

Retrotransposon *BARE-1* and Its Role in Genome Evolution in the Genus *Hordeum*

Carlos M. Vicent,^{a,1} Annu Suoniemi,^{a,1,2} Kesara Anamthawat-Jónsson,^b Jaakko Tanskanen,^a Alex Beharav,^c Eviatar Nevo,^c and Alan H. Schulman^{a,3}

^a Institute of Biotechnology, Viikki Biocenter, University of Helsinki, P.O. Box 56, Viikinkaari 9, FIN-00014 Helsinki, Finland

^b Faculty of Sciences, University of Iceland, Grensásvegí 12, Reykjavík 108, Iceland

^c Institute of Evolution, University of Haifa, Mount Carmel, Haifa 31905, Israel

The replicative retrotransposon life cycle offers the potential for explosive increases in copy number and consequent inflation of genome size. The *BARE-1* retrotransposon family of barley is conserved, disperse, and transcriptionally active. To assess the role of *BARE-1* in genome evolution, we determined the copy number of its integrase, its reverse transcriptase, and its long terminal repeat (LTR) domains throughout the genus *Hordeum*. On average, *BARE-1* contributes 13.7×10^3 full-length copies, amounting to 2.9% of the genome. The number increases with genome size. Two LTRs are associated with each internal domain in intact retrotransposons, but surprisingly, *BARE-1* LTRs were considerably more prevalent than would be expected from the numbers of intact elements. The excess in LTRs increases as both genome size and *BARE-1* genomic fraction decrease. Intrachromosomal homologous recombination between LTRs could explain the excess, removing *BARE-1* elements and leaving behind solo LTRs, thereby reducing the complement of functional retrotransposons in the genome and providing at least a partial “return ticket from genomic obesity.”

INTRODUCTION

Retrotransposons are genomic elements exhibiting a structure and life cycle very similar to that of the retroviruses (Adams et al., 1987; Doolittle et al., 1989; Grandbastien, 1992). Both of the two main classes of retrotransposons containing long terminal repeats (LTRs), the *copi*a-like and the *romani*- or *gypsy*-like elements, appear to be ubiquitous in the vascular plants (Flavell et al., 1992; Voytas et al., 1992; Suoniemi et al., 1998b). Unlike DNA transposons, retrotransposons replicate by transcription of genomic copies followed by reverse transcription and ultimate integration of the cDNA copy back into the genome (Boeke and Chapman, 1991). The replication strategy of retrotransposons offers the potential for explosive increases in copy number, were each new copy to generate many transcripts that would be integrated in turn as cDNA. The process, if not directly lethal due to insertional mutagenesis, would lead to a concomitant genome size increase as the number of retrotransposons increased. In plants, such increases would be heritable if they occurred in cells giving rise to pollen and ovules.

Over a broad range of organisms, retrotransposon copy number appears to be correlated with genome size. The small genome of the yeast *Saccharomyces cerevisiae*, consisting of 13×10^6 bp (<http://www.mips.biochem.mpg.de/proj/yeast/tables/inventy.html>), contains 51 full-length retrotransposons (Kim et al., 1998). The genome size of plants may be considered as the DNA content (1C) of the unreplicated haploid (1n) set of chromosomes (1x). Thus, a somatic nucleus of a diploid is 2n with 2x chromosomes containing 2C of DNA; for a hexaploid, the somatic nucleus also contains 2C of DNA but is 2n and 6x. Arabidopsis has the smallest plant genome known, consisting of 10^8 bp (Goodman et al., 1995). Approximately 30 retrotransposon families have been identified in Arabidopsis to date, each generally with one to three members (D. Voytas, personal communication). The fairly compact genome of rice comprises 4.3×10^8 bp (Kurata et al., 1997) and contains some 10^3 retrotransposons, estimated with a highly conserved probe (Hirochika et al., 1992). The large (1C of 3.2×10^{10} bp) genome of onion (genome sizes, unless otherwise cited, are from <http://www.rbgekew.org.uk/cval/database1.html>) contains 1 to 2×10^5 *copi*a-like retroelements (Pearce et al., 1996b), whereas within the very large genomes of the genus *Lilium* (1C of $\sim 1.4 \times 10^{11}$), the *gypsy*-like *del* family of elements alone numbers 10^6 in some species.

¹ These authors contributed equally to this work.

² Current address: Valio Ltd., P.O. Box 30, FIN-00039 Valio, Finland.

³ To whom correspondence should be addressed. E-mail alan.schulman@helsinki.fi; fax 358-9-708-59570.

Within the cereals of the tribe Triticeae, rye (1C of 3.8×10^{10} bp) has $\sim 10^5$ *copia*-like elements comprising 3.5% of the genome, with similarly high numbers in wheat, oats, and barley (Pearce et al., 1997). In cultivated barley (1C of 4.8×10^9), *BARE-1* (for barley retroelement 1) forms a major and active family of *copia*-like retroelements, dispersed on all chromosomes (Suoniemi et al., 1996a). We have demonstrated that *BARE-1* is transcribed in somatic tissues (Suoniemi et al., 1996b) from one of two promoters within well-conserved LTRs (Suoniemi et al., 1997) and furthermore contains functionally conserved reverse transcriptase (RT) priming sites (Suoniemi et al., 1997) and coding domains for GAG (putative capsid protein), aspartic proteinase, integrase (IN; Suoniemi et al., 1998a), and RT. Very few other plant retrotransposons have been demonstrated to be transcriptionally active (Hirochika, 1993; Lucas et al., 1995; Royo et al., 1996; Vernhettes et al., 1997; Takeda et al., 1998).

The genus *Hordeum* is widely distributed in both hemispheres, containing some 33 species and 46 taxa (von Bothmer et al., 1995) divided into meiotic-pairing groups designated the H, I, X, and Y genomes (Jacobsen and von Bothmer, 1992). The 1C genome sizes within the genus have been found to vary considerably, from 2.7 to 4.5×10^9 bp for the diploid species and up to 8.9×10^9 bp for the tetraploids (Kankaanpää et al., 1996). In view of the prevalence, conservation, and activity of *BARE-1* in barley and given the size of the genus, examination of *BARE-1* copy number in *Hordeum* spp may shed light on the role of the retrotransposon in genome evolution.

RESULTS

Strategy for Determining Copy Number

Earlier examinations of retrotransposon copy number (Joseph et al., 1990; Vershinin et al., 1990) made clear that apparent copy number depends on hybridization stringency, as expected if the genome contains elements differing in their divergence from the cloned type copy. All probes used here were subcloned from the type element *BARE-1a* (GenBank accession number Z17327) isolated from barley cultivar Bomi, schematically represented in Figure 1A. Hence, the genetic distance from barley might generate artifactually low copy number estimates in other *Hordeum* spp investigated.

To address this issue, we determined copy number with several distinct *BARE-1a* probes. One probe comprised part of the LTR upstream of the transcriptional start (Figure 1C). This segment includes transcriptional regulatory regions (Suoniemi et al., 1996b) and is well conserved (>90% identity) within barley and between barley and wheat (Schmidt and Graner, 1994; Suoniemi et al., 1997). The other two probes, for *rt* and *in* (Figure 1C), corresponded to regions encoding enzymes critical to the retrotransposon life cycle

and would be expected to be well conserved. All pairwise comparisons between 10 aligned *BARE-1 rt* domains (Gribbon et al., 1999; B. Gribbon, J. Tanskanen, A.H. Schulman, and A. Flavell, unpublished data) display an $86.4 \pm 0.7\%$ (se) similarity. The *Wis-2* retrotransposon family of wheat displays 2.4 to 12.7×10^{-8} substitutions per nucleotide per year (Matsuoka and Tsunewaki, 1996), and cotton *rt* displays a rate of 0.16 to 1.5×10^{-8} . We have estimated a comparable rate of 1.42×10^{-8} between the *in* domains of *H. spontaneum* and *Wis-2-1a* (Suoniemi et al., 1998a).

Data were collected from slot blot hybridizations. The washing conditions (final wash of 20 min in $0.2 \times$ SSC at 65°C, where $1 \times$ SSC is 0.15 M NaCl and 0.015 M sodium citrate) are 10°C below the calculated $\sim 75^\circ\text{C}$ melting temperature of the probes (Meinkoth and Wahl, 1984), permitting a mismatch of ~ 5 to 15% (Hyman et al., 1973). Less stringent washes ($2 \times$ SSC at 65°C) 25°C below the melting temperature did not increase the hybridization signal with the *in* probe in tests for barley, *H. murinum*, *H. euclaston*, and *H. pusillum*, and so the more stringent conditions were used for copy number determinations.

For interspecies comparisons, variations in the amounts of DNA on the filters bound and accessible to the probes were controlled for by hybridization of labeled barley total DNA. The control stringency was 18°C below the 83°C melting temperature calculated for barley DNA. Taking the synonymous substitution rate as 2 to 7×10^{-9} within the genus *Hordeum* and other grasses (Zhou et al., 1995; Gaut et al., 1996), the wheat genus *Triticum* as the nearest outgroup for *Hordeum*, and a divergence time of 9×10^6 years between the two (Ogihara et al., 1991; Ikeda et al., 1992) yields an estimated maximum of 6.3% substitutions per position between species in these genera. By comparison, the internal transcribed spacer of barley rDNA has diverged 10.5% from that of H-genome species *H. californicum* (Hsiao et al., 1995). Hence, the washing conditions should not have distorted the results for *Hordeum* spp genetically distant from barley. A set of controls was made to correct for differences in loading and hybridization accessibility of the blotted samples. In these controls, λ phage DNA was added to the extracted genomic DNAs at a copy number similar to that of the *BARE-1* elements before individual sample preparation. These experiments yielded results similar to those obtained by using total barley DNA.

