

Contrasting evolutionary trajectories of multiple retrotransposons following independent allopolyploidy in wild wheats

Natacha Senerchia¹, François Felber^{1,2} and Christian Parisod¹

¹Laboratory of Evolutionary Botany, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland; ²Musée et Jardins botaniques cantonaux, Avenue de Cour 14bis, 1007 Lausanne, Switzerland

Author for correspondence:

Christian Parisod

Tel: +41 032 718 2344

Email: christian.parisod@unine.ch

Received: 3 December 2013

Accepted: 9 January 2014

New Phytologist (2014)

doi: 10.1111/nph.12731

Key words: *Aegilops*, genome evolution, genome shock, pivotal-differential genome evolution, sequence-specific amplified polymorphism (SSAP), transposable elements (TEs).

Summary

- Transposable elements (TEs) are expectedly central to genome evolution. To assess the impact of TEs in driving genome turnover, we used allopolyploid genomes, showing considerable deviation from the predicted additivity of their diploid progenitors and thus having undergone major restructuring.
- Genome survey sequencing was used to select 17 putatively active families of long terminal repeat retrotransposons. Genome-wide TE insertions were genotyped with sequence-specific amplified polymorphism (SSAP) in diploid progenitors and their derived polyploids, and compared with changes in random sequences to assess restructuring of four independent *Aegilops* allotetraploid genomes.
- Generally, TEs with different evolutionary trajectories from those of random sequences were identified. Thus, TEs presented family-specific and species-specific dynamics following polyploidy, as illustrated by *Sabine* showing proliferation in particular polyploids, but massive elimination in others. Contrasting with that, only a few families (*BARE1* and *Romani*) showed proliferation in all polyploids. Overall, TE divergence between progenitors was strongly correlated with the degree of restructuring in polyploid TE fractions.
- TE families present evolutionary trajectories that are decoupled from genome-wide changes after allopolyploidy and have a pervasive impact on their restructuring.

Introduction

The last decades have highlighted the astonishing dynamics of plant genomes over evolutionary time, mainly in relation to prominent mechanisms such as retrotransposition, recombination and polyploidy (Kejnovsky *et al.*, 2009; Murat *et al.*, 2012). Polyploidy (i.e. hybridization between more or fewer divergent genomes, associated with duplication) is a recurrent process across angiosperms (Jiao *et al.*, 2011). Nascent and established polyploids show significant deviation from the expected additivity of their progenitor genomes, indicating major restructuring and epigenetic changes after the origin of polyploid lineages (Soltis & Soltis, 1999; Doyle *et al.*, 2008). Molecular mechanisms and evolutionary processes underlying such genome reorganization are not fully clear yet (Tayalé & Parisod, 2013), but the major fraction of plant genomes represented by transposable elements (TEs) seems to be specifically affected by polyploidy (McClintock, 1984; Fedoroff, 2012). ‘Genomic stress’ such as polyploidy may indeed activate TEs, and fast-evolving polyploids thus represent promising models to investigate the role of TEs in driving genome evolution (Comai *et al.*, 2003; Parisod & Senerchia, 2012; Levy, 2013).

Long terminal repeat (LTR) retrotransposons represent the predominant TEs in plants (Kumar & Bennetzen, 1999). Active LTR retrotransposons indeed increase their copy number through a copy-paste mode of transposition, affecting genome size and overall organization (Lynch & Conery, 2003; Ma & Bennetzen, 2004; Wang & Dooner, 2006). However, sequence elimination was highlighted as a major process in TE genome fractions (Vitte & Panaud, 2005), especially following polyploidy (Parisod *et al.*, 2010; Yaakov & Kashkush, 2011; Parisod & Senerchia, 2012). A balance between proliferation and partial elimination by ectopic and illegitimate recombination may thus lead to a high genome turnover (Bennetzen & Kellogg, 1997; SanMiguel *et al.*, 1998; Vitte & Panaud, 2005; El Baidouri & Panaud, 2013). Given that mutations accumulate among TE copies during and after transposition, the different TE lineages form large populations of repetitive sequences that evolve within genomes (Wicker *et al.*, 2007; Jurka *et al.*, 2011; Wicker, 2013). TE families (i.e. TE copies sharing at least 80% nucleotide similarity) may thus show distinct evolutionary dynamics (Senerchia *et al.*, 2013). However, the processes underlying the particular trajectories of TEs remain unclear and may be influenced by intrinsic TE

properties as well as processes acting at the host level (Tenailon *et al.*, 2010; Esnault *et al.*, 2011).

The wheat group belongs to the *Triticeae* clade of the grass family and presents a great diversity of species that evolved through homoploid divergence, hybridization and allopolyploidy. In particular, wild relatives of cultivated wheat (i.e. *Aegilops* species) independently evolved multiple di-, tetra- and hexaploid species that mostly co-occur in the Middle East (van Slageren, 1994). The reticulate evolutionary relationships among *Aegilops* taxa have been thoroughly inferred by genetic and cytogenetic analyses, although they remain poorly dated so far (reviewed in Baum *et al.*, 2012). *Aegilops* allopolyploids are young entities that mostly formed by combinations of diploid species with D or U genomes (Zohary & Feldman, 1962). Within such polyploid clusters, early cytogenetic evidence showed that chromosomes shared among polyploids (i.e. pivotal genome) remain largely unaltered as compared with their diploid donor, whereas the other genome (i.e. differential genome) presented considerable alteration (Zohary & Feldman, 1962; Feldman, 1965a,b,c). Zohary & Feldman (1962) postulated that this pattern results from gene flow among established allopolyploids, with common pivotal genomes providing substrates for homologous recombination and buffering the initial loss of hybrid fertility, while the differential genomes were accumulating restructuring events through homeologous recombination.

The overall genome organization of hybridizing *Aegilops* supports this hypothesis to a certain extent (Zohary & Feldman, 1962; Wang *et al.*, 2000), but hardly explains the balanced restructuring of both genomes reported in some species (Kimber *et al.*, 1988; Badaeva *et al.*, 2002, 2004) or that natural hybrids between species with so-called pivotal genomes are noticeably rare. Challenging the pivotal-differential hypothesis, recent studies used various molecular markers to show that asymmetrical restructuring of polyploid genomes may depend on other factors such as sequence types and genome sizes (reviewed in Bento *et al.*, 2011). Additional investigations are thus required to understand the processes underlying the evolution of genomes within species complexes such as *Aegilops*.

Aegilops allopolyploids present fast-evolving genomes, of which TEs make up *c.* 80% (Li *et al.*, 2004; Sabot & Schulman, 2009; Wicker & Buell, 2009; Senerchia *et al.*, 2013), thus offering promising models to assess the role of TEs in genome restructuring, and shedding light on the processes shaping genome architecture. Here, we focused on four diploid species, *Aegilops caudata* (genome C), *Aegilops comosa* (M), *Aegilops tauschii* (D) and *Aegilops umbellulata* (U) that recently combined into four natural allotetraploid species, *Aegilops crassa* (DM), *Aegilops cylindrica* (DC), *Aegilops geniculata* (UM) and *Aegilops triuncialis* (UC) (Fig. 1). Our aims were to investigate genome restructuring and assess the evolutionary trajectories of abundant LTR retrotransposon families after independent allopolyploidy events. Based on genome survey sequence (GSS) data from two *Aegilops* polyploid genomes, we designed complementary molecular fingerprint assays tracking restructuring events in random sequences as well as 17 LTR retrotransposon families. Comparisons between TEs and random sequences highlighted the proliferation of several TE

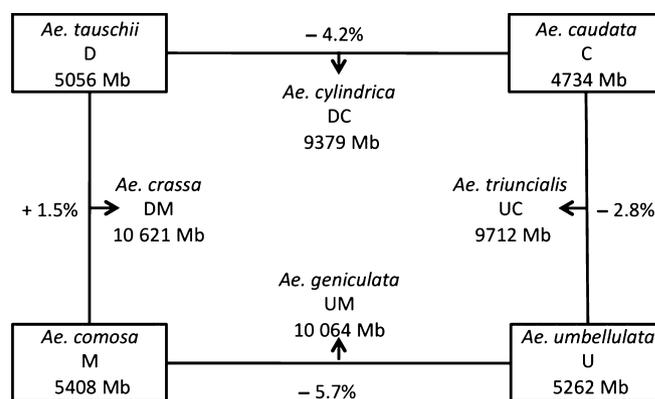


Fig. 1 Genome composition of investigated *Aegilops* diploid and derived allopolyploid species. Genome composition followed van Slageren (1994). Deviations of observed tetraploid genome size from expectations based on the addition of diploid progenitors, according to Eilam *et al.*, 2007, 2008 (using 1 pg of DNA = 978 Mb).

families and the predominant sequence deletion in others, indicating species-specific and TE-specific evolutionary trajectories following allopolyploidy. In particular, TE dynamics were associated with divergent insertions between progenitor species, suggesting that genome shocks at the origin of polyploid lineages had a long-lasting influence on genome organization.

