

Tree Genetics & Genomes

Genome-wide analysis of LTR-retrotransposon expression in leaves of *Populus x canadensis* water-deprived plants --Manuscript Draft--

Manuscript Number:	TGGE-D-16-00096R1	
Full Title:	Genome-wide analysis of LTR-retrotransposon expression in leaves of <i>Populus x canadensis</i> water-deprived plants	
Article Type:	Original Article	
Funding Information:	MIUR, Italy	Prof. Andrea Cavallini
Abstract:	<p>Retrotransposons represent a major component of plant genomes, however large-scale studies on their expression are rare. Massively parallel sequencing offers new analytical possibilities enabling a comprehensive study of retrotransposon RNA transcription. We evaluated the expression of long-terminal-repeat retrotransposons in leaves of two sister hybrids <i>Populus x canadensis</i> (<i>P. deltoides</i> x <i>P. nigra</i>), subjected to moderate or severe water deprivation, by mapping Illumina RNA-Seq reads onto a set of 958 unique full-length retrotransposons of <i>P. trichocarpa</i>. Detectable levels of transcription were ascertained for 140 retrotransposons in one hybrid and 182 in the other. The two hybrids showed different retrotransposon expression levels and these differences reduced at increasing drought levels. The number of expressed Gypsy elements in control and water-deprived plants was higher than those of Copia, as were their expression levels.</p> <p>The two hybrids showed different retrotransposon expression patterns following water deprivation. Such variations between hybrids were related to differential expression of a few genes involved in chromatin methylation and remodelling. Overall, our data indicate that even in genetically close individuals, large differences can occur in retrotransposon expression, with possible consequences for genome differentiation.</p>	
Corresponding Author:	Lucia Natali ITALY	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:		
Corresponding Author's Secondary Institution:		
First Author:	Tommaso Giordani	
First Author Secondary Information:		
Order of Authors:	Tommaso Giordani Rosa Maria Cossu Flavia Mascagni Fabio Marroni Michele Morgante Andrea Cavallini Lucia Natali	
Order of Authors Secondary Information:		
Author Comments:	Dear Sirs, with this letter I am sending you the revised version of an article by Tommaso Giordani, Rosa Maria Cossu, Flavia Mascagni, Fabio Marroni, Michele Morgante, Andrea Cavallini, and myself, titled "Genome-wide analysis of LTR-retrotransposon expression in leaves of <i>Populus x canadensis</i> water-deprived plants", to be submitted for	

publication in Tree Genetics and Genomes.

The submitted manuscript represents original work that is not being considered for publication, in whole or in part, in another journal, book, conference proceedings, or government publication with a substantial circulation; all previously published work cited in the manuscript has been fully acknowledged; the manuscript is one of a kind; all of the authors have contributed substantially to the manuscript and approved the final submission; no real or perceived conflicts of interest occur.

[Click here to view linked References](#)

Genome-wide analysis of LTR-retrotransposon expression in leaves of *Populus x canadensis* water-deprived plants

Tommaso Giordani¹, Rosa Maria Cossu^{1,2}, Flavia Mascagni¹, Fabio Marroni^{3,4}, Michele Morgante^{3,4}, Andrea Cavallini¹, Lucia Natali^{1*}

T. Giordani and R.M. Cossu contributed equally to this work.

Corresponding author:

Lucia Natali

lucia.natali@unipi.it

¹ Dept. of Agricultural, Food, and Environment, University of Pisa, Via del Borghetto 80, I-56124 Pisa, Italy

² School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, England

³ Institute of Applied Genomics, Udine, Italy

⁴ Dept. of Crop and Environmental Sciences, University of Udine, Udine, Italy

1 **Abstract** Retrotransposons represent a major component of plant genomes, however large-scale
2 studies on their expression are rare. Massively parallel sequencing offers new analytical
3 possibilities enabling a comprehensive study of retrotransposon RNA transcription. We evaluated
4 the expression of long-terminal-repeat retrotransposons in leaves of two sister hybrids *Populus x*
5 *canadensis* (*P. deltoides x P. nigra*), subjected to moderate or severe water deprivation, by mapping
6 Illumina RNA-Seq reads onto a set of 958 unique full-length retrotransposons of *P. trichocarpa*.
7 Detectable levels of transcription were ascertained for 140 retrotransposons in one hybrid and 182
8 in the other. The two hybrids showed different retrotransposon expression levels and these
9 differences reduced at increasing drought levels. The number of expressed *Gypsy* elements in
10 control and water-deprived plants was higher than those of *Copia*, as were their expression levels.
11 The two hybrids showed different retrotransposon expression patterns following water deprivation.
12 Such variations between hybrids were related to differential expression of a few genes involved in
13 chromatin methylation and remodelling. Overall, our data indicate that even in genetically close
14 individuals, large differences can occur in retrotransposon expression, with possible consequences
15 for genome differentiation.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Keywords** *Copia*; drought; *Gypsy*; poplar; retrotransposon transcription
32
33
34
35

36 Abbreviations

37 RE	Retrotransposon
38 LTR-RE	Long Terminal Repeat-Retrotransposon
39 RPKM	Reads per Kilobase per Million reads mapped.

Introduction

Transposable elements (TEs) form the mobile component of the genome, as they can change their location in chromosomes (transposition) through mechanisms operated by enzymes encoded by the transposable element itself. TEs are distinguished between class I retrotransposons (REs) and class II/DNA transposons, depending on their transpositional mechanism. Class I elements transpose through a replicative mechanism that involves an RNA intermediate ("copy and paste"). REs can be autonomous in transposition, because of the presence of ORFs that encode for transposition-related enzymes. Conversely, non-autonomous elements do not possess these ORFs, nevertheless they can use enzymes encoded by other elements to transpose (Wicker et al. 2007). The replicative mechanism of REs has allowed these elements, often accounting for hundreds of thousands of copies, to become the largest portion of genomes during eukaryote evolution (SanMiguel et al. 1998; Vicient et al. 1999). A striking example of RE amplification which doubled genome size (without polyploidization) has been reported in *Oryza australiensis* (Piegu et al. 2006).

Based on the presence of long terminal repeats (LTRs) at their ends, retrotransposons can be subdivided into LTR- and non-LTR-REs. Promoter elements, polyadenylation signals and enhancers are found in the LTRs and regulate RE transcription (Bennetzen 2000). The coding portion of LTR-REs is comprised of the *gag* and the *pol* domains. The former encodes virus-like particles (VLPs), the latter encodes enzymes necessary for processing RE transcripts, producing a double-stranded DNA and integrating such DNA into the host genome. LTR-REs have been further divided into two superfamilies, *Gypsy* and *Copia* (Voytas et al. 1992; Suoniemi et al. 1998), depending on the order of genes within the *pol* domain. In the last decade, LTR-REs lacking internal coding domains were described, and defined as TRIMs (Terminal-repeat Retrotransposons In Miniature; Witte et al. 2001) or LARDs (LARGE Retrotransposons Derivatives; Kalendar et al. 2004). TRIMs and LARDs can be identified only when the complete genome sequence, or at least, long DNA sequences are available. Their species-specific sequence and the absence of coding regions can explain their relative rarity in the literature. However, when surveying complete genomes and using structural features as diagnostics (for example the occurrence of LTRs, a primer binding site [PBS], and a poly-purine tract [PPT]), they have been shown to form a sizeable component in the TE fraction of the genome (Buti et al. 2011; Natali et al. 2013; Barghini et al. 2014; 2015).

Genome sequences of grasses showed that retrotransposition has determined extensive variations in genome structure, even within one and the same species (Brunner et al. 2005; Scherrer et al. 2005). It was suggested that these structural variations are involved in gene expression

1 regulation consequently affecting the phenotype and are subjected to selection. In this sense, the old
2 definition of repeated sequences as “junk” is considered obsolete. For example, allelic cis-
3 regulatory non-genic variation might have a role in heterosis, defined as the superior performance
4 of hybrids in comparison to their parents (Brunner et al. 2005; Stupar and Springer, 2006; Morgante
5 et al. 2007; Buti et al. 2013).

6
7
8
9 The transcription of retrotransposons has been reported in a number of plant species,
10 especially after exposure to various stresses (Grandbastien 2015). A survey of EST libraries of grass
11 species indicates that most LTR-REs are poorly constitutively transcribed (Meyers et al. 2001;
12 Vicient et al. 2001; Vicient and Schulman 2002; Ishiguro et al. 2014). Most of the studies relating
13 to LTR-RE expression refer to specific elements (Vicient et al. 2001; Rico-Cabanas and Martinez-
14 Izquierdo 2007; Ramallo et al. 2008; Buti et al. 2009; Kawakami et al. 2011) and/or are limited to
15 the discovery of LTR-RE sequences in transcriptomes (see for example Parchman et al. 2010; Lu et
16 al. 2013) and only a few comprehensive studies are specifically devoted to this class of sequences
17 (Meyers et al. 2001; Marcon et al. 2015; Jiang et al. 2016).

18
19
20
21
22
23
24
25 In a few cases, LTR-RE transcription has determined new insertions in the genome. For
26 example, *Tnt1* and *Tto1* in *Nicotiana* and *Tos17* in rice are not transcribed in standard culture
27 conditions and are induced to transpose by tissue culture (Hirochika 1993; Hirochika et al. 1996;
28 Grandbastien 1998). More recently, retrotransposition was reported for a *Copia* element of
29 sunflower for which RNA expression and subsequent insertion in the genome was shown (Vukich
30 et al. 2009).

31
32
33
34
35
36 It has been shown that barley *Copia* REs produce three distinct classes of transcripts (Chang
37 et al. 2013), of which two are capped and polyadenylated, and can be translated or spliced (to
38 produce a large amount of GAG-encoding transcripts); the third is not capped or polyadenylated,
39 and is destined for packaging and ultimate reverse transcription. Conversely, *Gypsy* elements of
40 *Drosophila* destined for reverse transcription have been shown to be capped and polyadenylated or
41 not, indifferently (Meignin et al. 2003).

42
43
44
45
46
47 Presumably, the activity of LTR-REs is limited by the host genome because of their
48 potential mutagenic action. The first mechanism of control of mobile elements relies on chromatin
49 structural alterations, since heterochromatin comprises “silent” DNA. Mechanisms underlying
50 chromatin packing in plants act through methylation of histones and cytosine residues in CG and
51 CHG combinations (Dieguez et al. 1998). The importance of epigenetic control of TEs is
52 emphasised by the role of RNA silencing which determines chromatin specific methylation and
53 RNA degradation (Slotkin and Martienssen 2007; Lisch 2009; Ito 2013). For example, a silencing
54
55
56
57
58
59
60
61
62
63
64
65

1 pathway driven by anti-sense small RNAs is responsible for REs silencing in a *Drosophila* germline
2 (Vagin et al. 2006).

3 Retrotransposon dynamics have been mainly investigated in grasses and other
4 monocotyledons. Dicotyledons have in general received less attention, despite their great economic
5 importance. Recently, we performed a survey of LTR-REs in the genome of *Populus trichocarpa*
6 (Natali et al. 2015). Based on the identification of conserved structural features, building multiple
7 alignments, and similarity searches, a number of putative full-length LTR-REs were collected.
8 LTR-RE fragments were by far more abundant in the genome than full-length elements, suggesting
9 that, during poplar evolution, amplification of these elements was followed by DNA depletion.
10 Retrotransposition occurred with increasing frequency following the separation of *Populus* sections,
11 with different waves of retrotransposition activity between Ty3-*Gypsy* and Ty1-*Copia* elements.
12
13
14
15
16
17
18
19

20 **Massively parallel sequencing procedures provide unprecedented levels of sequencing**
21 **coverage**, in a short time and at relatively low cost, allowing whole genome expression analyses.
22 We have applied such techniques to study the transcription of the entire set of poplar full-length
23 LTR-REs in two sister hybrids (i.e. obtained from the same parental trees) of *Populus x canadensis*
24 (*P. deltoides* x *P. nigra*) in control conditions and in plants subjected to water deprivation treatment.
25 As a reference, we used the available set of full-length LTR-REs of *P. trichocarpa* (Natali et al.
26 2015), which is phylogenetically closely related to both *P. deltoides* and *P. nigra*, having diverged
27 from them only 8-13 million years ago (Sterck et al. 2005).
28
29
30
31
32
33
34
35

36 **Materials and Methods**

37 **Plant materials**

38
39
40
41
42
43
44 The materials used in this work were the same as those described in Cossu et al. (2014) where gene
45 expression in water-deprived poplar plants was evaluated. Rooted cuttings of two sister hybrids
46 (DxN661200585 and DxN661200589, hereafter called hybrid 85 and 89, respectively) of *Populus x*
47 *canadensis* (*P. deltoides* x *P. nigra*), produced from the very same parents at INRA, Orleans
48 (France), in 1998, were cultivated in 20 x 20 cm² pots in the greenhouse, under natural daylight
49 conditions (750 $\mu\text{m} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, maximal photon flux density), with air temperature maintained at 17–
50 29 °C, and relative humidity from 55% to 90%.
51
52
53
54
55
56

57 In late spring 2011, some plants 50-70 cm in height were watered normally and others were
58 subjected to water deprivation by suspending watering for 8 days or 13 days. As transcriptome
59 responses to drought are shaped by the time of the day in a diverse range of plant species, including
60
61
62
63
64
65

