



Extensively Current Activity of Transposable Elements in Natural Rice Accessions Revealed by Singleton Insertions

Zhen Liu^{1,2}, Han Zhao³, Yan Yan², Ming-Xiao Wei², Yun-Chao Zheng², Er-Kui Yue², Mohammad Shah Alam², Kwesi Odel Smartt², Ming-Hua Duan^{4,5} and Jian-Hong Xu^{1,2*}

¹ Hainan Institute, Zhejiang University, Sanya, China, ² Zhejiang Key Laboratory of Crop Germplasm, Institute of Crop Science, Zhejiang University, Hangzhou, China, ³ Jiangsu Provincial Key Laboratory of Agrobiolgy, Institute of Biotechnology, Jiangsu Academy of Agricultural Sciences, Nanjing, China, ⁴ Zhejiang Zhengjingyuan Pharmacy Chain Co., Ltd., Hangzhou, China, ⁵ Hangzhou Zhengcaiyan Pharmaceutical Co., Ltd., Hangzhou, China

OPEN ACCESS

Edited by:

Chengzhi Liang,
Institute of Genetics
and Developmental Biology, Chinese
Academy of Sciences (CAS), China

Reviewed by:

Khalil Kashkush,
Ben-Gurion University of the Negev,
Israel
Zhou Mingbing,
Zhejiang Agriculture and Forestry
University, China

*Correspondence:

Jian-Hong Xu
jhxu@zju.edu.cn

Specialty section:

This article was submitted to
Plant Bioinformatics,
a section of the journal
Frontiers in Plant Science

Received: 22 July 2021

Accepted: 08 September 2021

Published: 28 September 2021

Citation:

Liu Z, Zhao H, Yan Y, Wei M-X,
Zheng Y-C, Yue E-K, Alam MS,
Smartt KO, Duan M-H and Xu J-H
(2021) Extensively Current Activity
of Transposable Elements in Natural
Rice Accessions Revealed by
Singleton Insertions.
Front. Plant Sci. 12:745526.
doi: 10.3389/fpls.2021.745526

Active transposable elements (TEs) have drawn more attention as they continue to create new insertions and contribute to genetic diversity of the genome. However, only a few have been discovered in rice up to now, and their activities are mostly induced by artificial treatments (e.g., tissue culture, hybridization etc.) rather than under normal growth conditions. To systematically survey the current activity of TEs in natural rice accessions and identify rice accessions carrying highly active TEs, the transposon insertion polymorphisms (TIPs) profile was used to identify singleton insertions, which were unique to a single accession and represented the new insertion of TEs in the genome. As a result, 10,924 high-confidence singletons from 251 TE families were obtained, covering all investigated TE types. The number of singletons varied substantially among different superfamilies/families, perhaps reflecting distinct current activity. Particularly, eight TE families maintained potentially higher activity in 3,000 natural rice accessions. Sixty percent of rice accessions were detected to contain singletons, indicating the extensive activity of TEs in natural rice accessions. Thirty-five TE families exhibited potentially high activity in at least one rice accession, and the majority of them showed variable activity among different rice groups/subgroups. These naturally active TEs would be ideal candidates for elucidating the molecular mechanisms underlying the transposition and activation of TEs, as well as investigating the interactions between TEs and the host genome.

Keywords: transposable element (TE), transposition activity, singleton insertion, unique insertion site, rice (*Oryza sativa*)

INTRODUCTION

Transposable elements (TEs) or transposons are mobile DNA sequences that can jump from one location to another within the genome. Since the discovery of the first transposon in maize more than 70 years ago (McClintock, 1950), TEs have been shown to be ubiquitous in almost all eukaryotic species and frequently constitute large fractions of plant and animal genomes

(Huang et al., 2012; Vitte et al., 2014). TEs are the major contributor to the variation in genome size between related species. The burst of three LTR-retrotransposon families results in massive genome expansion, which made the genome size of Australia wild rice (*Oryza australiensis*) twofold larger than that of Asian cultivated rice (*Oryza sativa*) (Piegu et al., 2006). As an important source of genetic variations, TEs have also been well documented to contribute to gene and genome evolution in a variety of ways, including insertion mutagenesis, chromosome rearrangement, and epigenetic regulation (Lisch, 2013; Bennetzen and Wang, 2014; Chuong et al., 2017; Song and Cao, 2017). Therefore, the study of TEs would give us a better understanding of genome evolution and genetic diversity.

