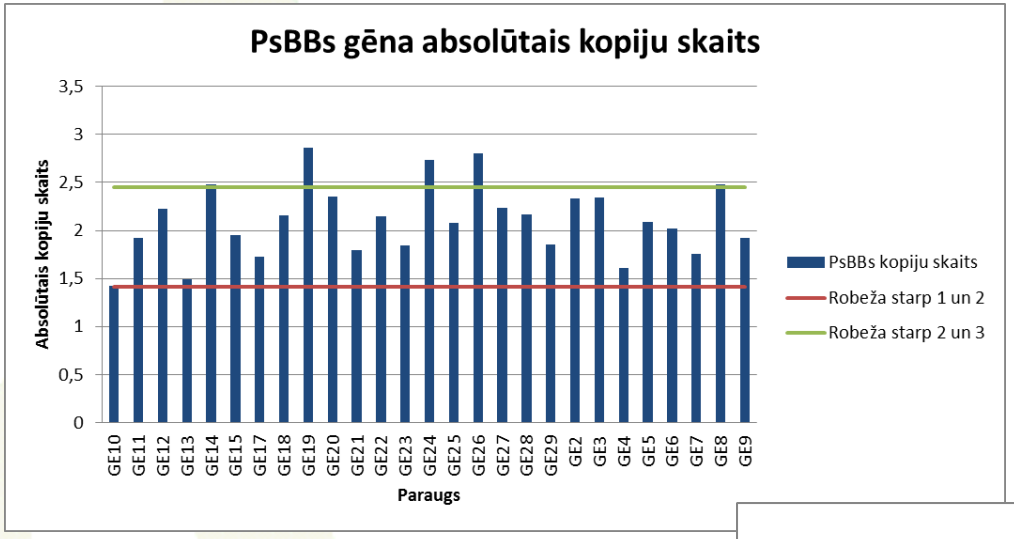
Candidate gene copy number variation analyses

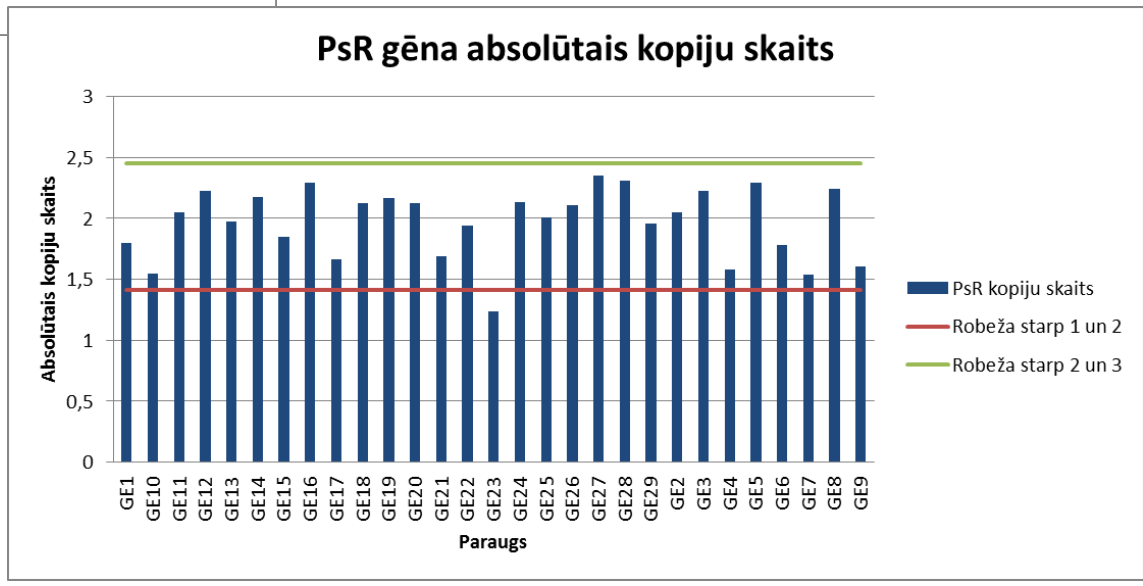


Endogēnās kontroles ΔCt_{TLP}=5,33 ; ΔCt_{GAPDH}=2,33;
ΔCt_{TLP}=4,42 ; ΔCt_{PsR}=0,92.

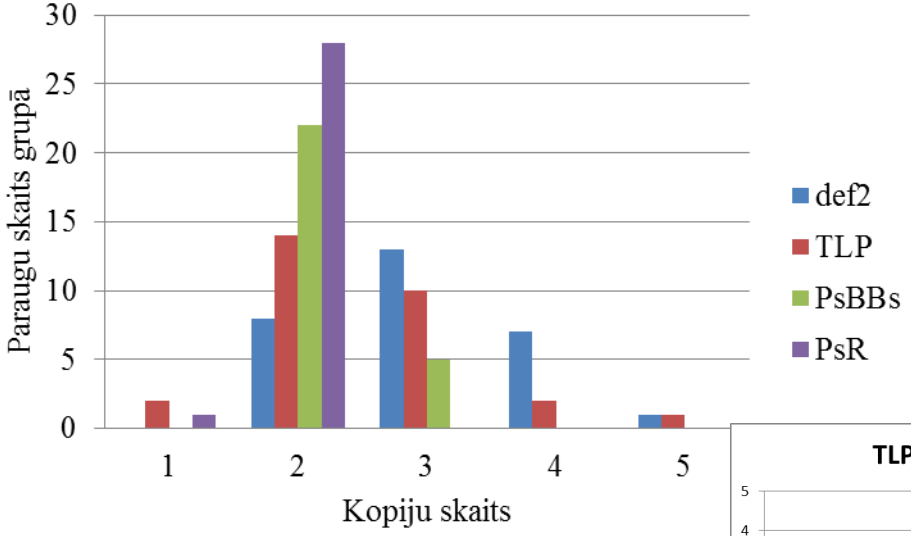


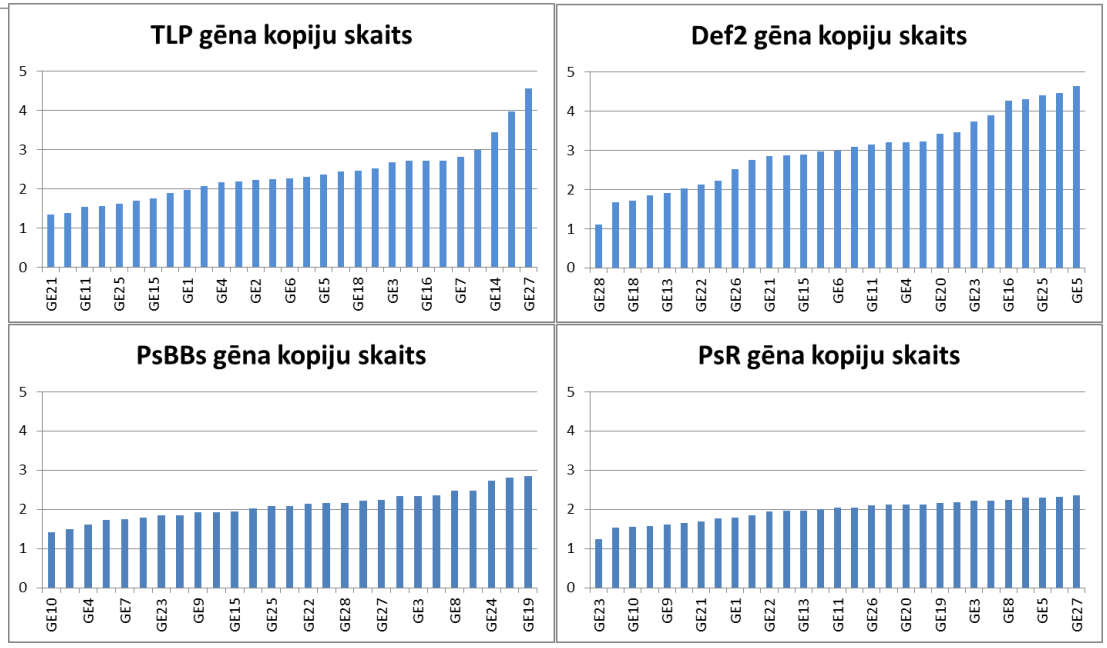
Candidate gene copy number variation analyses