BARE-1 Elements Are Conserved in Length in *Hordeum*

Calculation of the fraction of the genome occupied by a retrotransposon family requires knowledge of copy number, genome size, and retrotransposon size. *BARE-1a* is 12.09 kb but contains a 3.14-kb insert in the 3' LTR. Thus, without the insert, the *BARE-1a* element is 8932 bp in length, and earlier studies have indicated that *BARE-1* elements in barley are of this length (Suoniemi et al., 1996a), compared with

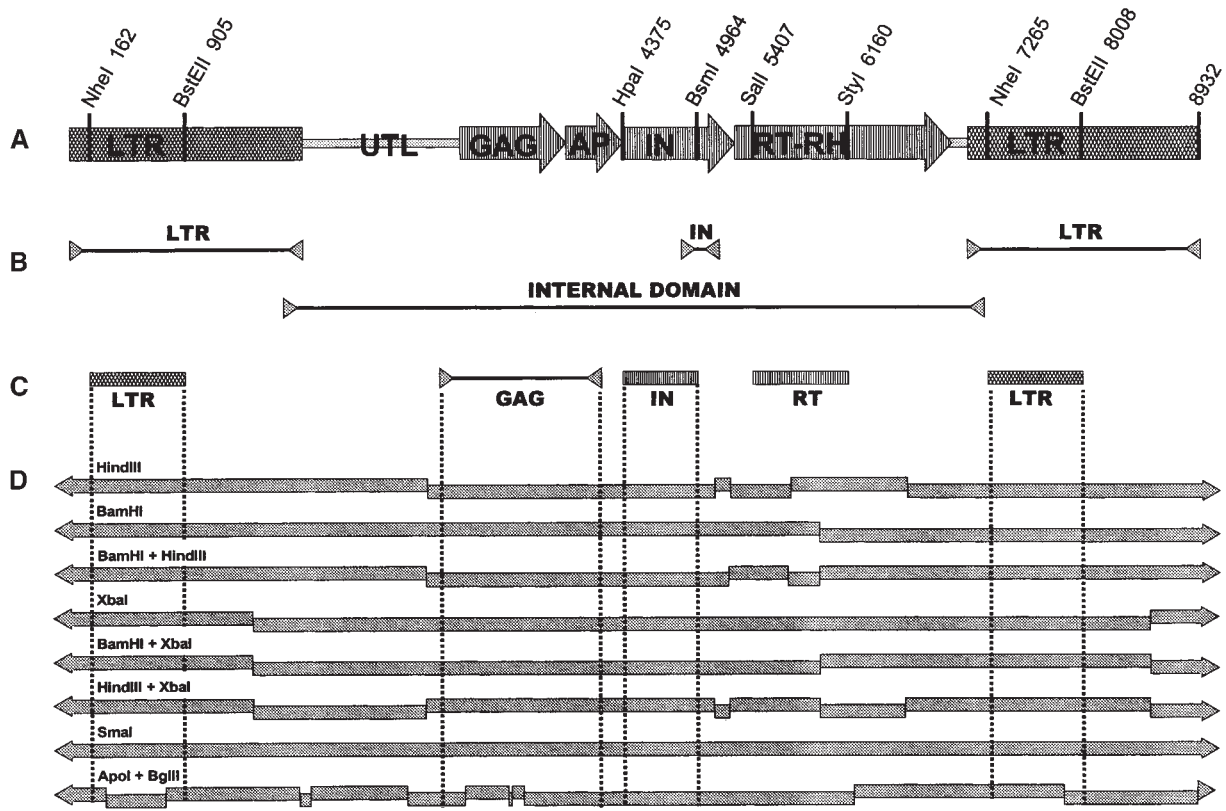


Figure 1. Organization of *BARE-1* and the Probes and Products Analyzed.

(A) Canonical *BARE-1* element. The LTRs are shown as hatched boxes and the segment of untranslated leader (UTL) outside the LTR as a gray bar. The protein-coding domains (arrows) include the putative capsid protein (GAG), aspartic proteinase (AP), integrase (IN), and reverse transcriptase–RNase H (RT–RH).

(B) The positions of the PCR products used in sequence and length divergence analyses. The inward-facing arrowheads diagrammatically represent the primers and are not drawn to scale.

(C) Probes used in filter and in situ hybridizations. The probes are positioned under the restriction sites shown on the map in (A); these sites were used to subclone the corresponding segments. The *gag* probe is indicated as in (B) because it was a PCR product.

(D) Restriction digests used in analysis of bacterial artificial chromosome (BAC) clones. The fragments predicted from *BARE-1a* are shown as shaded boxes; arrows indicate extension of a fragment into the flanking genomic region. Dotted lines indicate hybridization coverage of the probes in (C).

5334 bp for the *Tnt-1* element of tobacco (GenBank accession X13777). To examine the size of the *BARE-1* family in the genus *Hordeum*, we used primers to amplify the elements' LTRs and internal domains (Figure 1B) from genomic DNA. Figure 2A shows that throughout the genus, the *BARE-1* internal domain is conserved in length, the polymerase chain reactions (PCRs) yielding fragments of an estimated 5890 bp compared with the expected 5750 bp. The amplified LTR bands (Figure 2B) estimated at 1928 bp also match the predicted size of the *BARE-1a* LTRs (1809 bp), although one reaction (sample 26, *H. depressum*) also produced slightly shorter products. The conservation is significant in view of the unusually long untranslated leader (1.7 kb) and LTRs in *BARE-1a*.

The *in* and *rt* Regions as Estimators of *BARE-1* Copy Number

Estimates of *BARE-1* copy number per haploid genome equivalent obtained by using the LTR, *in*, and *rt* probes are presented in Table 1. The number of *in* copies per genome ranged from 4.3 to 19×10^3 , and *rt* copies ranged from 4.5 to 30×10^3 . Overall, the numbers of *in* and *rt* copies were very tightly correlated, the Pearson product moment correlation analysis yielding a coefficient (r_p) of 0.885 with a probability of error (P) of 2×10^{-10} . This would be expected because *in* and *rt* are adjacent components of the retrotransposons' genomes. Regarding the broad groups of species analyzed, the diploid I genome accessions, *H.*

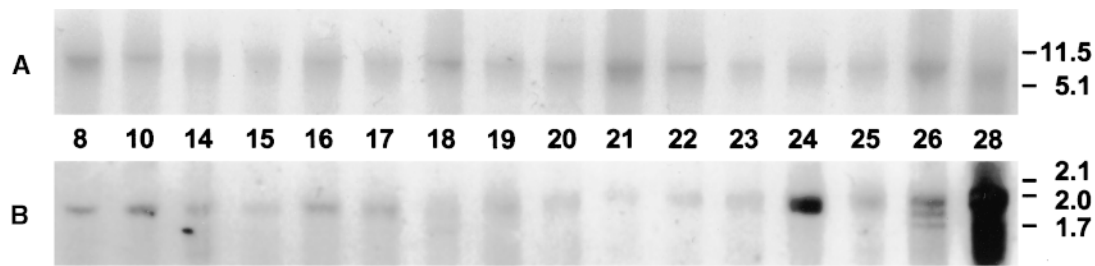


Figure 2. *BARE-1* Is Conserved in Length throughout the Genus *Hordeum*.

(A) PCR amplification of *BARE-1* internal domains.

(B) PCR amplification of LTRs.

Samples are numbered as in Table 1.

spontaneum and barley, showed 12 to 24×10^3 *rt* copies and 13 to 19×10^3 *in* copies per genome equivalent. By contrast, the basic genome (1x; 0.5n; 0.5C) of the autotetraploid *H. bulbosum* (II genome) contained only 3.1×10^3 (*rt*) to 5×10^3 (*in*) copies of *BARE-1*. The *H. murinum* (YY genome) tetraploid accessions contained 12 to 18×10^3 *BARE-1* copies, or 6 to 9×10^3 for the basic genome. The sole X genome representative, *H. marinum* subsp. *gussonianum*, had 8 to 10×10^3 copies of *BARE-1*. The greatest range of *in* and *rt* copy numbers was found in the American H genome accessions. In this set, 4.5 to 21×10^3 *rt* and 4.3 to 17×10^3 *in* copies per genome were detected. The two HH tetraploids, *H. depressum* and *H. jubatum*, contained 14 to 15×10^3 *rt* and 7 to 9×10^3 *in* copies in their basic genomes. For the genus as a whole, the *rt* probe detected $1.5 \pm 0.1 \times 10^4$ copies and the *in* probe detected $1.3 \pm 0.09 \times 10^4$ copies per genome.

We tested the reliability of the *in* and *rt* probes as estimators of *BARE-1* copy number by correlation analysis against genome size and genome type. The I genome species are most closely related to barley, and the species with Y, H, and X genomes are thought to follow in that order at increasing distance (Svitashev et al., 1994; Marillia and Scoles, 1996). A rapidly evolving *BARE-1* region might be expected to show an apparent copy number that decreases with increasing genetic distance from barley and therefore might be a bad estimator of true copy number. Applicability of this criterion would be limited if not only *BARE-1* copy number but also genome size were correlated with genetic distance from barley. Among the diploids, genome size was indeed negatively correlated with distance from barley ($r_p = -0.593$, $P = 0.003$), and the I and H genome accessions as a whole differed significantly in their genome sizes (Student's *t* test, $P = 0.009$). However, the correlation between genome type and *in* copy number was on the same order as for genome size ($r_p = -0.551$, $P = 0.002$) and not significant for *rt*. Furthermore, several H genome accessions showed

as many copies of *in* and *rt* as seen in barley. We interpret these data to show that the *in* and *rt* probes cannot be said to give artifactually low copy numbers for genomes of *Hordeum* spp that are distant from barley.

LTRs Are in Excess in Genomes of the Genus *Hordeum*

Each retrotransposon contains two LTRs, yielding an expected copy number ratio of 2:1 with respect to other regions such as *in* or *rt*. However, the measured number of LTRs (Table 1) was considerably higher than that of any other probe, ranging over 8.7 to $40.7 \times 10^4(C)^{-1}$. This indicates that genomes of *Hordeum* spp respectively contain 5.1 to 50 (average 14.3 ± 1.8) and 6.8 to 42 (average 16.5 ± 1.8) times more LTR than *rt* or *in* copies. These data are consistent with earlier observations. For barley cultivar Bomi, $9 \pm 0.6 \times 10^4$ copies is comparable to 5.9×10^4 previously estimated for the same cultivar (assuming a C of 4.53 pg) with a higher stringency wash ($0.1 \times \text{SSC}$, 65°C) (Manninen and Schulman, 1993) and with the 5.6 to 12.6×10^4 copies found, depending on the wash stringency, for a different cultivar (Vershinin et al., 1990).

Aside from the support of these earlier results, additional confirmation of the LTR excess was sought by several means. A set of dot blot hybridizations, in which the loading was controlled by using λ phage DNA, was performed, with nine individuals each from six stands of *H. spontaneum* growing in a single canyon at most 300 m from each other (Lower Nahal Oren, Mt. Carmel, Israel). Summing over the entire data set, an average of $15.2 \pm 0.4 \times 10^3$ (se) *in* copies were detected per genome equivalent, in the middle of the range seen in the broader slot blot data set. For the LTRs, the copy number was $81 \pm 2 \times 10^3$ (range of 37 to 118×10^3), and the LTR/*in* ratio was 3.8 to 7.6, in the range seen for several other barleys and *H. spontaneum* in the larger data set (Table 1). The LTR copy number in general shows

much greater variation in *H. spontaneum* than do *in* or *rt* numbers, which may have relevance for the mechanisms affecting their relative abundance (see Discussion).

Sequence Divergence Does Not Account for Apparent LTR Abundance

A trivial explanation for the differences in LTR abundance seen among *Hordeum* spp would be that LTRs are differentially more conserved than *in* or *rt* in species with high apparent LTR excess. To examine this, we first amplified a set of LTR segments by PCR under low primer annealing stringency from five species with varying degrees of LTR excess: barley cultivar Bomi (LTR/*in* = 6.8), *H. pusillum* (19.9), *H. euclaston* (22.4), *H. roshevitzii* (30.0), and *H. marinum* subsp *gussoneanum* (41.9). From each species, genomic DNA and

the pooled reaction product (a single band) from the LTR amplification were each digested with DpnII, blotted, and hybridized with the LTR probe used against the slot blots discussed above (data not shown). The similarity in digestion pattern and response between the genomic DNA and PCR products, digested with seven tetranucleotide-recognizing restriction enzymes, demonstrated that the PCR amplifications were representative of the LTRs present in the genome as a whole. The PCR products were cloned, and sequences of 19 clones from the five species were determined (GenBank accession numbers Y18767 to Y18785), aligned, and compared.