Transposable elements can be divided into two classes according to their DNA or RNA intermediates (Wicker et al., 2007). Class I element or retrotransposon is mediated by RNA intermediate during reverse transcription process *via* the “copy and paste” mode. Class II element or DNA transposon is directly excised from the original position and integrates into new genomic location *via* the “cut and paste” mode. Both TE classes contain autonomous and non-autonomous elements. The autonomous one has the ability to encode various enzymes required for (retro)transposition, which is responsible for the movement of themselves and the corresponding non-autonomous counterparts.

To date, a large number of TEs have been identified from eukaryotic genomes, but most of them have already lost mobility, because TEs accumulated lots of mutations or were partially deleted by combinations (Ma et al., 2004). On the other hand, the host genomes have evolved multiple defense mechanisms including DNA methylation, chromatin modification, and RNA interference to efficiently silence TE activity to avoid the adverse consequences (Lisch, 2009; Cui et al., 2013; Cui and Cao, 2014; Fultz et al., 2015). Nevertheless, a few TEs are still active, such as *Ac/Ds* and *Spm/dSpm* in maize (McClintock, 1950; Wessler, 1988), *Tto1* and *Tnt1* in tobacco (Grandbastien et al., 1989; Hirochika, 1993), *P* elements in *Drosophila* (Rubin et al., 1982), *Alu*, LINE-1 and SVA (SINE/VNTR/Alu) in humans (Mills et al., 2007). These active TEs have drawn more attention as they continue to create new insertions in the genome to drive intraspecies diversity. They not only cause deleterious phenotypes or diseases by altering gene expression and function when integrated into genes (Iskowitz et al., 2010; Lee et al., 2012), but also produce potentially beneficial variants and participate actively in adaptive evolution (Gonzalez et al., 2008; Schrader and Schmitz, 2019). In addition, these active TEs have served as valuable resources to elucidate the molecular mechanisms underlying the transposition and activation of TEs.

Likewise, transposition activity has always been an active field in rice genomics research. However, due to the scarcity of active TEs and the technical difficulty in their identification, only a dozen active TE families have been discovered up to now. *Tos17*, belonging to the *Copia* superfamily, is the first active retrotransposon isolated from rice (Hirochika et al., 1996), which exhibits high level of transposition activity under tissue culture and has been used to induce rice mutants for functional genomics studies (Hirochika, 2001). Subsequently, the first active

DNA transposon, *mPing* was discovered (Jiang et al., 2003). It is a non-autonomous tourist-like MITE (miniature inverted-repeat transposable element), derived from the perfect deletion of internal sequences of autonomous partner *Ping*. Its active mobilization has been demonstrated in three distinct situations: prolonged cell culture, anther-derived calli, and seeds exposed to γ -irradiation (Jiang et al., 2003; Kikuchi et al., 2003; Nakazaki et al., 2003). In the past decade, several other active TEs have been identified through characterization of spontaneous mutations or analysis of transcription profiles. These include *Copia* retrotransposons *Lullaby* (Picault et al., 2009), *Osr4*, *Osr13* (Cheng et al., 2015), *Osr7*, *Osr23* (Wang et al., 2009), *Gypsy* retrotransposons *Dasheng* (Cheng et al., 2015), LINE elements *Karma* (Komatsu et al., 2003), *LINE1-6_OS* (Cheng et al., 2015), and MITEs *dTok* (Moon et al., 2006), *nDart* (Tsugane et al., 2006), *nDaiZ* (Huang et al., 2009), *mGing* (Dong et al., 2012), *mJing* (Tang et al., 2019).