Materials and Methods

Plant material

Three accessions of the polyploids *Ae. crassa* (PI 487286, PI 542178, TA 1880), *Ae. cylindrica* (TA 2204 = AE 719, PI 486235, PI 554221), *Ae. geniculata* (TA 1800, PI 487221, PI 287737), *Ae. triuncialis* (PI 487246, PI 542345, PI 491442) as well as of diploid progenitor species *Ae. caudata* (IG 47965, IG 48080, IG 107317), *Ae. comosa* (AE 1376, AE 1260, AE 1378), *Ae. tauschii* (IG 47219, IG 46847, AE 528), and two accessions of *Ae. umbellulata* (IG 48082, IG 46964) were obtained from the Institutes für Pflanzengenetik und Kulturpflanzenforschung (AE), the United States Department of Agriculture (PI), the International Center for Agricultural Research in the Dry Areas (IG) and the Wheat Genetic and Genomic Resources Center (TA). These accessions have been characterized as genetically variable in previous studies (Badaeva *et al.*, 2002, 2004; Meimberg *et al.*, 2009). They have been maintained in germplasms by selfing and were thus assumed to be inbred. Plants were grown under controlled conditions (18°C, 18 h light) and DNA of 2-wk-old seedlings was extracted from fresh leaves disrupted in liquid nitrogen, following the standard DNeasy Plant Extraction Mini Kit protocol from Qiagen.

Genome Sequence Survey and selection of TE families

Forty nanograms of genomic DNA of one individual from both *Ae. cylindrica* (TA 2204 = AE 719) and *Ae. geniculata* (TA 1800) were mechanically shotgunned and sequenced on half a plate of the Roche 454 GS FLX titanium platform (service provided by

Microsynth, Balgach, Switzerland, following manufacturer's instructions). BLASTN of nonredundant reads against complete TREP database (wheat.pw.usda.gov/ITMI/Repeats/) classified TEs following Senerchia *et al.* (2013). Portions of reads matching to 300 bp at the 5' end of the LTR regions of 27 TE families were aligned using ClustalW and analyzed through molecular population genetics to distinguish active from quiescent families following Senerchia *et al.* (2013). We selected 16 LTR retrotransposon families showing evidence of recent transpositional activity (*BARE1*, *Barbara*, *Cereba*, *Claudia*, *Daniela*, *Danae*, *Derami*, *Fatima*, *Lila*, *Maximus*, *Nusif*, *Romani*, *Quinta*, *Sabine*, *WHAM*, and *Xalax*) as well as one quiescent TE family (*Egug*, selected as a negative control) and evaluated nucleotide diversity (π) along their LTRs within a sliding window of 5 bp using DnaSP v5 (Librado & Rozas, 2009). For each TE family, a sequence-specific amplified polymorphism (SSAP) primer was designed across the most conserved region of the alignments using Primer 3 (Rozen & Skaletsky, 2000).

Molecular fingerprint techniques

Amplified fragment length polymorphism (AFLP) and SSAP were carried out following the protocol of Parisod & Christin (2008). Briefly, digestion of genomic DNA with *EcoRI* and *MseI* (New England Biolabs Inc.) and then ligation of double-stranded adaptors with T4 ligase (Promega) were performed at 37°C in plates with randomly positioned individual DNA. After inactivation of enzymes at 65°C for 20 min, primers corresponding to adaptors were used for preselective amplification with GoTaq[®] DNA polymerase (Promega). Twenty-fold diluted PCR products were selectively amplified with a touchdown PCR using primers with three additional nucleotides. For AFLP, unlabeled *MseI* primers were used together with fluorescently labeled *EcoRI* primers, whereas SSAP used fluorescently labeled TE family-specific primers (Supporting Information, Table S1) instead of *EcoRI* primers. SSAP thus mostly amplifies the region encompassing the termini of retrotransposon insertions and their flanking genomic sequences up to a restriction site cleaved by *MseI*. To minimize experimental inconsistencies, digestion, ligation and preselective amplification were performed simultaneously and PCRs were carried out on thermocyclers with fixed ramp rate. Genotyping error rates (E_b) were estimated based on the replication of the whole procedure on five samples (i.e. 22% of the dataset) as $E_b = M_{\text{repl}}/n_{\text{bin}}$, where M_{repl} is the total number of mismatches between replicated samples and n_{bin} is the total number of scored bins (Bonin *et al.*, 2004). Of the 64 different *EcoRI*–*MseI* combinations of AFLP selective primers tested, and the eight *MseI* selective primer combinations tested for each TE family, only the most reproducible combinations were retained and fully analyzed (17 AFLP combinations and a mean of 4.5 SSAP combinations per TE family).

Selective PCR products amplified with FAM[™], VIC[®], NED[™] fluorescent dye were pooled with GeneScan[™] 500 LIZ[™] Size Standard and separated with an ABI 3500 capillary sequencer. Resulting electropherograms were visualized and scored with

GeneMapper 3.5 (Applied Biosystems) using AFLP default peak detection parameters. The scoring was manually checked and loci were recorded as present (1) or absent (0) in binary matrices.

AFLP and SSAP loci analysis

For each species, Nei's gene diversity was estimated, with standard error, for AFLP as well as SSAP for each TE family using the Bayesian method with nonuniform prior distribution of allele frequencies (assuming $FIS = 1$) as implemented in AFLPsurv (Vekemans, 2002).

For each locus j , frequencies of band presence (1) and absence (0) were estimated among accessions of the diploid progenitors ($D_1(1)_j$, $D_1(0)_j$ and $D_2(1)_j$, $D_2(0)_j$, respectively) as well as accessions of the derived allopolyploid ($A(1)_j$, $A(0)_j$). The expected polyploid profile (i.e. the additivity of the progenitors) was estimated as probabilities of band presence and absence based on the frequencies in diploid progenitors. The probability of band presence in the polyploid was calculated as $E(1)_j = D_1(1)_j \times D_2(1)_j + D_1(1)_j \times D_2(0)_j + D_1(0)_j \times D_2(1)_j$, whereas the probability of band absence was calculated as $E(0)_j = D_1(0)_j \times D_2(0)_j$. The observed polyploid profile was then compared with the expected polyploid profile assessing the probability that the locus j was: additive (i.e. identity between expected and observed polyploid profiles) as $E(1)_j \times A(1)_j + E(0)_j \times A(0)_j$; a new band (i.e. presence of a band in the observed polyploid profile despite absence in the expected profile) as $E(0)_j \times A(1)_j$; or a lost band (i.e. band

Table 1 Summary statistics of investigated long terminal repeat (LTR) retrotransposon families based on the sequencing of 300 bp at the 5' end of the LTR region in *Aegilops cylindrica* (CY) and *Aegilops geniculata* (GE)

TE family	Reads CY ^a	π CY ^b	Reads GE ^a	π GE ^b	Activity ^c
<i>Barbara</i>	11082 (1.66)	0.08	11677 (1.81)	0.06	+
<i>BARE1</i>	79188 (11.60)	0.06	89573 (13.41)	0.03	+
<i>Cereba</i>	5212 (0.78)	0.05	8263 (1.28)	0.06	+
<i>Claudia</i>	3841 (0.57)	0.08	4803 (0.74)	0.08	+
<i>Danae</i>	8129 (1.22)	0.19	9903 (1.53)	0.16	+
<i>Daniela</i>	12022 (1.80)	0.04	12587 (1.95)	0.1	+
<i>Derami</i>	5782 (0.87)	0.09	4319 (0.67)	0.13	+
<i>Egug</i>	5368 (0.81)	0.09	5538 (0.86)	0.1	–
<i>Fatima</i>	38357 (5.41)	0.04	38567 (5.97)	0.05	+
<i>Lila</i>	6779 (1.02)	0.14	4018 (0.62)	0.18	+
<i>Maximus</i>	10118 (1.52)	0.04	12798 (1.98)	0.06	+
<i>Nusif</i>	10822 (1.62)	0.11	12200 (1.89)	0.1	+
<i>Quinta</i>	1567 (0.23)	0.06	1906 (0.29)	0.06	+
<i>Romani</i>	10101 (1.51)	0.09	10436 (1.61)	0.09	+
<i>Sabine</i>	7141 (1.07)	0.06	274 (0.04)	na	+
<i>WHAM</i>	15715 (2.35)	0.11	15162 (2.35)	0.1	+
<i>Xalax</i>	7090 (1.06)	0.12	9819 (1.52)	0.15	+

^aNumber of reads from GSS with proportion of reads out of the total of sequences shown in parentheses.