The average sequence identity among and between clones from these five species for the LTR region and an *in* region (GenBank accession numbers Z80000 to Z80079) previously used in an extensive molecular evolution study (Suoniemi et al., 1998a) is shown in Table 2. The overall level

Table 1. BARE-1 Copy Number and Genome Share in *Hordeum* spp

Sample	Species, Subspecies (Provenance)	Accession	1C (pg)	Copies (10 ³ /1C)			Genome Share (<i>rt</i>)		
				<i>rt</i>	<i>in</i>	LTR	LTR/ <i>in</i>	%C	%ΔC
I genome diploids (2 <i>n</i> = 14)									
1	<i>H. vulgare</i> cv Bonus	CI 3947	3.96	15.4 ± 1.5	14.6 ± 1.6	170 ± 14	11.6	3.7	8.8
2	<i>H. spontaneum</i> (Sédé Boqér)	20-48	4.14	23.6 ± 4.9	14.7 ± 3.4	217 ± 12	14.8	5.5	13.5
3	<i>H. spontaneum</i> (Upper Gallilee)	SCI 77-1	4.18	12.2 ± 2.4	18.7 ± 1.8	207 ± 37	11.0	2.8	5.3
4	<i>H. spontaneum</i> (Atlit)	25-34	4.19	11.6 ± 1.2	14.1 ± 1.8	165 ± 17	11.7	2.7	4.9
5	<i>H. spontaneum</i> (Mount Hermon)	1-27	4.33	22.4 ± 8.3	18.4 ± 3.2	171 ± 20	9.3	5.5	11.0
6	<i>H. spontaneum</i> (Mount Meron)	9-40	4.44	15.7 ± 3.9	14.7 ± 2.1	189 ± 22	12.9	3.4	6.5
7	<i>H. spontaneum</i> (Mehola)	22-53	4.45	18.5 ± 3.6	15.3 ± 1.7	149 ± 35	9.7	4.0	8.0
8	<i>H. spontaneum</i> (China)	H 3174	4.51	18.0 ± 2.9	18.1 ± 2.2	180 ± 20	9.9	3.9	7.5
9	<i>H. spontaneum</i> (Machtésh)	31-22	4.51	13.0 ± 1.9	15.5 ± 2.2	169 ± 20	10.9	2.8	4.7
10	<i>H. vulgare</i> cv Bomi	Bomi	4.53	13.2 ± 2.2	13.2 ± 2.0	89.9 ± 6.4	6.8	2.8	4.8
11	<i>H. spontaneum</i> (Revivim)	18-27	4.68	19.7 ± 4.2	18.2 ± 0.8	152 ± 38	8.4	4.1	7.7
12	<i>H. vulgare</i> (India)	CI 1090	4.71	20.1 ± 1.8	18.3 ± 1.1	128 ± 14	7.0	4.1	7.6
I genome tetraploids (2 <i>n</i> = 28)									
13	<i>H. bulbosum</i>	H 136	7.40	6.1 ± 1.0	9.9 ± 0.5	69.6 ± 3.3	7.0	0.8	0.3
Y genome tetraploids (2 <i>n</i> = 28)									
14	<i>H. murinum</i> subsp <i>leporinum</i>	H 509	8.52	13.1 ± 1.1	12.6 ± 1.0	289 ± 50	22.9	1.5	1.4
15	<i>H. murinum</i> subsp <i>murinum</i>	H 217	9.18	11.9 ± 1.7	18.0 ± 1.6	310 ± 26	17.2	1.3	1.1
H genome diploids (2 <i>n</i> = 14)									
16	<i>H. euclaston</i>	H 1132	2.85	9.3 ± 3.0	8.3 ± 1.5	185 ± 33	22.4	3.2	58.8
17	<i>H. pusillum</i>	H 1906	2.77	4.5 ± 0.6	6.2 ± 0.9	123 ± 5.5	19.9	1.6	—
18	<i>H. brachyanterum</i>	H 1942	3.52	20.9 ± 3.3	17.5 ± 3.5	162 ± 39	9.3	5.7	20.9
19	<i>H. erectifolium</i>	H 1150	3.79	9.5 ± 0.6	5.9 ± 0.4	128 ± 18	21.6	2.4	4.7
20	<i>H. bogdonii</i>	H 4014	3.85	8.6 ± 0.4	4.5 ± 0.6	129 ± 22	28.5	2.1	3.6
21	<i>H. muticum</i>	H 958	3.89	6.3 ± 0.8	5.1 ± 1.0	86 ± 23	17.0	1.6	1.6
22	<i>H. stenostachys</i>	H 1108	3.97	9.5 ± 0.7	9.0 ± 1.7	203 ± 44	22.5	2.3	4.0
23	<i>H. patagonicum</i> subsp <i>santacruzense</i>	H 1240	4.05	13.2 ± 8.2	4.3 ± 0.8	87 ± 9	20.4	3.1	6.5
24	<i>H. roshevitzii</i>	H 7039	4.16	16.4 ± 4.3	12.0 ± 1.6	360 ± 86	30.0	3.8	8.2
25	<i>H. patagonicum</i> subsp <i>patagonicum</i>	H 1319	4.40	16.8 ± 7.9	4.3 ± 0.4	161 ± 14	38.0	3.7	7.3
H genome tetraploids (2 <i>n</i> = 28)									
26	<i>H. depressum</i>	H 2089	6.57	29.9 ± 2.6	18.8 ± 1.8	153 ± 21	8.1	4.4	6.5
27	<i>H. jubatum</i>	H 4159	7.63	27.8 ± 3.0	14.3 ± 1.6	151 ± 33	10.6	3.5	4.6
X genome diploids (2 <i>n</i> = 14)									
28	<i>H. marinum</i> subsp <i>gussoneanum</i>	H155	4.04	8.1 ± 1.8	9.7 ± 1.1	407 ± 170	41.9	1.9	2.8

Table 2. Average Sequence Identity between Regions of *BARE-1* LTRs and *in* Domains

Species	Sequence Identity (%)										
	<i>H. vulgare</i> cv Bomi		<i>H. roshevitzii</i>		<i>H. marinum</i>		<i>H. euclaston</i>		<i>H. pusillum</i>		
	LTR	<i>in</i>	LTR	<i>in</i>	LTR	<i>in</i>	LTR	<i>in</i>	LTR	<i>in</i>	
<i>H. vulgare</i> cv Bomi	93.1	91.4									
<i>H. roshevitzii</i>	83.1	86.5	94.5	92.2							
<i>H. marinum</i>	81.9	ND ^a	89.6	ND	90	ND					
<i>H. euclaston</i>	80.9	82.9	91.6	86.3	88	ND	90.2	91.5			
<i>H. pusillum</i>	80.8	80	90.7	88.6	87.2	ND	89.4	81.8	87.8	84.3	

^a ND, not determined due to lack of sequence data for *H. marinum* integrase.

of identity among LTR and *in* sequences is not significantly different (*t* test), whereas the specific level of identity for the LTR and *in* for any two-group comparison is highly correlated (Pearson; $P = 0.01$). For both *in* and LTR sequences, barley sequences are ~10% more similar to each other than to those from other *Hordeum* spp. There was no systematic difference between the LTRs and *in* for any of the nonbarley species investigated. The data thus indicate that sequence divergence and its effect on hybridization cannot generate the high apparent LTR excesses. The intensity of the PCR amplification of *BARE-1* LTRs (Figure 2B) mirrors the relative excess of LTRs seen in the slot blot hybridizations, with strong amplification from *H. roshevitzii* (sample 24) and *H. marinum* subsp. *gussoneum* (sample 28) and weak amplification from *H. muticum* (sample 21) and *H. patagonicum* subsp. *santacruce* (sample 23).

Barley Bacterial Artificial Chromosome Clones Contain Solo LTRs

To more directly address the association of *BARE-1* LTRs and internal domains, we analyzed a set of clones from a barley genomic library prepared in a bacterial artificial chromosome (BAC) vector and containing large inserts (kindly screened and given by A. Druka and A. Kleinhofs, Washington State University, Pullman). The library was constructed from barley cultivar Morex. Twenty clones were analyzed by diagnostic digestion with eight restriction enzymes or combinations thereof (Figure 1D) followed by DNA gel blotting and then hybridization with LTR, *gag*, and *in* probes (Figure 1C). The enzymes were chosen to generate from *BARE-1* either internal fragments, from one or both LTRs or from the coding domain, or fragments extending into the flanking DNA. This allowed both confirmation of the presence and an estimate of the numbers of LTRs and coding domains. Based on addition of band lengths from the digests, an estimated total of 1.35 Mb of genomic DNA was thereby assayed, as tabulated by clone in Table 3.

One part of these analyses, a BamHI digest probed for the LTR and *in*, is shown in Figure 3. In this digest (Figure 3A), a single site in the *rt* region is predicted for *BARE-1a*, generating two fragments extending into the flanking DNA. Each fragment contains an LTR but only one *in* domain (Figure 1D). A cryptic site for BamHI was revealed by the generation of internal *in* fragments in four of the clones (Figure 3C). Inspection of the *BARE-1a* sequence indicates a site at nucleotide 764 differing from the BamHI recognition motif by 1 bp and an additional nine sites between nucleotides 750 and 850 differing by 2 bp. All of the clones except one contained at least one LTR (Figure 1B). The digests and hybridizations are summarized in Table 3; 10 clones had solo LTRs with no internal domains, six had an excess of LTRs over internal domains, three had fewer than 2:1 LTRs per internal domain (as might occur in a *BARE-1* cut short by the edge of the insert), and one clone had the 2:1 ratio. An overall ratio of 4.5 LTR/*in* and 4.5 LTR/*gag* was found by pooling the data for the 20 clones. In these analyses, nested insertions of LTRs within LTRs would not have been detected, leading to a possible underestimation of the excess of LTRs.

In Situ Hybridizations Reflect *BARE-1* Copy Number Variations

Hybridizations with *in* and *rt* probes were made to metaphase chromosomes of barley, *H. euclaston*, and *H. pusillum* to investigate the distribution and relative size of the *BARE-1* family in these species. Although it is difficult to make in situ hybridization conditions identical among chromosome mounts so that results can be quantitatively compared, the images in Figure 4 are representative of the hundred or so cells examined for each hybridization and of the two to five independent hybridizations made for each species and probe combination. The chromosomes of barley (Figures 4A to 4C) appear larger than those of either *H. euclaston* (Figures 4D to 4F) or *H. pusillum* (Figures 4G to 4I), reflecting the >1.5-fold greater *C* value of barley. Both *rt*

(Figure 4B) and *in* (Figure 4C) hybridizations against barley, compared with the 4'-diamidino-2-phenylindole (DAPI)-stained controls (Figures 4A, 4D, and 4G), show the uniform and dispersed hybridization pattern observed earlier for *in* (Suoniemi et al., 1996a), whereby the hybridization sites are observed along the whole chromosomes except at the centromeres, telomeres, and nucleolar organizers. The *rt* and *in* signals, however, are less uniformly distributed in barley than in the other species (cf. Figures 4B and 4E; Figures 4C and 4F and 4I). This may reflect the higher copy number in barley, the fact that distal insertion sites are apparently preferred, and the tendency of additional retrotransposon copies to insert into existing ones (SanMiguel et al., 1996; Noma et al., 1997; Suoniemi et al., 1997). For both probes, the signal intensity of the in situ hybridizations mirrored the copy number estimates from the slot blot hybridizations, being considerably stronger in barley. As with the slot blots, this is not likely to be due to divergence of target sequences outside of barley because the stringency of hybridization and washing was ~77%. Furthermore, the *in* signal (Figure 4F versus 4I) is somewhat stronger, and the *rt* signal (Figure 4E versus 4H) is considerably stronger, in *H. euclaston* than in

Table 3. Copy Number of *BARE-1* Components in Genomic BAC Clones of *H. vulgare* cv Morex

BAC Clone	Insert Size ^a (kb)	Copy Number ^b		
		LTR	<i>in</i>	<i>gag</i>
1	90	3	0	1
2	64	0	1	1
3	78	3	0	0
4	28	2	0	0
5	53	6	2	1
6	65	2	1	1
7	84	3	0	0
8	103	4	1	1
9	54	3	0	0
10	34	2	0	0
11	55	2	0	0
12	57	3	1	1
13	74	2	2	2
14	75	1	1	1
15	64	3	0	0
16	63	1	0	0
17	69	1	0	0
18	95	4	1	1
19	69	2	0	0
20	90	3	1	1

^a Estimated from restriction digests.