The emergence of new insertion in the genome is the most powerful and direct evidence to prove active transposition. The availability of large amount of resequenced data makes it possible to identify new insertions from numerous varieties from different species (Quadrona et al., 2016; Carpentier et al., 2019; Fuentes et al., 2019; Macko-Podgorni et al., 2019; Dominguez et al., 2020; Liu et al., 2020a,b). To fully understand the current activity of TEs in natural rice accessions and identify rice accessions carrying highly active TEs, we analyzed the comprehensive transposon insertion polymorphisms (TIPs) profile of 60,743 TE loci genotyped in 3,000 diverse accessions across five varietal groups [indica, japonica, aus/boro, basmati/sadri (aromatic) and intermediate type (admix)] (Liu et al., 2020b) to detect singleton insertions, which were unique to a single accession of the whole population and represented the new insertion of TEs in the genome. It was found that eight families maintained potential high activity in 3,000 natural rice accessions, and nine accessions contributed more than 50 singletons. These naturally active TEs would be ideal candidates for elucidating the molecular mechanisms underlying the transposition and activation of TEs, and help us fully understand the dynamics of rice TEs and their contribution to genetic diversity.

MATERIALS AND METHODS

Screening and Filtering of Singleton Insertions

The comprehensive TIPs profile of 60,743 TE loci genotyped in 3,000 diverse accessions (including 1,760 indica, 843 japonica, 215 aus/boro, 68 basmati/sadri and 134 intermediate type accessions) was obtained by analyzing the resequencing data based on our previously developed pipeline ¹ (Liu et al., 2020b). A series of custom Perl scripts ² were written to check the presence/absence status of each insertion in 3,000 rice accessions, and we obtained the initial list of singleton insertions. Subsequently, these singletons were further filtered to

¹<http://ibi.zju.edu.cn/Rtrip/method.html>

²https://github.com/lanceliu2018/FrontPlantSci_OsaTE

TABLE 1 | The classification of 10,924 singletons identified from 3,000 natural rice accessions.

Class	Type	Superfamily	No. of Singletons
Retrotransposon	LTR		6,735
			6,712
		<i>Copia</i>	4,307
		<i>Gypsy</i>	2,390
	LINE		15
		<i>L1</i>	16
	SINE		7
<i>tRNA</i>		7	
DNA transposon	TIR		4,189
			4,138
		<i>PIF/Harbinger</i>	2,015
		<i>Tc1/Mariner</i>	932
		<i>Mutator</i>	525
		<i>hAT</i>	375
		<i>EnSpm/CACTA</i>	291
	Helitron		51
<i>Helitron</i>		51	

eliminate low-quality and redundant singletons in the following aspects (**Supplementary Figure 1**). Firstly, considering that all resequenced rice samples were from homozygous accessions, TE insertions detected as heterozygosity were considered to be ambiguous and be discarded. Secondly, for a particular singleton, the high proportion of missing calls across the population would increase the likelihood that the insertion was not a singleton. Therefore, if one singleton had missing calls in more than 10% of rice accessions, it would be excluded from our dataset. Thirdly, if one singleton had no sufficient supporting reads (≥ 3 reads), it was likely to possess potentially higher prediction error rates and would be removed. Besides, if one singleton had two records corresponding to the forward and reverse direction, only one record of them was retained in the subsequent analysis to avoid double counting. In this way, we obtained a filtered dataset consisting of high-confidence non-redundant TE insertions.

Integration Preference Analysis

To visualize the distribution of singletons from individual TE family in the genome, they were mapped to the 12 rice chromosomes *via* custom JavaScript scripts. Meanwhile, Perl scripts were written to calculate their density along the chromosome. Here, the genome was scanned in a 1 Mb sliding window with a 0.5 Mb increment. To compare the density between the chromosomal arm and pericentromere region, the annotation of the pericentromeric region in the pseudomolecules assembly was obtained from Rice Genome Annotation Project³. Then, the average values of all sliding windows located in the corresponding region were calculated. Finally, the paired *t*-test was used to test the significance of difference between two regions

³<http://rice.plantbiology.msu.edu/>

in the 12 rice chromosomes, and was carried out by the function *t* test in R (version 3.5.1).

