PNAS RESEARCH ARTICLE | PLANT BIOLOGY

Independent evolution of transposase and TIRs facilitated by recombination between *Mutator* transposons from divergent clades in maize

Charles T. Hunter[a,1], Donald R. McCarty[b], and Karen E. Koch[b]

Edited by Michael Freeling, University of California, Berkeley, CA; received March 31, 2023; accepted May 25, 2023

Nearly all eukaryotes carry DNA transposons of the Robertson's *Mutator* (*Mu*) superfamily, a widespread source of genome instability and genetic variation. Despite their pervasive impact on host genomes, much remains unknown about the evolution of these transposons. Transposase recognition of terminal inverted repeats (TIRs) is thought to drive and constrain coevolution of *MuDR* transposase genes and TIRs. To address the extent of this relationship and its impact, we compared separate phylogenies of TIRs and *MuDR* gene sequences from *Mu* elements in the maize genome. Five major clades were identified. As expected, most *Mu* elements were bound by highly similar TIRs from the same clade (homomorphic type). However, a subset of elements contained dissimilar TIRs derived from divergent clades. These "heteromorphs" typically occurred in multiple copies indicating active transposition in the genome. In addition, analysis of internal sequences showed that exchanges between elements having divergent TIRs produced new *mudra* and *mudrb* gene combinations. In several instances, TIR homomorphs had been regenerated within a heteromorph clade with retention of distinctive internal *MuDR* sequence combinations. Results reveal that recombination between divergent clades facilitates independent evolution of transposase (*mudra*), transposase-binding targets (TIRs), and capacity for insertion (*mudrb*) of active *Mu* elements. This mechanism would be enhanced by the preference of *Mu* insertions for recombination-rich regions near the 5′ ends of genes. We suggest that cycles of recombination give rise to alternating homo- and heteromorph forms that enhance the diversity on which selection for *Mu* fitness can operate.

MULE | TIR | Mu element | transposase | MuDR

Significance

A new contributor to the evolution of DNA-type transposable elements in the *MuDR*-class lies in their capacity to cycle between similar and dissimilar pairs of terminal inverted repeats (TIRs) that serve as targets for their own transposases. The extent of this cycling and the resulting diversity is enhanced by the preferential insertion of *MuDR*-type transposons in high-recombination zones near host genes.

Mutator (*Mu*) transposons of *Zea mays* (maize) comprise the most highly active transposon system thus far described in plants (1, 2) and provide a model for the study of DNA-type transposons in eukaryotes (3). Since their description by Donald Robertson (4), the *Mutator* (*MuDR*) system has led to the identification of a diverse superfamily termed *Mu-like* elements (*MULE*s). These are widespread among eukaryotes, occurring in fungal, plant, and animal genomes (5–7) where they are major contributors to genome evolution (8). More distant relatives of *MULE*s have been identified even in prokaryotic genomes (9, 10). The maize genome includes at least four *MULE* families, including the *MuDR* family, each with distinct TIRs (terminal inverted repeats) and autonomous element compositions (3, 11, 12). Here, we focus specifically on the *MuDR* family of elements in maize due to its accessibility, the presence of separate *mudra* and *mudrb* genes, and our experience using this system for large-scale transposon mutagenesis (13–16). Results may be broadly applicable to *MULE*s throughout the eukaryotes.

The *MuDR* transposon family is characterized by highly conserved TIR sequences between 150 and 250 nucleotides long, positioned at each terminus, and oriented in opposite directions (2, 3). The TIR sequences are shared between autonomous and non-autonomous elements and are critical for transposon functionality since they contain binding sites for the transposase enzyme that catalyzes transposition (2, 17, 18). Between the conserved TIRs, the internal sequences of *Mu* transposons are highly diverse. These regions can carry deletion derivatives of autonomous elements (19) or fragments derived from captured host sequences (20–22). The first *Mu* elements to be identified were named *Mu1* through *Mu19* based on the order of discovery and distinctive internal features (23–32). The abundance and locations of *Mu* elements can vary widely among maize lines, changing with *MuDR* activity and genetic background. Previous examinations of the B73 maize genome revealed 276 apparent *MuDR* family elements (22).

The *Mu* system has been an invaluable source of insertional mutagenesis, with several important characteristics making it particularly suitable for inducing loss-of-function alleles. These characteristics include high mutagenicity, a propensity for transposing into

Author affiliations: [a]Chemistry Research Unit, Center for Medical, Agricultural and Veterinary Entomology, United States Department of Agriculture - Agricultural Research Service, Gainesville, FL 32608; and [b]Horticultural Sciences Department, College of Agricultural and Life Sciences, University of Florida, Gainesville, FL 32611

Author contributions: C.T.H., D.R.M., and K.E.K. designed research; performed research; contributed new reagents/analytic tools; analyzed data; and wrote the paper.

Competing interest statement: The authors declare no conflicts of interest with this study. The use of trade name, commercial product, or corporation in this publication is for the information and convenience of the reader and does not imply an official recommendation, endorsement or approval by the U.S. Department of Agriculture or the Agricultural Research Service for any product or service to the exclusion of others that may be suitable. USDA is an equal opportunity provider and employer.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: Charles.hunter@usda.gov.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2305298120/-/DCSupplemental.

Published July 25, 2023.

genic sequences where loss-of-function mutations are likely, and straightforward detectability using simple molecular tools. Additionally, *MuDR* elements in maize form a two-component system in which selecting for activity of a few autonomous elements can facilitate control of many potentially transposable and nonautonomous elements (2, 14, 33).

The regulation of multiple *Mu* transposons by a single autonomous element is evident in *Mu*-active populations derived from Robertson's Mutator stock, where transposition of these elements is mediated by the *MuDR* transposase (2, 14, 28–30, 34, 35). Internal to the conserved TIRs of *MuDR* are the *mudra* and *mudrb* genes which encode the MURA and MURB proteins. Together, these catalyze transposition of the transposable *Mu* elements in *Mu*-active populations. Both MURA and MURB are required for generation of new germinal insertions (17, 18, 36). The MURA can mediate *Mu* element excision in the absence of MURB, but MURB appears essential for insertion of excised elements (3, 36, 37). Intriguingly, MURB appears to be a distinctive feature of the *MuDR* subfamily of *MULEs*. The *MuDR* TIRs contain promoters for the *mudra* and *mudrb* genes (29, 38). Although the specific mode of action for MURA/MURB has not been elucidated, it is likely similar to other class II *MULE* transposons such as *mutaI* from *Aedes aegypti* (6).

Key features of the *MuDR*-class transposable elements impact their host genome and distinguish them from other transposon systems. First, *MuDR*-class transposons are not preferentially targeted to linked sites (33, 34) such as observed for the *Ac*/*Ds* transposable-element system (39). Although *Mu* elements reportedly show some degree of sequence-dependent insertion (31, 40), they can generally insert into sites throughout the genome. Second, *MuDR*-type elements have a strong insertional bias for recombination-rich, genic sequences (33, 40–43), specifically near the 5′ regions of expressed genes (14, 31, 43, 44). This targeting tendency for *Mu* elements leads to frequent insertions in the 5′-untranslated region (UTR) and upstream regulatory sequences of genes where they can alter expression without disrupting coding sequences (13, 14, 44). *Mu* insertions in or near genes can have diverse effects, including increased or decreased gene expression (22, 45) along with increased rates of recombination and gene conversion (46, 47). These impacts contrast with the highly abundant retrotransposons, which do not preferentially target genic regions and appear to be recombinationally inert (48–50).

Here, we focus on the evolution of the *MuDR*-class transposons. We show that extensive recombination among five divergent clades of *Mu* transposons identified in the maize genome has facilitated semi-independent evolution of TIRs and internal genes for transposase (*mudra*) and capacity for insertion (*mudrb*). Phylogenetic analysis indicates that *Mu* elements evolve through a cycle in which divergent elements recombine to form active transposons with dissimilar TIRs (heteromorphs) followed by conversions that form elements with nearly identical TIRs (homomorphs). Recombination between *Mu* transposons is likely promoted by preferential insertion of *Mu* elements in compact, recombination-rich regions near transcription start sites of maize genes. We suggest that cycles of recombination give rise to alternating heteromorph forms that enhance the diversity on which selection for *Mu* fitness can operate.