^b Copy number is based on restriction digests with eight enzymes or combinations thereof, followed by DNA gel blotting of these digests and hybridization to LTR, *in*, or *gag* probes. The clones together comprised 1.353 Mb and contained 50 LTR, 11 *in*, and 11 *gag* sequences.

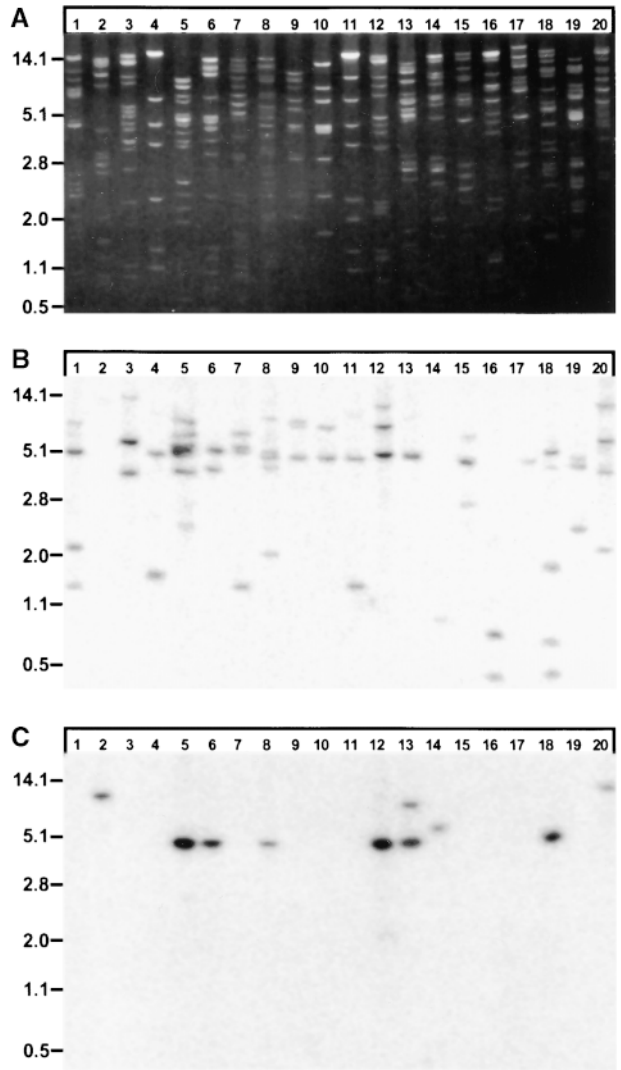


Figure 3. Barley Genomic BAC Clones Contain Solo LTRs.

(A) BamHI restriction digests of 20 BAC clones.
 (B) BamHI digest hybridized with an LTR probe.
 (C) BamHI digest hybridized with an *in* probe.
 Length markers indicated at left in (A) to (C) are in kilobases.

H. pusillum, in parallel with the copy number estimates for these species.

The Contribution of *BARE-1* to Genome Size in the Genus *Hordeum*

Plots of *BARE-1* copy number against basic genome size among the *Hordeum* spp accessions, shown in Figure 5,

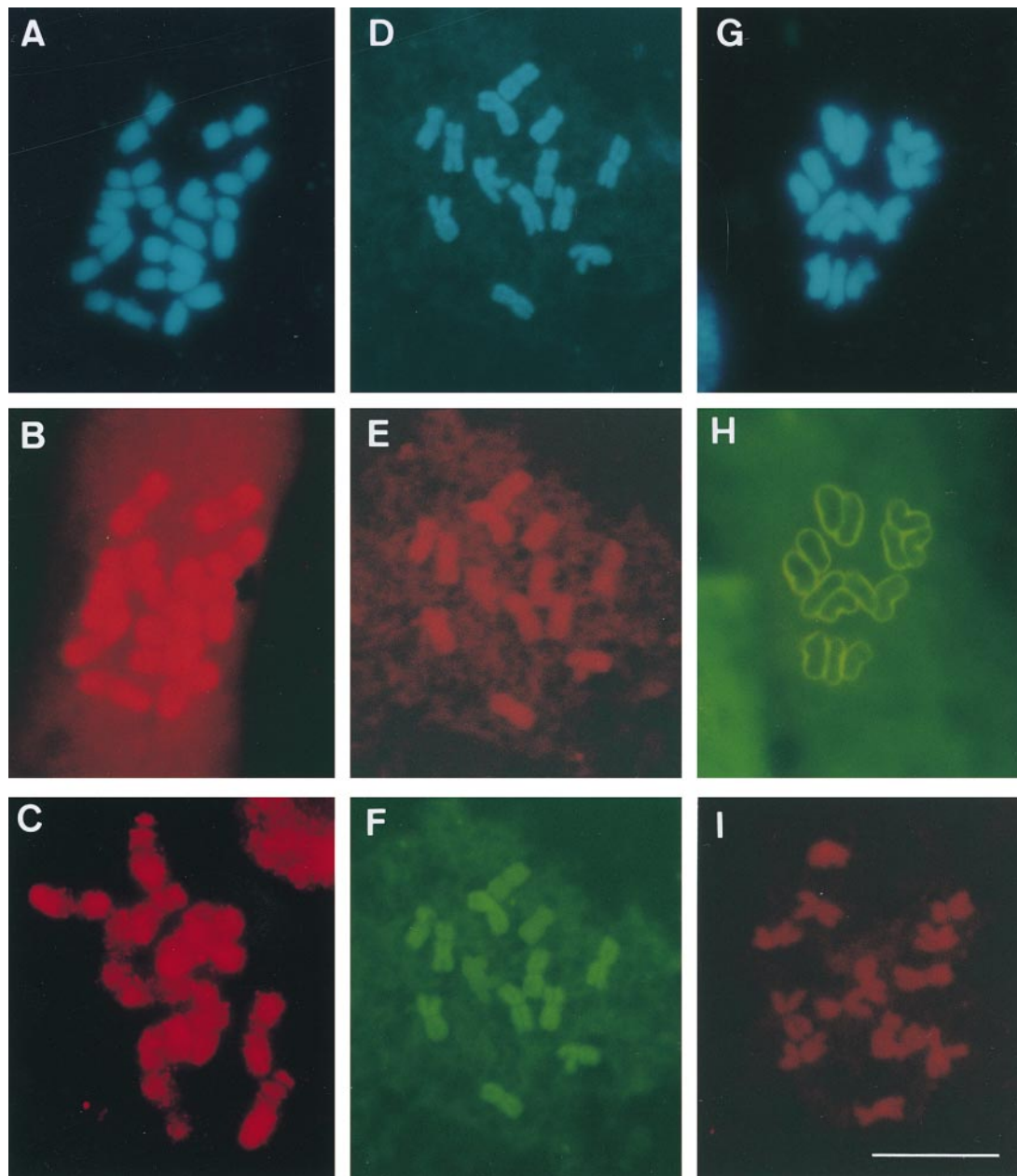


Figure 4. Strength of in Situ Hybridization of *BARE-1a* Probes to *Hordeum* spp Chromosomes Is Correlated with *BARE-1* Copy Number.

(A) to (C) show barley (*H. vulgare*) chromosomes. (D) to (F) show *H. euclaston* chromosomes. (G) to (I) show *H. pusillum* chromosomes. (A), (D), and (G) Blue, DAPI fluorescence from total DNA.

(B), (E), and (H) Hybridization to an *rt* probe.

(C), (F), and (I) Hybridization to an *in* probe.

Probe hybridization sites were detected by rhodamine (red) and fluorescein (green) fluorescence. Bar in (I) = 10 μm for (A) to (I).

have a positive and significant correlation for both the *rt* ($P = 0.01$; Figure 5A) and *in* ($P = 0.03$; Figure 5B) probes, linking *BARE-1* copy number to genome size. The barley accessions cluster at the high end of both the genome size and copy number distributions. Of the outliers below the distribution, three (samples 16, 17, and 18) are the diploids with the smallest genomes, respectively, *H. euclaston*, *H. pusillum*, and *H. brachyanterum*, and the other (sample 26) is the tetraploid *H. depressum*. These diploids showed fewer complete *BARE-1* elements for their genome size, whereas the tetraploid had a relatively larger basic genome for the *BARE-1* copy number. When the tetraploids are left out of the analysis, the regressions become both more predictive and more significant.

Based on these copy number estimates, the genome sizes previously determined for the same *Hordeum* spp accessions (Kankaanpää et al., 1996), and the observation that the *BARE-1* elements in each species have a length similar

to *BARE-1a* (8932 bp), the part of the genome comprising *BARE-1* was calculated (Table 1). The complement of the genome made up of full-length *BARE-1* was quite variable across the genus, constituting from 0.8 to 5.7% of the genome in the species examined by using the *rt* probe and 1.1 to 4.8% in the species examined by using the *in* probe. The *BARE-1* family contributed significantly less ($P = 0.002$) to the genomes of the H genome diploids than to barley, as determined by using the *in* probe, respectively comprising on average 2.1 and 3.5%. The same distinction was seen with the *rt* probe (2.9 vs. 3.7%), although it was not statistically significant. In the tetraploid *H. bulbosum* and *H. murinum* subspp, *BARE-1* formed significantly less of the genome (0.8 to 1.9%) than in the diploids as a whole, although this was not the case for the H genome tetraploids *H. depressum* and *H. jubatum*.

To assess the contribution of *BARE-1* to genome size growth in the genus *Hordeum* it would be necessary to make comparisons to the primitive state of the genome of the last common ancestor of the species in the genus. This ancestor is unknown, and the relative rates and effects of genome growth and shrinkage cannot be estimated a priori. Nevertheless, the *H. pusillum* genome, whether its features are in fact primitive or derived, is useful for comparison with the other species because it has the smallest genome analyzed, the fewest *BARE-1* copies, and the smallest percentage of the genome occupied by *BARE-1* (1.6%) in the genus (Table 1). On the basis of *H. pusillum*, a marginal %C of the genome has been calculated from the *rt* data for each accession (Table 1). This marginal share represents the fraction of the difference in total genome size that can be accounted for by the difference in the number of *BARE-1* copies. By using this method, a remarkable 59% of the difference in genome size between *H. pusillum* and *H. euclaston*, both similar H genome species of the section of the genus called *Anisolepis* (von Bothmer et al., 1995), can be accounted for by DNA of *BARE-1*. In another H genome species, *H. brachyanterum*, *BARE-1* contributed 21% of the calculated marginal difference in C value, but in *H. muticum* this was only 1.6%. The smallest differences calculated were for two tetraploid subspecies of *H. murinum* and tetraploid *H. bulbosum*.