^bNucleotide diversity (π) among copies within genome. na, not applicable.

^cActivity = TE evolutionary dynamics (recently active, +; quiescent, –), inferred from molecular population genetics (details in Senerchia *et al.*, 2013).

Table 2 Distribution of loci within and among *Aegilops* species

	<i>Ae. caudata</i> (C)	<i>Ae. comosa</i> (M)	<i>Ae. tauschii</i> (D)	<i>Ae. umbellulata</i> (U)	<i>Ae. crassa</i> (genome DM)				
	Div	Div	Div	Div	Div	S	A	L	N
AFLP	0.209 ± 0.006	0.191 ± 0.006	0.191 ± 0.006	0.180 ± 0.006	0.238 ± 0.007	63.57	55.11	38.34	6.46
<i>Barbara</i>	0.226 ± 0.015	0.172 ± 0.015	0.203 ± 0.016	0.194 ± 0.015	0.189 ± 0.015	65.77	50.84	31.90	16.73
<i>BARE1</i>	0.227 ± 0.008	0.272 ± 0.008	0.251 ± 0.008	0.203 ± 0.009	0.259 ± 0.008	59.47	58.07	30.71	11.09
<i>Cereba</i>	0.180 ± 0.010	0.200 ± 0.010	0.201 ± 0.011	0.138 ± 0.008	0.164 ± 0.009	65.24	54.28	32.66	12.80
<i>Claudia</i>	0.178 ± 0.009	0.180 ± 0.009	0.170 ± 0.008	0.115 ± 0.007	0.164 ± 0.008	54.95	52.13	35.51	12.16
<i>Danae</i>	0.179 ± 0.009	0.151 ± 0.009	0.159 ± 0.009	0.139 ± 0.009	0.160 ± 0.009	49.39	48.72	40.61	10.41
<i>Daniela</i>	0.193 ± 0.009	0.236 ± 0.009	0.177 ± 0.008	0.158 ± 0.007	0.173 ± 0.008	54.65	57.15	31.69	11.02
<i>Derami</i>	0.198 ± 0.013	0.238 ± 0.013	0.132 ± 0.012	0.195 ± 0.013	0.204 ± 0.013	53.33	50.04	33.74	15.83
<i>Egug</i>	0.173 ± 0.009	0.193 ± 0.010	0.183 ± 0.009	0.190 ± 0.010	0.184 ± 0.010	67.33	65.38	28.29	6.04
<i>Fatima</i>	0.234 ± 0.009	0.192 ± 0.009	0.154 ± 0.008	0.147 ± 0.008	0.172 ± 0.008	56.32	46.78	40.14	12.91
<i>Lila</i>	0.195 ± 0.009	0.170 ± 0.009	0.181 ± 0.009	0.156 ± 0.009	0.186 ± 0.009	60.17	60.98	28.65	10.14
<i>Maximus</i>	0.174 ± 0.012	0.249 ± 0.013	0.187 ± 0.013	0.174 ± 0.012	0.220 ± 0.013	56.95	51.94	36.84	10.84
<i>Nusif</i>	0.159 ± 0.010	0.169 ± 0.011	0.163 ± 0.011	0.189 ± 0.011	0.166 ± 0.011	67.35	58.14	30.81	10.70
<i>Quinta</i>	0.186 ± 0.010	0.220 ± 0.011	0.225 ± 0.011	0.183 ± 0.100	0.178 ± 0.010	61.88	56.91	31.26	11.54
<i>Romani</i>	0.118 ± 0.008	0.165 ± 0.009	0.158 ± 0.009	0.104 ± 0.008	0.172 ± 0.009	60.42	42.99	35.30	21.51
<i>Sabine</i>	0.257 ± 0.013	0.212 ± 0.014	0.172 ± 0.013	0.190 ± 0.013	0.151 ± 0.012	67.35	62.18	30.96	6.34
<i>WHAM</i>	0.163 ± 0.014	0.183 ± 0.015	0.188 ± 0.015	0.134 ± 0.011	0.161 ± 0.014	63.11	56.00	36.66	6.74
<i>Xalax</i>	0.227 ± 0.012	0.161 ± 0.011	0.244 ± 0.013	0.179 ± 0.011	0.183 ± 0.012	63.34	58.36	33.36	7.93

Nei's gene diversity (Div) within diploid and allotetraploid species (genome composition following van Slageren, 1994) is presented (\pm SE) for random sequences (amplified fragment length polymorphism, AFLP) as well as the 17 transposable element (TE) families (sequence-specific amplified polymorphism, SSAP). Mean proportions of shared bands among diploid progenitors (S, i.e. non shared bands, S : 1-S) and comparisons with the expected additivity of the diploids (A, additive bands; L, lost bands; N, new bands) are presented for each allotetraploid species.

absence in the observed polyploid profile despite predicted presence in the expected profile) as $E(1)_j \times A(0)_j$. Probabilities of additive, new and lost bands were calculated for each locus, summed across loci and divided by the total number of bands to obtain proportions of additive, new and lost band for AFLP and for each SSAP of each TE family. Proportions were calculated for each polyploid accession and averaged within species. The same procedure was applied to the four allopolyploids.

For each tetraploid species, one-way ANOVA followed by multiple comparisons with Tukey's honest significant differences (HSD) *post hoc* were carried out to assess TE families with proportions of new bands and ratios of lost bands divided by new bands significantly different from AFLP at 5% levels, using the package 'agricolae' on R cran (cran.r-project.org).

Correlation of nonshared bands between progenitors and new bands in polyploids

For each pair of diploid progenitors, the proportion of nonshared bands among progenitor species, taking multiple accessions into account, was estimated for each locus as $D_1(0)_j \times D_2(1)_j + D_1(1)_j \times D_2(0)_j$. It was then summed over loci and divided by the total number of loci for AFLP as well as SSAP for each TE family.

Correlations of the proportions of nonshared bands between progenitors and the proportions of new bands and nonadditive bands in the derived polyploids were assessed with a mixed effect linear model using the maximum likelihood method with the package 'lme4' and including TE families nested within

polyploid species as random effects. R^2 was defined using a function correlating fitted values to observed values in R cran (Nakagawa & Schielzeth, 2013).

Origin of lost bands in polyploids

The origin of lost bands was assessed for each polyploid on loci presenting band presence in only one progenitor (i.e. either the pivotal or the differential genome in the polyploid). First, Kruskal–Wallis tests tested significant difference in proportions of lost bands among pivotal and differential genomes in each polyploid. Then, differences in proportions of lost bands between the pivotal and differential genome for each the 17 TE families were tested by a prop.test on R cran.

Results

Candidate TE families

Genome survey sequence of *Ae. cylindrica* (genome DC) and *Ae. geniculata* (genome UM) by 454 sequencing identified > 70% of the genome as TEs, including > 165 LTR retrotransposon families (Senerchia *et al.*, 2013). Comparison of 300 bp at the 5' end of LTR regions of the 17 selected TE families identified seven families (*BARE1*, *Danae*, *Daniela*, *Fatima*, *Lila*, *Romani*, *Xalax*) with clear evidence of recent activity, whereas nine (*Barbara*, *Cereba*, *Claudia*, *Derami*, *Maximus*, *Nusif*, *Quinta*, *Sabine*, *WHAM*) showed signs of both old and recent transpositional activity (Table 1). By contrast, *Egug* was considered as quiescent and selected as a negative control. Based on GSS data, specific

Table 2 (Continued)