Results

_Mu_ Insertions Identified Bioinformatically in the B73 Genome. To explore and capture the diverse sequences of *MuDR*-family transposons in the B73 inbred, its genome (v4) was queried via sequential BLAST analyses using the terminal 150 bp TIR sequence of *MuDR*. Each new TIRidentified this way (481) was used to build a *Mu*-TIR database (Dataset S1) that included genome location, TIR sequence, and target site duplication (TSD) (the 9-bp direct repeat *Mu* elements generate upon insertion). TSDs were used to group TIRs located within 100,000 bases and positioned in opposite orientations together as putative left and right arms of single *Mu* elements. Nonexact TSDs that were recognizably similar were also used to help group TIRs of the same *Mu* element. TIRs that were missing their extreme termini (thus lacking TSDs) were also grouped if they were oriented in opposite directions and positioned within 10,000 bases of one another. This method for mining and informatically joining TIRs allowed 86.4% of the TIRs to be paired into single-element configurations (Table 1). In this way, 210 putative intact *Mu* elements were identified (*SI Appendix*, Table S1). The remaining TIRs 1) terminated at a gap in the genomic sequence, 2) appeared to have lost one recognizable TIR, or 3) occurred as tandem TIRs within a single element. To facilitate referencing of specific elements discussed in this study, each TIR was assigned a unique alphanumeric code to indicate chromosome number and orientation such that 1A-L designates the left (5′) TIR of the first *Mu* element on chromosome 1.

Table 1. The five major phylogenetic groups of TIRs from B73 are similar in size, although within-group sequence diversity and proportion of paired TIRs are variable.

Classification	No. of TIRs	Pairwise identity within group	Identical sites within group	TIRs paired
Group 1	88	80.1%	12.8%	95.5%
Group 2	132	75.0%	9.0%	90.1%
Group 3	90	82.2%	20.5%	96.7%
Group 4	84	72.7%	7.5%	72.6%
Group 5	54	74.2%	14.8%	77.8%
Combined	448	58.4	1.0%	86.4%

Group 4 and group 5 are least conserved and have fewer paired TIRs than other groups, indicating a more ancient origin.

Phylogenetic Relationships among the B73 _Mu_ Transposons. To classify the relatedness of *Mu* transposons, we analyzed the phylogeny of TIR sequences. Five major groups emerged from among the 448 bioinformatically identified TIR sequences of at least 100 bases. Sequences for the TIRs of the named *Mu* elements (*Mu1* through *Mu19*) were also included (Fig. 1). Group 1 included the named elements *Mu1* through *Mu19*. Connecting the left and right arms of individual *Mu* elements revealed that most had highly similar TIRs belonging to the same phylogenetic group, here termed homomorphic *Mu* elements. Not uncommon, however, were instances in which the two arms of a single *Mu* element belonged to phylogenetically distinct groups that we have termed heteromorphic *Mu* elements. Importantly, these heteromorphs have internal sequences and TSDs characteristic of intact, transposable elements. In addition, four independent groups of heteromorphic *Mu* elements all appear to have previously undergone germinally heritable transposition (see clusters of pink and blue arcs connecting the left and right arms of independent insertions in Fig. 1). Although there are several group 1 elements with TIRs phylogenetically distant from each other (including *Mu19*), these sequences are nonetheless similar enough to remain in the same clade and are thus considered homomorphic. Unpaired

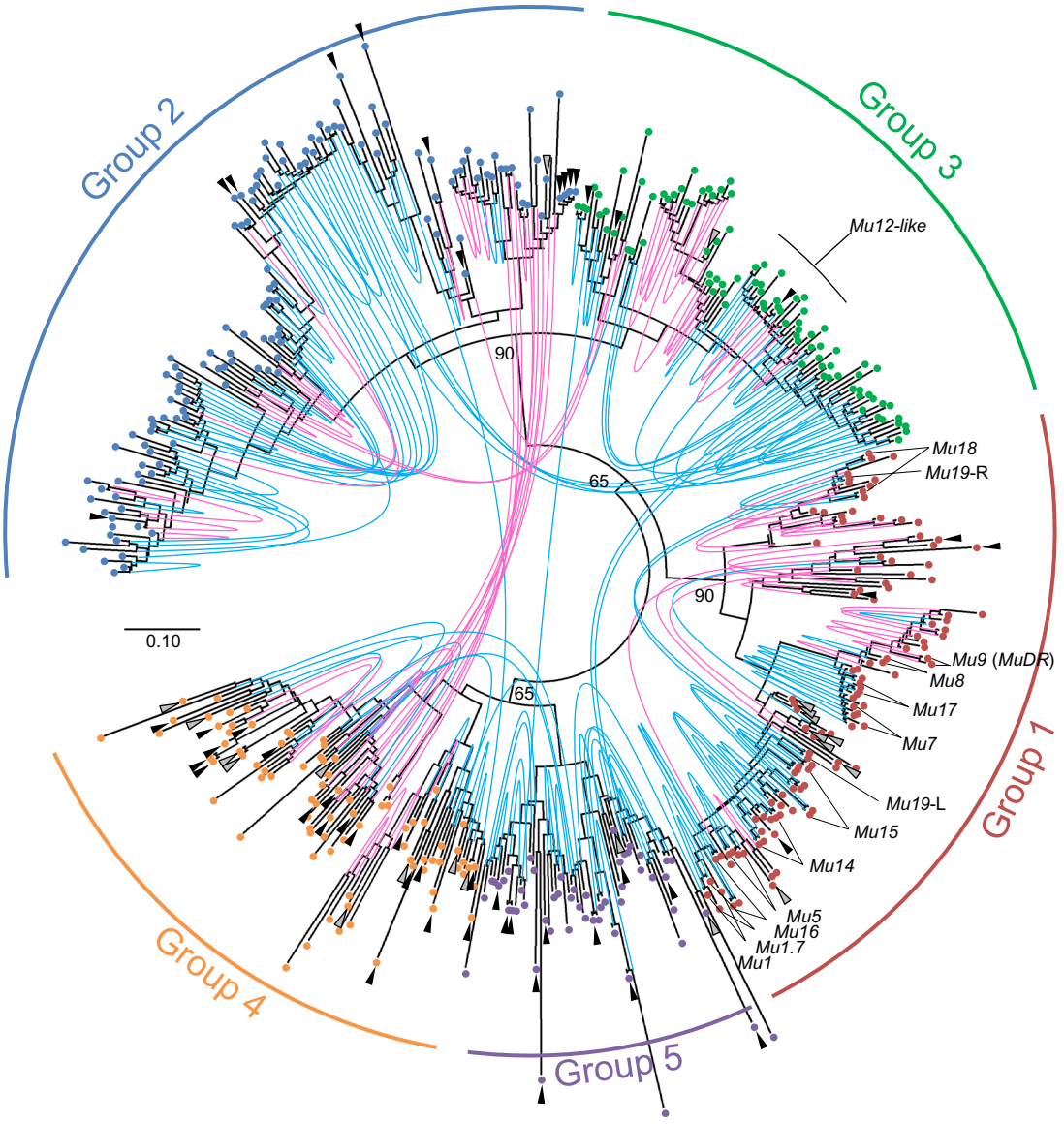

Fig. 1. *Mu*-element TIRs of maize are grouped phylogenetically and shown with overlaid arcs that link the left and right arms of individual transposons. Five distinct clades of TIRs emerge in a neighbor-joining tree generated by ClustalW alignment of named *Mu* elements (*Mu1* through *Mu19*) plus 448 bioinformatically identified TIRs from the B73 genome. Most TIRs of named *Mu* elements classify with group 1, whereas *Mu12* TIRs are more similar to those of group 3. Colored arcs connect the right and left arms of individual elements and are pink if the transposon includes *MuDR-like* (transposase) gene remnants and blue if they do not. Most TIR pairs from individual transposons cluster in the same phylogenetic group and define "homomorphic" elements. In contrast, "heteromorphic" elements have TIRs that lie in phylogenetically divergent groups. Triangles indicate unpaired TIRs that are either SOLO (black) or lack a partner in the phylogenetic tree due to sequence degradation (gray). Relevant bootstrap values are indicated at branches that define clades. Units show the numbers of base substitutions per site.