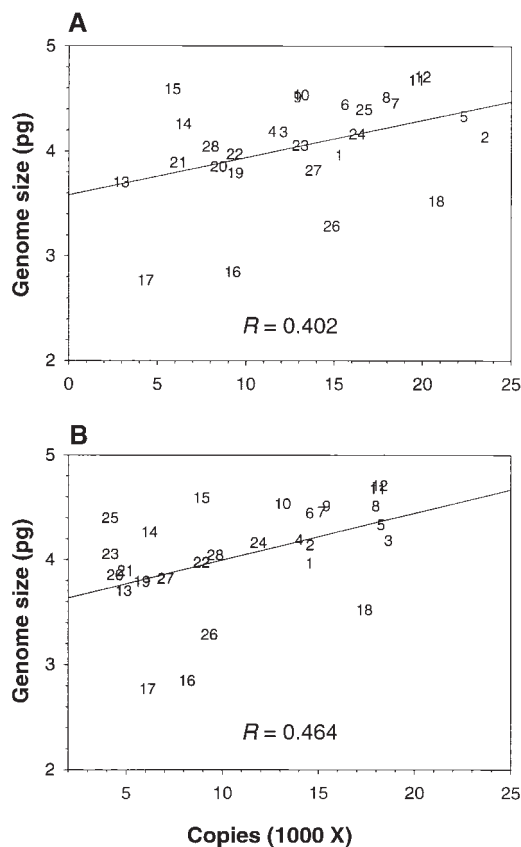


Figure 5. Genome Size Is Correlated with *BARE-1* Copy Number.

(A) Copy number determined with *rt* probe versus genome size.

(B) Copy number determined with *in* probe versus genome size.

The graphed numbers correspond to the sample numbers given in Table 1.

An Excess of LTRs Is Negatively Correlated with *BARE-1* Abundance

Unlike *in* or *rt*, the absolute LTR number shows no significant relationship to the basic genome size, as seen in Figure 6A. Nevertheless, for the diploids, the number of LTRs relative to *rt* or *in* (Table 1) is inversely correlated (Spearman rank order tests; *rt*, $r_s = -0.556$, $P = 0.006$; *in*, $r_s = -0.571$, $P = 0.005$) with genome size (Figure 6B). Furthermore, for the diploids, the proportion of the genome occupied by *BARE-1* (based on *rt*) is inversely related to the excess of LTRs over *in* copies (Figure 6C, $r_s = -0.528$, $P = 0.01$). The

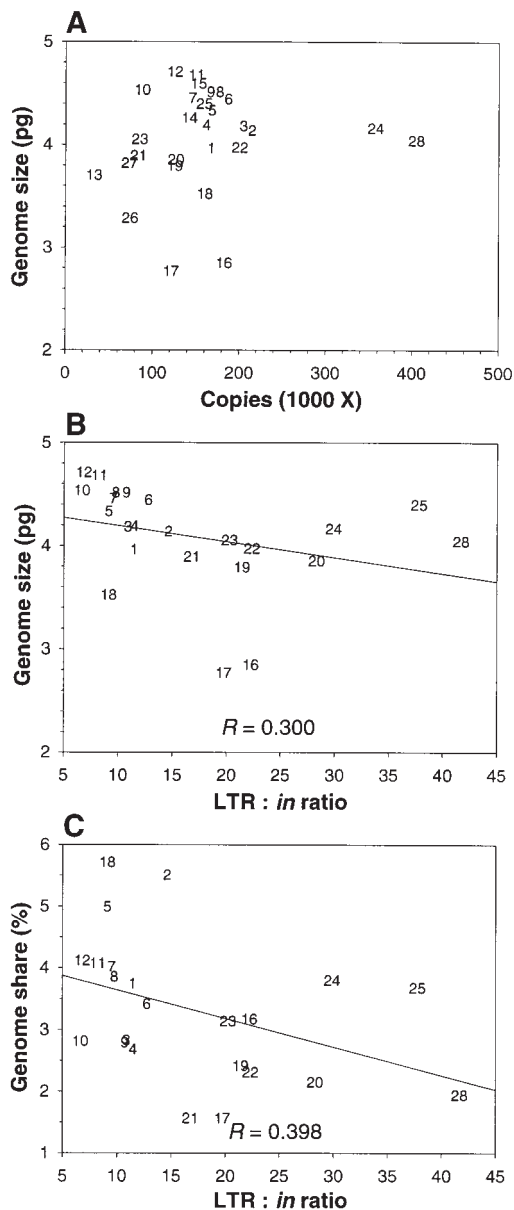


Figure 6. Genome Size as a Function of LTR Copy Number.

(A) LTR copy number versus genome size.

(B) Ratio of LTR number to *in* number versus genome size.

(C) Ratio of LTR number to *in* number versus fraction of the genome occupied by *BARE-1* (by *rt* number).

The graphed numbers correspond to the sample numbers given in Table 1.

outliers (samples 16, 17, and 18) in the correlation with genome size (Figure 6B) are the three diploid species with the smallest genomes. For a given proportional excess of LTRs, their genomes are otherwise considerably smaller than the general trend would predict. Taking the data together, the more LTRs relative to *BARE-1* elements, the smaller the contribution of *BARE-1* to the genome.

BARE-1 Copy Number and Genome Size in *H. spontaneum* Appear Correlated with Habitat

Analyses were made of the relationship of *BARE-1* copy number and genome size to environmental variables in the *H. spontaneum* populations for which detailed climatic and soil data were available. Genome size showed a high correlation ($|r_s| > 0.6$) with evaporation and alluvium soil type. The LTR number was strongly correlated ($|r_s| > 0.7$) with altitude, annual temperature, January and August temperature, the number of hot and dry days, and alluvium soil type. However, most correlations were insignificant, and the proportion of significant correlations may not be above that expected by chance. A trend is nevertheless clear, with effects being associated with variables of temperature, water availability, and soil type. In multiple regression, the coefficient of determination of *in* number was high ($R^2 = 0.874$, $P = 0.073$) and is explained by variables linked to hot, dry desert conditions. Furthermore, genome size appears higher in the desert, that is, it seems to increase with aridity. We interpret the lack of statistical significance at the 0.05 level as an effect of the small sample size.

DISCUSSION

The *C* value paradox, that variation in genome size (10^8 to 10^{11} bp in plants) is not correlated with the complexity of the organism, has long been apparent (Thomas, 1971). Comparisons of fine-structure maps within the family Poaceae (Ahn et al., 1993; Chen et al., 1997) sharpened the paradox by showing, across a >10-fold variation in genome size, that homologous chromosomes of grasses are syntenic, the order of mapped genes being generally preserved against great variation in genome size. Genes in the grasses appear to occur in islands separated by repetitive DNA (Barakat et al., 1997; Panstruga et al., 1998), repetitive DNA comprising >70% of the barley and maize genomes (Flavell et al., 1977; Barakat et al., 1997) and transposons comprising >50% of the maize genome (SanMiguel et al., 1996).

Together, these observations suggest that retrotransposons account for much of the size variation between genomes in the family Poaceae. Retrotransposons are "selfish" (Doolittle and Sapienza 1980; Orgel and Crick, 1980), possessing a self-contained system for replicatively increasing their copy number, which, in the absence of detrimental ef-

fects to the host, should lead to greater numbers of increasingly active copies. This and the lack of evidence for an equally efficient mechanism for removing repetitive DNA from the genome have sparked a discussion on the role of retrotransposons in unidirectionally ratcheting genome size upwards (Wessler et al., 1995; Bennetzen and Kellogg, 1997a, 1997b; Petrov, 1997; Voytas and Naylor, 1998).

To examine the prevalence of retrotransposon *BARE-1* in the genomes of *Hordeum* spp, we determined the copy number for its *in* and *rt* internal domains and the terminal LTRs throughout the genus by slot and dot blot hybridizations. The data here show that *BARE-1* copy number and genome size are positively correlated in the genus *Hordeum*, in contrast to what has been reported for the genus *Vicia* from more limited comparisons (Pearce et al., 1996a). Across the genus as a whole, *BARE-1* is present on average in $14 \pm 1 \times 10^3$ copies and in barley at $16.6 \pm 0.6 \times 10^3$ copies, conserved in length. The *in situ* hybridization data are consistent with the relative copy numbers calculated from the blotting data, the dispersed *BARE-1* elements in the three *Hordeum* spp examined displaying a relative intensity parallel with the detected copy number.

The data further show that *Hordeum* spp genomes contain a large excess (7- to 42-fold, mean 16 ± 2 -fold) of LTRs relative to the internal regions of *BARE-1*, departing from the expected 2:1 ratio of intact retrotransposons. The genome of barley cultivar Bomi, with one of the smallest excesses of LTRs found, contains $>6 \times 10^4$ LTRs than can be accounted for by the number of *BARE-1* elements. The BAC analyses examined 0.03% of the genome of one barley cultivar in detail; the dot and slot blot hybridizations surveyed the entire genomes of the accessions. The data taken together confirm the existence of solo LTRs in barley specifically and strongly suggest that excess LTRs detected in the other *Hordeum* spp are solo LTRs as well. Examples of solo LTRs have been found in mammals (Banville et al., 1992; Blusch et al., 1997), *Drosophila* (Geyer et al., 1988), yeast (Parket et al., 1995; Kim et al., 1998), and plants (Sentry and Smyth, 1989; SanMiguel et al., 1996; Noma et al., 1997; Bevan et al., 1998). However, an abundance such as that seen in the genus *Hordeum* has not been reported previously.

Homologous recombination between the LTRs of a single *BARE-1* element would remove the internal domain and leave behind a single recombinant LTR flanked by the direct repeats remaining from the original insertion site. Hence, the strongest support for the existence of solo LTRs and their derivation from formerly intact *BARE-1* elements would come from sequences of large contiguous genomic regions. Indeed, analysis of a fully sequenced contiguous stretch of 66 kb from chromosome 2H of cultivar Ingrid (K. Shirasu, A. Schulman, T. Lahaye, and P. Schulze-Lefert, manuscript in preparation) shows that this region contains four *BARE-1* units, derived from five retrotransposons. Of these units, two are solo LTRs that have resulted from recombination in a single element, and one consists of a pair of solo LTRs resulting from recombination within one *BARE-1* nested inside another. The

DNA gel blots as well as the PCRs show that these solo LTRs are present as such in the genome and are not a result of bacterial recombination in the BAC.