1 *Populus*, (Wilkins et al. 2009; Hamanishi and Campbell 2011), all materials were sampled at the
2 same time of day (11.00 a.m.). One fully expanded leaf per plant, 6-8 internodes from the apex was
3 collected per plant. Each leaf was then divided into two portions, one portion was used for RNA
4 isolation, the other to measure tissue hydration by determining the relative water content [RWC =
5 $100 (FW - DW) / (TW - DW)$], where FW is the fresh weight, DW the dry weight and TW the turgid
6 weight. The experimental design was as follows: 2 plants (biological replicates) x 3 treatments
7 (control [C], moderate [D1], and severe [D2] drought treatments) x 2 hybrids.
8
9
10
11
12

13 **DNA and RNA isolation and preparation of Illumina libraries**

14
15
16
17
18 Genomic DNA was extracted from leaves (0.5 g fresh weight) of single control plants of hybrid 85
19 and hybrid 89 as described by Doyle and Doyle (1989). Genomic libraries were prepared from 5 μ g
20 of genomic DNA using the Illumina PE DNA Sample Prep kit according to the manufacturer's
21 instructions. After spin column extraction and quantification, libraries were loaded on Cluster
22 Station to create CSMA (clonal single molecular array) and sequenced at ultra-high throughput on
23 the Illumina HiSeq2000 to produce 51-nt reads. Low quality bases, empty reads and adapter
24 sequences were removed by using CLC-BIO Genomic Workbench, version 7.0.3 (CLC-BIO). All
25 collected sequences were 50 nt long.
26
27
28
29
30
31
32

33 Total RNA was isolated from leaves (of single plants) with different RWC, according to the
34 method described by Logemann et al. (1987) followed by DNase I (Roche) treatments according to
35 the manufacturer's instructions to completely remove genomic DNA contamination. RNA was then
36 purified by phenol/chloroform extraction and precipitated following standard procedures.
37
38
39

40 RNA-Seq libraries were generated using the TruSeq RNA-Seq Sample Prep kit according to
41 the manufacturer's protocol (Illumina Inc., San Diego, CA). Poly-A RNA was isolated from total
42 RNA and chemically fragmented. First and second strand cDNA syntheses were followed by end
43 repair and adenosines were added to the 3' ends. Adapters were ligated to the cDNA and 200 ± 25
44 bp fragments were gel purified and enriched by PCR. The library was quantified using a
45 Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) and run on an Illumina HiSeq2000
46 (Illumina Inc.) using version 3 reagents. Single-read 50 nt sequences were collected. Low quality
47 bases, adapter sequences, and empty reads were removed by using CLC-BIO Genomic Workbench,
48 version 7.0.3 (CLC-BIO).
49
50
51
52
53
54
55
56
57

58 **Alignment of Illumina reads to the *P. trichocarpa* full-length LTR-REs and analysis of RE** 59 **abundance**

1 The occurrence and abundance of LTR-REs in the genome was analysed by aligning two random
2 subsamples of *Populus x canadensis* DNA reads of hybrids 85 and 89 (generated using an in-house
3 script, each subsample made by the same number of reads, corresponding to 2 genome-equivalents)
4 to *P. trichocarpa* full-length LTR-REs (Natali et al. 2015). Alignment to the complete dataset was
5 carried out using CLC-BIO Genomic Workbench, version 7.0.3, with the following parameters:
6 mismatch cost 1, deletion cost 1, insertion cost 1, similarity 0.9, and length fraction 0.9. Alignments
7 were also performed on a set of 33 well conserved, unique cDNAs of *P. trichocarpa* (Cossu et al.
8 2012), as a control.

9 Since CLC-BIO Genomic Workbench 7.0.3 distributes multireads (i.e. those reads that
10 match multiple distinct sequences) randomly, if all sequences of a lineage are taken together, the
11 total number of mapped reads (with respect to total genomic reads) reveals the effective redundancy
12 of that lineage. By contrast, the number of mapped reads to a single sequence cannot indicate its
13 abundance. For this reason, in other analyses, alignments were carried out to the same full-length
14 LTR-REs, one by one, using BWA (alignment via Burrows-Wheeler transformation) version
15 0.7.5a-r405 (Li and Durbin 2009) with the following parameters: alignment (bwa aln), seed (-l 12),
16 maximum edit distance (-n 4), maximum number of gap opens (-o 3), maximum number of gap
17 extensions (-e 3), mismatch penalty (-M 2), gap open penalty (-O 6), gap extension penalty (-E
18 3). The resulting single-end mappings were resolved via the “samse” module of BWA, and the
19 output was converted into a “bam” file using SAMtools version 0.1.19 (Li et al. 2009). SAMtools
20 was used to calculate the number of mapped reads for each alignment using the following
21 parameters: view, counting only (-c), filtering alignments (-F 4).

22 Analysis of expression of LTR-REs

23 CLC-BIO Genomic Workbench 7.0.3 was used to generate cDNA sequence alignments to the set of
24 958 *P. trichocarpa* full-length LTR-REs. The parameters used for alignments were the following:
25 mismatch cost = 1, deletion cost = 1, insertion cost = 1, similarity = 0.7, length fraction = 0.7. The
26 software assigns randomly non-uniquely matching reads (i.e. reads that align at more than one
27 position with an equally good score) to different sequences.

28 We considered all LTR-REs mapped by at least one read per million of reads in at least one
29 sample as expressed (Lu et al. 2013). The expression level of each sequence was calculated and
30 expressed as RPKM, as described in Mortazavi et al. (2008).

1 Expression profiles were evaluated considering RPKM values in C, D1, and D2 plants using
2 Baggerly's test (Baggerly et al. 2003). Weighted proportion fold changes among severely droughted
3 plants, moderately droughted plants and control plants were treated as follows: when values were
4 higher in droughted than in control plants (or in D2 than in D1 plants), they were reported as
5 positive, when they were higher in control than in droughted plants (or in D1 than in D2 plants)
6 they were reported as negative, thus leading to a '+' value in the case of above-control average
7 expression levels and a '-' value in the case of below-control average expression levels.
8

9 The weighted proportion fold changes between treatments were considered as significant
10 when the weight of a sample was at least two-fold higher or lower than another, with an FDR
11 corrected p-value ≤ 0.05 , according to Baggerly's test. Such conservative parameters allowed us to
12 establish significant differences between samples and between treatments, in spite of the low
13 number of non-specific matches assigned randomly by CLC-BIO Genomic Workbench 7.0.3 to
14 different matching sequences and of the small sample size (two replicates) used. Gene expression
15 profiles were subdivided into nine groups, those remaining constant, those increasing their
16 expression in D1 or in D2, or in both treatments, those reducing their expression in D1 or D2, or in
17 both treatments, those increasing their expression in D1 and reducing in D2 and vice versa.
18

30 **Isolation and analysis of consensus *P. x canadensis* LTRs**

31 Consensus 5'-LTR sequences of *P. x canadensis* hybrids 85 and 89 were obtained as follows. First,
32 5'-LTRs were extracted from the *P. trichocarpa* full-length LTR-REs using an internally developed
33 perl script. Then, consensus LTR sequences were obtained by mapping (using CLC-BIO Genomic
34 Workbench 7.0.3) Illumina DNA reads of the two hybrids to *P. trichocarpa* LTRs, with the
35 following parameters: Similarity = 0.7, Length fraction = 0.7, Insertion cost = 1, Deletion cost = 1,
36 Mismatch cost = 1.
37

38 Consensus 5'-LTR sequences of both hybrids were subjected to motif search by CLC-BIO
39 Genomic Workbench 7.0.3 using a list of putative drought responsive motifs (Supplementary
40 material 1) selected from the motif list downloaded from the PLACE website
41 (<http://www.dna.affrc.go.jp/PLACE/index.html>; Higo et al. 1999).
42

54 **Analysis of gene expression**

55 The *P. trichocarpa* unigene model database version 9.1, available at the Phytozome site
56 (<http://www.phytozome.net/poplar>) (Tuskan et al. 2006) was annotated using the web tool
57

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

BLAST2GO (<http://www.blast2go.com/b2ghome>). Eighty-two unigenes related to chromatin methylation and remodelling were selected (Supplementary material 2). CLC-BIO Genomic Workbench 7.0.3 was used to generate cDNA sequence alignments to these genes using the same parameters as above and RPKM was calculated for each gene.

Those genes mapped by at least one read per million of reads in at least one sample were considered to be expressed. The weighted proportion fold changes between RPKM in control leaves of the two hybrids and in leaves of moderately or severely droughted plants and control plants were considered as different when the RPKM was at least twofold higher or lower in one treatment than in the other, with an FDR corrected p-value ≤ 0.05 , according to Baggerly's test (Baggerly et al. 2003).

Reverse transcription polymerase chain reaction (RT-PCR) experiments were carried out on specific genes, selected according to RNA-Seq data. First-strand cDNA synthesis was performed with 3 μg of purified total RNA using MMLV reverse transcriptase RNase H⁻ (Solis Biodyne, Tartu, Estonia) according to the manufacturer's instructions. Forward and reverse specific primers were designed for four genes involved in chromatin remodelling (Potri.007G026700.1, Potri.013G006000.1, Potri.018G138000.1, Potri.T046100.1) and one gene with constant expression level (Potri.004G152900.1) (Supplementary material 3). Equal template amounts used in these experiments were verified by RT-PCR performed with this last gene by stopping cycle amplification during the exponential phase of PCR and checking the amounts of products by separating them in 2 % agarose and GelRedTM (Biotium, Hayward, CA) staining.

The PCR amplifications were carried out in a 15 μl volume by using 4 μl of 1:4 dilution of RT reaction as template and a HOT FIREPol DNA Polymerase (Solis Biodyne) in nonsaturating conditions. PCR involved a 95 °C step hold for 15 min, followed by 33 cycles (27 for the reference gene) at 95 °C for 30 s, 62°C (58°C for the reference gene) for 30 s, and 72 °C for 10 s (30 s for the reference gene). The PCR products were separated in 2 % agarose and GelRed stained. The RT-PCR experiments were repeated three times.

Results

LTR-retrotransposons in two *Populus x canadensis* hybrids

Two samples of 23,121,470 DNA reads of hybrids 85 and 89, each covering two genome equivalents (Table 1), were aligned to the *P. trichocarpa* set of full-length LTR-REs (Natali et al. 2015). All *P. trichocarpa* LTR-REs had DNA reads that mapped to them. The percentages of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

matched nucleotides resulting, were similar in the two hybrids, i.e. 14.76 vs. 13.05% (Table 1). As a control, the same read packages were aligned to a set of 33 unique, well conserved genes of *P. trichocarpa*, and the percentage of matched nucleotides were also similar, i.e 0.008 vs. 0.007% (data not shown). Mapping Illumina DNA reads of *P. trichocarpa* (Slavov et al. 2012) showed 13.38% nucleotides matching the set of LTR-REs and 0.008% matching the 33 genes (data not shown).

Fig. 1 shows the correlation between the number of reads matching LTR-REs in the hybrids 85 and 89. Only 6/546 *Gypsy* REs and no *Copia* RE showed a ratio higher than 2.0, between the two hybrids.

Expression level of LTR-REs in leaves of *Populus x canadensis* hybrids

We generated 76,635,449 Illumina cDNA reads, encompassing 3.9 Gb of sequence data (Table 2). Each condition (**control, moderate and severe drought**) was represented by at least 3.96 million reads, a tag density sufficient for quantitative analysis of gene expression (Rai et al. 2013).

We observed that a small portion (0.09-0.16%) of each RNA-seq read set matched the retrotransposons included in the dataset (Table 2). Separating *Gypsy* and *Copia* REs, the mean percentages were 0.112% (0.072-0.153%) and 0.010% (0.005-0.022%), respectively (data not shown). This indicated that *Gypsy* elements were far more highly expressed than *Copia* elements.

Expression analysis of LTR-REs in leaves of normally watered plants

We considered all those LTR-REs mapped by at least one read per million of reads in at least one sample as transcribed (Lu et al. 2013). The correspondence between RNA-Seq data and expression analyses by RT-PCR had already been established for these experiments (Cossu et al. 2014).

It is to be noted that CLC-BIO assigns randomly non-specifically matched reads, i.e. reads that align at more than one position with an equally good score. **However, in our experiments, the number of non-specific reads was low. For example, the average percentage of non-specific matches in control leaves was 7.15 ± 1.41 for LTR-REs.** Similar percentages were observed for all treatments (data not shown).

Fifty-five LTR-REs out of 958 (5.7%) were expressed in both hybrids, these belonged to the *Gypsy* (35 elements) and *Copia* (16) superfamilies, and to a group of unknown elements (4 elements). Forty-five LTR-REs (4.7%) were expressed in control leaves of one or other of the

1 hybrids (35 in hybrid 85 and 10 in hybrid 89). Finally, 858 LTR-REs (89.6%) were transcriptionally
2 inactive in control leaves of both hybrids.

3
4 The distribution of active LTR-REs in relation to their expression in normally watered plants
5 is reported in Fig. 2, keeping the two hybrids separate. *Gypsy* REs were by far the most active;
6 considering the 30 most expressed LTR-REs, in both hybrids, twenty-one REs belonged to the
7 *Gypsy* superfamily and six to the *Copia*; three of the most active elements were unclassified (data
8 not shown). The majority of LTR-REs transcribed in at least one hybrid (83 out of 100) were
9 expressed at comparable levels in both hybrids (42) or were more expressed in hybrid 85 than in
10 hybrid 89 (41) (Fig. 3). Only 12 LTR-REs were more highly expressed in hybrid 89 than in hybrid
11 85 and showed the highest mean expression levels (expressed as Reads per Kilobase per Million
12 reads mapped, RPKM) and the largest RPKM variability. No LTR-RE expression specific to hybrid
13 89 was found and only 5 LTR-REs were expressed specifically in hybrid 85 (Fig. 3).