TIRs were found in all major clades, though at greater frequency in groups 4 and 5. These include SOLOs arising from unpaired TIRs (13.6% of total) and those for which at least one member of a TIR pair was too short (<100 bases) to include in the phylogeny (6.9% of total). Table 1 summarizes the relatedness characteristics for each TIR group.

The TIRs of three previously named *Mu* elements were examined separately because full-length TIR sequences were not available for inclusion in the phylogenetic analysis. The 39 bases of available sequence for the TIRs of *Mu10*, *Mu11*, and *Mu12* (31) were compared to sequences in our database. Analysis of *Mu10* and *Mu11* indicated that both were homomorphs in group 1. The *Mu10*-L differed by only 1 bp from 2P-L, and *Mu10*-R differed by 2 bp from 3F-R. The *Mu11* was similar to *Mu19* in having comparatively distant left and right arms with *Mu11*-L differing from 7H-R by only 1 bp and *Mu11*-R differing from 4L-L by 6 bp. The TIRs of *Mu12* indicated that it was a group 3 homomorph whose identical left and right TIRs exactly matched 5C-R, 5C-L, 5N-R, and 10L-R.

Alignments of TIR sequences belonging to each phylogenetic clade were used to identify consensus sequences for the major groups (Fig. 2). The five TIR consensus sequences were used in BLAST queries to reexamine the B73 genome for TIRs that may have been missed in the initial capture. No additional B73 TIRs were identified. Alignment of the consensus sequences showed highly conserved regions between all five TIR groups, particularly within the most terminal bases and the predicted transposase binding site (3, 17). Pairwise identity was 80.3% when restricted to the predicted MURA binding sites as compared to 65.3% for the entire alignment. Phylogenetic analysis using the aligned consensus sequences showed well-supported separation for the five defined groups with groups 2 and 3 being distinct from

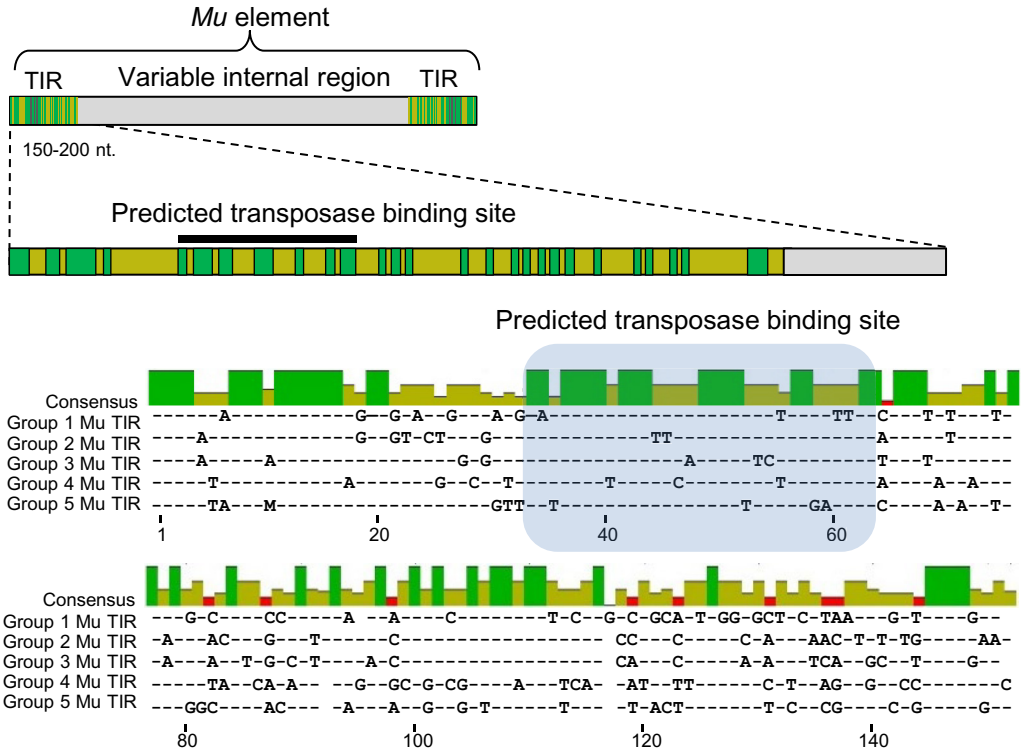

Fig. 2. Diagram of a *Mu* transposable element. The highly conserved TIRs can be used to identify *Mu* elements and categorize them phylogenetically. Alignment of consensus sequences for each of the five *Mu* TIR groups is shown. Dashes show agreement with the overall consensus sequence. The 20 most terminal positions are strongly conserved 83.5% pairwise identity) as are those of the predicted transposase binding site (80.3% pairwise identity) between positions 34 and 64 in all five groups, compared to 65.3% for the entire 150-bp sequence. Bright green indicates perfectly conserved regions, light green designates less conservation, and red denotes variable sites.

groups 1, 4, and 5 (*SI Appendix*, Fig. S1). The relative abundance of TIRs from each clade shows that no one group dominates in the B73 genome.

Analysis of *MuDR*-Derived Internal Sequences. Complete sequences of intact *Mu* elements from B73 (those with both left and right arms) were examined using BLAST queries for similarity to the *MuDR*-encoded genes *mudra* and *mudrb*. Numerous *Mu* elements in B73 showed sequence segments with high similarity to one or both of the *MuDR* genes (pink arc lines in Fig. 1). Indeed, at least 30% of intact *Mu* elements with TIRs in groups 1, 2, 3, and 4 contained fragments of *MuDR*-related genes. No element with a TIR from group 5 contained any identifiable *MuDR* gene remnants (Dataset S2). It would be unexpected to find functionally competent *MuDR* elements in stable inbred genomes, and most of the *mudra* and *mudrb* gene sequences in B73 *Mu* elements included only fragments of the canonical MURA- and MURB-encoding genes (Fig. 3). There were, however, several group 1 homomorphs with apparently intact, full-length coding sequences for MURA or MURB, although no element carried intact copies of both *mudra* and *mudrb*. In an element designated 6T, for example, the *mudra* gene had normal intron–exon borders and, if processed, would lead to translation of a predicted protein 97% identical to MURA. The *mudrb*-like coding sequence of the 6T element was disrupted by a 5-base insertion. In contrast, 12 other group 1 homomorphs carried full-length *mudrb*-like genes without frame-shifting disruptions. Translations of these genes would result in predicted proteins from 78 to 100% identical to MURB.

Phylogeny of *MuDR* Fragments and Evidence for Homomorph Formation from Heteromorphs. To determine the relationship among the *MuDR*-like gene fragments identified in the intact B73 *Mu* elements, those showing homology with the *MuDR* genes for at least half of their length (based on BLAST analyses) were collected for phylogenetic analyses. Homologous regions were extracted and aligned to generate neighbor-joining trees for *mudra* and *mudrb* gene remnants (Fig. 3). The *mudra* and *mudrb* related sequences grouped almost exclusively according to the phylogenetic clade of their TIRs. Interestingly, the *mudra*-like gene remnants of the group 2/4 heteromorphs clustered with those of group 4, while the *mudrb*-like gene remnants clustered with group 2. The same was observed for their adjacent TIRs. These heteromorphic *MuDR*-like elements demonstrate that *mudra* and *mudrb* genes from divergent phylogenetic clades can recombine and retain the capacity to transpose. Evidence for subsequent formation of homomorphs from these heteromorphs is most apparent in two examples (8K and 10A) where TIRs had become homomorphic for group 2, while carrying *mudra-like* sequences from group 2/4 heteromorphs.