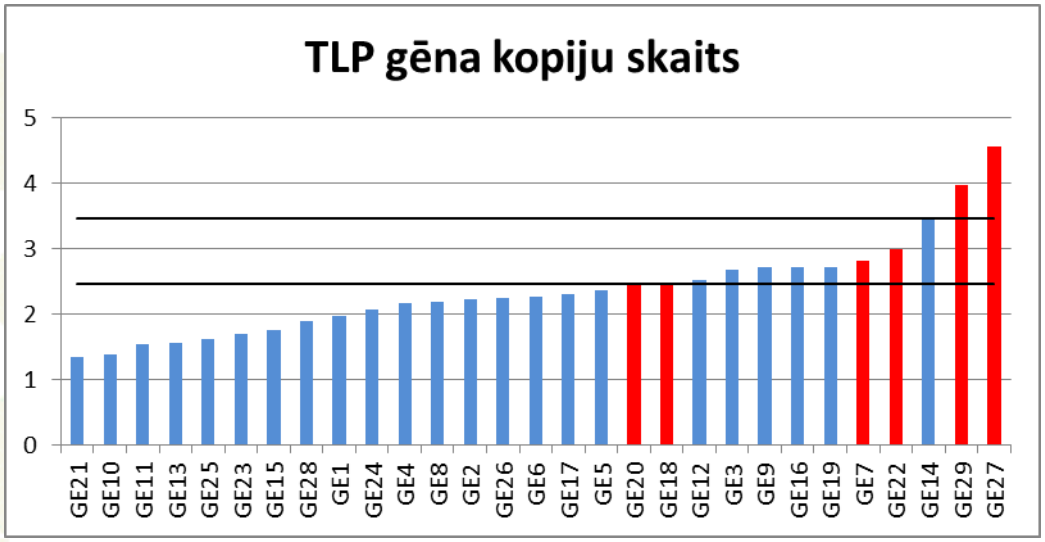


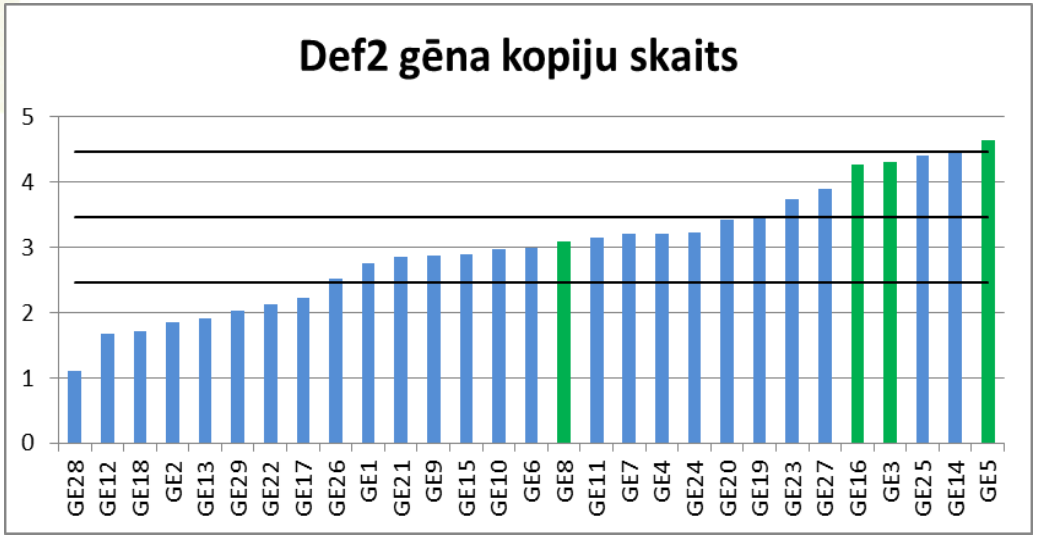
Candidate gene copy number variation analyses



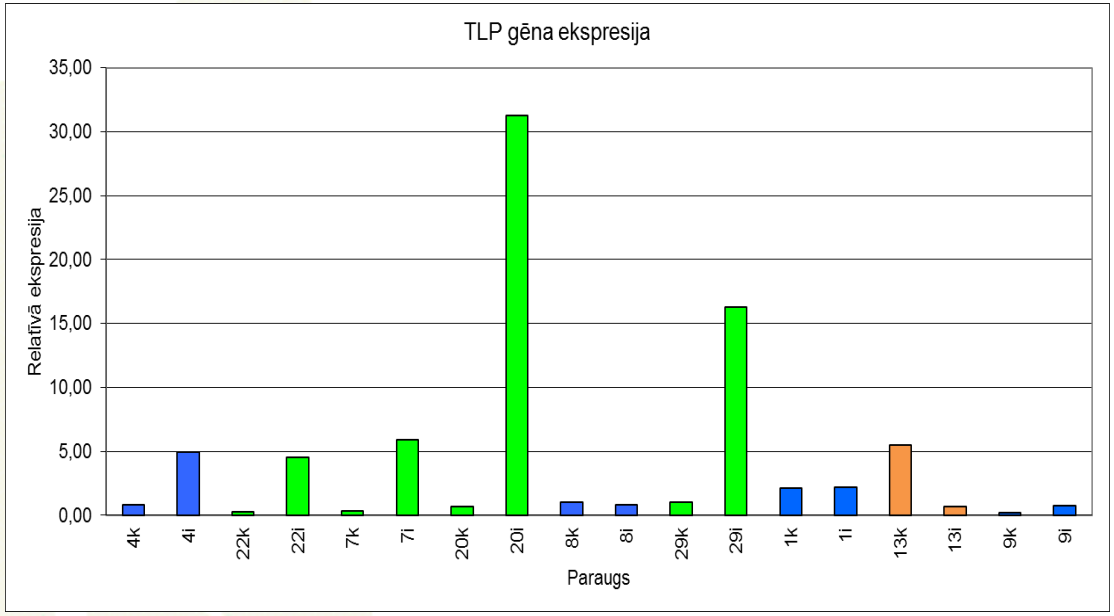


CNV and family structure





TLP gene expression



Paraugs	Relative gene expression	Control:inoculated
4k	0,78	
4i	4,90	6,30
22k	0,27	
22i	4,52	16,49
7k	0,30	
7i	5,89	19,30
20k	0,64	
20i	31,28	48,57
8k	1,00	
8i	0,82	0,82
29k	1,03	
29i	16,27	15,85
1k	2,09	
1i	2,20	1,06
13k	5,48	
13i	0,67	0,12
9k	0,16	
9i	0,74	4,60

Endogenous control - GAPDH

Candidate gene CNV

	Candidate gene	CNV analysis
Defense and PR genes	Thaumatin – like protein gene	CNV detected
	Defensin 2	CNV detected.
	Pinus sylvestris resistance gene	No CNV
	Pinosylvin synthase	Probably no CNV
	ACS1	Probably no CNV
	ACS2	Probably no CNV
	CYP	Possible CNV
	Defensin 1	Probably no CNV
	Dehydrin 2	Probably no CNV
Lignin biosynthesis	Pal1	Probably no CNV
	C4H	Possible CNV
	C3H	Probably no CNV
	4CL	Probably no CNV
	CCoAOMT	Possible CNV
	COMT	Probably no CNV
	CCR	Poor amplification
	CAD	Poor amplification
TF	MYB4	Probably no CNV
	MYB2	Possible CNV
	WRKY	Probably no CNV
	SCARECROW	Probably no CNV

Gene CNV

- Research is continuing to determine the extent of CNV in pine populations
- To date, mainly focussed on resistance genes
- Do different types of genes have different levels and extent of CNV?

Retrotransposon variation

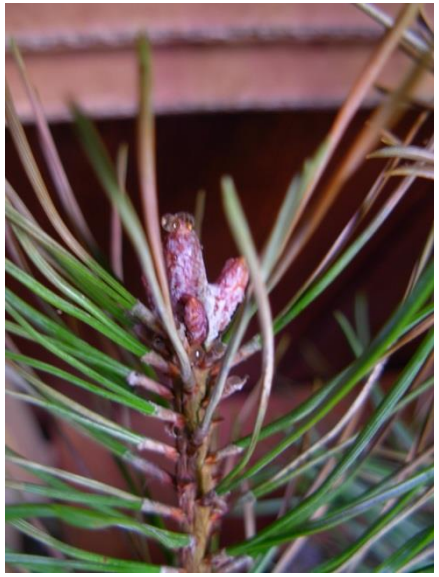
- mobile genetic elements
- Transposition via RNA intermediate
- the largest component of plant genomes (15-90 %)
- cluster formation

rearrangement of genome (instability), - functional mutations,- recombination process,- genome structure upkeep

- High level of expression observed in stress conditions & in developing tissues *(McClintock, 1984, Peschke et al.1987, Grandbastien et al. 1998, Hirochika et al. 1993, Poteau et al.1994, Ramallo et al., 2008).*

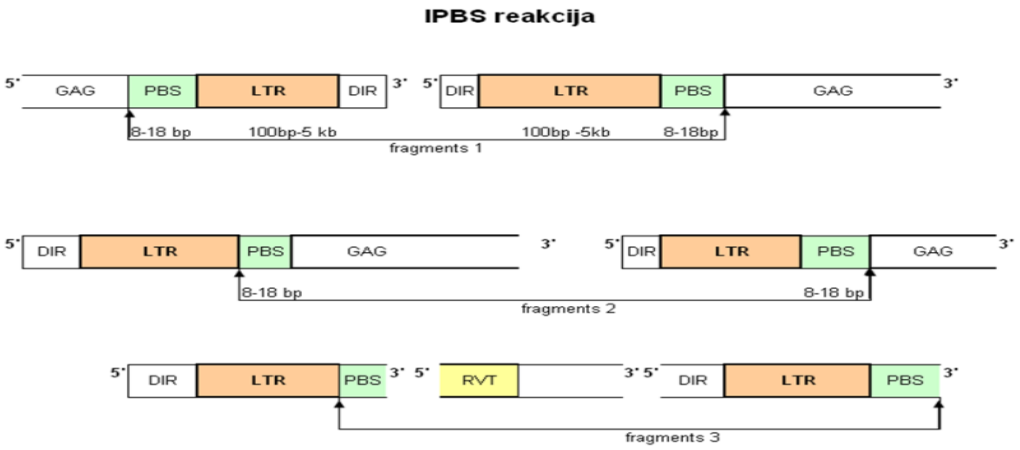
Adaptive selection for larger genomes with higher retrotransposon activity?