<i>Ae. cylindrica</i> (genome DC)					<i>Ae. geniculata</i> (genome UM)					<i>Ae. triuncialis</i> (genome UC)				
Div	S	A	L	N	Div	S	A	L	N	Div	S	A	L	N
0.161 ± 0.005	59.21	67.89	27.06	5.26	0.214 ± 0.006	69.61	62.06	33.14	4.65	0.215 ± 0.006	59.43	67.61	25.82	6.42
0.145 ± 0.012	54.75	66.47	26.20	6.80	0.173 ± 0.014	57.11	55.34	34.83	9.27	0.210 ± 0.015	58.29	59.87	32.84	6.76
0.149 ± 0.006	64.50	69.76	22.05	8.04	0.238 ± 0.008	55.55	57.60	29.06	13.20	0.216 ± 0.008	57.06	64.37	24.05	11.43
0.102 ± 0.007	61.90	66.54	27.11	6.07	0.146 ± 0.008	49.83	59.06	28.64	12.06	0.151 ± 0.009	52.94	63.55	28.98	7.19
0.113 ± 0.006	55.03	68.71	24.09	6.98	0.147 ± 0.007	46.21	64.58	26.54	8.67	0.151 ± 0.008	36.56	66.16	26.52	7.09
0.129 ± 0.008	54.08	67.85	25.30	6.57	0.144 ± 0.008	46.19	59.36	27.17	13.20	0.130 ± 0.008	49.95	60.90	29.82	9.01
0.110 ± 0.006	55.97	67.80	23.21	8.73	0.209 ± 0.008	50.52	59.24	30.74	9.85	0.147 ± 0.007	48.44	65.39	27.88	6.55
0.154 ± 0.012	50.56	54.72	33.26	11.57	0.202 ± 0.012	55.57	65.58	28.07	5.95	0.262 ± 0.015	50.62	56.59	38.77	4.21
0.131 ± 0.007	73.30	76.15	19.91	3.64	0.171 ± 0.009	66.76	71.51	23.00	5.19	0.160 ± 0.009	78.71	77.28	18.63	3.78
0.108 ± 0.006	48.46	64.40	26.10	9.31	0.167 ± 0.008	48.61	57.54	31.80	10.48	0.160 ± 0.007	68.76	65.52	26.13	8.18
0.111 ± 0.006	59.94	73.99	20.37	5.40	0.145 ± 0.008	55.10	66.02	24.86	8.76	0.161 ± 0.008	57.55	70.31	21.71	7.74
0.174 ± 0.012	48.53	63.73	30.93	4.91	0.205 ± 0.013	51.96	56.15	33.33	10.01	0.189 ± 0.012	54.43	64.70	27.36	7.52
0.099 ± 0.007	60.32	70.51	24.32	4.80	0.154 ± 0.009	55.40	68.36	24.89	6.42	0.170 ± 0.010	62.89	72.85	21.94	4.87
0.130 ± 0.008	63.38	71.09	21.06	7.59	0.152 ± 0.008	50.39	64.04	27.91	7.80	0.181 ± 0.009	55.46	66.87	24.70	8.17
0.077 ± 0.006	55.28	63.89	24.97	10.87	0.125 ± 0.007	49.45	50.91	29.75	19.11	0.131 ± 0.008	47.07	59.43	26.10	14.18
0.142 ± 0.009	58.34	78.91	13.87	6.76	0.066 ± 0.005	59.47	59.27	37.51	2.69	0.187 ± 0.012	59.52	73.02	20.48	5.89
0.107 ± 0.008	60.75	76.85	16.65	5.56	0.173 ± 0.014	54.98	68.69	23.72	6.95	0.151 ± 0.013	64.54	70.51	23.64	5.18
0.108 ± 0.006	57.30	74.28	18.18	7.19	0.160 ± 0.010	60.60	66.99	23.49	9.15	0.203 ± 0.012	55.90	63.67	32.40	3.57

primers for each of these 17 candidate TE families were designed in conserved regions at the 5' end of the LTR (Table S1) and used to perform SSAP assays representative of the diversity of TE copies within *Aegilops* genomes.

Genetic variation in *Aegilops* and genome restructuring in polyploids

Variable genetic diversity was assessed within *Aegilops* species, showing specific variation in AFLP as well as the different TE families (Table 2). Genetic diversity within diploid and polyploid species ranged from 0.104 for the TE *Romani* in *Ae. umbellulata* to 0.272 for *BARE1* in *Ae. comosa*, and from 0.099 for *Cereba* in *Ae. cylindrica* to 0.262 for *Derami* in *Ae. triuncialis*, respectively. Several TEs such as *BARE1* showed higher genetic diversity than genome-wide AFLPs, but no clear pattern differentiating diploid and tetraploid species was evident. Noticeably, *Sabine* showed a significantly higher diversity in *Ae. cylindrica* than in *Ae. geniculata*. Genetic divergence among species for AFLP as well as the different TE families showed considerable variation depending on the genome fraction considered (Notes S1). In particular, multiple Mantel tests correlating genetic distances between species as measured by AFLP and by SSAP of the different TEs were nonsignificant (after stringent Bonferroni correction for multiple testing), indicating TE-specific patterns of divergence among species.

Genome restructuring in four *Aegilops* allotetraploids was assessed by comparing the fingerprint profiles of multiple accessions for random sequences (i.e. using AFLP) and for the 17 candidate TE families (i.e. using SSAP) with the expected additivity of their respective diploid progenitor species. Comparison of AFLP with SSAP allows contrasting background genome restructuring and specific reorganization of TE fraction. Genotyping of

1224 AFLP loci (with an error rate of 3.36%) and 238–792 SSAP loci for each of the 17 TE families (with an error rate ranging from 1 to 6.6%; Tables S2, S3) from three accessions per species took genetic variation into account.

Combinational probabilities based on the frequencies of band presence and absence in the diploid progenitors were used to determine the expected polyploid profile at each locus and were then compared with the observed profile for each polyploid accession, assessing the probability that the locus was additive or lost, or that a new locus appeared in the polyploid (Table 2). Averaged over loci, the mean proportion for additive AFLP bands across polyploids was 63.1%, close to the overall proportion of SSAP bands for the 17 TE families (62.9%). The proportion of additive AFLP bands, however, varied considerably among the four allopolyploid species, indicating a species-specific degree of background restructuring following the different polyploidy events. In particular, *Ae. crassa* showed a higher degree of restructuring (i.e. nonadditive AFLP bands, 44.9%) than *Ae. cylindrica* (32.1%), *Ae. geniculata* (37.9%) and *Ae. triuncialis* (32.9%). Proportions of additive SSAP bands showed considerable variation among TE families, ranging from 43.0% for *Romani* in *Ae. crassa* to 78.9% for *Sabine* in *Ae. cylindrica*.

New bands in random sequences and in multiple TE families

Comparing degrees of restructuring from a large number of SSAP loci for each TE family with the degree of restructuring in random sequences (AFLP) allowed us to highlight TEs with specific evolutionary trajectories, while accounting for possible ambiguities of fingerprint methods (Petit *et al.*, 2010; Sarilar *et al.*, 2013). Tukey HSD tests assessed that 15 out of 17 TE families showed significantly higher proportions of new bands than

AFLP, indicating proliferation in at least one of the tetraploid species (Fig. 2a). The number of TEs showing higher proportions of new bands than AFLP varied among the tetraploids (Table S3). Nine TE families in *Ae. crassa*, 10 in *Ae. cylindrica* and 13 in *Ae. geniculata*, but only two TEs in *Ae. triuncialis*, showed such a pattern indicative of TE proliferation. However, few TEs showed consistent patterns of new bands in all polyploids. Only *BARE1* and *Romani* presented significantly higher proportions of new bands than AFLP in the four tetraploids, whereas *Egug* and *Nusifif* never exhibited significant differences. Most TE families instead showed species-specific patterns, with some TE families showing significantly more new bands in the three tetraploids (*Barbara*, *Claudia*, *Daniela*, *Fatima* and *Quinta* in *Ae. crassa*, *Ae. cylindrica* and *Ae. geniculata*); in two tetraploids (*Cereba* in *Ae. crassa* and *Ae. geniculata*; *Derami* in *Ae. crassa* and *Ae. cylindrica*; *Xalax* in

Ae. cylindrica and *Ae. geniculata*) or in only one tetraploid (*Danae*, *Lila*, *Maximus* and *WHAM* in *Ae. geniculata* or *Sabine* in *Ae. cylindrica*).

As expected, TE families with clear evidence of recent activity as assessed by 454 GSS (i.e. *Barbara*, *BARE1*, *Cereba*, *Claudia*, *Daniela*, *Danae*, *Derami*, *Fatima*, *Lila*, *Maximus*, *Quinta*, *Romani*, *Sabine*, *Wham* and *Xalax*) showed higher proportions of new SSAP bands indicative of transposition after polyploidy in *Ae. cylindrica* and/or *Ae. geniculata* (see later). By contrast, molecular fingerprints detected nonsignificantly higher proportions of new SSAP bands than AFLP for *Nusifif* and *Egug*. *Egug*, but also *Nusifif* to a certain extent, indeed showed quiescence based on GSS data (also see Senerchia *et al.*, 2013). Both high-throughput sequencing and fingerprinting approaches showed congruent results, but their respective advantages deserve further attention.