Enrichment Analysis for Rice Accessions Carrying Potentially Highly-Active Transposable Elements

To test whether the rice accession carrying potentially highly active (PHA) TEs were overrepresented in a particular varietal group, we performed enrichment analysis for the TE family showing high activity in multiple rice accessions. The odds ratio (OR), a parameter used to measure the degree of enrichment, was calculated by the following equation: $OR = [A/(B-A)]/[C/(3000-C)]$, where A represented the number of rice accession carrying the corresponding PHA TEs in a given varietal group, while B represented the total number of rice accession carrying the corresponding PHA TEs in 3,000 natural rice accessions; C represented the number of rice accessions owned by the variety group in 3,000 natural rice accessions. Then, the Fisher's exact test was used to test the significance of the degree of enrichment, and was implemented by the function *fisher* test in R (version 3.5.1).

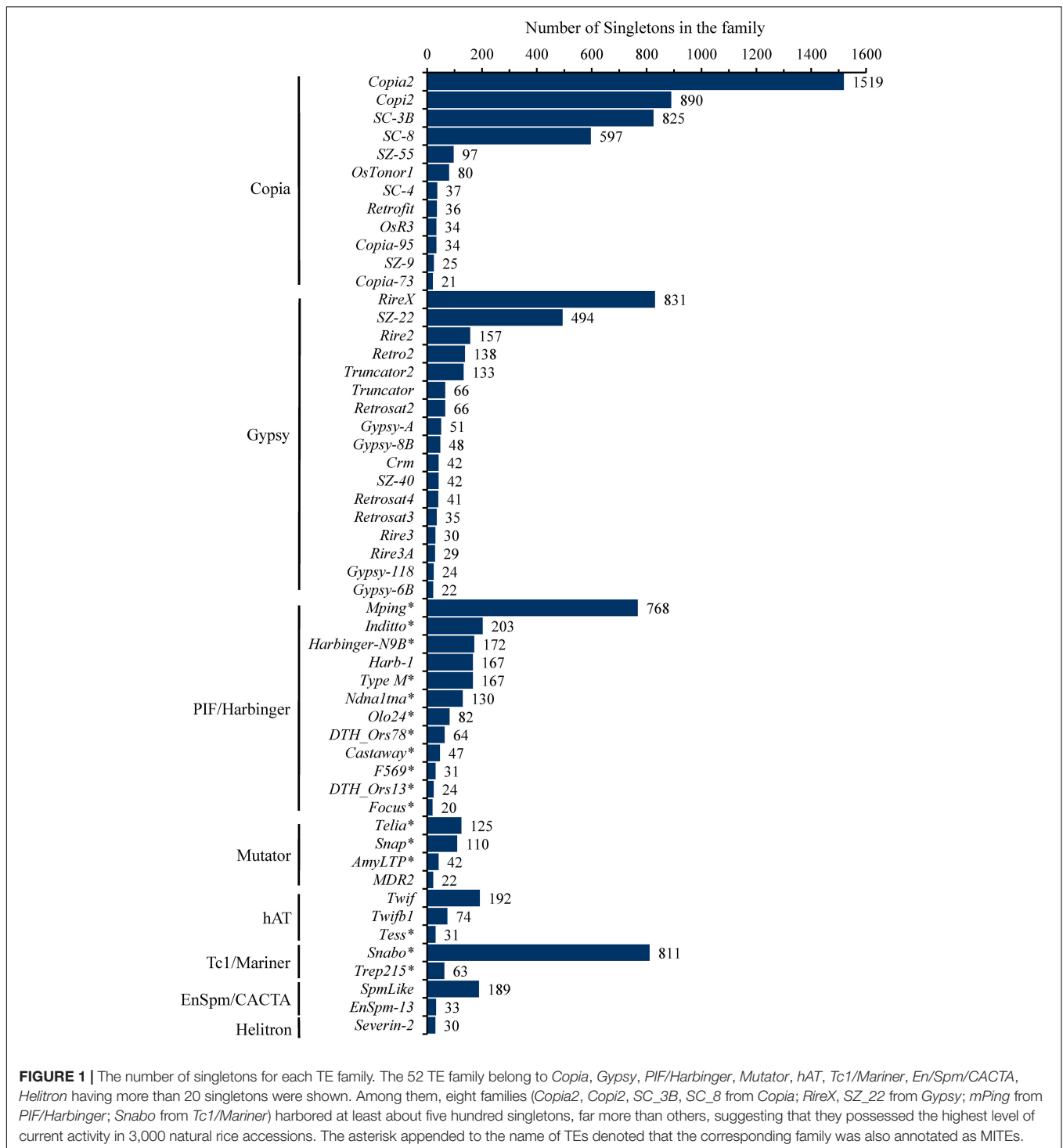
Detection of Target Site Duplications

In our dataset, about half of singleton loci were identified simultaneously from both the forward and reverse directions around the TE insertion. Moreover, all of them belong to the non-reference insertion, which was present in the resequenced genome but absent from the reference genome. Therefore, we could detect the TSD of these singletons by comparing the sequences of the two genotypes with or without TE insertion. Briefly, when the two flanking sequences from both ends of a non-reference insertion are mapped to the reference genome, an overlapping region consisting of several bases usually appears as shown in **Supplementary Figure 2**, which corresponds to the TSD of the non-reference insertion. Based on this idea, we calculated the overlapping region for each singleton according to their coordinates on the genome detected from the forward and reverse direction. Then, the corresponding sequences were extracted from the reference genome. To understand their conservation, the WebLogo 3 (Schneider and Stephens, 1990; Crooks et al., 2004) was used to generate the sequence logo for the TSDs of individual TE family.