TSD Divergence between *Mu* Elements with TIRs of Different Groups. The 9-bp TSD sequences were examined for each of the 210 intact *Mu* elements in the B73 genome to validate their identity as transposons and compare TIR-based clades (Dataset S2). Observed variations could reflect age of insertions since longer time spans after insertions occurred could allow more mutations to accumulate. Alternatively, the dissimilarities could be intrinsic characteristics of certain groups such that *Mu* elements from a given clade may be more prone to generating imperfect TSDs. Around 70% of group 1 homomorph TSDs were identical between the left and right bordering sequences. A similar frequency of perfect TSDs is reported for active *Mu* elements in previous studies (19, 22, 31, 44). Meanwhile, group

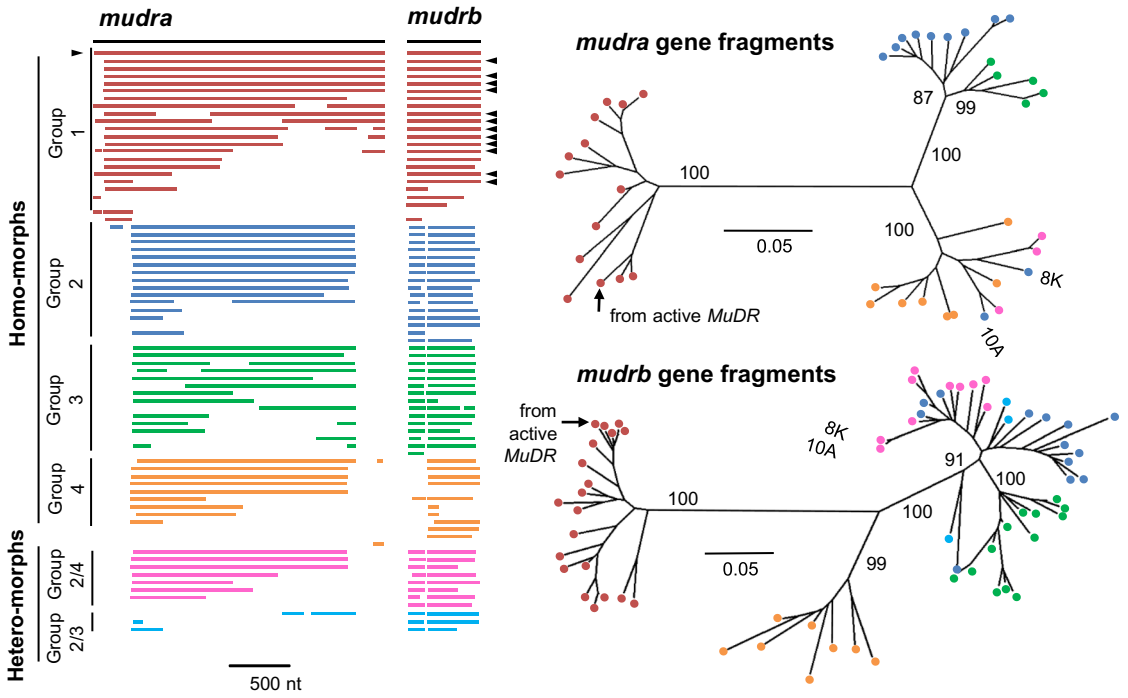

Fig. 3. Alignment and phylogenetic analyses of *mudra*- and *mudrb*-related sequences for 77 *Mu* elements in the B73 genome. Black triangles indicate group 1 homomorphic transposons with full-length *mudra*-like (1 element) or *mudrb*-like (17 elements) genes predicted to encode proteins highly similar to those of *MuDR* (97% for MURA and 78 to 100% for MURB). Homologous regions of *mudra* and *mudrb* sequences were aligned using MUSCLE to build neighbor-joining trees. The *mudr* gene remnants generally clustered by TIR identity for homomorphic elements. However, for group 2/4 heteromorphs, *mudrb* gene remnants clustered with TIRs from group 2, while those of *mudra* clustered with group 4. The group 2 homomorphs (elements 8K and 10A) have *mudra* gene remnants like those of group 4, providing evidence of heteromorph to homomorph conversion. Relevant bootstrap values are indicated. Units in phylogenies indicate the number of base substitutions per site.

5 homomorphs had only 31.6% perfect TSDs, supporting the likelihood of their ancient origin.

Analyses of NAM (Nested Association Mapping) Founders, Teosinte, and Tripsacum Genomes. To seek broader patterns and additional insights into *Mu* evolutionary dynamics, we analyzed the *Mu* elements of the 24 non-B73 founder lines used to generate the NAM population, two teosinte subspecies (*Z. mays subspecies parviglumis* and *mexicana*), and tripsacum (*Tripsacum dactyloides*). These genomes were BLAST queried using consensus sequences for the five B73 TIR groups. The resulting 10,166 TIR sequences were assigned group identities based on similarity to one of the five consensus sequences. Oppositely oriented TIRs within 32,000 bp of one another were designated as pairs. A total of 4,031 putatively intact *Mu* elements were identified in the non-B73 NAM founder genomes plus teosinte and tripsacum (Table 2 and Datasets S3 and S4).

Examination of the *Mu* elements compiled from the non-B73 NAM founder lines showed similar total numbers and proportions of TIR group identity to those found in B73. The occurrence of homo- vs. heteromorphic elements was also similar, with group 2/4 heteromorphs predominating over group 2/3, group 4/5, and group 3/4 in that order. As no group 3/4 heteromorphs had been identified in B73, the 20 identified in the NAM founder lines were examined for validity based on TSDs. Five unique group 3/4 heteromorphs with matching TSDs were verified and found as 1 to 6 copies in 18 of the NAM founder genomes. Results also identified an apparent group 1/2 heteromorph in both the Il14 and P39 inbred lines, but further analysis indicated that this sequence was unlikely to represent a transposed element (TSDs did not match). Still, the correctly oriented TIRs of group 1 and 2 for this element are in close proximity (2,520 bp) and could form a functional group 1/2 heteromorph if a transposase was able to initiate transposition.

The *Mu* elements identified in teosinte also appear in similar abundance and proportions to those in B73 and the NAM founders. As in maize, heteromorphs were abundant, those in groups 2/4 and 2/3 being the most so, followed by 4/5 and 3/5. There was a single apparent group 1/5 heteromorph in *spp. Parviglumis*, the only putative heteromorph in our analyses that included a group 1 component. This element appears to have resulted from transposition, with similar (though not perfect) TSDs and flanking sequences contiguous with homologous regions in the B73 genome. However, examination of the group 5–assigned TIR showed similar relatedness to group 1 and group 5, making its origin uncertain, and thus failing to provide conclusive evidence of a group 1-containing heteromorph.

The *Mu* element content of tripsacum varies in important ways from that of *Zea mays*. First, no *Mu* elements of tripsacum contain TIRs belonging to the three most abundant TIR groups of maize (groups 1, 2, and 3). Meanwhile, groups 4 and 5 are substantially larger in tripsacum than in maize or teosinte. This contrast supports the suggestion that groups 4 and 5 are the most ancient TIR clades in maize and that the *Zea*-specific groups 1, 2, and 3 have emerged relatively recently. Also, group 4/5 heteromorphic elements are relatively abundant in tripsacum, representing 15% of intact *Mu* elements identified there, compared to the 11% in maize and teosinte.

Phylogenetic Analysis of MuDR Fragments from Non-B73 NAM Founders. A nonredundant list of full-length *Mu* elements from the non-B73 NAM founder genomes was generated by condensing those with matching proximity and TSDs in multiple genomes into a single entry. The resulting list of 799 intact *Mu* elements

Table 2. The presence of *Mu* element homomorphs and heteromorphs in maize, teosinte, and tripsacum provides evolutionary context to the *MuDR* family

	Maize		Teosinte		Tripsacum
	Zea mays				
Mu-Clade	B73	NAM	*Z. mays spp parviglumis*	*Z. mays spp mexicana*	*Tripsacum dactyloides*
Homomorphs					
Group-1	46(22)	33 to 59	55	43	0
Group-2	53(16)	29 to 49	42	39	0
Group-3	39(15)	21 to 35	27	40	0
Group-4	28(10)	20 to 36	24	24	35
Group-5	18(0)	10 to 21	18	17	47
Heteromorphs					
Group 2 to 4	10(8)	7 to 13	8	5	0
Group 2 to 3	7(3)	2 to 12	8	5	0
Group 3 to 5	3(0)	0 to 4	1	2	0
Group 4 to 5	4(0)	2 to 6	6	4	14
Group 3 to 4	0	0 to 2	0	0	0
Group 1 to 5	0	0	1	0	0

The occurrence of heteromorphic elements in all genotypes analyzed indicates the frequency with which recombinations between distinct *MuDR* clades can occur, even prior to the differentiation of *Zea*. A single group 1/5 heteromorph was identified in teosinte *parviglumis*, the only group 1 TIR-containing heteromorph in our analysis. The absence of groups 1, 2, and 3 in tripsacum shows that these groups emerged since the split of *Zea* with *Tripsacum* and provides additional support for groups 4 and 5 being the most ancient. Numbers in parentheses indicate elements with remnants of *MuDR* genes for B73 elements.