14
15
16 Since in certain cases the occurrence of LTR-RE sequences in cDNA libraries can be related
17 to genomic DNA contamination, we analysed the RPKM value of each retroelement in relation to
18 its abundance in the two hybrids (Fig. 4). The most abundant LTR-REs were not (or were slightly)
19 expressed and correspondingly, the most expressed LTR-REs were poorly represented in the
20 genomes of both hybrids. This indicated that contamination by genomic DNA in the cDNA libraries
21 could be largely ruled out.

22 23 24 25 26 27 28 29 30 31 32 33 34 **Expression analysis of LTR-REs in leaves of plants subjected to water deprivation**

35
36
37 One hundred and forty and 182 LTR-REs were expressed in at least one culture condition in hybrids
38 85 and 89, respectively. A number of active LTR-REs were expressed in leaves of control,
39 moderately and severely droughted plants, in both hybrids (Table 3). In addition to this group, the
40 other major group was represented by LTR-REs expressed only in severely droughted plants (D2),
41 especially in hybrid 89 (86 vs. 36, i.e., 47.3 vs. 25.7%).

42
43
44
45
46
47 In all cases the most abundant LTR-REs were not (or only slightly) expressed (data not
48 shown), indicating that contamination by genomic DNA in the cDNA libraries was not present, as
49 for control plants.

50
51
52
53 Only 63 LTR-REs were expressed in leaves of droughted plants of both hybrids (Table 3),
54 i.e. the expression of all other LTR-REs were hybrid-specific. Forty-seven out of 86 elements
55 expressed in D2 in hybrid 89 were not expressed in hybrid 85, and 8 out of 36 elements expressed
56 in D2 in hybrid 85 were not expressed in hybrid 89, indicating large differences in LTR-RE
57 transcription between hybrids in droughted plants.

1
2
3
4
5
6
7
8
9
10
11
12
13
As the parents were heterozygous, the two poplar inter-specific hybrids were expected to be genetically different. Modulation of LTR-RE transcription in response to water deprivation was different between the two hybrids (Table 4). The total number of expressed LTR-REs was similar throughout the experiment in hybrid 85 and progressively increased in hybrid 89. The number of transcribed *Copia* REs slightly reduced in hybrid 85 and by contrast, progressively increased in hybrid 89. The number of transcribed *Gypsy* REs increased during the experiment, especially in hybrid 89. Unknown elements were few and their expression was apparently constant during the experiment (Table 4).

14
15
16
17
18
19
20
21
22
23
24
Concerning the expression level, in hybrid 85 the mean RPKM value increased after moderate drought and showed a marked reduction in severely droughted plants for *Copia* and *Gypsy* elements; similar patterns of expression was observed in hybrid 89 (Table 4). Unknown REs showed higher expression levels than *Copia* and *Gypsy* elements in control plants, their expression reduced in moderately droughted plants of hybrid 85 and maintained their expression levels in hybrid 89 (Table 4).

25
26
27
28
29
30
31
32
33
34
35
36
37
The expression profiles of transcriptionally active LTR-REs in the two hybrids were subdivided into nine clusters (Fig. 5) based on their expression modulation pattern. The majority of expressed LTR-REs were transcribed at constant rates in control and droughted plants (cluster e, 66.4% in hybrid 85 and 53.3% in hybrid 89). It is worth noting that LTR-REs activated by water deprivation (clusters a and b) were far more numerous in hybrid 89 than in hybrid 85 (34.6 vs. 2.9%, respectively); conversely, LTR-REs repressed by water deprivation (clusters h and i) were found only in hybrid 85.

38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
The expression of single elements was evaluated by pairwise comparisons of mean log RPKMs for each retrotransposon in leaves of C, D1, and D2 plants (Fig. 6). The correlation was highly significant in each condition tested, indicating the same retrotransposons were expressed at comparable levels in different genetic backgrounds. Keeping *Gypsy* and *Copia* elements separate, the correlation coefficient progressively increased from control plants to D1 through to D2 plants, especially for *Gypsy* elements. Interestingly, the correlation coefficient for *Copia* REs was not significant in the controls and at both levels of water deprivation whereas they were highly significant for *Gypsy* REs (Fig. 6).

53
54
55
56
57
58
59
60
61
62
63
64
65
Retroelements whose expression was induced, repressed, or unchanged (in D1 and/or D2) were analysed for the occurrence of sequence motifs putatively recognizable by transcription factors activated by drought, in their 5'-LTR (Supplementary material 1). For this analysis, *Populus x canadensis* consensus 5'-LTRs were obtained by mapping Illumina DNA reads of hybrids 85 and 89 to *P. trichocarpa* 5'RE-LTRs. We observed that the mean number of drought-related cis-

1 regulatory motifs was similar in LTRs of all expressed Res, whether activated or repressed (Table
2 5).

3 4 5 **Expression analysis of genes involved in chromatin methylation and remodelling** 6

7
8
9 A collection of 82 *P. trichocarpa* genes involved in chromatin methylation and remodelling
10 (Supplementary material 2) was prepared to study possible relationships between the activation of
11 these genes and the different response of the two hybrids in terms of RE expression. Of these genes,
12 73 were expressed, although generally at low level.
13
14

15
16 In the two hybrids under control conditions, 18 genes (24.7%, of which 5 with $p < 0.05$)
17 showed expression levels two-fold higher in hybrid 85 than in hybrid 89 and none were two-fold
18 more expressed in hybrid 89, indicating a generally stronger activation trend of this group of genes
19 in the hybrid 85 (Fig. 7a). Conversely, many genes were activated in leaves of droughted plants of
20 hybrid 89, while small changes in the expression levels were observed in droughted leaves of hybrid
21 85 (Fig. 7b).
22
23
24
25
26

27 We also specifically studied four genes whose involvement in DNA methylation and
28 activation/repression of REs in *Arabidopsis* has been reported in the literature, i.e. three *Decrease in*
29 *DNA Methylation* (*DDM1*, *DDM2.1* and *DDM2.2*) genes and one *Morpheus* (*MOM*) gene (Iwasaki
30 2014) (Fig. 8). In control leaves, *DDM1* was more highly expressed in hybrid 89 than in hybrid 85;
31 *MOM* was more highly expressed in hybrid 85. Concerning *DDM1*, RT-PCR experiments
32 confirmed this gene was more highly expressed in hybrid 89 than in hybrid 85 (Fig. 8), while the
33 other three genes did not show appreciable differences in expression (data not shown). Concerning
34 water deprivation treatments, no significant expression variability was observed in both hybrids, for
35 the four selected genes, compared to control leaves, except for *DDM2* and *MOM* which were more
36 highly expressed in D1 in hybrid 89 (data not shown).
37
38
39
40
41
42
43
44
45
46

47 **Discussion** 48 49 50

51 We used a LTR-RE dataset of *P. trichocarpa* containing 958 full-length retroelements (Natali et al.
52 2015) for the evaluation of LTR-RE expression in *Populus x canadensis* (*P. deltoides* x *P. nigra*).
53 We used the LTR-REs of *P. trichocarpa* because this species can be crossed with both *P. deltoides*
54 and *P. nigra* and diverged from them 8 -13 million years ago (Sterck et al. 2005). *P. trichocarpa*
55 LTR-REs could differ in sequence and in abundance from elements of *P. deltoides* and *P. nigra*.
56
57
58
59
60
61
62
63
64
65

1
2 However, previous data from the literature showed that sequence differences are rare among poplar
3 species, at least for genes (Maestrini et al. 2009; Cossu et al. 2012; 2014).

4 Mapping Illumina DNA reads to the *P. trichocarpa* LTR-REs showed nearly the same
5 percentage of matched nucleotides for the two hybrids (27.22 vs. 29.23%). All 958 LTR-REs were
6 shown to be present in the genomes of the two hybrids, as well as in that of *P. trichocarpa*. The
7 abundance of each LTR-RE was similar in the two genotypes analysed in this study. **Differences in**
8 **abundance were observed only for a few elements.** It is possible that such differences are in part
9 related to LTR-RE copy number variation between homologous chromosomes in the two parents.
10 Alternatively, it is possible that, after inter-specific hybridization, a few elements have been
11 subjected to amplification, as observed in newly formed inter-specific hybrids of other species (e.g.
12 Feldman and Levy 2012). It is known that inter-specific hybridization causes so called “genomic
13 shock” (McClintock 1984), leading to activation of transposons (Senerchia et al. 2015).
14 Generally, the expression of LTR-REs in plant species has been argued especially surveying EST-
15 and cDNA-libraries available in public databases (Vicent et al. 2001). **In these databases, most**
16 **elements resulted transcribed at low levels. In poplar also, our data showed low expression levels of**
17 **LTR-retrotransposons in all analysed plants.** In fact, in the various RNA-Seq samples, the
18 percentages of cDNA nucleotides that mapped to LTR-REs were around 0.1%, i.e. LTR-RE RNAs
19 represented only a small fraction of the transcriptome.
20
21

22 Low levels of transcription of repeated sequences are often attributed to DNA contamination
23 of RNA samples. **In the experiments described here, LTR-RE abundance and transcription were**
24 **totally uncorrelated. Abundant LTR-REs were not, or were only slightly expressed,** whereas rare
25 REs were actively transcribed. This suggests that the presence of RE sequences in the cDNA library
26 was not due to DNA contamination. **The lack of correlation between LTR-RE abundance and**
27 **transcription is not surprising because it is known that abundant elements are more easily**
28 **recognized and subjected to RNA silencing** (Meyers et al. 2001; Yamazaki et al. 2001; Lisch 2009).
29
30

31 In our experiments, *Copia* and *Gypsy* REs were transcribed at different levels, the least
32 active belonging to the *Copia* superfamily, in both hybrids and in all culture conditions tested in our
33 experiments. It is known that diversified silencing mechanisms occur in plants (Xie et al. 2004).
34 These results indicate that different silencing mechanisms may differentially affect the two RE
35 superfamilies.
36
37

38 Hybrids 85 and 89 showed differences in the expression level of a number of LTR-REs in
39 control plants. In fact, although the correlation between LTR-REs expression in the two hybrids
40 were highly significant, a number of retroelements were differentially transcribed in the two
41 hybrids. In general, more LTR-REs were constitutively transcribed (and at higher rates) in hybrid
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

85 than in hybrid 89. These differences were accompanied by differences in the expression of a number of genes involved in chromatin methylation and remodelling. It is presumed that differences were related to genetic differences between hybrids, for example to trans-acting factors more active in one hybrid than in the other. For example, the *DDMI* gene was more highly expressed in hybrid 89 than in hybrid 85. It is known that, in *Arabidopsis*, *DDMI* allows the methylation status to be maintained, resulting in transposon inactivation (Mirouze and Paszkowski 2011). The lower *DDMI* transcript accumulation in hybrid 85 could explain the generally higher LTR-RE expression in this hybrid.

If the different pattern of RE expression in the leaf of the two sister hybrids were to also occur in the vegetative meristem, this would imply that, during evolution, the same hybridization event could determine different results in terms of genome structure of the resulting interspecific hybrid. In fact, if RE expression were followed by retrotranscription and insertion in a new site, copy number variations of those elements would be produced during evolution, contributing to genome structure alteration and consequently to the generation of new species. For example, it is known that *Helianthus anomalus*, *H. deserticola*, and *H. paradoxus* were derived from interspecific hybridization between *H. annuus* and *H. petiolaris*, but the LTR-RE portions of their genome are very different (Ungerer et al. 2009).

It remains to be determined whether the occurrence of LTR-RE transcripts in poplar leaves has any biological function. LTR-RE transcripts could be involved in the insertion of new copies in the genome. If the transposition mechanism (involving three types of transcripts) reported in barley (Chang et al. 2013) were general, the polyadenylated transcripts analysed in our experiments could lead to retrotransposition only in the presence of other uncapped and not polyadenylated transcripts. Much data obtained in other species suggests that LTR-RE transcripts have acquired specific cell functions during their evolution, especially in the production of siRNAs involved in transcriptional silencing of REs and, eventually, nearby genes through the RNA-directed DNA methylation pathway (Slotkin and Martienssen 2007; Lisch 2009). In maize, siRNAs, 21- and 22-nt in length, are derived from distinct retrotransposon families, and are differentially accumulated between genotypes (Barber et al. 2012). In soybean, another study indicated the 22 nt long siRNAs resulting from different retrotransposon families, as a major component in the silencing of most REs (Zabala et al. 2012).