Scots pine ramets and stress initiation

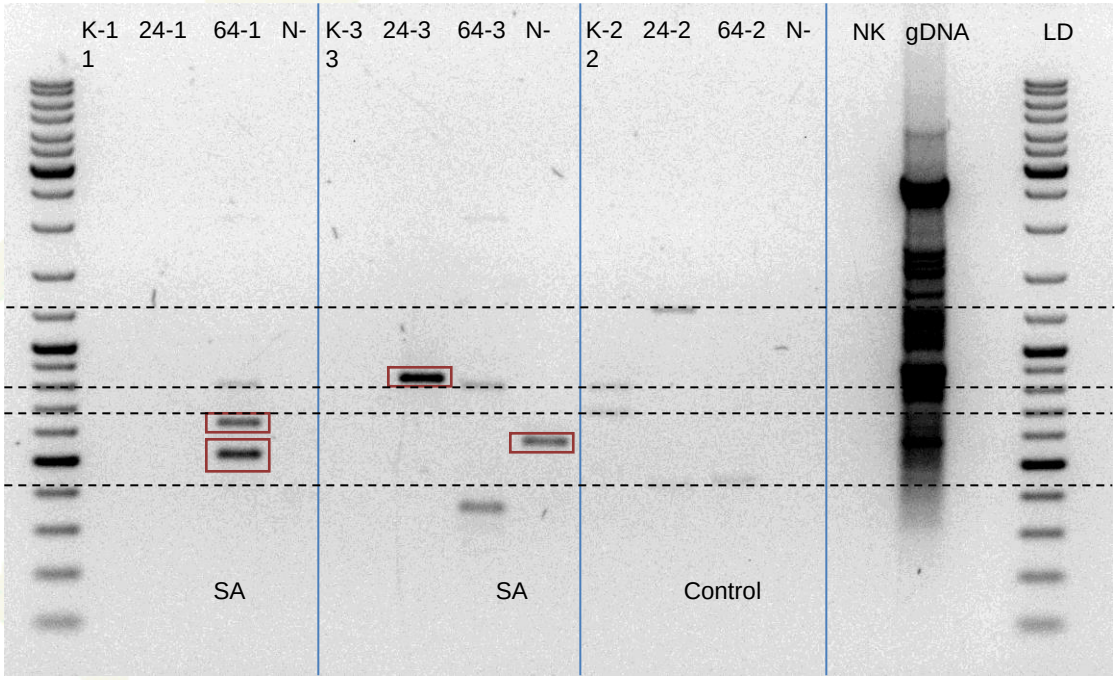


- ✓ Heat stress
- ✓ insect damage
- ✓ Abscisic acid & Salicylic acid

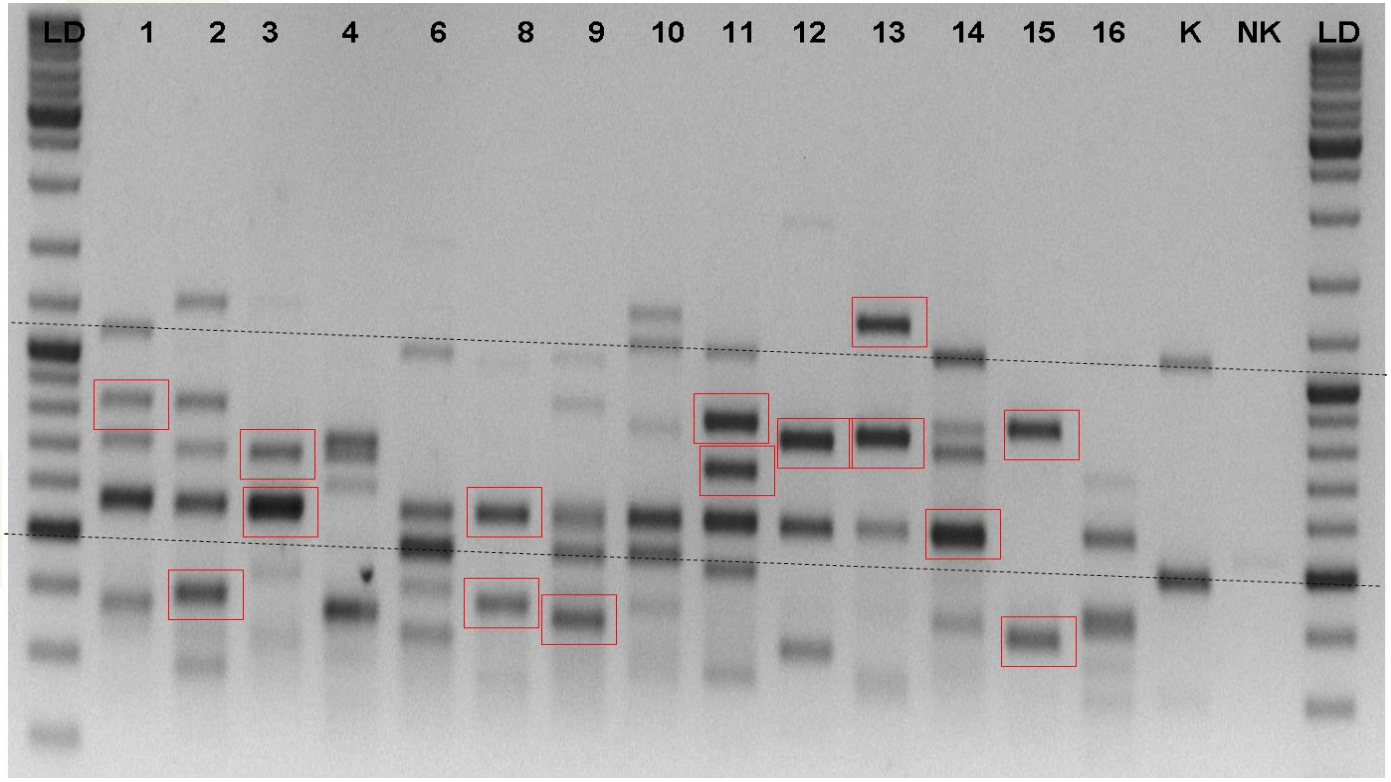
Pine Woolly Aphid
(*Pineus pini*)



iPBS markers
(Kalendar et.al., 2010);



Inter PBS amplification. Lanes 24, 64, N indicates fragments amplified from cDNA samples from SA treated trees. lane K is the cDNA sample from the same tree before stress. Unstressed tree were damaged with sampling and sprayed with same solution without SA. lane gDNS shows amplification with genomic DNA of the same ramet, and the last lane is size marker GeneRuler DNA Ladder Mix (*Fermentas*). Excised fragments are indicated.



Non-specific iPBS amplification after insect Pine Woolly Aphid damage. 1-16 stressed samples, K-unstressed pine ramet, NK- Negative control.

Sequence analysis

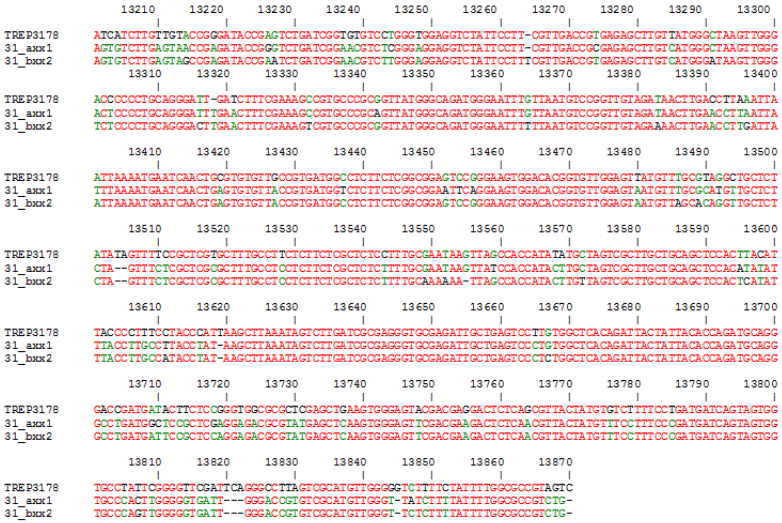
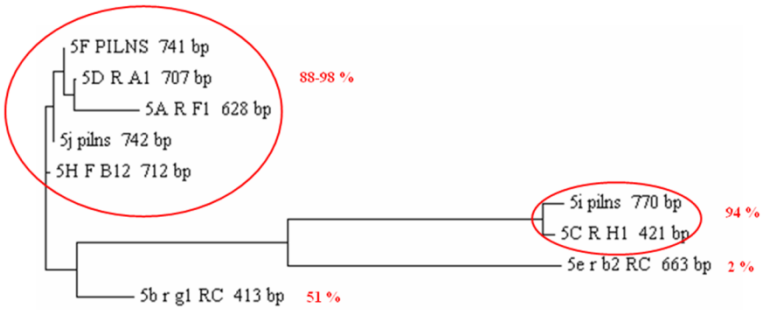