RESULTS

Abundant Unique Insertion Sites Belong to Various Transposable Element Families

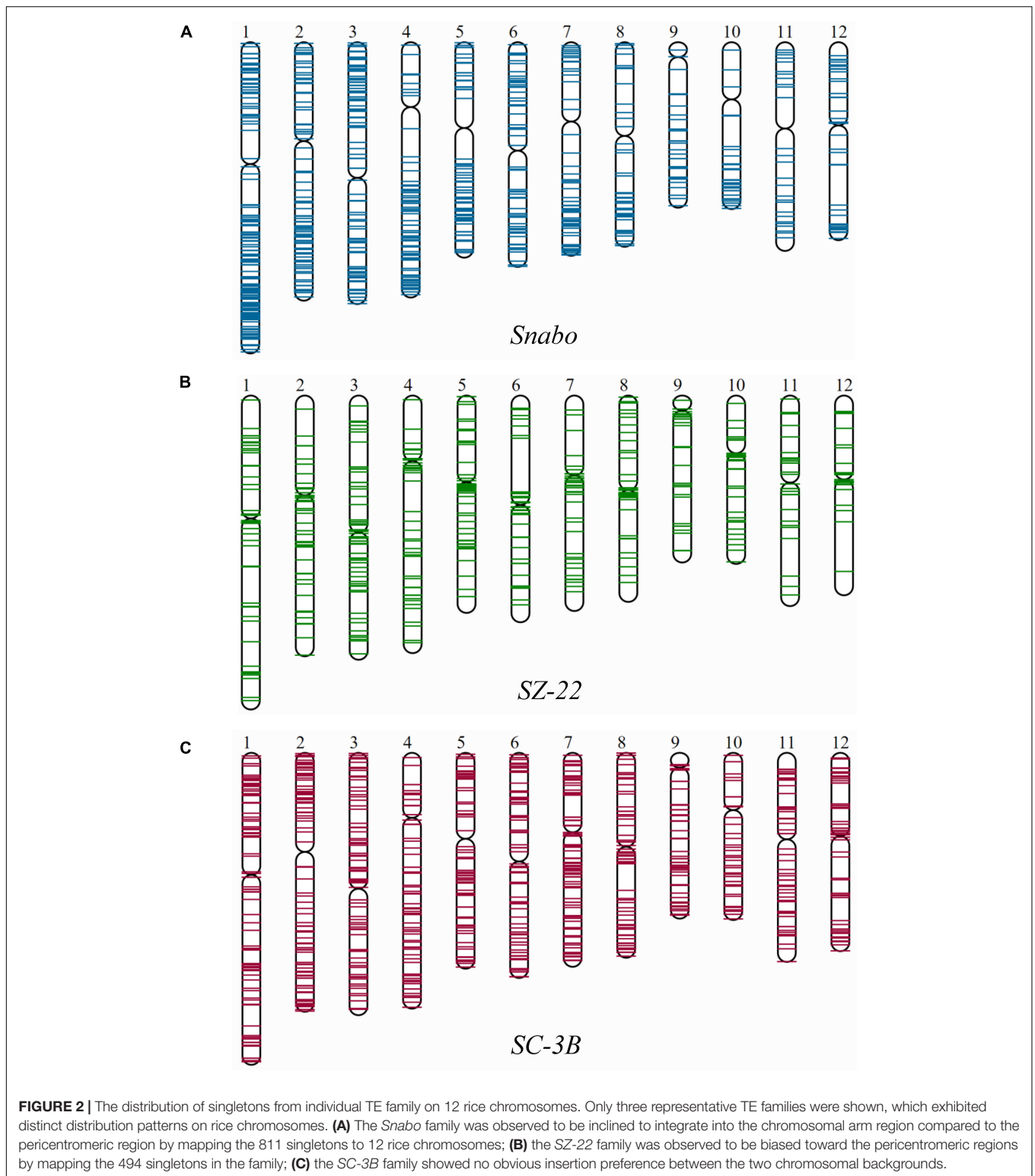
In the TIPs profile, 23,036 TE insertions were found to be only present in a single rice accession, which were called singletons. Subsequently, the low-quality and redundant singletons were removed through four strict filtering steps (**Supplementary Figure 1**; See section "Materials and Methods"). Finally, 10,924 high-confidence singletons were retained for further analysis (**Supplementary Table 1**). The hierarchical classification showed that more singletons resulted from active transposition of



retrotransposons than from DNA transposons (6,735 versus 4,189) (Table 1). All 11 TE superfamilies were identified in our dataset (Table 1), but the number of singletons varied substantially across different superfamilies, implying distinct current activity. The loci of *Copia* were approximately twice more than *Gypsy* (4,307 versus 2,390; Table 1) despite that both belong to LTR type. For TIR elements, *PIF/Harbinger* exhibited

the highest singletons number (2,015), while *EnSpm/CACTA*, by contrast, had the lowest (291 singletons).

These singletons were further classified into 251 TE families, with the number of singletons in each family ranging from 1 to 1,519 (Supplementary Table 2). In particular, eight families, including four *Copia* (*Copia2*, *Copi2*, *SC-3B*, *SC-8*), two *Gypsy* (*RireX*, *SZ-22*), one *PIF/Harbinger* (*mPing*) and one *Tc1/Mariner*



(*Snabo*), had the maximum singletons and accounted for 61.7% (6,735/10,924) of the total singletons (**Figure 1** and **Supplementary Table 2**), suggesting that they probably possess the highest current activity in 3,000 natural rice accessions. In

addition, 53 of 251 families possessed at least 20 singletons and together contributed 91.9% (10,041/10,924) of the total singletons (**Figure 1** and **Supplementary Table 2**). Of them, 12, 17, and 12 were from *Copia*, *Gypsy* and *Harbinger* superfamilies,