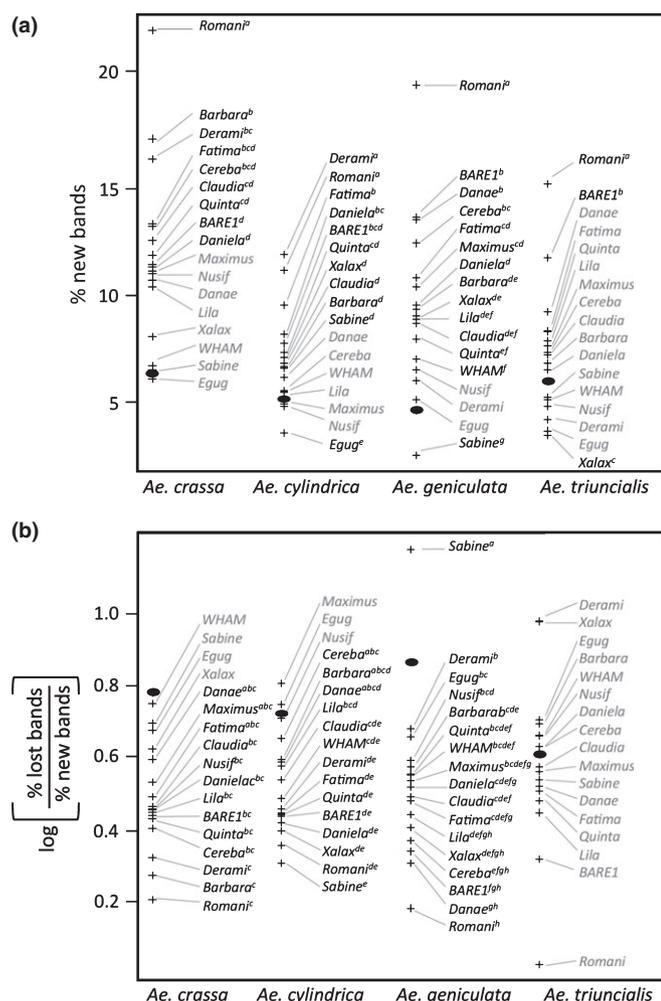


Fig. 2 Deviation from the expected additivity of progenitors in *Aegilops* allopolyploid species. Proportions of deviating bands for the 17 transposable element (TE) families investigated with sequence-specific amplified polymorphism (SSAP) and random sequences (amplified fragment length polymorphism, AFLP). (a) Proportions of new bands; (b) ratio between proportions of lost and new bands (log-scaled). TE families (crosses) presented in gray show nonsignificantly different proportions from the AFLP (black ellipse), as assessed by Tukey tests. TE families differing significantly from the AFLP (in black) and not sharing small letters show significantly different proportions.

Balance of lost/new bands in random sequences and in multiple TE families

The proportion of lost bands divided by the proportion of new bands offered an integrative proxy of the evolutionary trajectories of TE families towards either sequence deletion (> 1) or proliferation (< 1) when compared with AFLP. Ratios for AFLP as well as TE family were usually larger than one, indicating that sequence loss was the major restructuring process following polyploidy (Fig. 2b). Tukey HSD tests assessed that all 17 TE families had significantly different ratios from the AFLP ratio in at least one of the four polyploids (Table S3). The number of TEs with significantly different ratios of lost/new band proportions from the AFLP ratio varied considerably among polyploids. None of the TE families displayed significantly different ratios from the AFLP ratio in *Ae. triuncialis*, whereas 12 and 10 TE families offered evidence of proliferation with significantly lower ratios than the AFLP ratio in *Ae. crassa* and *Ae. cylindrica*, respectively. In *Ae. geniculata*, all TEs presented significantly different ratios from AFLP, with 16 families showing a lower ratio indicating proliferation, whereas *Sabine* presented a higher ratio indicating considerable sequence loss after polyploidy. TE families thus showed species-specific long-term evolutionary trajectories following independent polyploidy events.

Most TEs with significantly different ratios of proportions of lost/new bands compared with AFLP showed evidence of proliferation, but also revealed noticeable exceptions. In particular, *Sabine* showed considerable variation among tetraploid species, with clear evidence of proliferation in *Ae. cylindrica* (ratio of 2.05, significantly lower than AFLP), whereas considerable band loss was reported in *Ae. geniculata* (ratio of 13.94, significantly higher than AFLP). These results were firmly supported by random GSS, which identified *Sabine* as the TE presenting the largest difference in number of reads, with > 7000 out of 667 485 reads in *Ae. cylindrica* (i.e. 1.06% of the genome), whereas as few as 274 out of 646 327 reads (0.04%) were detected in *Ae. geniculata*. Semiquantitative PCR using specific primers designed in the 300 bp of the LTR of *Sabine* on the same DNA extracts further confirmed the contrasting abundance of TEs in the allopolyploid species (data not shown).

Correlation between progenitor divergence and restructuring after polyploidy

A linear mixed model nesting TE families within polyploid species as random effects assessed that the proportion of nonshared bands between pairs of diploid progenitors was significantly associated with the proportion of new bands in the derived polyploid (Fig. 3a; maximum likelihood linear mixed effect model; t -value = 4.18, $P < 0.01$, $R^2 = 0.72$). Thus, for each TE family, the more divergent the arrangements of TE insertions among diploid genomes being merged, the more new SSAP bands we saw that were indicative of transposition for that specific TE in the derived allotetraploid.

A similar model assessed that the proportion of nonshared bands between pairs of diploid progenitors was significantly associated with the proportions of nonadditive bands for each TE family (Fig. 3b, linear mixed effect model; t -value = 5.73, $P < 0.01$, $R^2 = 0.74$). Thus, the higher the divergence between the respective diploid progenitors, the higher the proportion of nonadditive bands (i.e. new bands and lost bands) in the derived polyploid.

Origin of lost bands following polyploidy

The proportion of lost bands originating from the pivotal vs the differential genome was assessed on a subset of loci that were specific to each progenitor species (Table S4). The proportions of lost bands were significantly higher in the differential than in the pivotal genome in *Ae. crassa* (Kruskal–Wallis test, $\chi^2 = 69.04$, $P < 0.001$) for all TEs except *Quinta* (Fig. 4). Similarly, *Ae. triuncialis* ($\chi^2 = 29.77$, $P < 0.001$) and *Ae. cylindrica* ($\chi^2 = 4.39$, $P = 0.036$) showed significantly higher band loss from the differential genome, but only specific TEs significantly followed this pattern (*Claudia*, *Danae*, *Fatima*, *Nusif* and *WHAM* for *Ae. triuncialis*, and *BARE1*, *Maximus* and *Nusif* for *Ae. cylindrica*). By contrast, *Ae. geniculata* showed significantly higher proportions of lost bands from the pivotal genome

($\chi^2 = 50.81$, $P < 0.001$), with significant differences for *Cereba*, *Maximus*, *Romani*, *Sabine* and *WHAM*.

Discussion

Restructuring of polyploid genomes

Combinations of different diploid genomes in four established *Aegilops* allotetraploids examined here with two complementary molecular fingerprint methods tracking either random (AFLP) or TE sequences (SSAP) revealed considerable departure from the expected additivity of their progenitors (Table 2). Despite a relatively low number of surveyed accessions per species, our quantitative handling of the presence/absence of hundreds of loci through a probabilistic approach sheds light on genome evolution in wild wheats, taking intraspecific variation into account. This approach offers significant novelties as compared with the traditional approach comparing single accessions through direct count and could be straightforwardly extended to large numbers of samples across the distribution ranges of species.

Genome restructuring highlighted here is congruent with GSS data (Senerchia *et al.*, 2013) and coherent with the otherwise reported instability of the *Triticum–Aegilops* genomes over evolutionary time (Badaeva *et al.*, 2002, 2004, 2007; Brenchley *et al.*, 2012; Yaakov *et al.*, 2013). It further highlights allopolyploidy as a major process eliciting drastic genome changes (McClintock, 1984; Comai, 2000; Liu & Wendel, 2003) as otherwise reported in cultivated wheat (Kashkush *et al.*, 2003; Kraitshtein *et al.*, 2010; Feldman & Levy, 2012).