Figure 3. Multiple sequence alignment of 3.1 group fragments with LTR *Gypsy*, *Erika* TE (data base number: TREP3178) nucleotide sequences.




```
>gnl|PTREP|PTREP173 DNA transposon, TIR, CACTA,"DTC_Conan_consensus-1";
(TREP3415) ORF2 protein Length = 1124

 Score =  274 bits (701), Expect(2) = 2e-89
 Identities = 170/228 (74%), Positives = 188/228 (82%), Gaps = 2/228 (0%)
 Frame = +2

Query: 101 PRCRTWFYAHGGA-LDQKTGKISWKASLKGAKQKIIDTIEESRRGEFMPNRENDELTCAL 277
           PRCRTWFYAHGG  LD KTG +S KA L GA    ++  IEE+R G F PNRENDELT AL
Sbjct: 277 PRCRTWFYAHGGGELDPKTGNVSTKACLNGADDALLVAIEEARSGVFQPNRENDELTRAL 336

Query: 278 GNPEYQGRTRGKGVIPWYEGFSDWNDDYRSRARKKMEEEKKRKLEEEQRKQDAERLQGLE 457
           GNPE+ GRTRGKG IPWYEGFSDWN DYR+RARKK+ EEKKRK+EEEQRK D ERLQGLE
Sbjct: 337 GNPEHPGRTRGKGAIPWYEGFSDWNADYRTRARKKIAEEKKRKMEEEQRKLDYERLQGLE 396

Query: 458 ARHADLALKFRQQQQ*IDSLSQERGSQQRQQQADDRPALDSTVPSMPRSSVGSAPGD-TL 634
           A  A+LA KF++QQ+ IDSLSQ+RGSQQ QQ ADD PALD+T PSMPRSSVGSAPGD  +
Sbjct: 397 ASQAELAAKFQRQQEQIDSLSQQRGSQQLQQLADD-PALDTTAPSMPRSSVGSAPGDAVV 455

Query: 635 LDTYPVDDIIEDTNCELHSKMKNISMKVADGVAFPVTPRATYHCIPIP 778
           LD YPVDDI E+TNCELH KMKNISMKVAD VAF  +P AT+HC PIP
Sbjct: 456 LDRYPVDDITENTNCELHFKMKNISMKVADAVAFSNSPEATFHCNPIP 503
```

Searches were done in the

- NCBI data bases nr, EST, refRNA, ref_genomic
- (http://www.ncbi.nlm.nih.gov/BLAST/);
- GrainGenes Triticeae Repeat Sequence Database blastn, blastx (http://wheat.pw.usda.gov/ITMI/Repeats/blastrepeats3.html)
- Gypsy Database 2.0 LTR, genome, cores (http://gydb.org/index.php/Blast) (Llorens et al., 2011).
- Repbase blastn (http://www.girinst.org)

Classification of mobile genetic elements

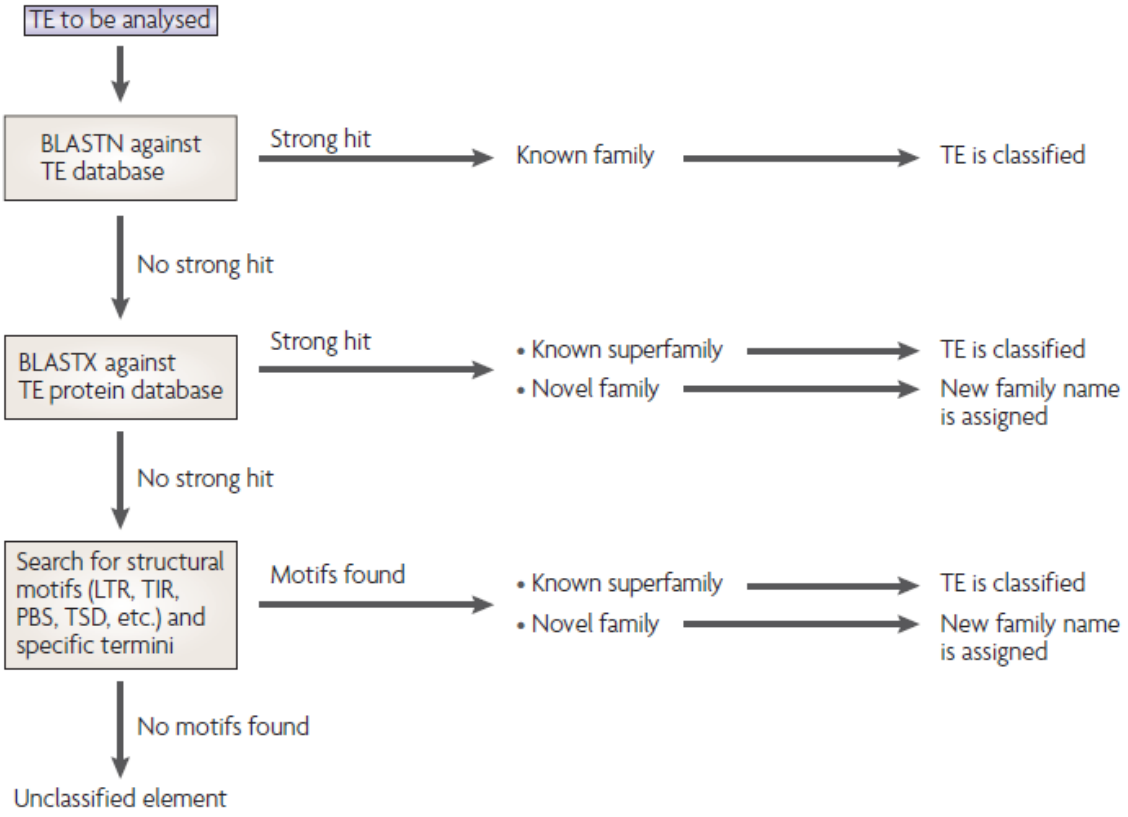


Figure 3 | **Step by step transposable element (TE) classification.** LTR, long terminal repeat; PBS, primer binding site; TIR, terminal inverted repeat; TSD, target site duplication.

Wicker et al., 2007

Results of similarity searches

No.	Order, Superfamily	Element name	Data base	Score/ E-value/ Identities	TE Host Organism	EST
T9	LTR, Gypsy	*Carmilla*	TREP	75.8/ 9e-15/ 85%	*Triticum turgidum*	4
T10-2	LTR, Gypsy	*Laura*	TREP	686/ 0.0/ 90%	*Triticeae*	19
T5	Non-LTR, LINE	*Persephone*	TREP	56/ 6e-08/ 86%	*Hordeum vulgare*	896
T49	DNS, CACTA	*Caspar*	TREP	609/ e-174/ 95%	*Triticeae*	2
T3-2	LTR, Gypsy	*Erika*	TREP	676/ 0.0/ 89%	*Triticeae*	75
T7	LTR, Gypsy	*Wham*	TREP	722/ 0.0/ 93%	*Triticeae*	63
T10-3[1]	LTR, Copia	*Copia-2 TA-I*	Repbase	387/ 2e-110/ 84%	*Triticeae*	3
B5-4	LTR, Copia	*Maximus*	TREP	236/ 2e-62/ 82%	*Triticum aestivum*	57
B12-3	LTR, Gypsy	*Ogre*	Gydb	80/ 3e-15/ 80%	*Pisum sativum*	56
B12-5	LTR, Gypsy	*Ogre*	Gydb	697/ 0.0/ 95%	*Pisum sativum*	0
B16-6	LTR, TRIM	*Cassandra*	TREP	184/ 5e-47/ 90%	*Hordeum vulgare*	48
B21-1	LTR, Gypsy	*BAGY-2*	TREP	60/ 3e-09/ 82%	*Hordeum vulgare*	22
Bv11-4[1]	LTR, Copia	*Copia-2 TA-I*	Repbase	392/ 4e-112/ 84%	*Triticeae*	4
Bv12-3[2]	ERV/ERV3	*MSTB*	Repbase	102/ 2e-26/ 86%	*Mammals*	1
SA4-2	LTR, Copia	*Angela*	TREP	210/1e-54/ 84%	*Triticeae*	66
SA5-1	LTR, Gypsy	*Gypsy-3*	Repbase	740/ 0.0/ 75%	*Hordeum vulgare*	27
SA13-1[2]	LTR, ERV	*MstII*	Repbase	396/1e-114/ 85%	*Mammals*	1
SA10-1[3]	LTR, Copia	*Copia-18BD*	Repbase	318/ 4e-90/ 87%	*Brachypodium*	10277
ABA3-3	LTR, Gypsy	*Ogre*	Gydb	556/ e-159/ 92%	*Pisum sativum*	0
ABA12-1[3]	LTR, Copia	*Copia-18BD*	Repbase	351/ 7e-100/ 89%	*Brachypodium*	10277