(Dataset S10) was used to analyze *MuDR* gene fragments by querying them with 1000 of *mudra* (exon 2) and 444 bases of *mudrb* (entirety of exon 2). Results revealed 252 recognizable fragments of *mudra* and 269 of *mudrb*. Phylogenetic analysis of those with greater than 50% coverage of the query sequences led to well-supported trees wherein *MuDR gene* fragments from homomorphic elements grouped almost exclusively together (Fig. 4). As in B73, the group 2/4 heteromorphs had *mudra* sequences similar to those of group 4 homomorphs and *mudrb* fragments that clustered with group 2 homomorphs. Additional evidence for a hetero- to homomorph conversion emerged in a group 4 homomorph that carried a *mudrb* gene remnant most similar to the group 2/4 heteromorphs (phylogenetically grouping with group 2 rather than group 4).

Discussion

Diversity of the *Mu* Family in Maize. A central question addressed by this research is the extent to which evolution of *Mu* transposons is constrained by the relationship between *MuDR* transposases and their binding sites on TIRs since this will define how far either type of sequence can change without coordinate modification of the other. Although diversification of TIR sequences is thought to minimize recognition by host silencing systems, any such alterations must be balanced with conservation of transposase binding sites. Here, we were able to address the broader relationship between *MuDR* specificity and TIR diversity by comparing separate phylogenies constructed from transposons with demonstrated historical activity in the maize genome. Findings indicate cross-mobilization between phylogenetically divergent clades of *MuDR*-class elements and highlight a previously unappreciated level of recombination within and between them. This strategy is consistent with the known preference for *MuDR*-type transposons to insert into recombination-rich zones immediately upstream of genic regions in host DNA.

A key issue is how efficiently a canonical *MuDR* can cross-mobilize *Mu* elements with TIRs of divergent clades. It is unclear whether the *Mu* element transpositions leading to the 262 insertions identified in B73 were dependent on one or more divergent autonomous elements subsequently lost from the B73 genome or whether there may be a low level of background *Mu* activity. Abundant data are available from *Mu* tagging populations such as UniformMu (13, 44, 51) where transpositions typically rely on a single, active *MuDR* to control *Mu* activity. In these instances, where a group 1 *MuDR* mediates movement, the vast majority of new *Mu* insertions result from a few highly conserved group 1 homomorphic elements (*Mu1*, *Mu1.7*, *Mu8*, and *MuDR*). Mobilization of elements outside group 1 is thus rare.

Possible exceptions are two *Mu12*-element insertions belonging to group 3 that were captured as alleles of *glossy8* in an active *Mu* population (31). Evidence for two independent insertions implies that *Mu12* transposition did occur in that population. There are at least two possibilities. 1) The group 3 *Mu12s* could have been mobilized by a group 1 *MuDR* if such events were otherwise limited by chromatin state associated with nonautonomous elements in the genome rather than by transposase specificity per se. If so, then a change in the chromatin status of a *Mu12* element in the genome of *glossy8* may have been responsible. 2) *Mu12* could have been mobilized by a divergent group 3 *MuDR* that was also active in the population.

Even if *MuDR* has low levels of activity on TIRs outside group 1 and can mediate cross-mobilization, a strong preference is still evident for nonautonomous elements with which it shares TIRs. Although intact sequences are lacking for *MuDR* relatives in the reference genome for B73, more complete versions could nonetheless exist in other backgrounds and direct transposition of more divergent TIRs. It is also possible that *MuDR* transposases may function in multimeric complexes as observed for other type II transposable elements such as the *hAT* superfamily (52, 53). If so, then subunits produced from divergent *MuDRs* could theoretically act together as heterodimers or multimers in the mobilization of heteromorphic elements.

Heteromorphic *Mu* Elements. Transposition is evident for heteromorphic *Mu* elements based on the presence of unique TSDs and multiple copies of highly similar heteromorphic elements (Fig. 1). Since transposition is thought to depend on recognition

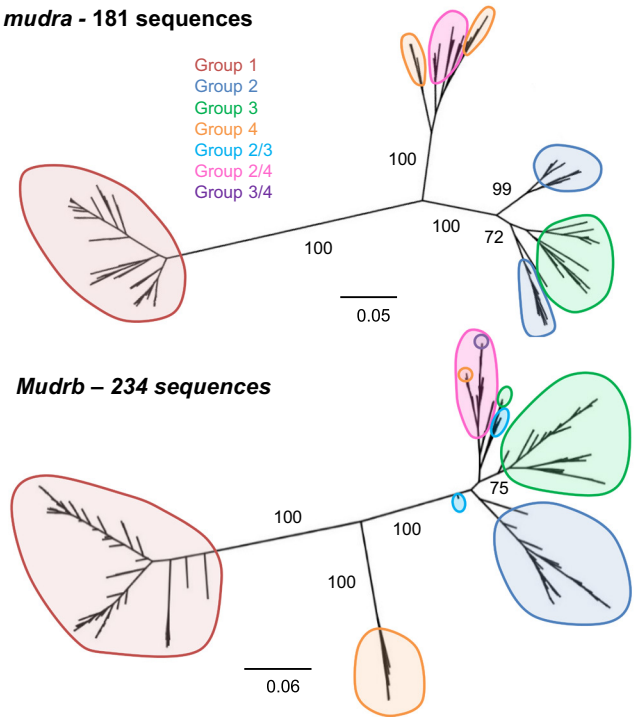

Fig. 4. Extended phylogenetic analysis of *mudra*- and *mudrb*-related sequences in genomes of the 24 non-B73 NAM founder lines. Relationships were generally similar to those in B73 (Figure 3) with *mudra*- and *mudrb*-homologous regions clustering with TIR group identity (indicated by shaded areas). Analyses also supported the B73 observation that *mudra* gene remnants in group 2/4 heteromorphs clustered with group 4 homomorphs, while *mudrb* gene remnants clustered with group 2 homomorphs. Analysis of *mudrb*-like sequences in NAM founder lines also revealed combinations that included a *mudrb* remnant from a group 4 homomorph that clustered with group 2/4 heteromorphs (circled with orange inside the pink shaded region of the *mudrb* tree). This observation provided another possible example of a heteromorph to homomorph conversion. Relevant bootstrap values are indicated. Units show the number of base substitutions per site.

of specific binding sites in the TIR by the transposase enzyme (2, 17, 39), any element with divergent TIRs would likely constitute a weaker target. However, it is unclear what degree of homology between the left and right TIRs is important for transposition of an element. The MURA transposase encoded by *mudra* seems to have strong specificity for the group 1 TIR sequences, a clade of *MuDR*-class transposons with imperfect, but moderately high similarity between the right and left TIRs While we are aware of no instances where a true group-spanning heteromorph has been active in modern lines, we can look to the *glossy8* alleles in which five mutants arose from insertions of *Mu11* elements with dissimilar TIRs (31). The TIRs of *Mu11* are almost identical to those of *Mu19*, an element lacking evidence of transposability in *Mu*-active populations (32). The *Mu19* and *Mu11* elements have TIRs that cluster in group 1 but are nonetheless phylogenetically distant from each other. The apparent activity of *Mu11* provides the best current evidence for a transposon with somewhat divergent arms causing new insertions and gene disruption in a *Mu*-active population.

Transposable heteromorphic elements require an adjustment in our thinking regarding the age of *Mu* transposons because TIR similarity does not serve as a proxy for time passage after insertion. In the B73 genome, TSDs are conserved to a similar degree for homo- and heteromorphic elements of the same overall TIR clade (Dataset S2). Homomorphic elements within group 1 or group 2 tended to have the most well-conserved TSDs, suggesting that they were the most recently active. In contrast, relatively few perfect TSDs were observed for elements within group 4 or group 5, regardless of whether these were homomorphic or heteromorphic, consistent with a more ancient origin for these insertions.