Intrachromosomal homologous recombination has been demonstrated and analyzed in dicotyledonous plants with integrated model substrates (reviewed in Puchta and Hohn, 1996). The recombination is thought to proceed by either double-strand break repair or single-strand annealing at rates of 10^{-5} to 10^{-6} homologous recombination events per cell division (Puchta and Hohn, 1996). The rate is additive for transgenes in allelic positions, and it is dependent on the plant species, configuration, and genomic position of the recombination substrate, plant organ, and particular progeny (Puchta and Hohn, 1996), although perhaps not on the degree of methylation at CG and CXG nucleotide motifs (Puchta et al., 1992). Intrachromosomal recombination between repetitive sequences is thought to be infrequent, limiting the loss of intervening genes (Hu et al., 1998). However, double-stranded breaks can efficiently induce intrachromosomal recombination between flanking homologous sequences in maize (Athma and Peterson, 1991) and yeast (Parket et al., 1995). Whereas no direct data are available, the degree of polymorphism seen with an anchored PCR method (Vaughn et al., 1997) indicates that the *BARE-1* insertion frequency is in the range seen for intrachromosomal recombination, $<4 \times 10^{-5}$ events per element per generation.

The ratio between the number of LTRs and full-length elements may reflect the balance of the relative rates of *BARE-1* propagation and inter-LTR recombination. These rates need not be either constant or consonant; the maize genome has apparently experienced an explosive increase in retrotransposon numbers in at least part of the genome over the last 3 million years (SanMiguel et al., 1998), although only two solo LTRs were found in a region of abundant, nested retrotransposons (SanMiguel et al., 1996). Recombination between LTRs would be expected to reduce the complement of functional retrotransposons in the genome, limiting but not eliminating the contribution of the *BARE-1* family to a genome size increase. Consistent with this idea, the excess of LTRs observed relative to the *BARE-1* numbers increases as both the genome size and the portion of the genome occupied by *BARE-1* decrease. Selective pressure for a small genome in the presence of countervailing activity by *BARE-1* would lead to accumulation of excess of LTRs through recombination. The yeast genome, which is thought to be under pressure for compactness, represents an example of the accumulation of solo LTRs, containing 51 full-length Ty retrotransposons but 280 solo LTRs or LTR fragments (Kim et al., 1998), with high variability in the number per chromosome.

The opposite selective pressures may be found under hot, dry desert conditions, which multivariate analysis shows are strongly and significantly correlated with increasing genome size and *BARE-1* copy number. The data suggest that increases in both genome size and genetic polymorphism (Nevo et al., 1979; Nevo and Beiles, 1988) in dry environments might be adaptive in the genus *Hordeum* as it is in the

pinus (Wakamiya et al., 1996) and concomitantly associated with either propagation of *BARE-1* or inheritance of new copies. In this regard, transcription or transposition of various retrotransposons has been shown to be linked to biotic and abiotic stresses (Wessler, 1996; Grandbastien, 1998; Takeda et al., 1998).

In summary, recombination may provide a partial "return ticket from genomic obesity" (Bennetzen and Kellogg, 1997a), at least from the genome bloat due to retrotransposons, providing that the frequency relative to retrotransposition is sufficiently high. Although accumulation of the 1.8-kb solo *BARE-1* LTRs that result from recombination would still contribute to genome size increase, additional recombination between nearby LTRs could limit it. The tendency of retrotransposons to insert into other retrotransposons, at least in grass genomes (SanMiguel et al., 1996; Noma et al., 1997; Suoniemi et al., 1997), and into regions of microsatellites (Kalendar et al., 1999; Ramsay et al., 1999) would simultaneously provide an additional means of inactivating retrotransposons and reduce the risk of gene deletion through the recombinational loss of genomic DNA intervening between solo LTRs.

The overall importance to genome evolution in the genus *Hordeum* of *BARE-1* propagation and loss depends on the number and dynamics of the other families of retrotransposons and repetitive sequences in the genome, the latter known to undergo rapid change in several grass genomes (Saghai Maroof et al., 1990; Ceccarelli et al., 1992; Hueros et al., 1993; SanMiguel et al., 1998). The three smallest genomes among the accessions examined all have smaller genomes than is typical for their *BARE-1* copy number and relative LTR excess. This suggests that these genomes have been under strong downward selection for size and that other forms of repetitive DNA contributing to the total *C* value have been depleted more successfully than has *BARE-1*. Conversely, compared with the smallest (not necessarily representing the ancestral) genome measured in the genus, *BARE-1* contributes an average of $8 \pm 2\%$ and a range of 0.3 to 59% to the difference between the size of this genome and others in the genus. The magnitude of this range and the variation in LTR abundance suggest that the amount of activity and impact of *BARE-1* on the genome has been highly lineage-specific. The selective pressures and mechanisms modulating this "genome war" between the cell and the retrotransposon remain to be clarified.

METHODS

Plant Materials

Sources and provenances of the *Hordeum* spp accessions are detailed elsewhere (Kankaanpää et al., 1996). They comprised 27 accessions and 17 species, including nine accessions of *H. spontaneum* (wild barley), an Indian and an African landrace of cultivated

barley (*H. vulgare*), and one commercially bred barley. Seeds were germinated and then grown in a controlled environment chamber as described previously (Suoniemi et al., 1996b). For in situ hybridizations, seeds derived from a cross between barley cultivars Alexis and Regata were used.

DNA Preparation

DNA was isolated from leaves, essentially as detailed before (Manninen and Schulman, 1993), and then treated with RNase ONE ribonuclease (Promega, Madison, WI) followed by phenol and chloroform extractions and ethanol precipitation. Total genomic and plasmid DNA was quantified both spectrophotometrically and after agarose gel electrophoresis as ethidium bromide fluorescence under UV light. Polaroid type 665 negatives were scanned, and densitometry was performed on the scanned image using Tina 2.09 software (Isotopenmessgeräte, Straubenhardt, Germany) against a known amount of a PstI fragment from λ phage DNA. All plasmids were cloned in *Escherichia coli* JM109.

Polymerase Chain Reactions and Sequencing

The primers used for the amplification by polymerase chain reaction (PCR) of the *BARE-1* internal domain were previously described (Suoniemi et al., 1997) and corresponded to bases 1685 to 1704 in the left long terminal repeat (LTR) and 7214 to 7235 in the right LTR. The PCR reactions were performed using the Expand Long Template PCR System (Roche Molecular Biochemicals, Mannheim, Germany) with buffer 1, as described by the supplier, using 10 ng genomic DNA, 0.2 mM each dNTP, and 1 pmol μL^{-1} each primer in a final volume of 50 μL . The mix was overlaid with paraffin oil. The reaction mixtures were heated to 95°C for 5 min; and then subjected to seven cycles of 94°C for 30 sec, 43°C for 2 min, a ramp of +1°C (2 sec) $^{-1}$ to 72°C, and 72°C for 4 min. This was followed by 41 cycles of 94°C for 30 sec, 60°C for 2 min, a ramp of +1°C (2 sec) $^{-1}$, and 72°C for 4 min. The reaction was completed by a 10-min incubation at 72°C. Reactions were performed in a Minicycler (MJ Research, Waltham, MA) thermal cyclers.

For amplification of genomic LTRs, primers at the ends of the LTRs were used (N referring to equal amounts of A, T, G, and C in the primer preparation at that position): forward, 5'-NNTGTTGGAATTATGCCCTAGAGGCAA-3' (GenBank accession number Z17327, bases 309 to 333); reverse, 5'-NNTGTGGGGAACGTCGCATGGGAAAC-3' (GenBank accession number Z17327, bases 2113 to 2137). The reaction conditions were as for the internal domain (above) except that the extension time was 2 min at 72°C. For cloning and sequencing LTRs of various species for sequence comparisons, we used the same set of primers as previously. The PCR mix was the same as used previously, but the reaction cycles were different. The reaction mixtures were first heated to 95°C for 5 min, followed by 21 cycles of 94°C for 30 sec, 40°C for 2 min, and 72°C for 2.5 min. Reactions were completed with one incubation at 72°C for 10 min. The PCR products were purified from agarose gels (QIAEX II; Qiagen, Hilden, Germany) and cloned (pGEM-T vector system; Promega). Sequencing reactions on plasmid minipreps were performed with Sequenase v2.0 (Amersham Pharmacia Biotech, Uppsala, Sweden) and analyzed under standard conditions with an automated system (ALF; Amersham Pharmacia Biotech).

Sequences were aligned and compared as previously described (Suoniemi et al., 1998a).

Slot and Dot Blots

Slot blots were prepared by filtration through a vacuum manifold (Hoeffer PR600; Amersham Pharmacia Biotech), applying 10 to 100 ng of sample genomic DNA with herring sperm DNA used to maintain a constant DNA load. The DNA was cross-linked to filters under UV light. Herring sperm DNA at 100 ng per well also served as a negative control. Isolated plasmids (0.1 to 10 ng per well) containing the fragment used for hybridization probes served as positive controls on each filter. The LTR probe (NheI-BstEII, 743 bp) was from the untranscribed region. The other probes were for *in* (HpaI-BsmI, 589 bp) and *rt* (Sall-StyI, 703 bp). All the probes were random-primed (Rediprime or Megaprime; Amersham Pharmacia Biotech) and ³²P-labeled. Filters were hybridized in 50% formamide, 1.25 × SSPE (1 × SSPE is 58.8 mM Na₂HPO₄, 61.2 mM NaH₂PO₄, 0.6 M NaCl, 60 mM sodium citrate, pH 6.8), 5 × Denhardt's solution (1 × Denhardt's solution is 0.02% [w/v] Ficoll 400, 0.02% [w/v] polyvinylpyrrolidone averaging 360 kD, 0.2% BSA), 0.5% SDS, and 20 μg mL⁻¹ herring sperm DNA overnight at 42°C. Hybridized filters were washed successively with 2 × SSC (1 × SSC is 0.15 M NaCl and 0.015 sodium citrate), 0.1% SDS (10 min, 25°C), twice in 2 × SSC, 0.1% SDS (10 min, 65°C), and once in 0.2 × SSC (20 min, 65°C). Bound radiation was quantified by exposure of an imaging plate for 2 to 16 hr followed by scanning on either a PhosphorImager (model SI; Molecular Dynamics, Sunnyvale, CA) or a BAS PhosphorImager (model 1500; Fuji Photo Film, Tokyo, Japan).

Isolation and Characterization of Bacterial Artificial Chromosome Clones

A barley genomic bacterial artificial chromosome (BAC) library constructed from barley cultivar Morex (Y. Yu, J. Tompkins, D. Frisch, R. Waugh, R. Brueggeman, D. Kudrna, A. Kleinhofs, and R. Wing, personal communication) was screened using a full-length LTR probe. Twenty positive clones were selected, and BAC DNA was prepared by standard methods (Sambrook et al., 1989). The BAC DNA fragments were separated on agarose gels, denatured, and transferred to a Hybond N+ (Amersham Pharmacia Biotech) membrane using standard methods (Sambrook et al., 1989). Hybridization conditions were as for the slot blot hybridizations. The hybridized filters were washed twice in 2 × SSC, 0.1% SDS for 10 min at 25°C and twice in 2 × SSC, 0.1% SDS for 10 min at 65°C. The sizes of the BAC clone inserts were estimated on the basis of the restriction fragments obtained in restriction digests.

Copy Number Calculations

Genomic copy number was calculated from the hybridization response of the genomic DNA compared with the control plasmids on the blots as follows: copies ng⁻¹ = genomic PSL ng⁻¹ × plasmid copies × plasmid PSL⁻¹, where PSL stands for photo-stimulated luminescence units, the output unit for exposure of the PhosphorImager screens. The copies ng⁻¹ were converted to copies genome⁻¹ using the genome size data available for the same accessions of *Hordeum* spp (Kankaanpää et al., 1996). For each sample, the mean response of three to six hybridizations on at least three separate filters was used to determine copy number.