Concerning retrotransposon transcription in response to drought treatments, it appeared stable for most elements. Many elements were constitutively expressed, in the leaves of both control and water-deprived plants. Interestingly, the two hybrids showed a different pattern of expression during water deprivation treatment, i.e. the number of expressed LTR-REs increased in hybrid 89

1 and remained stable (apart from a small reduction in D1) in hybrid 85. Moreover, many
2 retroelements were differentially transcribed in the two hybrids during drought treatments. Different
3 expression modulation patterns are presumably related to differences in regulatory sequences or
4 silencing mechanisms between hybrids. Another hypothesis is that different RE expression patterns
5 are related to their co-localization with genes involved in drought response.
6

7 Differences between the hybrids were also seen in the number of expressed elements during water
8 deprivation between *Copia* and *Gypsy* superfamilies. In fact, the number of expressed *Copia* REs
9 increased in hybrid 89 and decreased in hybrid 85; the number of expressed *Gypsy* elements
10 increased in D2 in both hybrids.
11

12 We showed differences in transcript accumulation of genes involved in chromatin
13 methylation and remodelling between the two hybrids, analogous to those observed for LTR-RE
14 transcription, with a number of genes activated in droughted plants of hybrid 89 only. The activated
15 genes in droughted hybrid 89 plants included *DDM2* and *MOM*. It was hypothesized that activation
16 of these genes constituted a response mechanism to the drought-related activation of LTR-REs,
17 because their activity should increase DNA methylation of LTR-REs (Mirouze and Paszkowski
18 2011). Our data for control leaves indicated that at least one *DDM* gene (*DDMI*) is involved in the
19 general maintenance of the chromatin status, and its expression level is related to differences in
20 LTR-RE expression between the two hybrids. On the other hand, probably via different molecular
21 pathways, other *DDM* genes (*DDM2*) are activated in response to drought to counteract LTR-RE
22 activation. This confirms the existence of different mechanisms of controlling LTR-RE
23 transcription acting over the long term (for some generations) or transiently, as a stress response
24 (Pecinka and Mittelsten Scheid 2012).
25

26 Searching for drought-related cis-regulatory motifs in the consensus LTRs of REs showed
27 the occurrence of many ABRE and DRE motifs in the consensus LTRs of both hybrids. It is known
28 that these motifs are crucial for expression of genes activated by dehydration (Yamaguchi-
29 Shinozaki and Shinozaki 2006). However, a similar number of these drought related cis-regulatory
30 elements were present in LTRs of all expressed elements in both induced and repressed REs. This
31 result suggests that activity of retrotransposons should be related to the presence in LTRs of cis-
32 regulatory motifs other than ABRE and DRE, or to differences in condensation/decondensation of
33 chromatin (determined by siRNAs).
34

35 In conclusion our data provides evidence of the low expression, in poplar hybrids, of a large
36 set of LTR-REs, especially belonging to the *Copia* superfamily, suggesting that LTR-RE activity is
37 strongly controlled at the transcriptional level. However, the observed small differences in LTR-RE
38 abundance between hybrids and the enhanced transcription of a few elements suggest that LTR-REs
39

1 contributed and are probably still contributing to the genome differentiation in *Populus* species.
2 This study also shows that differences in LTR-RE expression occur between genetically close
3 individuals (two sister hybrids) and suggests that they can be related to the activity of a few genes,
4 involved in chromatin remodelling.
5
6
7
8
9

10 **Acknowledgements** This research work was supported by PRIN-MIUR, Italy, Project “Verso la
11 delucidazione delle basi molecolari dell’eterosi nelle piante coltivate: variazione cis-regolatoria ed
12 espressione genica in ibridi di pioppo.”
13
14

15
16 Thanks are due to Dr. Catherine Bastien (INRA-UAGPF, Orleans, France) for providing the
17 hybrid plants; to Pierfrancesco Miscali (DiSAAA-a, Pisa, Italy) for technical collaboration on data
18 handling and treatment; to Dr. John A. Walsh (Warwick University, UK) for critical reading of the
19 manuscript.
20
21
22
23
24

25 **Data archiving statement** All LTR-RE sequences are available in the Department of Agriculture,
26 Food and Environment of the University of Pisa repository website
27 (<http://www.agr.unipi.it/ricerca/plant-genetics-and-genomics-lab/sequencerepository.html>).
28
29

30 All genomic DNA and cDNA raw Illumina sequences used in this work are available at the
31 NCBI Sequence Read Archive under the accession numbers SRPXXXXXX (Submission: to be
32 ultimated) and SRP024267 (Submission: *Populus x canadensis* RNAseq).
33
34
35

36 The global analysis of LTR-RE and gene expression is available in an excel file in the
37 Department of Agriculture, Food, and Environment of University of Pisa repository website
38 (<http://www.agr.unipi.it/Sequence-Repository.358.0.html>), in which each LTR-RE and each gene
39 was represented by its absolute expression level in control, moderately dehydrated and severely
40 dehydrated leaves.
41
42
43
44
45
46

47 **References**

- 48
49
50
51 Baggerly KA, Deng L, Morris JS, Aldaz CM (2003) Differential expression in SAGE: Accounting
52 for normal between-library variation. *Bioinformatics* 19:1477-1483
53
54 Barber WT, Zhang W, Win H, Varala KK, Dorweiler JE, Hudson ME, Moose SP (2012) Repeat
55 associated small RNAs vary among parents and following hybridization in maize. *Proc Nat*
56 *Acad Sci USA* 109:10444-10449
57
58
59
60
61
62
63
64
65

- 1
2 Barghini E, Natali L, Cossu RM, Giordani T, Pindo M, Cattonaro F, Scalabrin S, Velasco R,
3 Morgante M, Cavallini A (2014) The peculiar landscape of repetitive sequences in the olive
4 (*Olea europaea* L.) genome. *Genome Biol Evol* 6:776-791
- 5 Barghini E, Natali L, Giordani T, Cossu RM, Scalabrin S, Cattonaro F, Šimková H, Vrána J,
6 Doležel J, Morgante M, Cavallini A (2015) LTR retrotransposon dynamics in the evolution of
7 the olive (*Olea europaea*) genome. *DNA Res* 22:91–100
- 8
9
10 Bennetzen JL (2000) Transposable elements contributions to plant gene and genome evolution.
11 *Plant Mol Biol* 42:251-269
- 12
13
14 Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A (2005) Evolution of DNA sequence
15 nonhomologies among maize inbreds. *Plant Cell* 17:343–360
- 16
17
18 Buti M, Giordani T, Vukich M, Gentzbittel L, Pistelli L, Cattonaro F, Morgante M, Cavallini A,
19 Natali L (2009) HACRE1, a recently inserted copia-like retrotransposon of sunflower
20 (*Helianthus annuus* L.). *Genome* 52:904–911
- 21
22
23 Buti M, Giordani T, Cattonaro F, Cossu RM, Pistelli L, Vukich M, Morgante M, Cavallini A, Natali
24 L (2011) Temporal dynamics in the evolution of the sunflower genome as revealed by
25 sequencing and annotation of three large genomic regions. *Theor Appl Genet* 123:779-791
- 26
27
28 Buti M, Giordani T, Vukich M, Pugliesi C, Natali L, Cavallini A (2013) Retrotransposon-related
29 genetic distance and hybrid performance in sunflower (*Helianthus annuus* L.). *Euphytica*
30 192:289–303
- 31
32
33
34 Chang W, Jääskeläinen M, Li S, Schulman AH (2013) BARE retrotransposons are translated and
35 replicated via distinct RNA pools. *PLoS ONE* 8:e72270
- 36
37
38 Cossu RM, Buti M, Giordani T, Natali L, Cavallini A (2012) A computational study of the
39 dynamics of LTR retrotransposons in the *Populus trichocarpa* genome. *Tree Genet Genomes*
40 8:61-75
- 41
42
43 Cossu RM, Giordani T, Cavallini A, Natali L (2014) High-throughput analysis of transcriptome
44 variation during water deficit in a poplar hybrid: a general overview. *Tree Genet Genomes*
45 10:53–66
- 46
47
48 Dieguez MJ, Vaucheret H, Paszkowski J, Mittelsten Scheid O (1998) Cytosine methylation at CG
49 and CNG sites is not a prerequisite for the initiation of transcriptional gene silencing in plants,
50 but it is required for its maintenance. *Mol Gen Genet* 259:207-215
- 51
52
53
54 Doyle JJ, Doyle JL (1989) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15
- 55
56
57 Feldman M, Levy AA (2012) Genome evolution due to allopolyploidization in wheat. *Genetics*
58 192:763-774
- 59
60
61
62
63
64
65

- 1 Grandbastien MA (1998) Activation of plant retrotransposons under stress conditions. *Trends Plant*
2 *Sci* 3:181-189
- 3 Grandbastien MA (2015) LTR retrotransposons, handy hitchhikers of plant regulation and stress
4 response. *Biochim Biophys Acta* 1849:403–416
- 5 Hamanishi ET, Campbell MM (2011) Genome-wide responses to drought in forest trees. *Forestry*
6 84:273–283
- 7 Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements
8 (PLACE) database. *Nucl Acids Res* 27:297–300
- 9 Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. *EMBO J* 12:2521-
10 2528
- 11 Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996) Retrotransposons of rice
12 involved in mutations induced by tissue culture. *Proc Nat Acad Sci USA* 93:7783-7788
- 13 Ishiguro S, Ogasawara K, Fujino K, Sato Y, Kishima Y (2014) Low temperature-responsive
14 changes in the anther transcriptome's repeat sequences are indicative of stress sensitivity and
15 pollen sterility in rice strains. *Plant Physiol* 164:671-682
- 16 Ito H (2013) Small RNAs and regulation of transposons in plants. *Genes Genet Syst* 88:3–7
- 17 Iwasaki M (2014) Identification of genes preventing transgenerational transmission of stress
18 induced epigenetic states. *Proc Nat Acad Sci USA* 111:8547-8552
- 19 Jiang S, Cai D, Sun Y, Teng Y (2016) Isolation and characterization of putative functional long
20 terminal repeat retrotransposons in the *Pyrus* genome. *Mobile DNA* 7:1
- 21 Kalendar R, Vicent CM, Peleg O, Anamthawat-Jonsson K, Bolshoy A, Schulman AH (2004)
22 LARD retroelements: Conserved, non-autonomous components of barley and related
23 genomes. *Genetics* 166:1437-1450
- 24 Kawakami T, Morgan TJ, Nippert JB, Ochltree TW, Keith R, Dhakal P, Ungerer MC (2011)
25 Natural selection drives clinal life history patterns in the perennial sunflower species,
26 *Helianthus maximiliani*. *Mol Ecol* 20:2318–2328
- 27 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform.
28 *Bioinformatics* 25:1754-1760
- 29 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J et al (2009) The sequence alignment/map format
30 and SAMtools. *Bioinformatics* 25:2078-2079
- 31 Lisch D (2009) Epigenetic regulation of transposable elements in plants. *Ann Rev Plant Biol* 60:43–
32 66
- 33 Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant
34 tissues. *Anal Biochem* 163:16–20

- 1
2 Lu X, Chen D, Shu D, Zhang Z, Wang W, Klukas C, Chen LL, Fan Y, Chen M, Zhang C (2013)
3 The differential transcription network between embryo and endosperm in the early developing
4 maize seed. *Plant Physiol* 162:440–455
- 5 Maestrini P, Cavallini A, Rizzo M, Giordani T, Bernardi R, Durante M, Natali L (2009) Isolation
6 and expression analysis of low temperature-induced genes in white poplar (*Populus alba*). *J*
7 *Plant Physiol* 166:1544-1556
- 8
9
10 Marcon HS, Domingues DS, Costa Silva J, Junqueira Borges R, Filippi Matioli F, de Mattos Fontes
11 MR, Marino CL (2015) Transcriptionally active LTR retrotransposons in *Eucalyptus* genus
12 are differentially expressed and insertionally polymorphic. *BMC Plant Biol* 15:198
13
14
15
16 McClintock B (1984) The significance of responses of the genome to challenge. *Science* 226:792-
17 801
- 18
19
20 Meignin C, Bailly JL, Arnaud F, Dastugue B, Vaury C (2003) The 5' untranslated region and Gag
21 product of Idefix, a long terminal repeat retrotransposon from *Drosophila melanogaster*, act
22 together to initiate a switch between translated and untranslated states of the genomic mRNA.
23 *Mol Cell Biol* 23:8246-8254
- 24
25
26
27 Meyers BC, Tingey SV, Morgante M (2001) Abundance, distribution, and transcriptional activity of
28 repetitive elements in the maize genome. *Genome Res* 11:1660-1676
- 29
30
31 Mirouze M, Paszkowski J (2011) Epigenetic contribution to stress adaptation in plants. *Curr Opin*
32 *Plant Biol* 14:267–274
- 33
34
35 Morgante M, De Paoli E, Radovic S (2007) Transposable elements and the plant pan-genomes. *Curr*
36 *Opin Plant Biol* 10:149-155
- 37
38
39 Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying
40 mammalian transcriptomes by RNA-Seq. *Nature Meth* 5:621–628
- 41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