There are multiple possible mechanisms by which heteromorphic *Mu* elements could form (Fig. 5*A*). One scenario is that a transposase could bind to distal arms of two, physically close, but unrelated *Mu* elements and create a hybrid element. A similar possibility is that one transposon could insert into an unrelated element and generate a new hybrid element if a transposase bound one TIR from each of the original elements. In both instances, internal TIR copies would be present in the new hybrid element, although they could be quickly lost as internal deletions common in *Mu* elements (3). A third possibility is that independent insertions in the same gene (which occur frequently) are present in a heterozygous state and lead to heteromorphic elements through recombination or gene conversion. The frequent clustering of independent *Mu* insertions at 5′ ends of genes would facilitate such events. Finally, heteromorphic elements could arise if one TIR (or both) accumulated mutations independent of the other, while maintaining transposase recognition. However, the absence of such intermediates in our analyses makes this scenario seem unlikely. Each of these possibilities necessitates that a transposase recognizes and binds both of the divergent TIRs of the newly formed heteromorph.

A striking feature of *Mu* elements is that the left and right TIRs of subfamily members typically have greater similarity to each other than with TIRs of the autonomous *MuDR*. Indeed, the left and right TIRs of group 1 *MuDR* are nearly identical, differing by only 1 base, whereas they share only 85% identity with TIRs of the highly active nonautonomous *Mu1* transposons. This suggests that near identity of the left and right TIRs is maintained by periodic formation of homomorphs and not predominantly by selection for transposase recognition of TIRs. Regeneration of symmetry appears to occur regularly among *Mu* elements in the B73 genome. Each phylogenetic group contains examples of homomorphic elements with highly similar TIRs that have fewer than five mismatches (Dataset S1). One group 3 element (10P in our analysis) has identical TIRs that extend over 1,100 bases into the center of the transposon. Elements with highly similar left and right TIRs may represent relatively recent TIRduplication events, as the left and right arms have not accumulated mutations since formation of their homomorphic symmetry. Fixation of TIR variants through formation of homomorphs is likely an important mechanism driving coevolution of transposase and binding sites in the TIR.

A mechanism for regaining TIR symmetry need not be common but should allow a single TIR sequence to be duplicated such that the resulting *Mu* element has identical left and right arms oriented in opposite directions (Fig. 5*B*). One way this could occur is through gene conversion via homologous recombination of double-stranded breaks (DSB) resulting from excision of a single TIR, wherein the unexcised TIR would serve as a template during

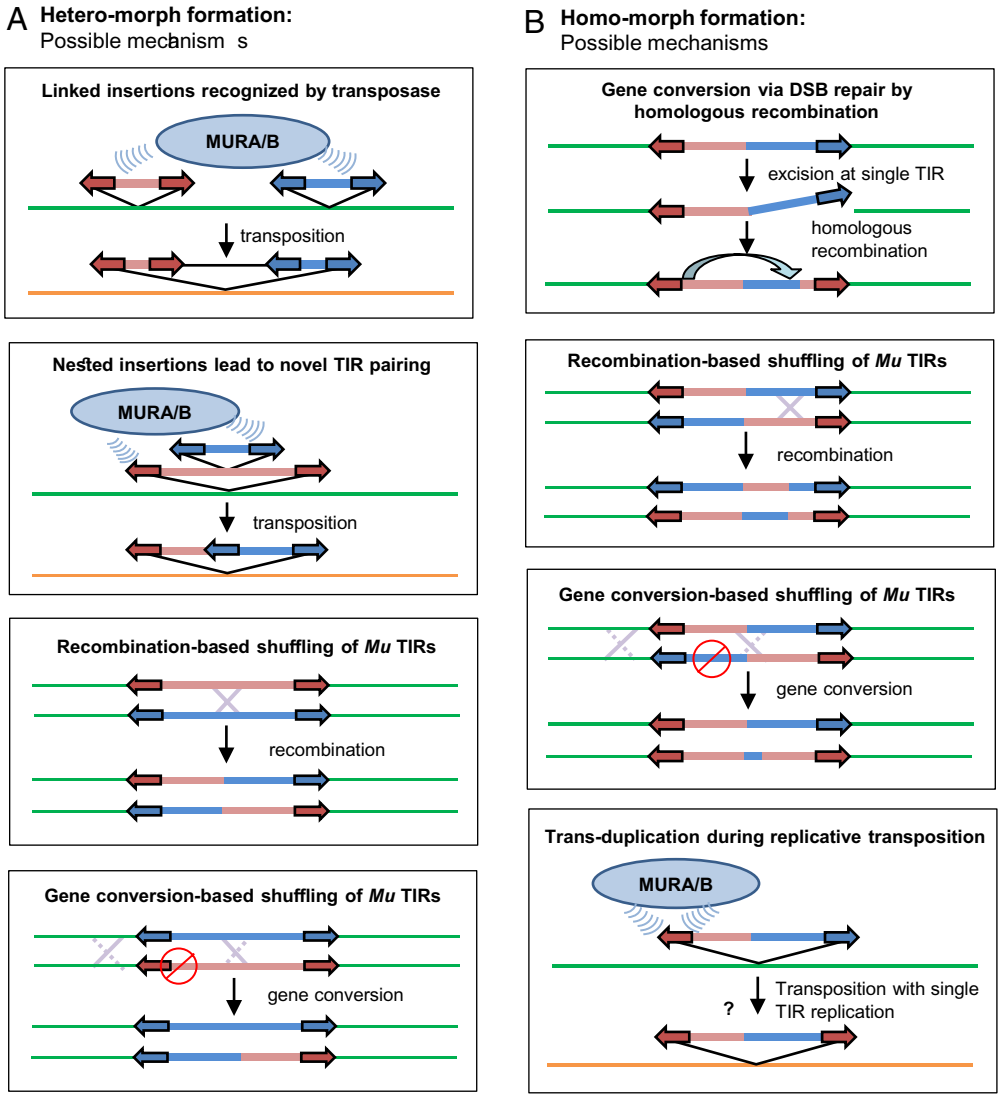

Fig. 5. Possible mechanisms of formation for (*A*) heteromorphic and (*B*) homomorphic *Mu* elements. Heteromorphs could form during transposition of closely linked or nested *Mu* elements with different TIRs in instances where dissimilar TIRs are bound by MURA/MURB transposases or via recombination or gene conversion between distinct insertions in allelic locations on different chromosomes. Heteromorphs could be converted to homomorphs via similar mechanisms, including recombination or gene conversion between allelic insertions, gene conversion via homologous recombination after excision of a single TIR, and possibly during replicative transposition.

repair and thus form a homomorphic element. This would be similar to gene conversions observed between cis-linked inverted repeats in *Arabidopsis* following DSBs induced by the HO endonuclease (54). This mechanism could account for the long and nearly perfect TIRs that extend far into the center of some *Mu* elements such as 10P. Homomorphs could also form from heteromorphs in scenarios similar to that suggested for allelic, heterozygous insertions that generate heteromorphs via recombination, gene conversion, or possibly via an unknown mechanism involving transduplication of a single TIR during transposition.

A potential evolutionary mechanism for the *MuDR* class of transposons is supported by phylogenetic evidence for cycling between homo- and heteromorphic elements. If a newly formed heteromorphic element can transpose, despite having divergent TIRs, then it can also form a homomorph with capacity to initiate a new clade. Such sources of variability among TIR sequences could also facilitate diversification of transposase affinities for particular TIRs, further aiding the delineation of distinct groups of autonomous elements. A similar process may have contributed to advent of the new Zea-specific clades, and the group 2/4 heteromorphs were positioned to do so. This clade is distinctive among those new to Zea because it carries ancestral genes (*mudra*'s from group 4). The overall extent of interconversion and cycling would be enhanced by the dense targeting of *MuDR*-class insertions to regions near the 5′ ends of genes. In this recombination-rich region of the host genome, pairing of nearby independent allelic insertions could promote recombination at meiosis. A high recombination frequency in these locales thus favors diversification of transposon sequences in addition to enabling highly expressed genes to be parasitized without disrupting their function (44).