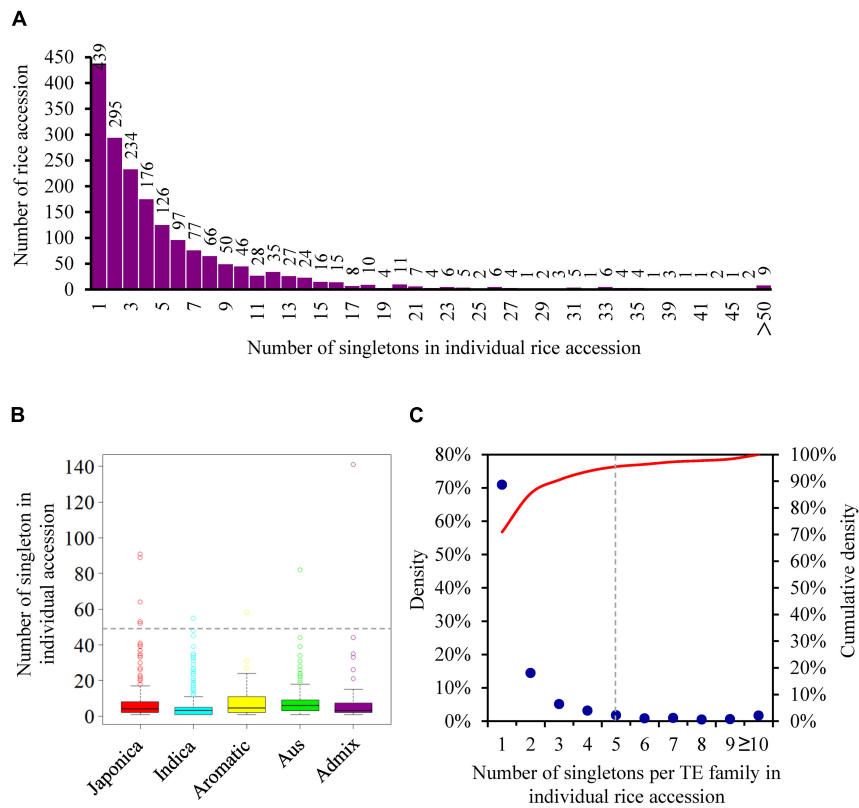


FIGURE 3 | The distribution of singletons among different rice accessions. **(A)** The distribution of the number of singletons in individual rice accession; **(B)** the distribution of the number of singletons in rice accession from five distinct groups. Each boxplot shows the distribution of the number of singletons owned by each rice accession in the corresponding group. The median represented the middle number of singletons among all rice accession from the same group. The number corresponding to the gray dashed line is 50; **(C)** the distribution of the number of singletons per TE family in individual rice accession. The red curve represented the cumulative density.

respectively, while *Mutator*, *hAT*, *Mariner*, *CACTA*, *Helitron* superfamilies had only 1 to 4 families (Figure 1). Notably, 17 of 24 DNA transposon families belonged to MITEs, and accounted for 80.3% (2,890/3,597) of DNA transposons, indicating that most of the active DNA transposons were short non-autonomous elements in rice genome.

Integration Preference of Transposable Elements Revealed by Singletons

Since the singleton was newly inserted into the genome, it was almost impossible to be affected by the selection or chromosome recombination. Therefore, the distribution of singletons on chromosomes could truly reflect the transposition characteristics. Fifty-three TE families harboring at least 20 singletons were then mapped on 12 rice chromosomes, and divided into three categories: pericentromere preference, chromosomal arm preference and unbiased distribution (Figure 2 and Supplementary Table 3). There were 38 (71.7%) TE families were significantly inclined to the chromosomal arm region, which included 79.3% (7,960/10,041) singletons [71.8% (4,630/6,444) retrotransposons and 92.6% (3,330/3,597) DNA transposons]. For instance, the *Snabo* family, whose average

density in the chromosomal arm region was 6.4 times higher than that in the pericentromere region (Figure 2A; p -value < 7.31E-07, paired t -test). By comparison, only four families were biased toward the pericentromere region (Supplementary Table 3). A typical case was the *SZ-22* family, whose average density in pericentromere region was 5.6 times higher than that in the chromosomal arm region (Figure 2B; p -value = 1.60E-03, paired t -test). Interestingly, all the four families belonged to the *Gypsy* superfamily, accounting for 30.7% (691/2,249) of the *Gypsy* singletons. The remaining 11 families had no significant insertion preference between the two chromosomal backgrounds (Supplementary Table 3), although they might be unevenly distributed along the chromosomes, such as *SC-3B* (Figure 2C; p -value = 0.1291, paired t -test).