- 1 Pecinka A, Mittelsten Scheid O (2012) Stress-induced chromatin changes: a critical view on their
2 heritability. *Plant Cell Physiol* 53:801–808
- 3 Piegu B, Guyot R, Picault N, Roulin A, Saniyal A, Kim H, Collura K, Brar DS, Jackson S, Wing
4 RA, Panaud O (2006) Doubling genome size without polyploidization: Dynamics of
5 retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice.
6 *Genome Res* 16:1262-1269
- 7 Rai HD, Mock KE, Richardson BA, Cronn RC, Hayden KJ, Wright JW, Knaus BJ, Wolf PG (2013)
8 Transcriptome characterization and detection of gene expression differences in aspen
9 (*Populus tremuloides*). *Tree Genet Genomes* 9:1031-1041
- 10 Ramallo E, Kalendar R, Schulman AH, Martínez-Izquierdo JA (2008) *Remel*, a *Copia*
11 retrotransposon in melon, is transcriptionally induced by UV light. *Plant Mol Biol* 66:137-150
- 12 Rico-Cabanas L, Martinez-Izquierdo JA (2007) CIRE1, a novel transcriptionally active Ty1-*copia*
13 retrotransposon from *Citrus sinensis*. *Mol Genet Genomics* 277:365-377
- 14 SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of
15 intergene retrotransposons of maize. *Nature Genet* 20:43–45
- 16 Scherrer B, Isidore E, Klein P, Kim JS, Bellec A, Chalhoub B, Keller B, Feuillet C (2005) Large
17 intraspecific haplotype variability at the Rph7 locus results from rapid and recent divergence
18 in the barley genome. *Plant Cell* 17:361-374
- 19 Senerchia N, Felber F, Parisod C (2015) Genome reorganization in F1 hybrids uncovers the role of
20 retrotransposons in reproductive isolation. *Proc Royal Soc B Biol Sci* 282:DOI:
21 10.1098/rspb.2014.2874
- 22 Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, Lipphardt MF
23 et al (2012) Genome resequencing reveals multiscale geographic structure and extensive
24 linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytol* 196:713-725
- 25 Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the
26 genome. *Nature Rev Genet* 8:272-285
- 27 Sterck L, Rombauts S, Jansson S, Sterky F, Rouzé P, Van de Peer Y (2005) EST data suggest that
28 poplar is an ancient polyploid. *New Phytol* 167:165–170
- 29 Stupar RM, Springer NM (2006) *Cis*-transcriptional variation in maize inbred lines B73 and Mo17
30 leads to additive expression patterns in the F1 hybrid. *Genetics* 173:2199–2210
- 31 Suoniemi A, Tanskanen J, Schulman AH (1998) *Gypsy*-like retrotransposons are widespread in the
32 plant kingdom. *Plant J* 13:699-705

- 1
2 Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S,
3 Rombauts S, Salamov A et al (2006) The genome of black cottonwood, *Populus trichocarpa*
4 (Torr. & Gray). Science 313:1596-1604
- 5 Ungerer MC, Strakosh SC, Stimpson KM (2009) Proliferation of Ty3/*gypsy*-like retrotransposons in
6 hybrid sunflower taxa inferred from phylogenetic data. BMC Biology 7:40
- 7
8 Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway
9 silences selfish genetic elements in the germline. Science 313:320-324
- 10
11
12 Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH
13 (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*.
14 Plant Cell 11:1769–1784
- 15
16
17
18 Vicient CM, Jaaskelainen MJ, Kalendar R, Schulman AH (2001) Active retrotransposons are a
19 common feature of grass genomes. Plant Physiol 125:1283-1292
- 20
21
22 Vicient CM, Schulman AH (2002) *Copia*-like retrotransposons in the rice genome: Few and
23 assorted. Genome Lett 1:35-47
- 24
25
26 Voytas DF, Cummings MP, Konieczny A, Ausubel FM, Rodermeil SR (1992) *Copia*-like
27 retrotransposons are ubiquitous among plants. Proc Nat Acad Sci USA 89:7124-7128
- 28
29 Vukich M, Giordani T, Natali L, Cavallini A (2009) *Copia* and *Gypsy* retrotransposons activity in
30 sunflower (*Helianthus annuus* L.). BMC Plant Biol 9:150
- 31
32
33 Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante
34 M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for
35 eukaryotic transposable elements. Nature Rev Genet 8:973–982
- 36
37
38 Wilkins O, Waldron L, Nahal H, Provart NJ, Campbell MM (2009) Genotype and time of day shape
39 the *Populus* drought response. Plant J 60:703–715
- 40
41
42 Witte CP, Le QH, Bureau T, Kumar A (2001) Terminal-repeat retrotransposons in miniature
43 (TRIM) are involved in restructuring plant genomes. Proc Nat Acad Sci USA 98:13778-
44 13783
- 45
46
47 Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE,
48 Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants.
49 PLoS Biology 2:642-652
- 50
51
52 Yamaguchi-Shinozaki K, Shinozaki K (2006) Transcriptional regulatory networks in cellular
53 responses and tolerance to dehydration and cold stresses. Ann Rev Plant Biol 57:781-803
- 54
55
56 Yamazaki M, Tsugawa H, Miyao A, Yano M, Wu J, Yamamoto S, Matsumoto T, Sasaki T,
57 Hirochika H (2001) The rice retrotransposon *Tos17* prefers low-copy-number sequences as
58 integration targets. Mol Genet Genomics 265:336-344
- 59
60
61
62
63
64
65

Zabala G, Campos E, Varala KK, Bloomfield S, Jones SI, Win H, Tuteja JH, Calla B, Clough SJ,
Hudson M, Vodkin LO (2012) Divergent patterns of endogenous small RNA populations
from seed and vegetative tissues of *Glycine max*. BMC Plant Biol 12:177

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

FIGURE LEGENDS

1
2
3 **Fig. 1** Pairwise comparison of number of mapped reads on each of the 958 *P. trichocarpa* full-
4 length LTR-REs between Illumina read sets of hybrids 85 and 89. The same number of reads was
5 used for both hybrids
6
7

8
9
10 **Fig. 2** Distribution of 206 LTR-REs (transcribed in hybrid 85 and/or 89) in leaves of control plants
11 of hybrids 85 and 89, according to their expression value (RPKM)
12
13

14
15 **Fig. 3** Box and whiskers plot of LTR-RE expression differences (calculated as RPKM) between
16 control leaves of hybrids 85 (H85) and 89 (H89). Transcribed retrotransposons of the two hybrids
17 were subdivided into 5 classes, those more highly expressed in hybrid 85 than in hybrid 89 or *vice*
18 *versa*, those equally expressed in both hybrids, those expressed only in hybrid 85, or on hybrid 89.
19 LTR-REs were considered more highly expressed in one hybrid than in the other when the RPKM
20 ratio was higher than 2. The boxes represent the 25% to 75%, whiskers represent the range of
21 RPKM, and lines in the box represent the mean values of the distribution. For each class, the
22 number of LTR-REs is reported
23
24
25
26
27
28
29
30

31
32 **Fig. 4** Relationship between RPKM expression values (in control plants) of each element of the set
33 of 958 *P. trichocarpa* LTR-REs and the respective number of mapped DNA reads in hybrids 85 and
34 89
35
36
37

38
39 **Fig. 5** Schematic representation of the expression patterns of nine LTR-RE (indicated by letters (a)-
40 (i)) observed in leaves of *P. x canadensis* hybrids (85 and 89) comparing control conditions (C) to
41 moderate (D1) and to severe drought (D2). For each pattern, the number of LTR-REs is reported in
42 hybrids 85 and 89. Only expressed LTR-REs were counted. The total number of expressed LTR-
43 REs per hybrid is reported in parentheses
44
45
46
47
48
49

50
51 **Fig. 6** Two-dimensional representations of LTR-RE expression estimated by log RPKM in leaves
52 from plants of two *P. x canadensis* hybrids normally watered (Control), moderately (Drought 1) and
53 severely droughted (Drought 2). For each treatment, the correlation coefficient between hybrid 85
54 and 89 are reported. Data are reported for all expressed LTR-REs or keeping separated *Copia* and
55 *Gypsy* elements
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
Fig. 7 Distribution of RPKM fold changes between hybrid 89 (H89) and 85 (H85) of 82 poplar genes involved in chromatin methylation and remodelling. When values were higher in H89 than in H85, they were reported as positive, when they were higher in H85 than in H89, they were reported as negative. (a) Distribution of RPKM fold changes between droughted (moderately, D1 or severely, D2) and control (C) plants of the same poplar genes as above, in H85 and H89. (b) When values were higher in droughted than in control plants, they were reported as positive, when they were higher in control than in droughted plants, they were reported as negative. Fold change values significant at $p \leq 0.05$ are indicated in red. The lines in the distributions represent the mean values for each fold change

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
Fig. 8 (a) Mean and standard deviation of RPKM values of *Decrease in DNA Methylation (DDM)* 1.1, 2.1, 2.2, and *Morpheus* in leaves of normally watered plants of hybrid 85 (H85/3 and H85/4) and 89 (H89/6 and H89/8). The Phytozome code for each gene is reported in parentheses. (b) RT-PCR analysis of *DDMI.1*. The *CLC protease - proteolytic subunit 6* encoding gene (Potri.004G152900.1) was used as control as in Cossu et al. (2014). The RPKM value is reported below each sample

Table 1 Mapping Illumina genomic DNA reads on 958 *P. trichocarpa* LTR-REs

Genotype	Nr. of reads	Average length (nt)	Coverage	Nr. of mapped reads	% of mapped reads
Hybrid 85	23,121,470	50	2.00x	3,411,873	14.76
Hybrid 89	23,121,470	50	2.00x	3,017,590	13.05

Table 2 Relative water content (RWC) in leaf samples of *P. x canadensis* plants normally watered (C) or subjected to moderate (D1) or severe drought (D2), RNA-Seq data and percentage of nucleotides matched to the set of 958 *P. trichocarpa* LTR-REs for each sample

Sample	RWC	Nr. of reads	Nr. of nucleotides	% nucleotides matched to LTR-REs
Hybrid 85/3 (C)	95.51	15,327,554	781,705,254	0.12
Hybrid 85/4 (C)	92.58	4,316,064	220,119,264	0.09
Hybrid 89/6 (C)	95.75	8,927,114	455,282,814	0.10
Hybrid 89/8 (C)	95.40	3,963,712	202,149,312	0.10
Hybrid 85/12 (D1)	86.31	5,487,345	279,854,595	0.12
Hybrid 85/24 (D1)	85.64	5,003,314	255,169,014	0.15
Hybrid 89/10 (D1)	84.89	6,123,484	312,297,684	0.15
Hybrid 89/15 (D1)	86.30	7,355,080	375,109,080	0.15
Hybrid 85/42 (D2)	54.78	3,985,186	203,244,486	0.11
Hybrid 85/45 (D2)	61.83	5,715,359	291,483,309	0.16
Hybrid 89/20 (D2)	52.78	4,575,904	233,371,104	0.15
Hybrid 89/35 (D2)	59.69	5,855,333	298,621,983	0.13
Total		76,635,449	3,908,407,899	0.13

Table 3 Number of expressed LTR-REs in each of the three conditions studied, i.e. in leaves from *P. x canadensis* hybrids 85 and 89, normally watered (C), moderately (D1) and severely droughted (D2) plants. The number of LTR-REs showing the same pattern of expression in both hybrids is also reported. The numbers in parentheses indicate the percentages

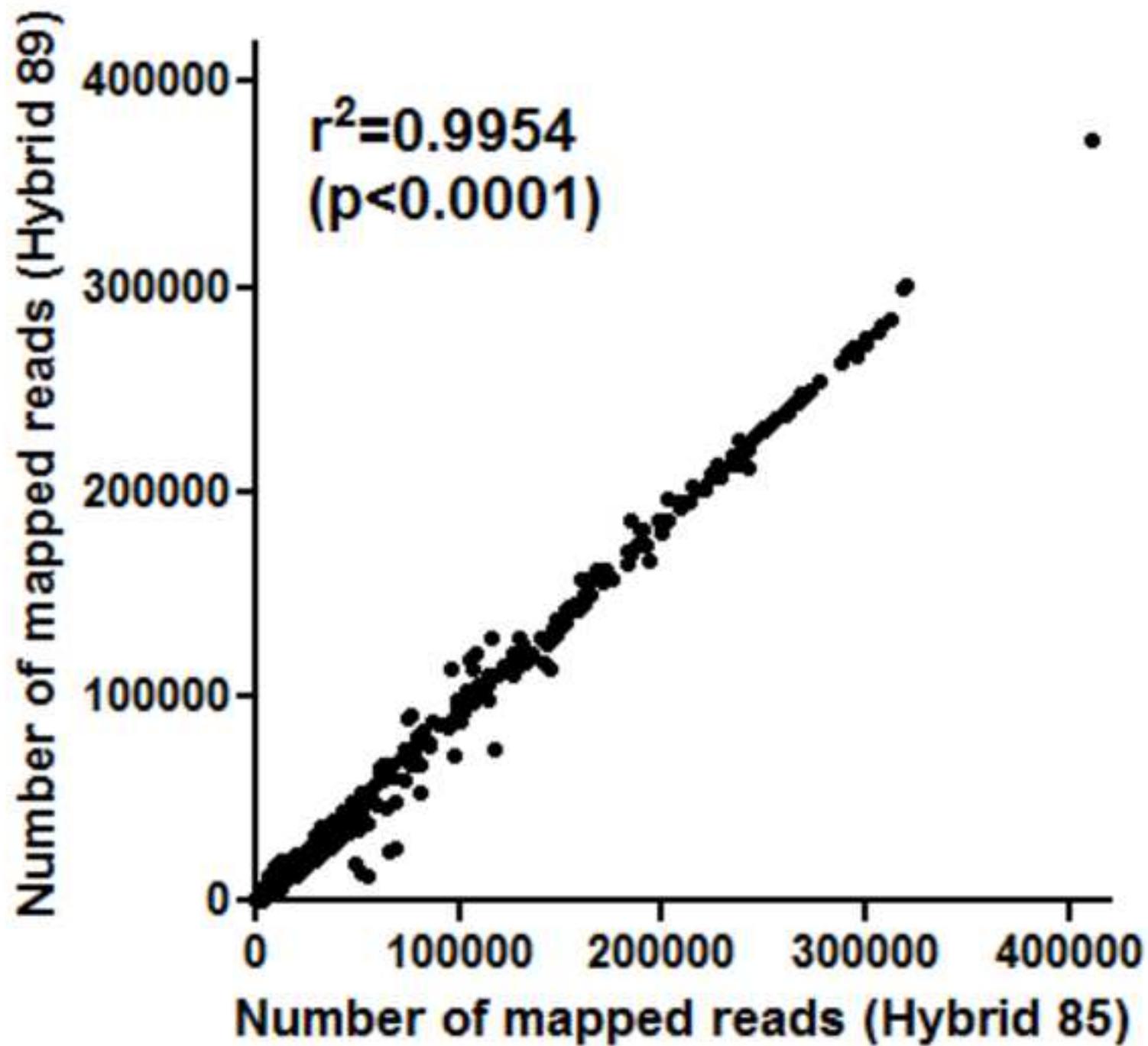
Conditions in which LTR-REs are expressed	Number of expressed LTR-REs		
	Hybrid 85	Hybrid 89	Both hybrids
C-D1-D2	47 (33.6)	56 (30.8)	37 (58.7)
C	17 (12.1)	2 (1.1)	0 (0.0)
D1	11 (7.9)	10 (5.5)	1 (1.6)
D2	36 (25.7)	86 (47.3)	22 (34.9)
C-D1	18 (12.9)	5 (2.7)	2 (3.2)
C-D2	8 (5.7)	2 (1.1)	0 (0.0)
D1-D2	3 (2.1)	21 (11.5)	1 (1.6)
Total	140	182	63