MuDR-Related Sequences in Maize Genomes. The presence of *MuDR*-like sequences in elements without group 1 TIRs indicates a common ancestry of internal sequences that predate emergence of *MuDR* in its current form. The different TIR-defined groups could thus house their own active *MuDR*-like elements. These could either function exclusively on elements with similar TIRs (as *MuDR* tends to do with group 1 elements) or have broader specificity across TIR groups.

The possibility that additional elements in B73 might encode functional transposases is indicated by the apparently intact *mudra*-like gene of the 6T element and the 12 intact *mudrb*-like genes. These sequences show high degrees of predicted amino acid sequence similarity with the *MuDR*-encoded proteins (98% for the MURA of 6T and up to 100% for the MURBs) suggesting recent common ancestry. Although no single element contains both intact *mudra*- and *mudrb*-like genes, the possibility exists that the two sequences need not be encoded on the same element to have some level of efficacy. Either way, the high degree of homology in genic (*mudra* and *mudrb*) vs. intergenic sequences of *MuDR* derivatives like 6T (~50%) indicates a selective pressure for retention of coding integrity. Although the 6T sequence is probably a nonfunctional h*MuDR* at present (as described in ref. 19), the element is likely to have retained activity for at least some period after its divergence from *MuDR*. The same is true for the many elements outside of group 1 with *MuDR*-like sequences.

Analysis of *MuDR*-like remnants in the B73 genome (Fig. 3) and that of other NAM founders (Fig. 4) provides additional insight into mechanisms underlying formation of hetero- and homomorphs. The phylogenetic clustering of *mudra*-like fragments in group 4 and *mudrb*-like fragments in group 2 could have arisen from a cross-over or gene conversion between nearby elements as outlined in Fig. 5. The emerging 2/4 heteromorph retained transposability and proliferated across the maize genome as an active element. Formation of homomorphs from such heteromorphs is supported by phylogenetic clustering of *mudra* remnants from group 2 homomorphs (the 8K and 10A elements) with those of group 4. These homomorphs likely arose from a group 2/4 heteromorph (such as described above) being converted to a homomorph but retaining internal sequences (like *mudra*) originating from group 4. Similar observations were evident among elements identified in the non-B73 NAM founder lines (Fig. 4). These examples demonstrate the propensity of *Mu* elements to exchange TIRs and internal genes between distinct populations of *Mu* elements and provide a novel source of variation upon which selection for *Mu* viability can act.

Additional Considerations for the History of *Mu* Transposons. Intriguing commonalities are evident among the *Mu* element clades classified by the TIR group. Each clade is a similar size and has a comparable diversity of elements. Four of the five groups include *MuDR*-related sequences. Also, all clades other than group 1 are joined by heteromorphic elements. Finally, each of these clades includes strongly homomorphic elements. These collective features are consistent with shared evolutionary paths that involved similar levels of historical transposition, conversion, and generation of symmetrical TIRs. In this respect, the especially high activity of *MuDR* and nonautonomous group 1 elements observed in Robertson's Mutator stocks may be an exception rather than the norm for *MuDR* family elements and *MULE*s in general. If other *Mu* clades have undergone bursts of activity similar to those in *MuDR*-active lines, those elements are no longer evident in the landscape of NAM founders, teosinte, or tripsacum.

Methods

Bioinformatic Identification of *Mu* Elements. The B73 reference genome (version 4) was mined for *Mu* insertions by repeated BLAST analyses that began with a single *Mu* TIR from *MuDR*. Individual returns from each BLAST output (cutoff at e = 10^−6) were annotated and deposited in a database that included TIR sequence and chromosome position. This process was repeated, using more divergent *Mu* TIR sequences, until no additional unique transposons were identified. Genomic sequences surrounding each *Mu* were then individually examined to extract 9-bp target-site duplications. Once individual TIRs were paired, full-length sequences were manually extracted and compiled. The resulting dataset was screened for *mudra*- and *mudrb*-related sequences via local BLAST analyses. The NAM founders, teosinte subspecies, and tripsacum genomes were searched in the same way using consensus sequences from each of the 5 groups as queries. Identified TIRs were classified based on which query sequence yielded the highest BLAST score.

Phylogenetic Analysis of TIRs. Bioinformatically identified TIRs were aligned using Clustal W (55) in the Geneious8 (geneious.com) software interface (Dataset S5) with the IUB cost matrix, a gap open cost of 15, and a cap extend cost of 6.66. The TIRs included 448 sequences with at least 100 bp from previously unnamed insertions, together with 150 bases from the left and right termini of named *Mu* elements (*Mu1*, *Mu1.7*, *Mu3*, *Mu4*, *Mu5*, *Mu7*, *Mu8*, *Mu9*, *Mu14*, *Mu15*, *Mu16*, *Mu17*, *Mu18*, and *Mu19*). Less than 100 bp were available for the remaining 33 TIRs, so these were omitted to avoid nonoverlapping sequences. A neighbor-joining tree was created with MEGAX (56) using the pairwise deletion option, the Maximum Composite Likelihood method (57), and 1,000 bootstrap replications. The phylogenetic tree was manually annotated to show the connectedness of paired TIRs, color-coded group identities, and the presence of unpaired TIRs.

***MuDR* Gene Remnant Analyses.** Full-length sequences for the 210 intact B73 *Mu* elements were assembled in a BLAST database. Coding sequences for *mudra* and *mudrb* were used in queries of this database to determine which *Mu* elements contained *mudra*- and *mudrb*-like sequence (Datasets S6 and S7). Those with homology to the complete *mudra* or *mudrb* genes were examined for introns, and putative coding sequences were extracted and translated to determine the capacity to encode potentially functional proteins. Elements with greater than 50% of *mudra* or *mudrb* coverage were used to generate phylogenies for *mudra*- and *mudrb*-like genes. The regions homologous to *mudra* and *mudrb* for each element were extracted and aligned using MUSCLE (58) within the Geneious8 software interface. The MUSCLE alignments (Datasets S8 and S9) were done using UPGMB clustering, with kmer4_6 distance measure for iteration 1, and pctid_kimura for subsequent iterations. Clustal W was also used for the sequence weighting scheme, with free end gaps, an anchor spacing of 32, a minimum length of 24, a minimum best of 90, a gap open score of −1, a margin of 5, a minimum smoothed of 90, a window size of 5, and a multiplier of 1.2. The same approach was used to construct alignments of *mudra* and *mudrb* gene remnants from the NAM founder genomes except the BLAST queries were limited to the most highly conserved 1,000 bases of exon 2 from *mudra* and the entirety of exon 2 from *mudrb*. The MUSCLE alignments (Datasets S11 and S12) were done using the same parameters employed for the B73 alignments. Neighbor-joining trees were created with MEGAX using the parameters described for the TIR tree.

Data, Materials, and Software Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. This work was funded by USDA, ARS, SEA-CMAVE research project 6036-21000-011-00D to C.T.H., and NSF-PGRP project 1748105 to D.R.M. and K.E.K.

1. M. Alleman, M. Freeling, The Mu transposable elements of maize: Evidence for transposition and copy number regulation during development. *Genetics* **112**, 107–119 (1986).
2. D. Lisch, Mutator transposons. *Trends Plant Sci.* **7**, 498–504 (2002).
3. D. Lisch, Mutator and MULE transposons. *Microbiol. Spectr.* **3**, Mdna3-0032-2014 (2015).
4. D. S. Robertson, Characterization of a mutator system in maize. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **51**, 21–28 (1978).
5. F. Chalvet *et al.*, Hop, an active Mutator-like element in the genome of the fungus Fusarium oxysporum. *Mol. Biol. Evol.* **20**, 1362–1375 (2003).
6. K. Liu, S. R. Wessler, Functional characterization of the active Mutator-like transposable element, Muta1 from the mosquito *Aedes aegypti*. *Mob. DNA* **8**, 1 (2017).
7. M. Dupeyron *et al.*, Evolution of Mutator transposable elements across eukaryotic diversity. *Mob. DNA* **10**, 12 (2019).
8. M. Sahebi *et al.*, Contribution of transposable elements in the plant's genome. *Gene* **665**, 155–166 (2018).
9. J. A. Eisen, M. I. Benito, V. Walbot, Sequence similarity of putative transposases links the maize Mutator autonomous element and a group of bacterial insertion sequences. *Nucleic Acids Res.* **22**, 2634–2636 (1994).
10. R. Guérillot *et al.*, The diversity of prokaryotic DDE transposases of the mutator superfamily, insertion specificity, and association with conjugation machineries. *Genome Biol. Evol.* **6**, 260–272 (2014).