Aegilops is a radiating genus, whose polyploids most likely originated during the Pleistocene (Baum *et al.*, 2012). Despite questionable dating of polyploidy events (Doyle & Egan, 2010), *Aegilops* allopolyploids were assumed to be of comparably recent origin. Accordingly, our results indicate variable degrees of genome restructuring as illustrated by *Ae. crassa* presenting a higher proportion of nonadditive bands than *Ae. cylindrica*, *Ae. geniculata* or *Ae. triuncialis*. Noticeably, *Ae. crassa* presents an

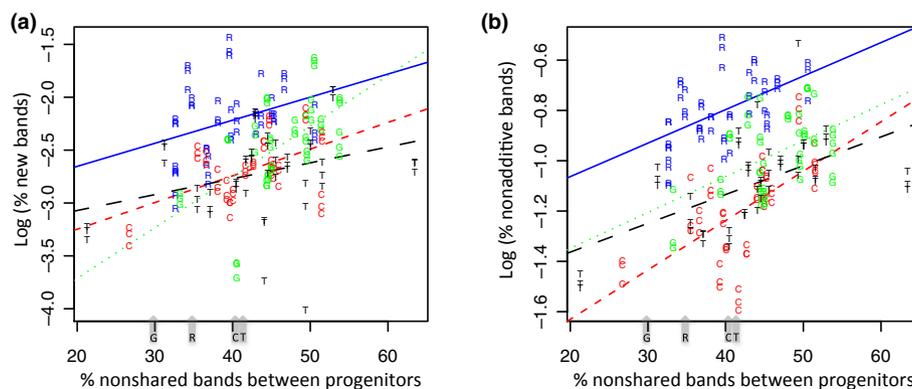


Fig. 3 Association of divergence between progenitor species and restructuring in derived *Aegilops* polyploids. General linear mixed model associating proportions of nonshared bands between the progenitors with the proportion of new bands (a) or the proportions of lost and new bands (b) in the allopolyploids for each transposable element (TE) family. Gray arrows indicate proportions of nonshared bands between the progenitor species for random sequences (amplified fragment length polymorphism, AFLP), whereas letters and lines show proportions and relationships, respectively, for the 17 TE families in *Aegilops crassa* (R, solid lines), *Aegilops cylindrica* (C, short dashed), *Aegilops geniculata* (G, dotted) and *Aegilops triuncialis* (T, long dashed).

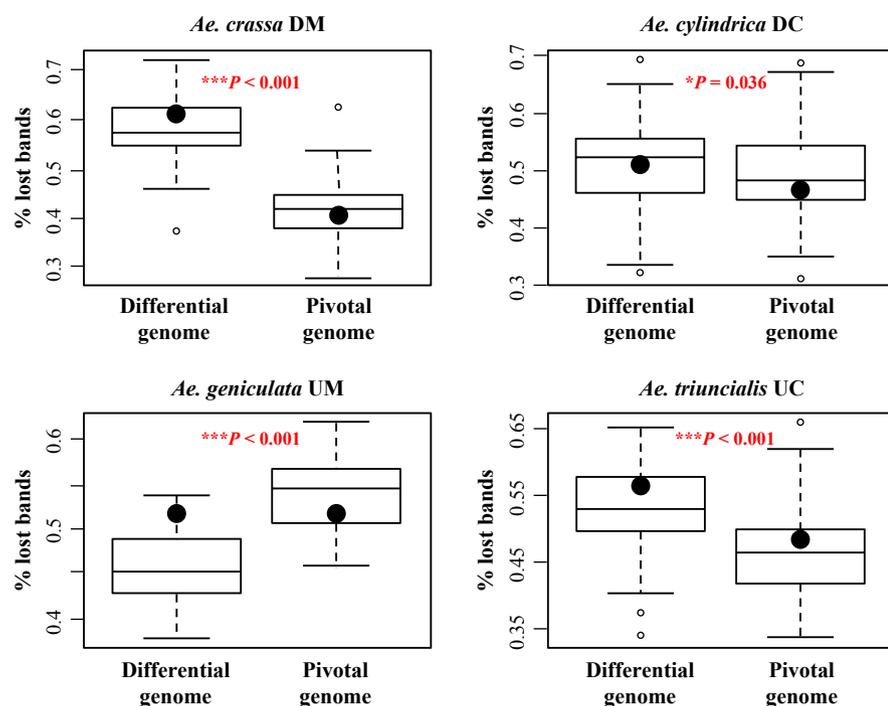


Fig. 4 Proportion of lost bands originating from the differential and the pivotal genome in the allotetraploid species *Aegilops crassa* (DM), *Aegilops cylindrica* (DC), *Aegilops geniculata* (UM) and *Aegilops triuncialis* (UC). Significant differences were assessed by Kruskal–Wallis tests. Medians and quartiles are shown as boxes, whereas whiskers show minimum and maximum values within 1.5 interquartile ranges. Open circles, outlier data; closed circles, proportions of lost random sequences (amplified fragment length polymorphism, AFLP).

overall distinctive evolutionary trajectory, as it is the only polyploid examined here showing genome upsizing (Eilam *et al.*, 2008). This species presents an atypically restricted distribution range for a polyploid *Aegilops* (Kilian *et al.*, 2011), but it remains unknown to what extent such genome dynamics is associated with processes acting at the level of natural populations (Bonchev & Parisod, 2013).

As expected under the pivotal-differential hypothesis proposed by Zohary & Feldman (1962) to explain chromosomal repatterning in polyploid species of the U-cluster, D and U genomes showed overall lower amounts of band loss than differential genomes (here, C and M) in all polyploids except *Ae. geniculata*. In agreement with this, detailed cytogenetic work has shown substantial modification of both parental genomes in *Ae. geniculata*, but also in *Ae. crassa* (Kimber *et al.*, 1988; Badaeva *et al.*, 2002, 2004). The present results only partially match expectations raised by the pivotal-differential hypothesis and further indicate that restructuring does not consistently affect the larger parental genome as suggested by Bento *et al.* (2011). Accordingly, processes other than long-term gene flow between species and intergenomic recombination also drive genome reorganization in established polyploids (Feldman *et al.*, 2012). In particular, the specific trajectories of repeated sequences reported here suggest that the intrinsic TE dynamics played a significant role in restructuring polyploid genomes.

TE evolutionary trajectories following allopolyploidy

This study tracks several LTR retrotransposon families at a genome-wide scale, thus offering a powerful comparative framework to contrast their dynamics, particularly when compared with random sequences. Large numbers of SSAP loci were indeed scored for the 17 TE candidate families and systematically

compared with AFLP proportions to minimize potential bias owing to: polymorphic loci segregating among populations of the diploid species; nonadditive bands appearing following molecular changes at the insertion site and modifying the size of the amplified product rather than being indicative of transposition events; and the sensitivity of the *EcoRI* restriction enzyme to rare cytosine methylation states (Cervera *et al.*, 2003). Nonadditive SSAP bands could indeed find their origin in more complex scenarios than transposition and deletions, as shown by traditional studies that cloned and sequenced polymorphic SSAP and that reported new/lost bands corresponding to transposition/deletion events only to a certain extent (Petit *et al.*, 2010; Sarilar *et al.*, 2013). In this report, we provide an original solution circumventing these limitations by relying on quantitative comparisons between SSAP (i.e. tracking TE insertions) and AFLP (i.e. tracking random sequences). As these techniques share very similar features, molecular events that are not TE-specific (e.g. chromosomal rearrangements, deletions or introgression) result in nonadditive bands in both SSAP and AFLP profiles, whereas TE-specific events are apparent in corresponding SSAP profiles only. Such a probabilistic approach does not designate specific bands, but highlights TEs showing significantly higher proportions of nonadditive SSAP bands than AFLP profiles as a result of a predominance of TE-specific events (e.g. reorganization in TE-rich genome regions such as heterochromatin; Le Rouzic *et al.*, 2007; Parisod *et al.*, 2009). In loosely compartmentalized genomes such as *Aegilops*, SSAP vs AFLP patterns can thus be chiefly interpreted as evolutionary trajectories of the corresponding TEs (e.g. proliferation).

This study assessed TE families with significantly higher proportions of new SSAP bands than new AFLP bands and offered convincing support for proliferative transposition events after allopolyploidy. Noticeably, few TE families showed consistent proliferation in all polyploids. Only *BARE1*, a TE known to be

active in *Triticeae* (Vicent *et al.*, 1999), and *Romani* always presented evidence of continual transpositional activity, whereas *Nusif* and *Egug* were seemingly quiescent. Other TE families presented evidence of proliferation after particular allopolyploidy events, but remained quiescent after others (Fig. 2).

High proportions of lost bands highlighted sequence deletion as a major restructuring process following polyploidy. In particular, the balance between lost and new bands confirmed proliferation of a majority of TE families as compared with random sequences, but indicated that deletion of TE sequences was usually predominant. Such evolutionary trajectories are in line with the general downsizing of polyploid genomes (Leitch & Bennett, 2004; Leitch *et al.*, 2008) as reported in *Ae. cylindrica*, *Ae. geniculata* or *Ae. triuncialis* (Ozkan *et al.*, 2003; Eilam *et al.*, 2007, 2008). Accordingly, the predominance of band loss reported here matches predictions raised by the increase/decrease model of genome size evolution as a balance between transposition and small deletions resulting from illegitimate recombination (Ma *et al.*, 2004; Parisod *et al.*, 2010; El Baidouri & Panaud, 2013). Consequently, TE genome fractions showed considerable turnover after independent polyploidy events, supporting genome divergence whose precise evolutionary causes and consequences have still not been fully investigated.