Results of similarity searches

Similarity with protein domains							
1b	termo	LTR, copia	*HORPIA2*	polyprotein	TREP, blastx	135/3e-34/41%/61%	*Hordeum vulgare*
10jb	termo	LTR, copia	*Maximus*	polyprotein	TREP, blastx	84.7/8e-19/59%/73%	*Triticum aestivum*
283	termo	LTR, Gypsy	*Geneva*	GAG	TREP, blastx	54.7/ 2e-09/ 29%/43%	*Hordeum vulgare*
3a	termo	LTR, Gypsy	*Sabrina*	polyprotein	TREP, blastx	53,5/3e-09/ 28%/ 47%	*Triticum turgidum*
7-4	bio	LTR, Gypsy	*Ogre*	AP	Gydb, blastx	107/2e-25/58%/74%	*Pisum sativum*
15-3	bio	LTR, Gypsy	*Cereba*	polyprotein	TREP, blastx	50/4e-08/38%/58%	*Hordeum vulgare*
19-3	bio	LTR, Gypsy	*CRM*	AP	Gydb, blastx	60.5/3e-11/36%/58%	*Zea mays*
v3-1	bio	LTR, Gypsy	*Athila4-1*	GAG	Gydb, blastx	167/2e-42/46%/65%	*Arabidopsis thaliana*
V11-4	bio	LTR, Copia	*Maximus*	polyprotein	TREP, blastx	84/1e-18/59%/73%	*Triticum aestivum*
2220sa14a	SA	LTR, Gypsy	*Ogre*	RT	Gydb, blastx	74/1e-16/46%/68%	*Pisum sativum*
2242aba3-3	ABA	LTR, Gypsy	*Ogre*	GAG, ORF2	Gydb, blastx	37/2e-3478/67%	*Pisum sativum*

Plasmids isolated:~500

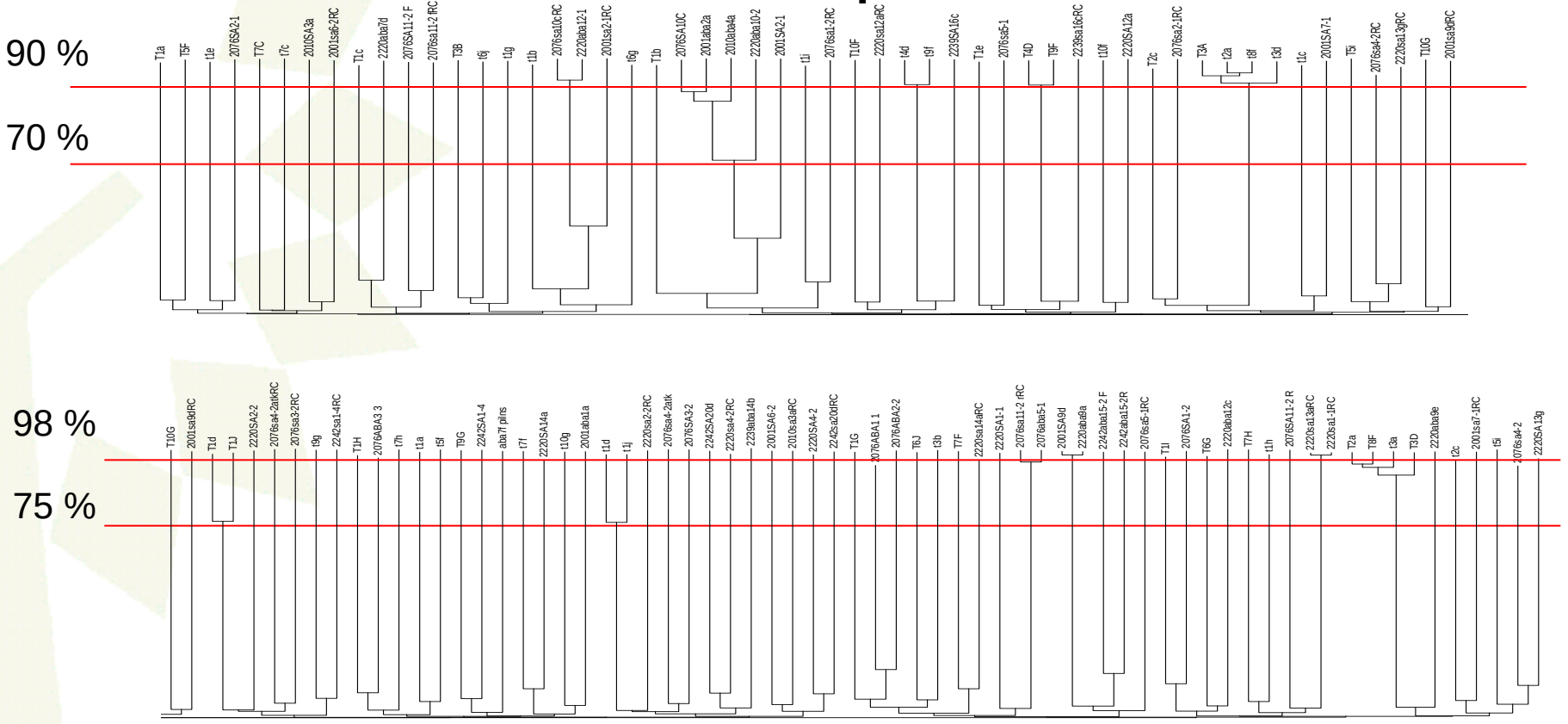
Sequences obtained: ~400

Unique fragments: 125

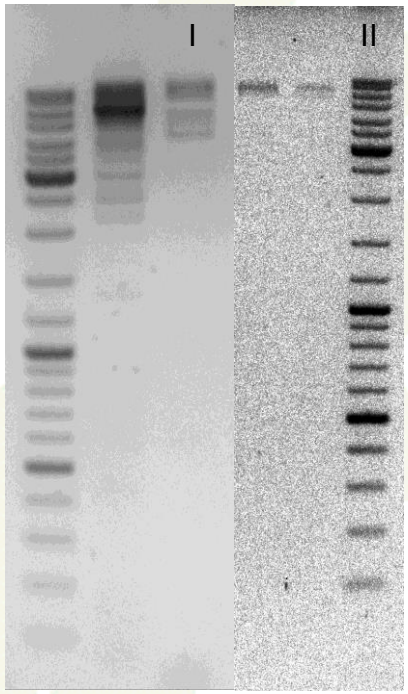
No similarity or similar to un-annotated seq in any db: 63

Similar to known ME: 32

Multiple sequence alignment of heat stress, ABA and SA expressed nucleotide sequences