11. Z. Xu *et al.*, Jittery, a Mutator distant relative with a paradoxical mobile behavior: Excision without reinsertion. *Plant Cell* **16**, 1105–1114 (2004).
12. Y. Li, L. Harris, H. K. Dooner, TED, an autonomous and rare maize transposon of the mutator superfamily with a high gametophytic excision frequency. *Plant Cell* **25**, 3251–3265 (2013).
13. A. M. Settles *et al.*, Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. *BMC Genomics* **8**, 116 (2007).
14. D. R. McCarty *et al.*, Mu-seq: Sequence-based mapping and identification of transposon induced mutations. *PLoS One* **8**, e77172 (2013).
15. C. T. Hunter *et al.*, Phenotype to genotype using forward-genetic Mu-seq for identification and functional classification of maize mutants. *Front. Plant Sci.* **4**, 545 (2014).
16. C. Marcon *et al.*, BonnMu: A sequence-indexed resource of transposon-induced maize mutations for functional genomics studies. *Plant Physiol.* **184**, 620–631 (2020).
17. M. I. Benito, V. Walbot, Characterization of the maize Mutator transposable element MURA transposase as a DNA-binding protein. *Mol. Cell Biol.* **17**, 5165–5175 (1997).
18. M. N. Raizada, V. Walbot, The late developmental pattern of Mu transposon excision is conferred by a cauliflower mosaic virus 35S -driven MURA cDNA in transgenic maize. *Plant Cell* **12**, 5–21 (2000).
19. G. N. Rudenko, V. Walbot, Expression and post-transcriptional regulation of maize transposable element MuDR and its derivatives. *Plant Cell* **13**, 553–570 (2001).
20. J. L. Bennetzen, P. S. Springer, The generation of Mutator transposable element subfamilies in maize. *Theor. Appl. Genet.* **87**, 657–667 (1994).
21. N. Jiang *et al.*, Pack-MULE transposable elements mediate gene evolution in plants. *Nature* **431**, 569–573 (2004).
22. N. Jiang *et al.*, Pack-Mutator-like transposable elements (Pack-MULEs) induce directional modification of genes through biased insertion and DNA acquisition. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1537–1542 (2011).
23. J. L. Bennetzen, Transposable element Mu1 is found in multiple copies only in Robertson's Mutator maize lines. *J. Mol. Appl. Genet.* **2**, 519–524 (1984).
24. L. P. Taylor, V. Walbot, Isolation and characterization of a 1.7-kb transposable element from a mutator line of maize. *Genetics* **117**, 297–307 (1987).
25. L. E. Talbert, G. I. Patterson, V. L. Chandler, Mu transposable elements are structurally diverse and distributed throughout the genus Zea. *J. Mol. Evol.* **29**, 28–39 (1989).
26. P. S. Schnable, P. A. Peterson, H. Saedler, The bz-rcy allele of the Cy transposable element system of Zea mays contains a Mu-like element insertion. *Mol. Gen. Genet.* **217**, 459–463 (1989).
27. D. Fleenor *et al.*, Nucleotide sequence of the maize Mutator element, Mu8. *Nucleic Acids Res.* **18**, 6725 (1990).
28. P. Chomet *et al.*, Identification of a regulatory transposon that controls the Mutator transposable element system in maize. *Genetics* **129**, 261–270 (1991).
29. R. J. Hershberger, C. A. Warren, V. Walbot, Mutator activity in maize correlates with the presence and expression of the Mu transposable element Mu9. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10198–10202 (1991).
30. M. M. Qin, D. S. Robertson, A. H. Ellingboe, Cloning of the Mutator transposable element MuA2, a putative regulator of somatic mutability of the a1-Mum2 allele in maize. *Genetics* **129**, 845–854 (1991).
31. C. R. Dietrich *et al.*, Maize Mu transposons are targeted to the 5′ untranslated region of the gl8 gene and sequences flanking Mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome. *Genetics* **160**, 697–716 (2002).
32. B. C. Tan *et al.*, Identification of an active new mutator transposable element in maize. *G3 (Bethesda)* **1**, 293–302 (2011).
33. A. M. Settles, S. Latshaw, D. R. McCarty, Molecular analysis of high-copy insertion sites in maize. *Nucleic Acids Res.* **32**, e54 (2004).
34. D. Lisch, P. Chomet, M. Freeling, Genetic characterization of the Mutator system in maize: Behavior and regulation of Mu transposons in a minimal line. *Genetics* **139**, 1777–1796 (1995).
35. D. Burgess *et al.*, Silencing of Mutator elements in maize involves distinct populations of small RNAs and distinct patterns of DNA methylation. *Genetics* **215**, 379–391 (2020).
36. D. Lisch *et al.*, Functional analysis of deletion derivatives of the maize transposon MuDR delineates roles for the MURA and MURB proteins. *Genetics* **151**, 331–341 (1999).
37. M. N. Raizada, G. L. Nan, V. Walbot, Somatic and germinal mobility of the RescueMu transposon in transgenic maize. *Plant Cell* **13**, 1587–1608 (2001).
38. M. N. Raizada, M. I. Benito, V. Walbot, The MuDR transposon terminal inverted repeat contains a complex plant promoter directing distinct somatic and germinal programs. *Plant J.* **25**, 79–91 (2001).
39. T. P. Brutnell, Transposon tagging in maize. *Funct. Integr. Genomics* **2**, 4–12 (2002).
40. A. D. Cresse *et al.*, Mu1-related transposable elements of maize preferentially insert into low copy number DNA. *Genetics* **140**, 315–324 (1995).
41. J. Fernandes *et al.*, Genome-wide mutagenesis of Zea mays L. using RescueMu transposons. *Genome Biol.* **5**, R82 (2004).
42. E. Vollbrecht *et al.*, Genome-wide distribution of transposed dissociation elements in maize. *Plant Cell* **22**, 1667–1685 (2010).
43. S. Liu *et al.*, Mu transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet.* **5**, e1000733 (2009).
44. X. Zhang *et al.*, Transposable elements employ distinct integration strategies with respect to transcriptional landscapes in eukaryotic genomes. *Nucleic Acids Res.* **48**, 6685–6698 (2020).
45. M. L. Robbins *et al.*, A Mutator transposon insertion is associated with ectopic expression of a tandemly repeated multicopy Myb gene pericarp color1 of maize. *Genetics* **178**, 1859–1874 (2008).
46. B. Lowe, J. Mathern, S. Hake, Active Mutator elements suppress the knotted phenotype and increase recombination at the Kn1-O tandem duplication. *Genetics* **132**, 813–822 (1992).
47. J. Mathern, S. Hake, Mu element-generated gene conversions in maize attenuate the dominant knotted phenotype. *Genetics* **147**, 305–314 (1997).
48. H. Fu, Z. Zheng, H. K. Dooner, Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1082–1087 (2002).
49. H. Yao *et al.*, Molecular characterization of meiotic recombination across the 140-kb multigenic a1-sh2 interval of maize. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6157–6162 (2002).
50. H. K. Dooner, L. He, Maize genome structure variation: Interplay between retrotransposon polymorphisms and genic recombination. *Plant Cell* **20**, 249–258 (2008).
51. D. R. McCarty *et al.*, Steady-state transposon mutagenesis in inbred maize. *Plant J.* **44**, 52–61 (2005).
52. E. Rubin, G. Lithwick, A. A. Levy, Structure and evolution of the hAT transposon superfamily. *Genetics* **158**, 949–957 (2001).
53. A. B. Hickman *et al.*, Structural basis of hAT transposon end recognition by Hermes, an octameric DNA transposase from Musca domestica. *Cell* **158**, 353–367 (2014).
54. M. Chiurazzi *et al.*, Enhancement of somatic intrachromosomal homologous recombination in Arabidopsis by the HO endonuclease. *Plant Cell* **8**, 2057–2066 (1996).
55. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
56. S. Kumar *et al.*, MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
57. K. Tamura, M. Nei, S. Kumar, Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11030–11035 (2004).
58. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).