The dynamics of multiple LTR retrotransposons revealed both species-specific and TE-specific dynamics in response to allopolyploidy as a common trigger. The remarkably contrasting evolutionary trajectory of *Sabine* in different polyploids is particularly illustrative of the variable evolutionary trajectories of TEs after polyploidy. Several lines of evidence indeed revealed that this retrotransposon presented significant proliferation in *Ae. cylindrica*, but massive elimination in *Ae. geniculata*, whereas it was apparently quiescent in the other polyploids. The molecular mechanisms underlying such a deletion of *Sabine* sequences in *Ae. geniculata* remain elusive, as illegitimate recombination is unlikely to achieve such specific elimination of interspersed TE insertions (Devos *et al.*, 2002; El Baidouri & Panaud, 2013). The evolutionary processes driving contrasting evolutionary trajectories of TEs in different species deserve further attention.

Determinism of allopolyploid genome evolution

Despite species-specific and TE-specific evolutionary trajectories, TE dynamics following allopolyploidy seems nonrandom. Differences in arrangements of TE insertions between progenitors, rather than genome-wide differentiation, showed a significant association with restructuring of corresponding TEs in the polyploids. In other words, for a given TE family, the higher the divergence between the diploid genomes being merged, the higher the turnover of the corresponding TE fraction (i.e. new bands indicative of transposition and nonadditive bands indicative of genome changes) in established polyploids examined here.

Several nonexclusive processes may account for TE dynamics after allopolyploidy (reviewed in Parisod *et al.*, 2010). The present results highlighting TE divergence among progenitors as an influential factor in the TE dynamics in allopolyploids largely support the genome shock hypothesis. Allopolyploidy expectedly

weakens key epigenetic processes repressing TEs, leading to their activation (Bourc'his & Voinnet, 2010; Parisod & Senerchia, 2012). Consequently, the merging of specific parental genomes with divergent TEs at the origin of a particular allopolyploidy event could result in genomic conflicts that would drive nonrandom restructuring of TE families (Tayalé & Parisod, 2013), as reported here in all polyploids examined.

Conclusions

The evolutionary trajectories of TEs do not necessarily match the dynamics of genomic backgrounds (e.g. Le Rouzic *et al.*, 2007; Parisod *et al.*, 2012). TEs can indeed show intrinsic dynamics through proliferation in response to various triggers, such as genome shocks, but their dynamics are also influenced by extrinsic factors such as sequence deletion (Kejnovsky *et al.*, 2009; Murat *et al.*, 2012; El Baidouri & Panaud, 2013) or evolutionary processes acting at the host level (Tenailon *et al.*, 2010; Bonchev & Parisod, 2013).

Transposable elements generally followed predictions raised by the pivotal-differential hypothesis in *Ae. crassa*, suggesting that extrinsic factors have shaped the evolutionary trajectories of TEs to a large extent. However, most TEs were evenly removed from both parental genomes and showed TE-specific trajectories as compared with random sequences in the majority of polyploids. Furthermore, the present study revealed long-term restructuring of TE genome fractions that is remarkably coherent with expectations raised by intrinsic properties of TEs in response to the initial genome shock of allopolyploidy. Intrinsic and extrinsic processes may act in concert in driving the evolutionary trajectories of TEs (Ma & Gustafson, 2005), but their relative importance remains elusive. As polyploidy-induced genome shocks probably result in TE activation and genome restructuring during the first generations after the origin of new polyploid species (Ha *et al.*, 2009; Parisod & Senerchia, 2012), resynthesized polyploids could be compared with established ones to distinguish processes acting after initial genome shock and gradual changes occurring throughout the life span of host species. In that respect, polyploid speciation is a promising model to investigate the multiple factors controlling TE dynamics.

Acknowledgements

We thank R. A. Slobodeanu for statistical advice, as well as M.-A. Grandbastien, A. Tayalé and anonymous reviewers for valuable comments on the manuscript. This work was funded by the Swiss National Science Foundation through the National Centre of Competence in Research 'Plant Survival' and a grant (PZ00P3-131950 to C.P.).

References

- Badaeva ED, Amosova AV, Muravenko OV, Samatadze TE, Chikida NN, Zelenin AV, Friebe B, Gill BS. 2002. Genome differentiation in *Aegilops*. 3. Evolution of the D-genome cluster. *Plant Systematics and Evolution* **231**: 163–190.

- Badaeva ED, Amosova AV, Samatdze TE, Zoshchuk SA, Shostak NG, Chikida NN, Zelenin AV, Raupp WJ, Friebe B, Gill BS. 2004. Genome differentiation in *Aegilops*. 4. Evolution of the U-genome cluster. *Plant Systematics and Evolution* 246: 45–76.
- Badaeva ED, Dedkova OS, Gay G, Pukhalskiy VA, Zelenin AV, Bernard S, Bernard M. 2007. Chromosomal rearrangements in wheat: their types and distribution. *Genome* 50: 907–926.
- Baum BR, Edwards T, Mamuti M, Johnson DA. 2012. Phylogenetic relationships among the polyploid and diploid *Aegilops* species inferred from the nuclear 5S rDNA sequences (Poaceae: *Triticeae*). *Genome* 55: 177–193.
- Bennetzen JL, Kellogg EA. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell* 9: 1509–1514.
- Bento M, Gustafson JP, Viegas W, Silva M. 2011. Size matters in *Triticeae* polyploids: larger genomes have higher remodeling. *Genome* 54: 175–183.
- Bonchev G, Parisod C. 2013. Transposable elements and microevolutionary changes in natural populations. *Molecular Ecology Resources* 13: 765–775.
- Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13: 3261–3273.
- Bourc'his D, Voinnet O. 2010. A small-RNA perspective on gametogenesis, fertilization, and early zygotic development. *Science* 330: 617–622.
- Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D *et al.* 2012. Analysis of the breadwheat genome using whole-genome shotgun sequencing. *Nature* 491: 705–710.
- Cervera MT, Ruiz-Garcia L, Martinez-Zapater JM. 2003. Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics* 268: 832–833.
- Comai L. 2000. Genetic and epigenetic interactions in allopolyploid plants. *Plant Molecular Biology* 43: 387–399.
- Comai L, Madlung A, Josefsson C, Tyagi A. 2003. Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids? *Philosophical Transactions of the Royal Society B* 358: 1149–1155.
- Devos KM, Brown JKM, Bennetzen JL. 2002. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Research* 12: 1075–1079.
- Doyle J, Flagel L, Paterson A, Rapp R, Soltis D. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* 42: 443–461.
- Doyle JJ, Egan AN. 2010. Dating the origins of polyploidy events. *New Phytologist* 186: 73–85.
- Eilam T, Anikster Y, Millet E, Manisterski J, Feldman M. 2008. Nuclear DNA amount and genome downsizing in natural and synthetic allopolyploids of the genera *Aegilops* and *Triticum*. *Genome* 51: 616–627.
- Eilam T, Anikster Y, Millet E, Manisterski J, Sag-Assif O, Feldman M. 2007. Genome size and genome evolution in diploid *Triticeae* species. *Genome* 50: 1029–1037.
- El Baidouri M, Panaud O. 2013. Comparative genomic paleontology across plant kingdom reveals the dynamics of TE-driven genome evolution. *Genome Biology and Evolution* 5: 954–965.
- Esnault C, Palavesam A, Pilitt K, O'Brochta DA. 2011. Intrinsic characteristics of neighboring DNA modulate transposable element activity in *Drosophila melanogaster*. *Genetics* 187: 319–331.
- Fedoroff NV. 2012. Transposable elements, epigenetics, and genome evolution. *Science* 338: 758–767.
- Feldman M. 1965a. Further evidence for natural hybridization between tetraploid species of *Aegilops* Section *Pleionathera*. *Evolution* 19: 162–174.
- Feldman M. 1965b. Fertility of interspecific F1 hybrids and hybrid derivatives involving tetraploid species of *Aegilops* Section *Pleionathera*. *Evolution* 19: 556–562.
- Feldman M. 1965c. Chromosome pairing between differential genomes in hybrids of tetraploid *Aegilops* species. *Evolution* 19: 563–568.
- Feldman M, Levy AA. 2012. Genome evolution due to allopolyploidization in wheat. *Genetics* 192: 763–774.
- Feldman M, Levy AA, Fahima T, Korol A. 2012. Genomic asymmetry in allopolyploid plants: wheat as a model. *Journal of Experimental Botany* 63: 5045–5059.
- Ha M, Lu J, Tian L, Ramachandran V, Kasschau K. 2009. Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proceedings of the National Academy of Sciences, USA* 106: 17835–17840.
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS *et al.* 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–113.
- Jurka J, Bao W, Kojima K. 2011. Families of transposable elements, population structure and the origin of species. *Biology Direct* 6: 44.
- Kashkush K, Feldman M, Levy AA. 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nature Genetics* 33: 102–106.
- Kejnovsky E, Leitch IJ, Leitch AR. 2009. Contrasting evolutionary dynamics between angiosperm and mammalian genomes. *Trends in Ecology and Evolution* 24: 572–582.
- Kilian B, Mammen K, Millet E, Sharma R, Graner A, Salamini F, Hammer K, Özkan H. 2011. *Aegilops*. In: Kole C, ed. *Wild crop relatives: genomic and breeding resources, cereals*. Berlin, Heidelberg, Germany: Springer-Verlag, 1–17.
- Kimber G, Sallee PJ, Feiner MM. 1988. The interspecific and evolutionary relationships in *Triticum ovatum*. *Genome* 30: 218–221.
- Kraitshtein Z, Yaakov B, Khasdan V, Kashkush K. 2010. Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics* 186: 801–812.
- Kumar A, Bennetzen JL. 1999. Plant retrotransposons. *Annual Review of Genetics* 33: 479–532.
- Le Rouzic A, Dupas S, Capy P. 2007. Genome ecosystem and transposable elements species. *Gene* 390: 214–220.
- Leitch IJ, Bennett MD. 2004. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* 82: 651–663.
- Leitch IJ, Hanson L, Lim KY, Kovarik A, Chase MW, Clarkson JJ, Leitch AR. 2008. The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). *Annals of Botany* 101: 805–814.
- Levy AA. 2013. Transposons in plant speciation. In: Fedoroff NV, ed. *Plant transposons and genome dynamics in evolution*. Ames, IA, USA: John Wiley and Sons Ltd., 165–176.
- Li WL, Zhang P, Fellers JP, Friebe B, Gill BS. 2004. Sequence composition, organization, and evolution of the core *Triticeae* genome. *Plant Journal* 40: 500–511.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Liu B, Wendel JF. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Molecular Phylogenetics and Evolution* 29: 365–379.
- Lynch M, Conery JS. 2003. The origins of genome complexity. *Science* 302: 1401–1404.
- Ma JX, Bennetzen JL. 2004. Rapid recent growth and divergence of rice nuclear genomes. *Proceedings of the National Academy of Sciences, USA* 101: 12404–12410.
- Ma JX, Devos KM, Bennetzen JL. 2004. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Research* 14: 860–869.
- Ma XF, Gustafson JP. 2005. Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenetics and Genome Research* 109: 236–249.
- McClintock B. 1984. The significance of responses of the genome to challenge. *Science* 226: 792–801.
- Meimberg H, Rice K, Milan N, Njoku C, McKay J. 2009. Multiple origins promote the ecological amplitude of allopolyploid *Aegilops* (Poaceae). *American Journal of Botany* 96: 1262–1273.
- Murat F, Van de Peer Y, Salse J. 2012. Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. *Genome Biology and Evolution* 4: 917–928.
- Nakagawa S, Schielzeth H. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Method in Ecology and Evolution* 4: 133–142.