full length retrotransposon isolation



Vsp,
CaiI,
MspI
AdeI

→ Klenow

↓ Cl-IAA

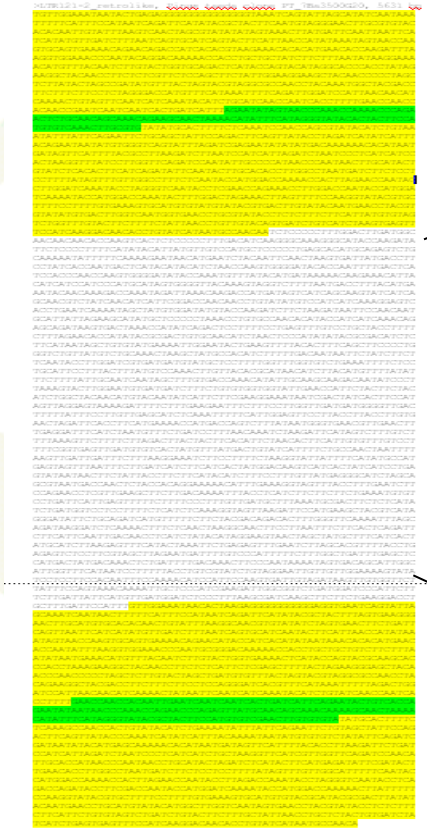
↓ A-tailing

↓ Ligation in TA-vector&transformation → 136 plasmids, 32 true inserts

~5 similar to retro domains in protein level

Long distance PCR

BAC clone analysis



>121-3_internal_RC, 2558 bp

AATGGAATCAAAGCAGGTCCTTCAGAAGAAGCTTGATCGCTTAAAGGGAGATCCATCAAC
ATGAATAATCAAGAAAGAATGGGATCCATCAGACCGCCAATCTTCGATGGGAGCAATTTT
GTTTACTGGAAATAAGAACCACAACTTATCTACAATCACTTGGAATGGATGTTTTGGAAA
TTGTGGAAGGAGGATATACTTTTCCAACAACAATTCCTACAGATACGACAGGTAAAAAGG
ATTATGAAACCAATGCCAATGCTGTCACTATTTTATTGGGAAGTTTGTCAAAATCAGAGT
TTGTCATAGTCATGGAGCTCAAATCAGCCAAAGAAAAATGGGACAAAATCATTCTAAGCT
ACGAAGGAGACTCTCAGGTAAAACGTGCTAAGATTCAAACTCTCAGAATTTAGTATGAAA
CTCTTAAGATGCATAGTGATGAAAGCATAGCTAGTTACTTCCTATGTATAGATGAGGTTG
TCAATTGAATGAAGAATCTGAGTGAAGAAATTAAGGAAGTTGCCTTAGTTGAGAAAGTTT
TGAGATCCTTATCTGCTAAATTTTGAACCCAAAGTGTCTGTCGTAGAAGAAAAACATGAT
CTGCAGAATATCCCTATGACGTAGCTTCATGGAATCTTAACTACCTTTGAGATGAGAAAA
GGAGGACCATCAGATATGAGAGAAGTCGCATTTAAAGCATCAACAAAGGGAAAAGAAAAA
CTCAATGAATCAGGACACATTTCAGAAGAAGAAGATGAGGTAAATTTTGTCAAGAAGCTT
CAACGAGGTTCTGGAAGATTCAAAGGTAAACTACCTTTCAAATGTTTTTCCTGTGGTAGA
GTTGGTCATTACGCTGCTAGATGCCCTCATAACAAAGGAAAGATGTATGAAGAAGGTAAT
AGAAGTTATTATACTCAGGATGATAGTGATGACTTGTCCATAGTGATGAAGATGATCAAG
AAATTAAACTACTCATGGCATATGAAAATAATACCTTAGAAAAGGAAGATTTCCTTAAAG
AAATCAATCAACTTAAAATTAGTTGGCAGAAAATGATACAGTCATAAACATAGTGACACA
TCAACTCACCGAAAAGGACAAACACAATGAAGTGTTAGAATGTGAAGTAGTAAGTCTAAG
AAAAGAACTTTAAAAGACAAAGACTATGAATCTTAGATTTGTTAAAGGATCAGAAACATT
AGATGAAATCCTCAAAGTTCAACGTTCACCCATTATAAAGACTGGTCATGGTTTTCATGA
AGGTGAATCTAGTTCACAAGGTAAGGTAAGGAACTCCAATGAAAAATTTGAGATGCTCAA
CAAGGAAATAAAAAGTCAACCCCATCATCAACCAAGGAAAGAAATTCTTCAAAGAAAATC
TTTTACTCCTAACTATGGAAGTGATAGTCGATTATTTCCTTCGAAGAATGATATTGTACA
TGTTGTAGCCAGATGTAGAAGTAGAATGGTTCAATACCACCACACAGAAAGATCATCACA
TTCCAAGTACTTTAAGGGATATTGTCTTGCTTGCAATATGTTTGGTCACAAAGCTATTGA
TTGCAATAAAAGAAATATAAAACATGTAAGATGTTATGCGTGTAACAAGTTTGGACATAA
AGTAAAGGAATGCAGGAGAAAATTTCAGACACCAAACCAAAAGGAGCATACATCATCACG
ATCCAAGGTATTGAAGAAGATAGAATTATTGTCAAAAAGATGTGGCATAGCTTAGTTTGC
AGACATAACAGACCCAGGGGAAGCTGAAAGTGTAAAACTTCAGTATTCCAATTTTCATAC
ACAGCTATTATGAAGAGATGTCGCGTATATATGGGAGTTAGATGTTGCACAGTCGCGTAT
ATGGTGTTCTAAAGAAAAGGTAGCAGGACAAACTCAGGAAAAGGAGTCTGATATGGTTTA
GTCACTTATCTGCTCTGTTTGATGATGGTATGTGTTGGCACAGGTTTAGGGGGAGCATAT
GCTTCTAATAATGCATTGTTGGAATTATCTTAGAAGATCTTGGTACATATCCACATAGCT
ATTTTGATTCAGGTGACTCCTTTGATGATGGACATATACAGGTTGTTGGTCCGAATGATG
TTGATAGACGTTGCTGATGATACTTGCTGATGACTATCATGGTCTTGTTTAATCTATTTG
GTCTTTGTTGTATTTCATGTAAAGGTCATTAAAAGACCTACTTTGTAACCCCACTATGCA
TGGGATGGATGATGTAATGTTTCTTGTTTTTATCATGTATAAACATTTGGTATATCCCCA
CTTGGTTGGGTGGATGAGTCAAAATTGGTGTATCCCCACTTGGTTAGATGTATGTATGAG
TCATTGGTGATAGGAAGGTCATAATCACTTAGTTGAATTGTAGATTCATGTTATTCTTTT
GAAAAATATTTTTGCAGACTCTGCATGTGCTCAGGGGGAGCATGGACAACATAATGTATA
TGAAAAGGGGAGAATATCTTGGTATGCCCTTTGCCCTTGATGTCAAAGGGGGAGAGTGAC

NCBI EST

Predicted proteins in RC with StarORF

NCBI EST

Are retrotransposons active in the pine genome?

- Or what are they doing?

Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE*-1 retrotransposon dynamics in response to sharp microclimatic divergence

Ruslan Kalendar*, Jaakko Tanskanen*, Sirkka Immonen†, Eviatar Nevo‡, and Alan H. Schulman*§

*Institute of Biotechnology, University of Helsinki, Plant Genomics Laboratory, Viikki Biocenter, P.O. Box 56, FIN-00014 Helsinki, Finland; †Agricultural Research Centre, Myllytie 10, FIN-31600 Jokioinen, Finland; and ‡Institute of Evolution, University of Haifa, Mount Carmel, 31905 Haifa, Israel

Edited by Susan R. Wessler, University of Georgia, Athens, GA, and approved March 28, 2000 (received for review January 3, 2000)

The replicative spread of retrotransposons in the genome creates new insertional polymorphisms, increasing retrotransposon numbers and potentially both their share of the genome and genome size. The *BARE*-1 retrotransposon constitutes a major, dispersed, active component of *Hordeum* genomes, and *BARE*-1 number is positively correlated with genome size. We have examined genome size and *BARE*-1 insertion patterns and number in wild barley, *Hordeum spontaneum*, in Evolution Canyon, Lower Nahal Oren, Mount Carmel, Israel, along a transect presenting sharply differing microclimates. *BARE*-1 has been sufficiently active for its insertional pattern to resolve individuals in a way consonant with their ecogeographical distribution in the canyon and to distinguish them from provenances outside the canyon. On both slopes, but especially on the drier south-facing slope, a simultaneous increase in the *BARE*-1 copy number and a decrease in the relative number lost through recombination, as measured by the abundance of solo long terminal repeats, appear to have driven the *BARE*-1 share of the genome upward with the height and dryness of the slope. The lower recombinational loss would favor maintenance of more full-length copies, enhancing the ability of the *BARE*-1 family to contribute to genome size growth. These local data are consistent with regional trends for *BARE*-1 in *H. spontaneum* across Israel and therefore may reflect adaptive selection for increasing genome size through retrotransposon activity.

Retrotransposons resemble retroviruses in their structure and life cycle (1, 2). They are ubiquitous (3–5) and contribute a large proportion of the total repetitive DNA of some plant genomes (6). Retrotransposons are mobilized by a replicative mechanism that has the capacity to generate and insert many new daughter copies into the genome, thereby increasing genome size (7). The error-prone nature of their replication by reverse transcriptase (8), the mutagenic potential of their transpositional integration (9), and the effects of their accumulation and recombination (10) together suggest that active retrotransposons may be major contributors to genome diversification in the plants. Genomic changes induced by retrotransposons can be tracked by the joints between the flanking DNA and the conserved retrotransposon termini created upon integration. Marker techniques based on PCR amplification between retrotransposons and flanking DNA recently have been developed (11–13).

Accumulated data indicate that retrotransposons in plants (14–16), animals, and fungi respond to various forms of stress. When stress factors in the environment vary ecogeographically, retrotransposon prevalence and insertion patterns may vary accordingly. The immediate wild ancestor of cultivated barley (*Hordeum vulgare*), *Hordeum spontaneum*, is ideal for analyzing retrotransposon insertions and their role in the genome because of the presence of a large and active retrotransposon family and the availability of well-studied wild populations distributed in diverse habitats (17–21). The *BARE*-1 family of retrotransposons comprises on average 14×10^3 copies in the genomes of *Hordeum* species (10). Members of this family are transcriptionally (22) and translationally (23) active, encoding both a polyprotein (24, 25) and processing signals (26), which are functionally conserved. The *BARE*-1 copy number is positively correlated with both genome size and habitat aridity (10), factors that are themselves correlated (27) regionally in *H. spontaneum*.

We have examined the role of the *BARE*-1 retrotransposon in genome diversification in individuals at the Evolution Canyon microsite, Lower Nahal Oren, Mount Carmel, Israel (28–30). This 400-m-wide erosion gorge (see Fig. 4, which is published as supplementary data on the PNAS web site, www.pnas.org), dating from the Plio-Pleistocene era, presents north- and south-facing slopes (NFS and SFS, respectively) with common geologies and macroclimates but microclimates sharply differing in solar irradiation and aridity. Biotically, the NFS is Eurasian and the SFS is Afro-Asian within the Mediterranean context (28, 30). We have examined *H. spontaneum* along a north-south transect across the canyon slopes to test whether regional patterns (10) can be detected locally. The *BARE*-1 copy number and patterns of insertional polymorphism, as well as total genome size, were determined for accessions from the canyon.