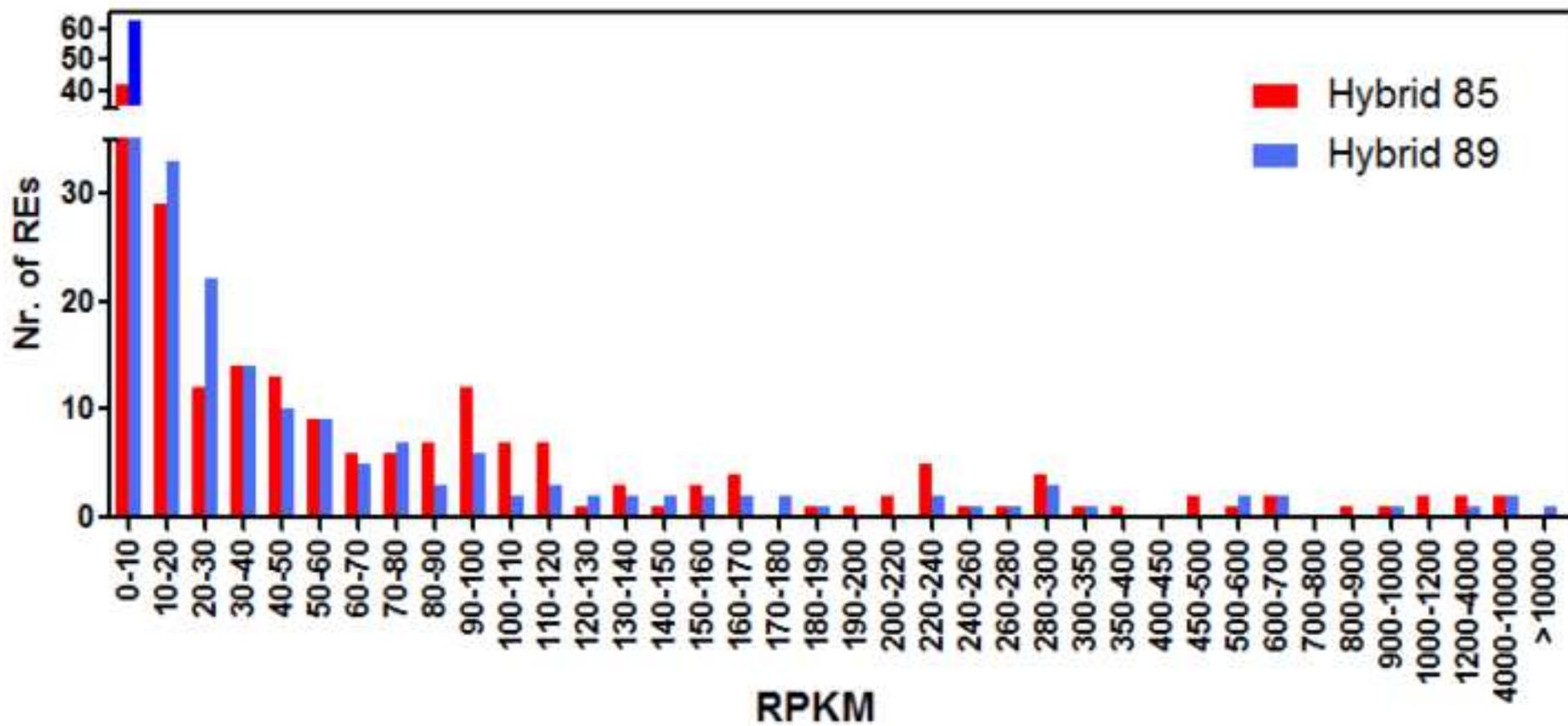
Table 4 Number of LTR-REs expressed in leaves of control (C), moderately (D1) and severely (D2) droughted plants of hybrids 85 and 89. The mean RPKM is reported for expressed LTR-REs only and for all elements of each superfamily

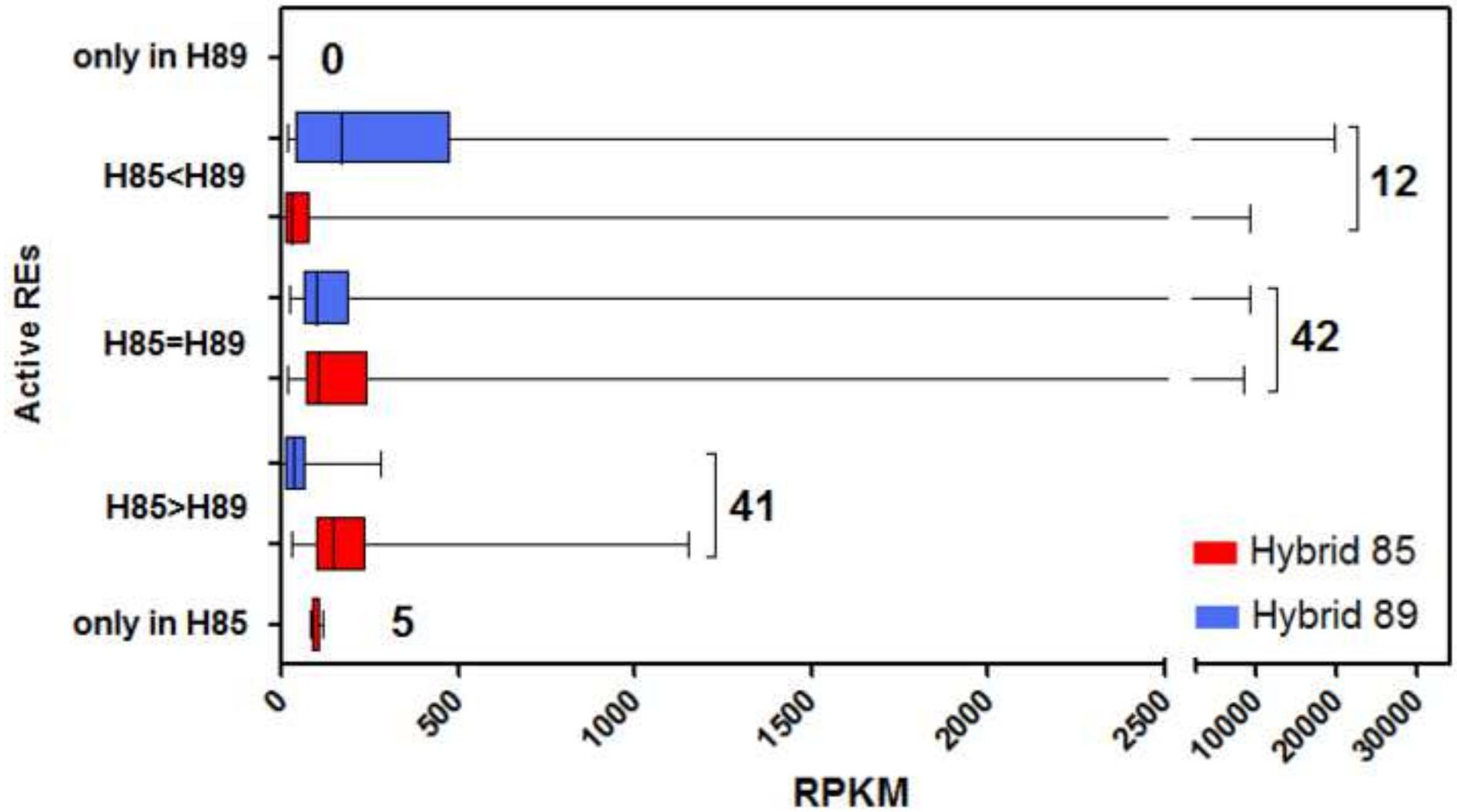
Hybrid	Superfamily	Nr. of LTR-REs in dataset	Control		D1		D2	
			Nr. of LTR-REs	Mean RPKM	Nr. of LTR-REs	Mean RPKM	Nr. of LTR-REs	Mean RPKM
85	<i>Copia</i>	394	36 (9.1%)	143.6	32 (8.1%)	166.0	31 (7.9%)	146.3
	<i>Gypsy</i>	546	49 (9.0%)	625.3	42 (7.7%)	1275.2	58 (10.6%)	936.2
	Unknown	18	5 (27.8%)	1037.7	5 (27.8%)	634.8	5 (27.8%)	681.6
	Total	958	90 (9.4%)	455.5	79 (8.2%)	785.4	94 (9.8%)	662.1
89	<i>Copia</i>	394	17 (4.3%)	124.9	34 (8.6%)	146.1	82 (20.8%)	149.0
	<i>Gypsy</i>	546	44 (8.1%)	850.5	52 (9.5%)	1040.1	77 (14.1%)	490.4
	Unknown	18	4 (22.2%)	1476.7	6 (33.3%)	1320.3	6 (33.3%)	1400.4
	Total	958	65 (6.8%)	699.3	92 (9.6%)	728.0	165 (17.2%)	353.8

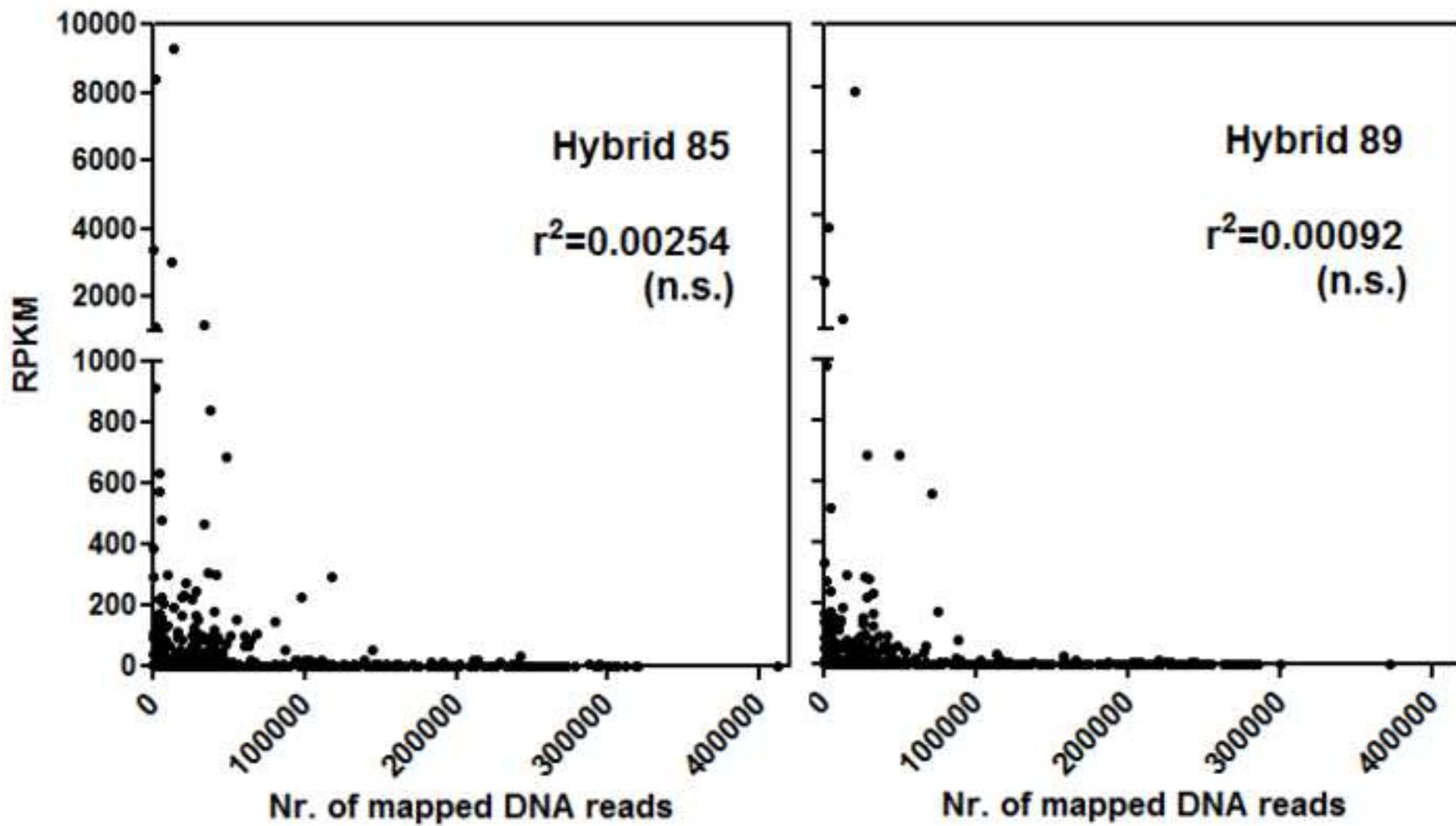
Table 5 Mean number (\pm SE) of drought related cis-regulatory motifs in the consensus LTRs of *P. x canadensis* hybrids 85 and 89, obtained mapping *P. trichocarpa* LTRs with genomic DNA reads of hybrids 85 and 89, respectively (see Materials and Methods). Three groups of LTR-REs were analysed, whose expression was a) induced, b) repressed, or c) maintained in leaves subjected to moderate (D1) and/or severe (D2) drought. (n.s.) not significant according to one-way ANOVA