- Ozkan H, Tuna M, Arumuganathan K. 2003. Nonadditive changes in genome size during allopolyploidization in the wheat (*Aegilops-Triticum*) group. *Journal of Heredity* 94: 260–264.
- Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhou B, Grandbastien MA. 2009. Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*. *New Phytologist* 184: 1003–1015.
- Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhou B, Grandbastien M-A. 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist* 186: 37–45.
- Parisod C, Christin P-A. 2008. Genome-wide association to fine-scale ecological heterogeneity within a continuous population of *Biscutella laevigata* (Brassicaceae). *New Phytologist* 178: 436–447.
- Parisod C, Mhiri C, Lim KY, Clarkson JJ, Chase MW, Leitch AR, Grandbastien M-A. 2012. Differential dynamics of transposable elements during long-term diploidization of *Nicotiana* section Repandae (*Solanaceae*) allopolyploid genomes. *PLoS ONE* 7: e50352.
- Parisod C, Senerchia N. 2012. Responses of transposable elements to polyploidy. In: Grandbastien M-A, Casacuberta JM, eds. *Plant transposable elements*. Berlin, Heidelberg, Germany: Springer-Verlag, 147–148.
- Petit M, Guidat C, Daniel J, Denis E, Montoriol E, Bui QT, Lim KY, Kovarik A, Leitch AR, Grandbastien MA *et al.* 2010. Mobilization of retrotransposons in synthetic allotetraploid tobacco. *New Phytologist* 186: 135–147.
- Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* 132: 365–386.
- Sabot F, Schulman AH. 2009. Genomics of transposable elements in the *Triticeae*. In: Muehlbauer GJ, Feuillet C, eds. *Genetics and genomics of the Triticeae*. New York, NY, USA: Springer-Verlag, 387–405.
- SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. 1998. The paleontology of intergenic retrotransposons of maize. *Nature Genetics* 20: 43–45.
- Sarilar V, Palacios PM, Rousselet A, Ridet C, Falque M, Eber F, Chevre A-M, Joets J, Brabant P, Alix K. 2013. Allopolyploidy has a moderate impact on restructuring at three contrasting transposable element insertion sites in resynthesized *Brassica napus* allotetraploids. *New Phytologist* 198: 593–604.
- Senerchia N, Wicker T, Felber F, Parisod C. 2013. Evolutionary dynamics of retrotransposons assessed by next generation sequencing in wild relatives of wheat. *Genome Biology and Evolution* 5: 1010–1020.
- van Slageren MW. 1994. Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub & Spach) Eig (*Poaceae*). *Wageningen Agricultural University Papers* 97.
- Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology & Evolution* 14: 348–352.
- Tayalé A, Parisod C. 2013. Natural pathways to polyploidy in plants and consequences for genome reorganization. *Cytogenetic and Genome Research* 140: 79–96.
- Tenaillon M, Hollister J, Gaut B. 2010. A triptych of the evolution of plant transposable elements. *Trends in Plant Science* 15: 471–478.
- Vekemans X. 2002. *AFLP-SURV version 1.0*. Distributed by the author. Belgium: Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles.
- Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH. 1999. Retrotransposon *BARE-1* and its role in genome evolution in the genus *Hordeum*. *Plant Cell* 11: 1769–1784.
- Vitte C, Panaud O. 2005. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenetics and Genome Research* 110: 91–107.
- Wang JB, Wang C, Shi SH, Zhong Y. 2000. Evolution of parental ITS regions of nuclear rDNA in allopolyploid *Aegilops* (Poaceae) species. *Hereditas* 133: 1–7.
- Wang Q, Dooner HK. 2006. Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proceedings of the National Academy of Sciences, USA* 103: 17644–17649.
- Wicker T. 2013. So many repeats and so little time: how to classify transposable elements. In: Grandbastien M-A, Casacuberta JM, eds. *Plant transposable elements*. Berlin, Heidelberg, Germany: Springer-Verlag, 1–15.
- Wicker T, Buell CR. 2009. Gene and repetitive sequence annotation in the *Triticeae*. In: Muehlbauer GJ, Feuillet C, eds. *Genetics and genomics of the Triticeae*. New York, NY, USA: Springer-Verlag, 407–425.
- Wicker T, Sabot F, Van Hua A, Bennetzen J, Capy P, Chalhou B, Flavell A, Leroy P, Morgante M, Panaud O *et al.* 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* 8: 973–982.
- Yaakov B, Kashkush K. 2011. Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. *Genome* 54: 42–49.
- Yaakov B, Meyer K, Ben-David S, Kashkush K. 2013. Copy number variation of transposable elements in *Triticum-Aegilops* genus suggests evolutionary and revolutionary dynamics following allopolyploidization. *Plant Cell Reports* 32: 1615–1624.
- Zohary D, Feldman M. 1962. Hybridization between amphidiploids and evolution of polyploids in wheat (*Aegilops-Triticum*) group. *Evolution* 16: 44–61.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Sequences of adaptors, preselective and selective primers used in this study

Table S2 Mean number of loci in the various diploid and tetraploid *Aegilops* species

Table S3 Multiple comparisons of proportions of new bands, and ratio between the proportions of lost bands and new bands in the different *Aegilops* polyploids

Table S4 Proportions of lost bands originating from the differential vs the pivotal genome of each *Aegilops* tetraploid species

Notes S1 Structure of genome-wide loci and TE genome fractions among four diploid and allotetraploid *Aegilops* species.

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.