Materials and Methods

Plant Materials. Spikes from individual *H. spontaneum* plants were collected at six stations located along a 300-m north-south transect across the NFS and SFS of Evolution Canyon (28). The stations previously described (29, 31) as NFS (stations 5–7) and SFS (stations 1–3) are referred to here as: NH (north high), NM (north middle), NL (north low), SL (south low), SM (south middle), and SH (south high). From each station, seeds of 10 individual plants, separated by at least 1 m from each other, were used as the samples. The seeds were grown to seedlings for preparation of DNA and nuclei.

Retrotransposen-Microsatellite Amplified Polymorphism (REMAP) Amplification. DNAs were prepared by the cetyltrimethylammonium bromide method (32). For REMAP PCR amplification, primers facing outward from the long terminal repeats (LTRs)

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: *in*, integrase; LTR, long terminal repeat; NFS, north-facing slope; NH, north high; NM, north middle; NL, north low; REMAP, retrotransposen-microsatellite amplified polymorphism; SFS, south-facing slope; SH, south high; SL, south low; SM, south middle; SSR, simple sequence repeat.

See commentary on page 6250.

§To whom reprint requests should be addressed. E-mail: alan.schulman@helsinki.fi.

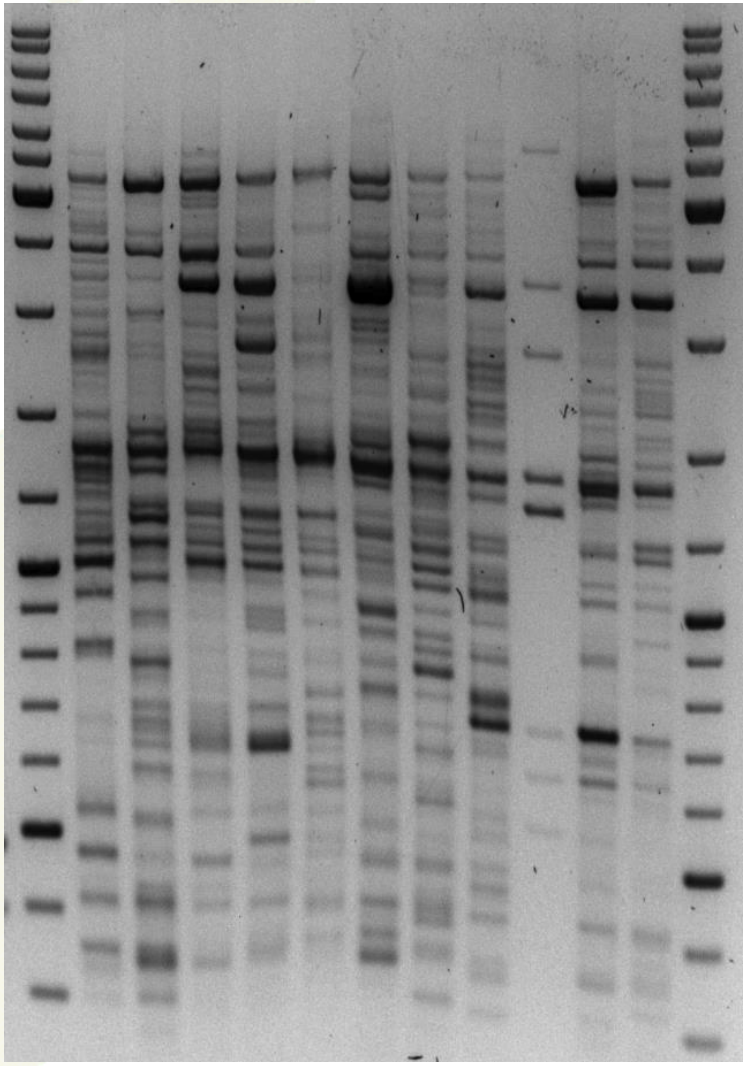
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.110587497. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.110587497

EVOLUTION

- Development of nine pine specific retrotransposon markers
- Identification of a pine stand growing in divergent conditions
  - On a slope, dry conditions at the top, sheltered and wet conditions at the bottom
- Tested with SSR markers for population divergence

Inter Retrotransposon amplification

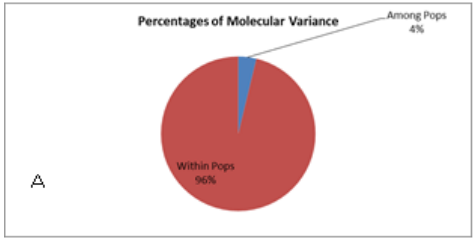


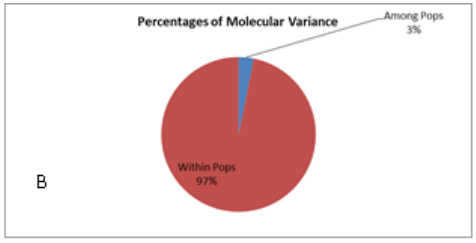
Nine newly developed Retrotransposon Markers

184 fragments were analyzed

150 samples from one natural pine stand growing in highland, slope and lowland.

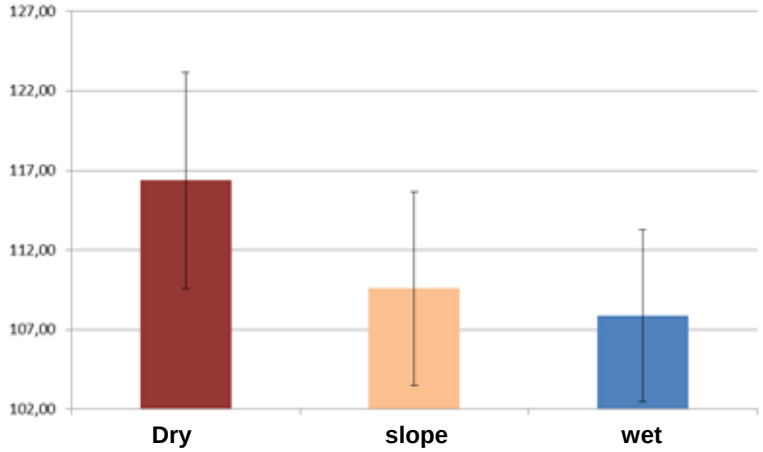
On average,more fragments were found in the samples from trees growing in highland