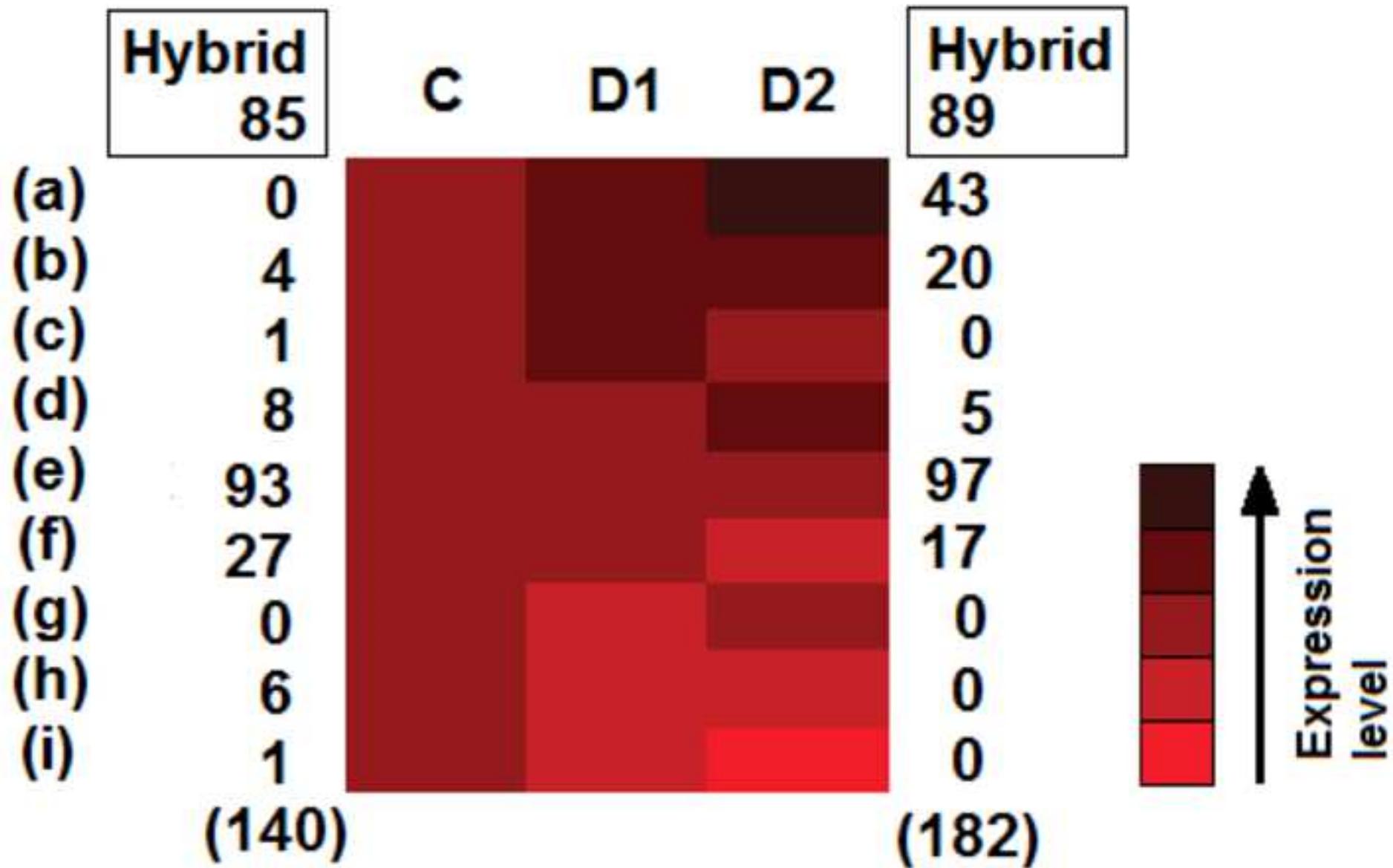
Hybrid	RE group	Nr. of expressed REs	Mean nr. of drought-related cis-regulatory motifs		
			All	ABRE	DRE
85	Induced	13	17.6 \pm 4.4	12.5 \pm 2.6	8.5 \pm 2.4
	Same expression level	93	26.2 \pm 2.7	14.7 \pm 1.2	12.9 \pm 1.4
	Repressed	34	30.9 \pm 5.2 (n.s.)	17.3 \pm 2.7 (n.s.)	13.6 \pm 2.2 (n.s.)
89	Induced	68	25.0 \pm 2.8	15.1 \pm 1.6	12.4 \pm 1.5
	Same expression level	97	24.3 \pm 2.6	13.5 \pm 1.1	11.5 \pm 1.3
	Repressed	17	23.1 \pm 4.5 (n.s.)	13.6 \pm 3.3 (n.s.)	9.1 \pm 2.3 (n.s.)