For the slot blots, corrections were made for variations in sample binding and accessibility by hybridization response of each filter and sample to a labeled total DNA probe. For this purpose, the raw copy number was normalized relative to barley based on the samples' hybridization response to random-primed labeled total cultivar Bomi DNA. Sample 12 (Table 1) was present on every filter. All the blots were hybridized together so that the probe concentration was identical for all filters. The control hybridizations were washed under conditions 18°C below the melting temperature of 83°C, calculated for barley DNA on the basis of 0.043 M Na⁺ in the washing solution and an average 59.1% G + C nucleotide content in barley coding DNA (Nakamura et al., 1997).

A set of dot blot hybridizations with sample 12 was used to determine absolute *BARE-1* copy number in this accession and thereby convert the relative copy numbers in the other accessions to absolute numbers. The DNA of cultivar Bomi was also present as a control on these dot blot filters. Copy number could not be determined by direct comparison of the hybridization response of the control plasmids and genomic DNA against the labeled total DNA probe because the plasmid control spots saturated the available *BARE-1* fragments in the genomic probe, giving gross underestimations of the amount of bound control plasmid DNA.

In addition, this method was assessed with a set of control dot blot hybridizations in which the accession DNA was spiked with λ phage DNA before individual samples were pipetted. For each microgram of genomic plant DNA, 139.4 ng λ phage DNA was added, giving 1.2 × 10⁴ copies of λ phage per barley genome equivalent. Dot blots were prepared with multiple replicates and both 1 and 10 ng of genomic DNA per sample. The same filter was probed in series with integrase (*in*), LTR, and λ phage probes as above for the slot blots. Hybridization response to the *in* and LTR probes was corrected to the average value for the λ phage DNA hybridization response and copy number calculated as above.

In Situ Hybridization

The clones used for probes were as described above for slot blots. Probes were labeled by nick translation with either rhodamine-4-dUTP (Amersham Pharmacia Biotech) or biotin-16-dUTP (Roche Molecular Biochemicals), following a protocol described previously (Ananthawat-Jónsson et al., 1996). Root tips were collected from germinating *Hordeum* spp seedlings when the roots were 1 to 2 cm long and then treated in ice water for 25 hr before fixation in 3:1 absolute ethanol:glacial acetic acid. The fixed root tips were digested with cellulase and pectinase, after which chromosome preparations were made either by the squash method (Schwarzacher and Leitch, 1994) or by protoplast dropping (Busch et al., 1996).

Chromosomes were treated and in situ hybridization was performed according to a standard procedure (Ananthawat-Jónsson et al., 1996). Hybridizations were performed at a stringency of 77% in 50% formamide, 2 × SSC at 37°C overnight and then washed at a stringency of 76% in 40% formamide, 2 × SSC at 42°C for 10 min. Each experiment (DNA probe and species) was repeated two to five times with independently prepared probes and chromosome preparations. Rhodamine-labeled probes were detected directly, whereas biotin-labeled probes were detected with ExtrAvidin-FITC (Sigma). The hybridization signal was visualized and photographed with epifluorescence microscopy, with >100 cells per preparation examined to locate representative chromosome sets.

Statistical Analyses

Copy Number

The DNA size estimations were made with TableCurve V. 2.02 (SPSS Science, Chicago, IL). Linear regression analyses for copy number data were made with SigmaPlot version 4.0 (SPSS Science). Parametric correlation analyses were made with the Pearson product moment test (yielding r_p and P) and nonparametric tests with the Spearman rank order correlation tests (r_s and P) as implemented in SigmaStat version 2.0 (SPSS Science). P values in the text represent the likelihood of falsely rejecting the null hypothesis that the variables are not correlated. The variables r_p and r_s are the correlation coefficients for the Pearson and Spearman tests, respectively. Values in the text are expressed as means and standard errors (SE).

Correlations among Environmental Factors, Genome Size, and BARE-1 Copy Number

Spearman rank correlations (SAS Institute, 1996) were made between geographic, climatic, and edaphic factors and the genome size, *rt*, *in*, and LTR copy number and LTR:*in* ratio for seven Israeli populations of *H. spontaneum*. Only significant correlations were corrected by the Bonferroni test for the number of correlations that were attempted. Sample 3, accession number SCI 77-1 from the Upper Galilee, was not included in this test due to missing ecological data. Ecogeographical data and sites are detailed elsewhere (Table 1 and Figure 1 in Nevo et al., 1979). We considered environmental variables in the following categories: geographical, including longitude, latitude, and altitude; climatic means, including annual, January, and August temperature, seasonal temperature difference, daily temperature difference, Sharav or the number of hot and dry days, the number of tropical days, and evaporation; and moisture conditions, including annual rainfall, the number of rainy days, the number of dewy nights in summer, annual humidity, humidity at 14:00 hours, and the Thornthwaite moisture index. We also considered four edaphic dummy variables, one for each soil type: alluvium, loess, sandy loam, and terra rossa.

Multiple Regression Analysis

A test of the best predictors of *C*, *rt*, *in*, LTR, and LTR/*in* of seven Israeli populations of *H. spontaneum* was conducted by stepwise multiple regression analysis (SAS Institute, 1996) using these characteristics as dependent variables and geographic, climatic, and edaphic factors as independent variables. The environmental variables we used included those that were used in the Spearman rank correlation, excluding three variables: Sharav, number of dewy nights in summer, and terra rossa soil type.

ACKNOWLEDGMENTS

We thank Anne-Mari Narvanto for her continuing excellent technical assistance. We are grateful to Arnis Druka and Andris Kleinhofs (Department of Crop and Soil Sciences, Washington State University, Pullman, WA) for screening and donating BAC clones for further anal-

ysis. The research reported here was supported by grants from the Academy of Finland Genome Research Program and the European Union Directorate for Biotechnology research program on Molecular Tools for Biodiversity. E.N. thanks the Israel Discount Bank Chair of Evolutionary Biology and the Ancell-Teicher Research Foundation for Genetics and Molecular Evolution for financial support.

Received April 16, 1999; accepted June 2, 1999.

REFERENCES

- Adams, S.E., Mellor, J., Gull, K., Sim, R.B., Tuite, M.F., Kingsman, S.M., and Kingsman, A.J. (1987). The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian retroviral proteins. *Cell* **49**, 111–119.
- Ahn, S., Anderson, J.A., Sorrells, M.E., and Tanksley, S.D. (1993). Homeologous relationships of rice, wheat, and maize chromosomes. *Mol. Gen. Genet.* **241**, 483–490.
- Anamthawat-Jónsson, K., Heslop-Harrison, J.S., and Schwarzacher, T. (1996). Genomic in situ hybridization for whole chromosome and genome analysis. In *In Situ Hybridization Laboratory Companion*, M. Clark, ed (Weinheim, Germany: Chapman and Hall GmbH), pp. 1–24.
- Athma, P., and Peterson, T. (1991). *Ac* induces homologous recombination at the maize *P* locus. *Genetics* **128**, 163–173.
- Banville, D., Rotaru, M., and Boie, Y. (1992). The intracisternal A particle derived solo LTR promoter of the rat oncomodulin gene is not present in the mouse genome. *Genetica* **86**, 85–97.
- Barakat, A., Carels, N., and Bernardi, G. (1997). The distribution of genes in the genomes of Gramineae. *Proc. Natl. Acad. Sci. USA* **94**, 6857–6861.
- Bennetzen, J.L., and Kellogg, E.A. (1997a). Do plants have a one-way ticket to genome obesity? *Plant Cell* **9**, 1509–1514.
- Bennetzen, J.L., and Kellogg, E.A. (1997b). Reply. *Plant Cell* **9**, 1901–1902.
- Bevan, M., et al. (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* **391**, 485–488.
- Blusch, J.H., Haltmeier, M., Frech, K., Sander, I., Leib-Mosch, C., Brack-Werner, R., and Werner, T. (1997). Identification of endogenous retroviral sequences based on modular organization: Proviral structure at the SSAV1 locus. *Genomics* **43**, 52–61.
- Boeke, J.D., and Chapman, K.B. (1991). Retrotransposition mechanisms. *Curr. Opin. Cell Biol.* **3**, 502–507.
- Busch, W., Herrmann, R.G., Houben, A., and Martin, R. (1996). Efficient preparation of plant metaphase spreads. *Plant Mol. Biol. Rep.* **14**, 149–155.
- Ceccarelli, M., Falistocco, E., and Cionini, P.G. (1992). Variation of genome size and organization within hexaploid *Festuca arundinacea*. *Theor. Appl. Genet.* **83**, 273–278.
- Chen, M., SanMiguel, P., de Oliveira, A.C., Woo, S.-S., Zhang, H., Wing, R.A., and Bennetzen, J.L. (1997). Microcolinearity in *sh2*-homologous regions of the maize, rice, and sorghum genomes. *Proc. Natl. Acad. Sci. USA* **94**, 3431–3435.
- Doolittle, R.F., Feng, D.-F., Johnson, M.S., and McClure, M.A.