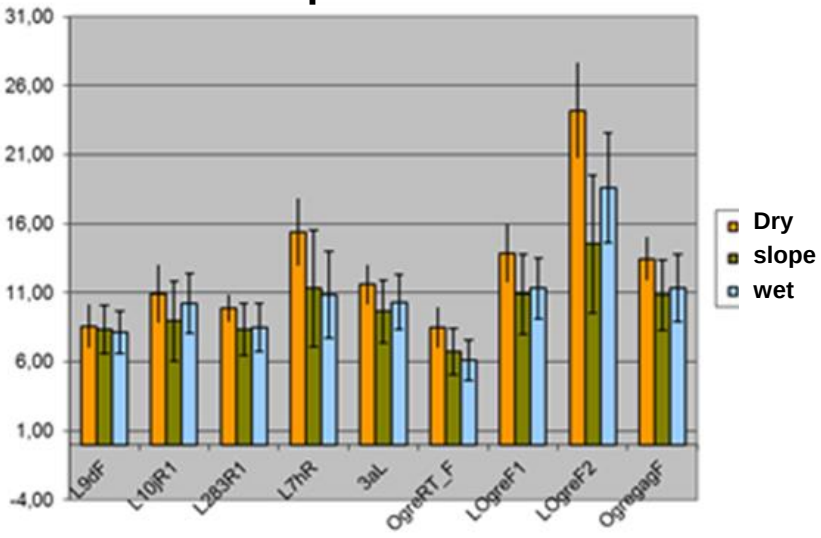


Differences in average number of amplification bands

Differences by populations

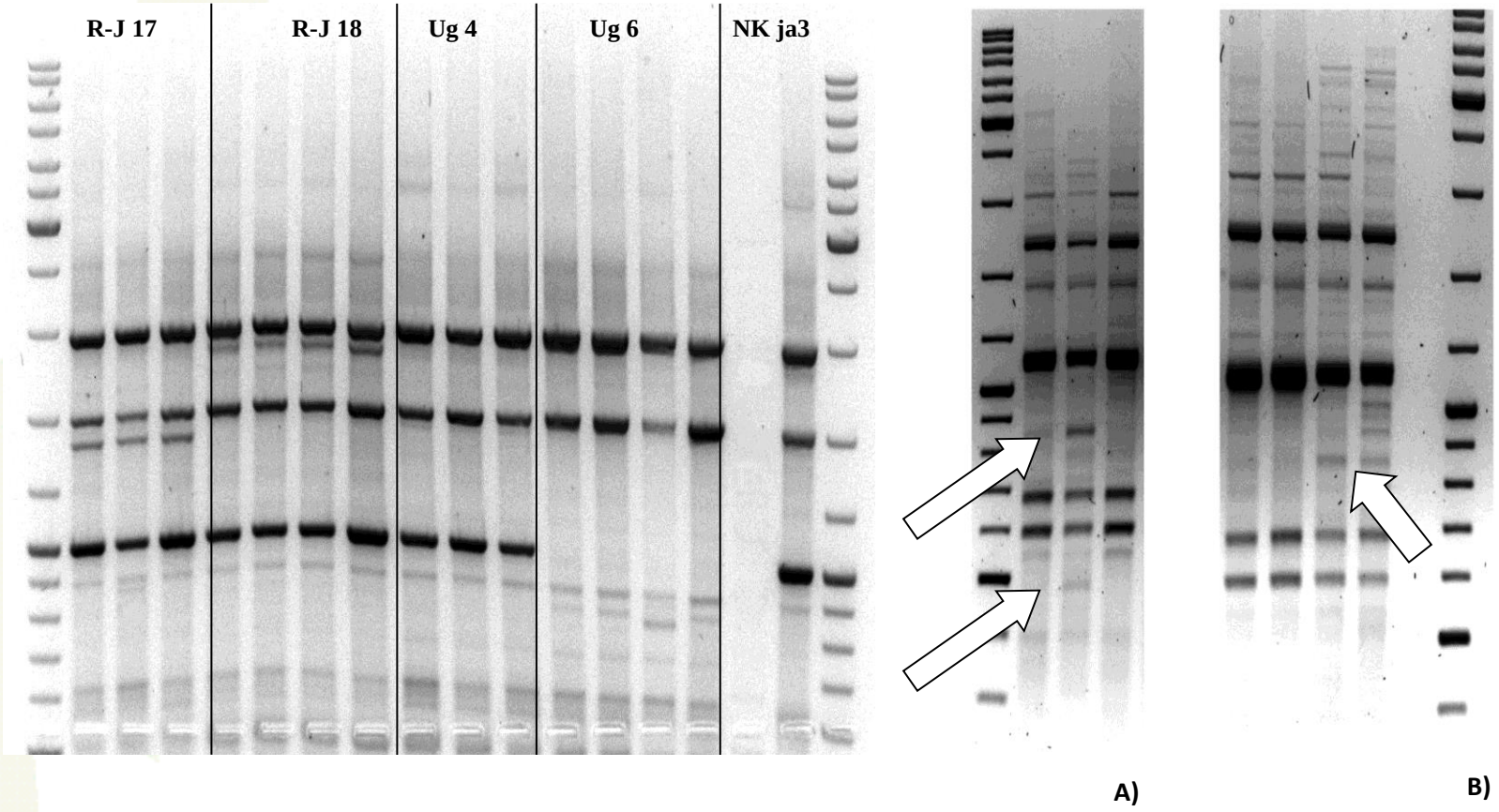


Average values by marker with the mean square deviation



Population	the arithmetic mean	Deviation	Mean square error	Difference	Error of average difference	Predetermined deviation	F-criteria
	Xvid	δ	Sx	d(Xvk-Xvn/z)	Sd	t>3, būtiska atšķirība	F, 0,05, df(40)=>1,51
dry	116,37	6,80	1,10	-	-	-	-
slope	109,59	6,08	0,92	6,78	1,42	4,77	1,25
wet	107,89	5,40	0,80	8,48	1,34	6,33	1,58

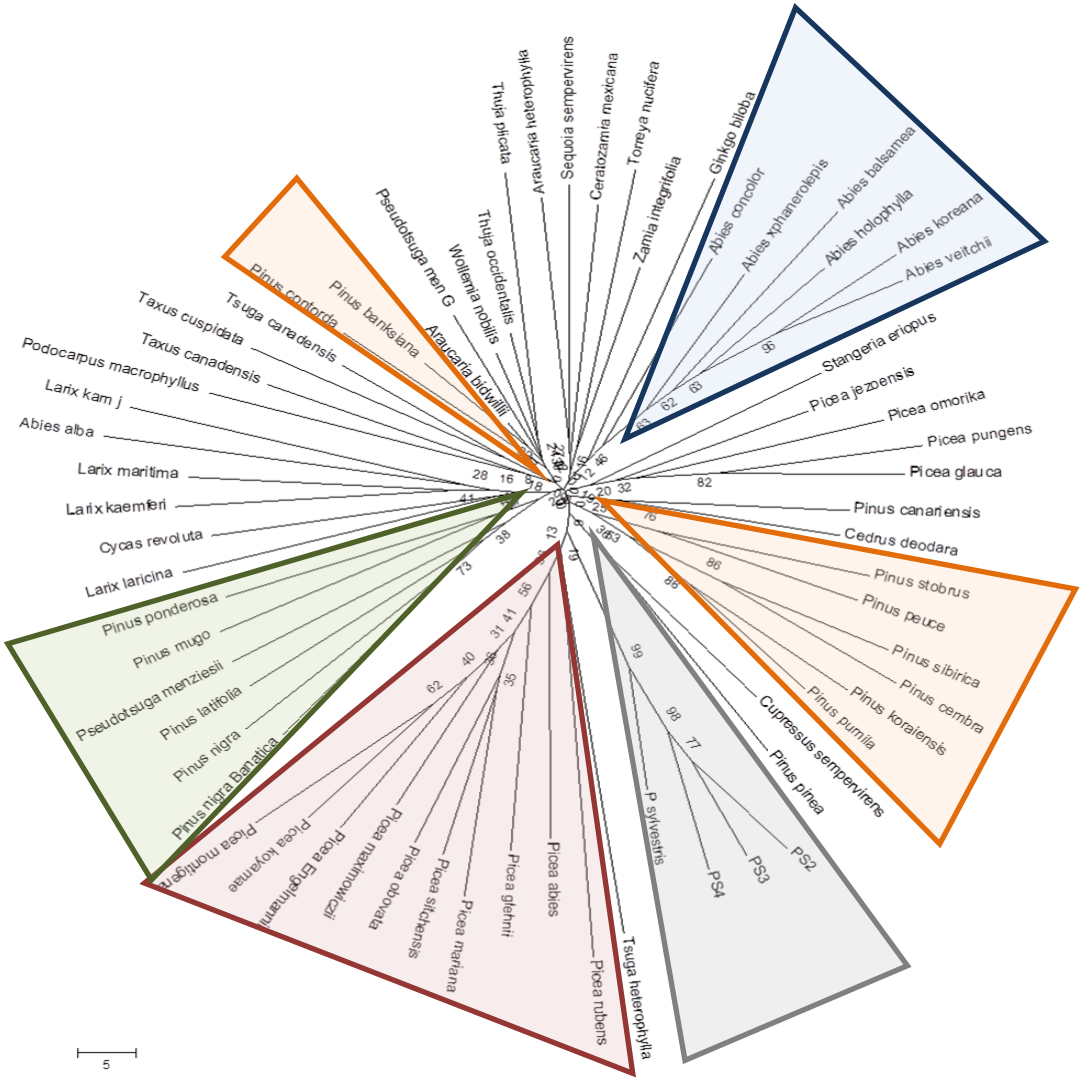
Somaclonal variation detected by IRAP



Ramets of four different pine clones growing in two seed plantations were genotyped with nine pine IRAP markers. Somoclonal variation was detected between ramets of the same clone. No variation was identified with SSR markers

Phylogenetic analyses

Retrotransposon-based markers were developed and used for phylogenetic analyses of a wide range of gymnosperm species.



Conclusions

- Expression of retrotransposon-like sequences under stress conditions
  - more fragments amplified after heat stress and insect damage than after application of SA and ABA
  - Different sequnces expressed under different stress conditions

- Retrotransposon-like sequences are consistently transcribed under various stress conditions as analyzed sequences show similarity with EST database sequences

- Some of the analyzed sequences were similar to several known active retroelement proteins (*BARE-1, Tnt-1*) which could indicate that these retroelements possess transposable activity

- Specific primers designed to the identified sequences multiply multiple fragments from the pine genome

- Pine trees growing in a natural stand in dry conditions contain an increased number of amplified fragments, which is in concordance with previous studies and may reflect proliferation of retrotransposons in their genomes.

- CNV in pine – correlation with gene expression levels
- Retrotranspon-like sequences expressed in pine
  - Differentially expressed under various stress conditions
  - More retrotransposon-like sequences in ‘stressed’ populations
- Preliminary results – need to be confirmed and expanded

Thank you!

- Ilze Veinberga
- Angelika Voronova
- Vilnis Šķipars
- Ilze Gaile
- Baiba Krivmane
- Viktorija Beļeviča
- Krista-Kanberga Siliņa
- Linards Ļubinskis
- Anna Korica
- Krišs Bitenieks

ESF projects:

2009/0228/1DP/1.1.1.2.0/09/APIA/VIAA/035

2009/0200/1DP/1.1.1.2.0/09/APIA/VIAA/146

ERAF
EIROPAS REĢIONĀLĀS ATTĪSTĪBAS FONDS

Ieguldījums tavā nākotnē!