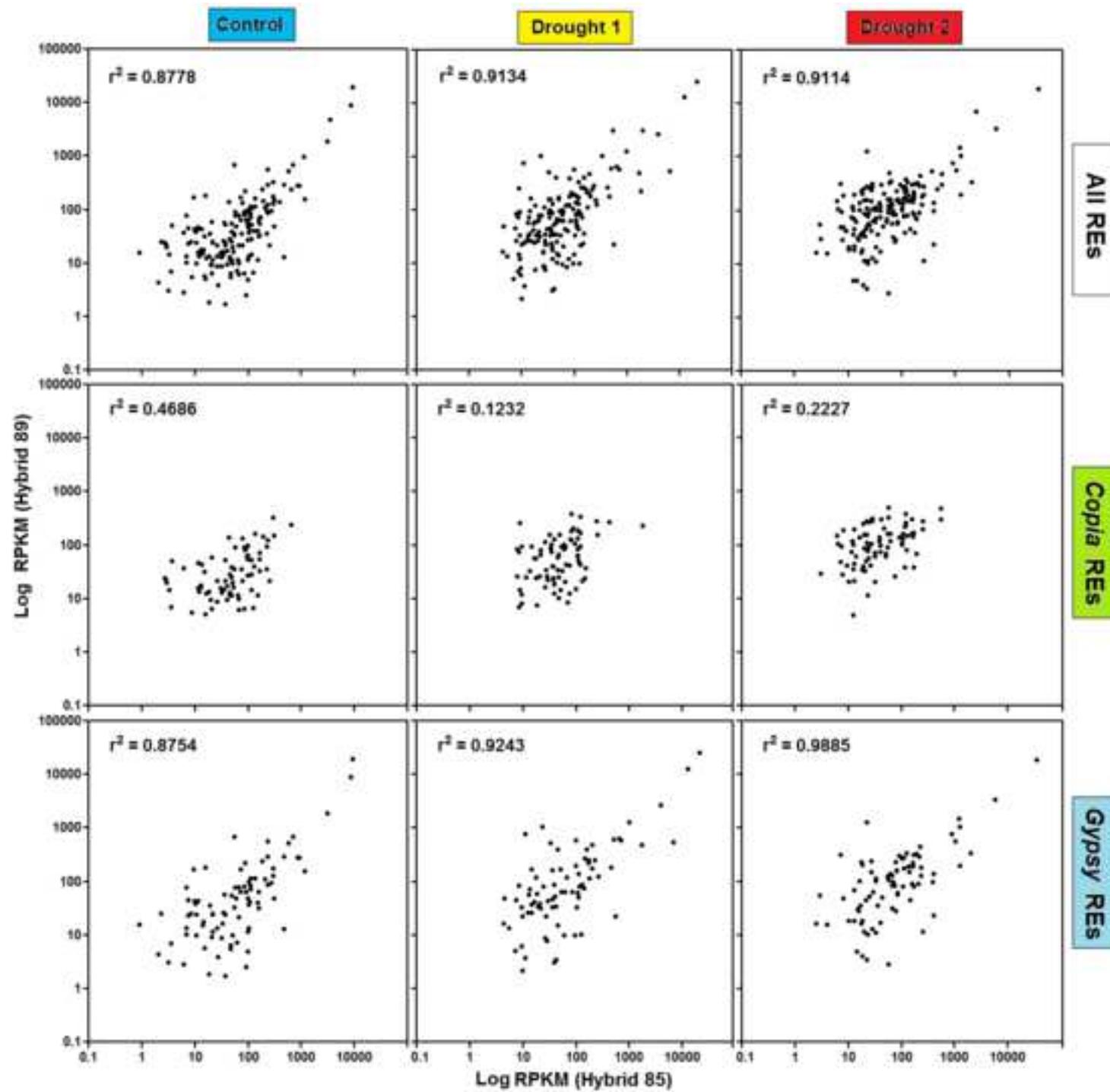


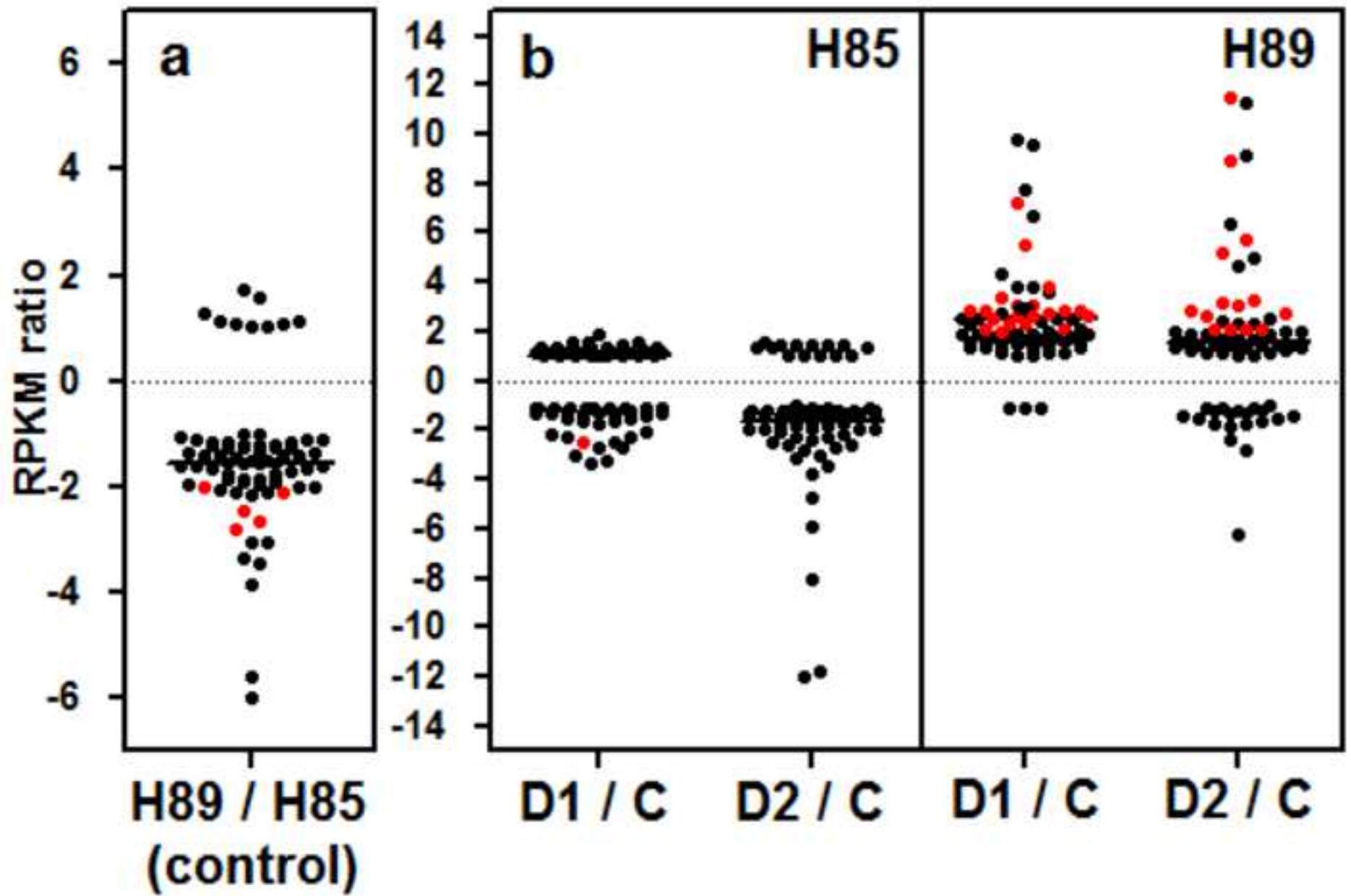


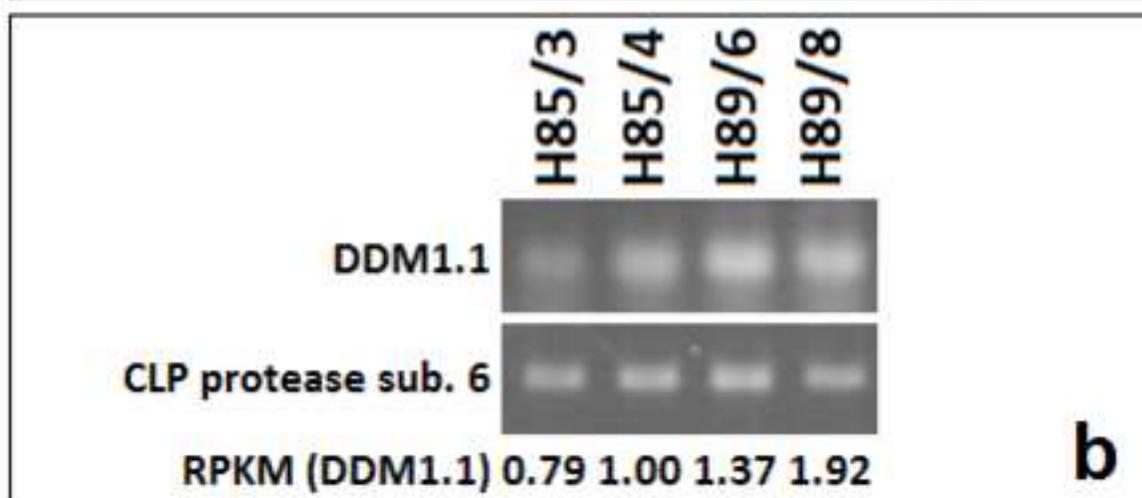
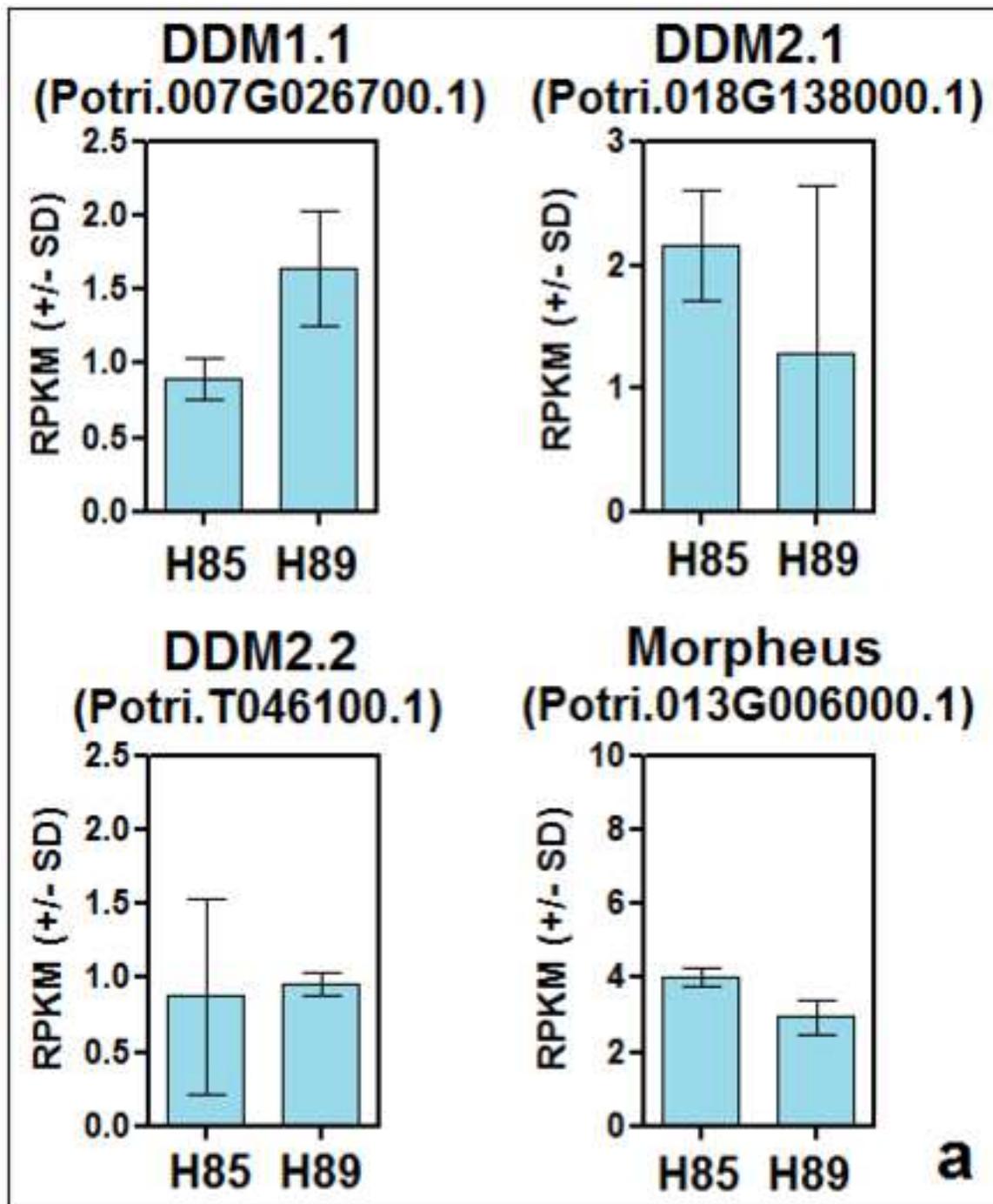


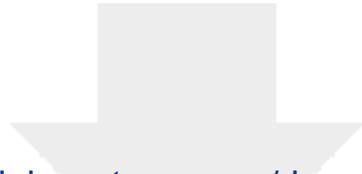




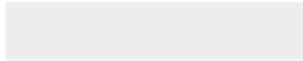








Click here to access/download
Supplementary Material
Giordani Suppl mat 1.pdf



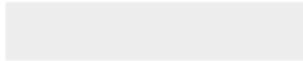


Click here to access/download
Supplementary Material
Giordani Suppl mat 2.pdf





Click here to access/download
Supplementary Material
Giordani Suppl mat 3.pdf



TGGE-D-16-00096 – Responses to reviewers

Comments for the Author:

Both reviews were overall positive, but they did raise some issues that will need to be addressed by the authors before the manuscript can be accepted for publication. -Assoc. Editor

RESPONSE OF THE AUTHORS: First of all, we wish to thank the associate editor and the reviewers for the suggested changes. In this document we answer to all suggestions of the reviewers.

Reviewer #1: This paper describes an RNA-seq experiment aimed at quantifying retrotransposon expression in two clones of a poplar hybrid. In general the paper is well-written and the data well analyzed and presented. The authors find about 100 LTR-REs expressed in each clones, with only about half of these overlapping between the clones. This is not an area for which there isn't a great deal of data in poplar, and so the paper should be of interest. A few specific comments follow.

- What was the motivation for studying expression of REs under drought? Do we expect leaky expression of these elements as a function of high transcriptional activity associated with the drought response and perhaps associated chromatin remodelling?

R.: The aim of our work was to study variability of retrotransposon expression in two highly related genotypes (i.e. two sister hybrids). We decided to study leaves of plants under control and stress conditions because it is known the retrotransposons are often activated by stress. As a stress we choose water deficit, because this work is related to a project on poplar response to drought.

We wish to point out that – in the manuscript – differences in the expression of chromatin remodeling associated genes were found between regularly watered hybrids (see Figure 8).

- What of all the other mRNA transcripts that would have been found in this experiment? I assume the authors are holding these data for an additional paper, but could they be used to inform our understanding of why these REs were expressed? For example, are the REs highly co-expressed with transcripts that are co-located with the RE in the genome?

R.: Actually, a paper on mRNA expression in the very same plants was already published (TGGE CITAZIONE) and is regularly cited in the manuscript. We agree with reviewer 1 that REs might be co-expressed with genes that are located in their proximity, in fact we are preparing a manuscript in which we will analyse the influence of chromosomal location on the expression of genes and repeated sequences, also in relation to the DNA methylation status (analyzed by bisulfite sequencing). So, we would prefer not using this kind of data here. However, we inserted a sentence at this regard in the discussion.

- The authors note that LTR-REs may be involved in generation of siRNAs. Is there any way to predict whether the specific REs found in this study could have this function? This would add to the biological interest of the study.

R.: Unfortunately we have no data on this topics. The involvement of LTR-REs in generation of siRNAs is merely a speculation based on results obtained in other species.

- The authors note on p13-14 that differences in abundance of specific LTR-REs between the two clones may be explained by copy number variation. If this is the case, one would expect a predictable increase in expression as a function of the number of copies (for example, a doubling of the number of reads if an RE was present in two copies in one clone vs one copy in the other clone). Did the authors check whether this was the case?

R.: We checked the expression of differently abundant REs in the two hybrids and found no relation between abundance and expression level of a RE. Actually, Figure 4 clearly shows that RE expression is in general unrelated (or inversely related) to the abundance of the RE in the genome.

- ABRE and DRE promoter motifs were found in the LTR sequence, which could explain drought related

expression. However, these are short motifs and probably occur by chance in many areas of the genome. Indeed, the results suggest these are unrepresented among the induced REs, which means they probably are not related to drought induced expression of the REs (as the authors note). It would be valuable to complete an exhaustive search of known promoter elements to see if any other motifs are enriched. As it stands, the authors statement that because ABRE/DRE are not enriched, expression must be mostly due to chromatin dynamics is not well supported.

R.: We agree with the reviewer and changed the text accordingly (in red): the expression of REs is not related to the presence of ABRE/DRE motifs but it might be related to other drought responsive motifs or to chromatin dynamics.

Reviewer #2: This paper provides important insights into the expression patterns of LTR transposons in a model tree system. The combination of DNA and RNA sequencing works quite well, since it enables the researchers to tease apart various aspects of what is going on in the genome and transcriptomes of these poplar hybrids. I do not have any major problems with the paper and I only list a handful of minor suggestions, that hopefully improves the manuscript.

Overall impression: This paper provides important insights into the expression patterns of LTR transposons in a model tree system. The combination of DNA and RNA sequencing works quite well, since it enables the researchers to tease apart various aspects of what is going on in the genome and transcriptomes of these poplar hybrids. I do not have any major problems with the paper and I only list a handful of minor suggestions, that hopefully improves the manuscript.

Major comment

1. In several places in the paper you mention the redundancy of LTR-REs. It is unclear to me what that means. Examples: Page 7, line 1; page 11, line 34; page 13, line 49. Page 14, line 24. I strongly advise you to clarify this more in your paper what you mean by redundancy.

R.: We changed the word “redundancy” with “abundance”, which is a more suitable term to indicate the occurrence of multiple copies (complete or partial) of an element.

Minor comments and suggestions

2. Page 2 (Abstract), line 7. I would remove “or not” in “...subjected or not...”

3. Page 2 line 9. I would remove “exemplar” in “...onto an exemplar...”

4. Page 5, line 20. I would recommend substituting “Next Generation Sequencing” with “Massively parallel sequencing”. Personally, I think that it is a better word for this technology.

5. Page 7, line 14. It is unclear why you use two different approaches for mapping reads to your reference sequences. You state that you use CLC and BWA, but you don't give reasons for this. A sentence explaining why would be appropriate here.

6. Page 8, line 26, change hybrid(s) to hybrids.

7. Page 9, line 48. How come you have exactly the same numbers of reads in both samples. Did you subsample the reads or how this was achieved?

8. Page 9, line 52. The line: “All *P. trichocarpa* LTR-RE were mapped”. What do you mean? Do you mean the reads mapped to all of the reference LTR contigs from *P. trichocarpa*, i.e. that all contigs had reads that mapped to them from the RNA or DNA reads? Please clarify.

9. Page 10, line 12. Change “...Each condition (control and moderate...” to “...Each condition (control, moderate...”

10. Page 10, line 38. Change “...reads was low: for example...” to “...reads was low. For example...”

11. Page 13, last word of line 5. Change level to levels.

12. Page 13, line 20. Change *Arabidopsis* to *Arabidopsis* (italicize).

13. Page 14, line 11. Sentence starting with: "Generally, the expression...". The meaning of this sentence is unclear. Do you mean that it has been debated in the literature whether LTR-REs are expressed? Rewrite to clarify.

14. Page 4, line 59. Italicise Arabidopsis.

R.: All the requested changes were made. Where requested, explanations were added to the text. All related changes in the manuscript are in red.