- (1989). Origins and evolutionary relationships of retroviruses. *Quart. Rev. Biol.* **64**, 1–30.
- Doolittle, W.F., and Sapienza, C.** (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**, 601–603.
- Flavell, A.J., Dunbar, E., Anderson, R., Pearce, S.R., Hartley, R., and Kumar, A.** (1992). *Ty1-copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Res.* **20**, 3639–3644.
- Flavell, R.B., Rimpau, J., and Smith, D.B.** (1977). Repeated sequence DNA relationships in four cereal genomes. *Chromosoma* **63**, 205–222.
- Gaut, B.S., Morton, B.R., McCaig, B.C., and Clegg, M.T.** (1996). Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. *Proc. Natl. Acad. Sci. USA* **93**, 10274–10279.
- Geyer, P.K., Green, M.M., and Corces, V.G.** (1988). Reversion of a *gypsy*-induced mutation at the yellow (*y*) locus of *Drosophila melanogaster* is associated with the insertion of a newly defined transposable element. *Proc. Natl. Acad. Sci. USA* **85**, 3938–3942.
- Goodman, H.M., Ecker, J.R., and Dean, C.** (1995). The genome of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **92**, 10831–10835.
- Grandbastien, M.-A.** (1992). Retroelements in higher plants. *Trends Genet.* **8**, 103–108.
- Grandbastien, M.-A.** (1998). Activation of retrotransposons under stress conditions. *Trends Plant Sci.* **3**, 181–187.
- Gribbon, B.M., Pearce, S.R., Kalendar, R., Schulman, A.H., Jack, P., Kumar, A., and Flavell, A.J.** (1999). Phylogeny and transpositional activity of *Ty1-copia* group retrotransposons in cereal genomes. *Mol. Gen. Genet.* **261**, 883–891.
- Hirochika, H.** (1993). Activation of tobacco retrotransposons during tissue culture. *EMBO J.* **12**, 2521–2528.
- Hirochika, H., Fukuchi, A., and Kikuchi, F.** (1992). Retrotransposon families in rice. *Mol. Gen. Genet.* **233**, 209–216.
- Hsiao, C., Chatterton, N.J., Asay, K.H., and Jensen, K.B.** (1995). Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. *Genome* **38**, 211–223.
- Hu, W., Timberman, M.C.P., and Messing, J.** (1998). Interchromosomal recombination in *Zea mays*. *Genetics* **150**, 1229–1237.
- Hueros, G., Loarce, Y., and Ferrer, E.** (1993). A structural and evolutionary analysis of a dispersed repetitive sequence. *Plant Mol. Biol.* **22**, 635–643.
- Hyman, R.W., Brunovskis, I., and Summers, W.C.** (1973). DNA base sequence homology between coliphages T7 and ϕ 11 and between T3 and ϕ 11 as determined by heteroduplex mapping in the electron microscope. *J. Mol. Biol.* **77**, 189–196.
- Ikeda, T.M., Terachi, T., and Tsunewaki, K.** (1992). Variations in chloroplast proteins and nucleotide sequences of three chloroplast genes in *Triticum* and *Aegilops*. *Jpn. J. Genet.* **67**, 111–123.
- Jacobsen, N., and von Bothmer, R.** (1992). Supraspecific groups in the genus *Hordeum*. *Hereditas* **116**, 21–24.
- Joseph, J.L., SENTRY, J.W., and Smyth, D.R.** (1990). Interspecies distribution of abundant DNA sequences in *Lilium*. *J. Mol. Evol.* **30**, 146–154.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A., and Schulman, A.H.** (1999). IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. *Theor. Appl. Genet.* **98**, 704–711.
- Kankaanpää, J., Mannonen, L., and Schulman, A.H.** (1996). The genome sizes of *Hordeum* species show considerable variation. *Genome* **39**, 730–735.
- Kim, J.M., Vanguri, S., Boeke, J.D., Gabriel, A., and Voytas, D.F.** (1998). Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.* **8**, 464–478.
- Kurata, N., Umehara, Y., Tanoue, H., and Sasaki, T.** (1997). Physical mapping of the rice genome with YAC clones. *Plant Mol. Biol.* **35**, 101–113.
- Lucas, H., Feuerbach, F., Kunert, K., Grandbastien, M.-A., and Caboche, M.** (1995). RNA-mediated transposition of the tobacco retrotransposon *Tn1* in *Arabidopsis thaliana*. *EMBO J.* **14**, 2364–2373.
- Manninen, I., and Schulman, A.H.** (1993). *BARE-1*, a *copia*-like retroelement in barley (*Hordeum vulgare* L.). *Plant Mol. Biol.* **22**, 829–846.
- Marillia, E.F., and Scoles, G.J.** (1996). The use of RAPD markers in *Hordeum* phylogeny. *Genome* **39**, 646–654.
- Matsuoka, Y., and Tsunewaki, K.** (1996). Wheat retrotransposon families identified by reverse transcriptase domain analysis. *Mol. Biol. Evol.* **13**, 1384–1392.
- Meinkoth, J., and Wahl, G.** (1984). Hybridization of DNA immobilized on solid supports. *Anal. Chem.* **138**, 267–284.
- Nakamura, Y., Gojobori, T., and Ikemura, T.** (1997). Codon usage tabulated from the international DNA sequence databases. *Nucleic Acids Res.* **25**, 244–245.
- Nevo, E., and Beiles, A.** (1988). Genetic parallelism of protein polymorphism in nature: Ecological test of the neutral theory of molecular evolution. *Biol. J. Linn. Soc.* **35**, 229–245.
- Nevo, E., Zohary, D., Brown, A.H.D., and Haber, M.** (1979). Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum*, in Israel. *Evolution* **33**, 815–833.
- Noma, K., Nakajima, R., Ohtsubo, H., and Ohtsubo, E.** (1997). *RIRE1*, a retrotransposon from wild rice *Oryza australiensis*. *Genes Genet. Syst.* **72**, 131–140.
- Ogihara, Y., Terachi, T., and Sasakuma, T.** (1991). Molecular analysis of the hot spot region related to length mutations in wheat chloroplast DNAs. I. Nucleotide divergence of genes and intergenic spacer regions located in the hot spot region. *Genetics* **129**, 873–884.
- Orgel, L.E., and Crick, F.H.C.** (1980). Selfish DNA: The ultimate parasite. *Nature* **284**, 604–607.
- Panstruga, R., Büschges, R., Piffanelli, P., and Schulze-Lefert, P.** (1998). A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. *Nucleic Acids Res.* **26**, 1056–1062.
- Parket, T., Inbar, O., and Kupiec, M.** (1995). Recombination of *Ty* elements in yeast can be induced by a double-stranded break. *Genetics* **140**, 67–77.
- Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Kumar, A., and Flavell, A.J.** (1996a). The *Ty1-copia* group of retrotrans-

- posons in *Vicia* species: Copy number, sequence heterogeneity and chromosomal localisation. *Mol. Gen. Genet.* **205**, 305–315.
- Pearce, S.R., Pich, U., Harrison, G., Flavell, A.J., Heslop-Harrison, J.S., Schubert, I., and Kumar, A. (1996b). The *Ty1-copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chrom. Res.* **4**, 357–364.
- Pearce, S.R., Harrison, G., Heslop-Harrison, J.S., Flavell, A.J., and Kumar, A. (1997). Characterization and genomic organization of *Ty1-copia* group retrotransposons in rye (*Secale cereale*). *Genome* **40**, 1–9.
- Petrov, D. (1997). Slow but steady: Reduction of genome size through biased mutation. *Plant Cell* **9**, 1900–1901.
- Puchta, H., and Hohn, B. (1996). From centiMorgans to base pairs: Homologous recombination in plants. *Trends Plant Sci.* **1**, 340–348.
- Puchta, H., Kocher, S., and Hohn, B. (1992). Extrachromosomal homologous DNA recombination in plant cells is fast and is not affected by CpG methylation. *Mol. Cell. Biol.* **12**, 3372–3379.
- Ramsay, L., Macaulay, M., Carle, L., Morgante, M., degli Ivanisovich, S., Maestri, E., Powell, W., and Waugh, R. (1999). Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. *Plant J.* **17**, 415–425.
- Royo, J., Nass, N., Matton, D.P., Okamoto, S., Clarke, A.E., and Newbigin, E. (1996). A retrotransposon-like sequence linked to the S-locus of *Nicotiana glauca* is expressed in styles in response to touch. *Mol. Gen. Genet.* **250**, 180–188.
- Saghai Maroof, M.A., Allard, R.W., and Zhang, Q. (1990). Genetic diversity and ecogeographical differentiation among ribosomal DNA alleles in wild and cultivated barley. *Proc. Natl. Acad. Sci. USA* **87**, 8486–8490.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., and Bennetzen, J.L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**, 765–768.
- SanMiguel, P., Gaut, B.S., Tikhoniv, A., Nakajima, Y., and Bennetzen, J.L. (1998). The paleontology of intergene retrotransposons in maize. *Nat. Genet.* **20**, 43–45.
- SAS Institute. (1996). *SAS User's Guide: Statistics*. Version 6.09. (Cary, NC: SAS Institute).
- Schmidt, D., and Graner, A. (1994). Genomic organization and sequence diversity of the long terminal repeat (LTR) of the *BARE-1* retrotransposon family in barley. *Barley Genet. Newsl.* **24**, 24–27.
- Schwarzacher, T., and Leitch, A.R. (1994). Enzymatic treatment of plant material to spread chromosomes for in situ hybridization. In *Methods in Molecular Biology*, Vol. 28, P.G. Isaac, ed (Totowa, NJ: Humana Press), pp. 153–160.
- Sentry, J.W., and Smyth, D.R. (1989). An element with long terminal repeats and its variant arrangements in the genome of *Lilium henryi*. *Mol. Gen. Genet.* **215**, 349–354.
- Suoniemi, A., Anamthawat-Jónsson, K., Arna, T., and Schulman, A.H. (1996a). Retrotransposon *BARE-1* is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. *Plant Mol. Biol.* **30**, 1321–1329.
- Suoniemi, A., Narvanto, A., and Schulman, A. (1996b). The *BARE-1* retrotransposon is transcribed in barley from an LTR promoter active in transient assays. *Plant Mol. Biol.* **31**, 295–306.
- Suoniemi, A., Schmidt, D., and Schulman, A.H. (1997). *BARE-1* insertion site preferences and evolutionary conservation of RNA and cDNA processing sites. *Genetica* **100**, 219–230.
- Suoniemi, A., Tanskanen, J., Pentikäinen, O., Johnson, M.S., and Schulman, A.H. (1998a). The core domain of retrotransposon integrase in *Hordeum*: Predicted structure and evolution. *Mol. Biol. Evol.* **15**, 1135–1144.
- Suoniemi, A., Tanskanen, J., and Schulman, A.H. (1998b). *Gypsy*-like retrotransposons are widespread in the plant kingdom. *Plant J.* **13**, 699–705.
- Svitashev, S., Bryngelsson, T., Vershinin, A., Pedersen, C., Säll, T., and von Bothmer, R. (1994). Phylogenetic analysis of the genus *Hordeum* using repetitive DNA sequences. *Theor. Appl. Genet.* **89**, 801–810.
- Takeda, S., Sugimoto, K., Otsuki, H., and Hirochika, H. (1998). Transcriptional activation of the tobacco retrotransposon *Tto1* by wounding and methyl jasmonate. *Plant Mol. Biol.* **36**, 365–376.
- Thomas, C.A. (1971). The genetic organization of chromosomes. *Annu. Rev. Genet.* **5**, 237–256.
- Wakamiya, I., Price, H.J., Messina, M.G., and Newton, R.J. (1996). Pine genome size diversity and water relations. *Physiol. Plant.* **96**, 13–20.
- Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B.T., and Powell, W. (1997). Genetic distribution of *BARE-1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet.* **253**, 687–694.
- Vernhettes, S., Grandbastien, M.-A., and Casacuberta, J.M. (1997). In vivo characterization of transcriptional regulatory sequences involved in the defence-associated expression of the tobacco retrotransposon *Tnt1*. *Plant Mol. Biol.* **35**, 673–679.
- Vershinin, A.V., Salina, E.A., Solovyov, V.V., and Timofeyeva, L.L. (1990). Genomic organization, evolution, and structural peculiarities of highly repetitive DNA of *Hordeum vulgare*. *Genome* **33**, 441–449.
- von Bothmer, R., Jacobsen, N., Baden, C., Jørgensen, R.B., and Linde-Laursen, I. (1995). *An Ecogeographical Study of the Genus Hordeum*. (Rome: International Plant Genetic Resources).
- Voytas, D.F., and Naylor, G.J.P. (1998). Rapid flux in plant genomes. *Nat. Genet.* **20**, 6–7.
- Voytas, D.F., Cummings, M.P., Konieczny, A.K., Ausubel, F.M., and Rodermeil, S.R. (1992). *copia*-like retrotransposons are ubiquitous among plants. *Proc. Natl. Acad. Sci. USA* **89**, 7124–7128.
- Wessler, S.R. (1996). Turned on by stress. *Plant retrotransposons*. *Curr. Biol.* **6**, 959–961.
- Wessler, S.R., Bureau, T.E., and White, S.E. (1995). LTR-retrotransposons and MITEs: Important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.* **5**, 814–821.
- Zhou, J., Kilian, A., Warner, R.L., and Kleinhofs, A. (1995). Variation of nitrate reductase genes in selected grass species. *Genome* **38**, 919–927.