When integrating into new positions, TEs typically result in the target sites duplication (TSD), which represent the hotspots of TE insertions. To explore the mechanism underlying the biased distribution, TSDs of 5,849 singletons were extracted from both forward and reverse directions flanking the insertion sites (Supplementary Figure 2). The results showed that TSDs of 12 families in *Harbinger* superfamily had the conserved sequences of TAA/TTA (Supplementary Table 4 and Supplementary Figure 3). The *Telia* family in the

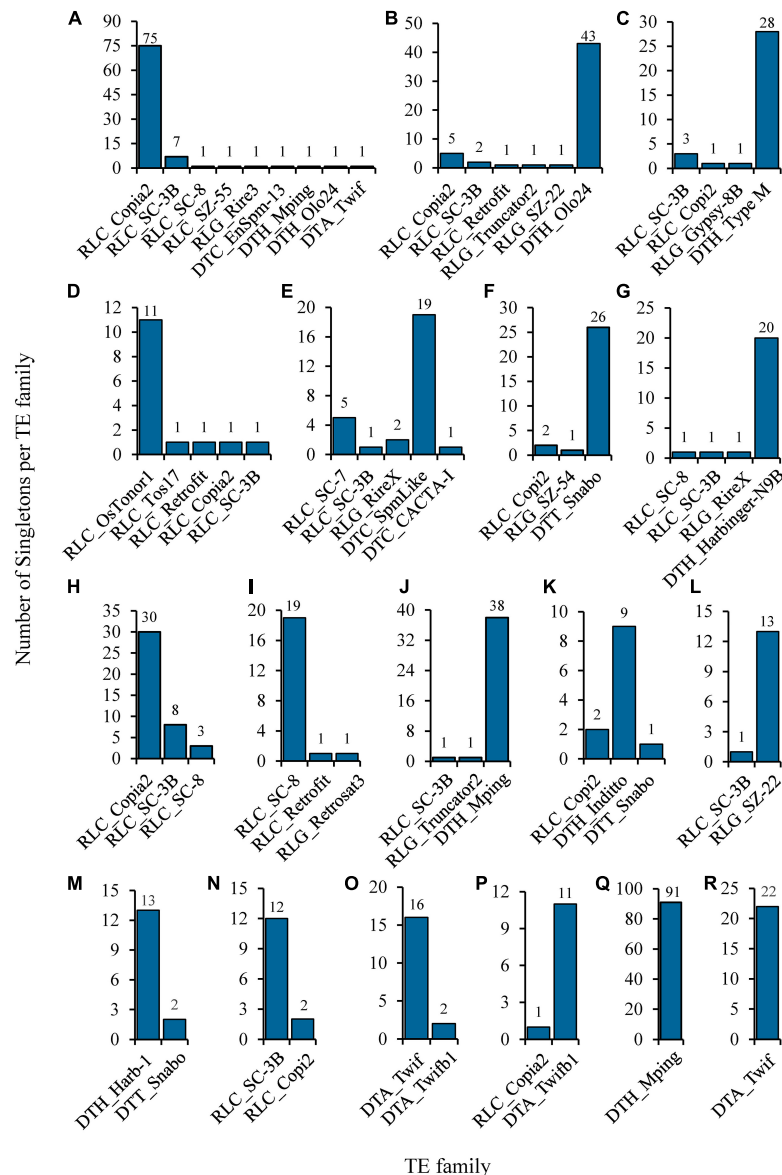
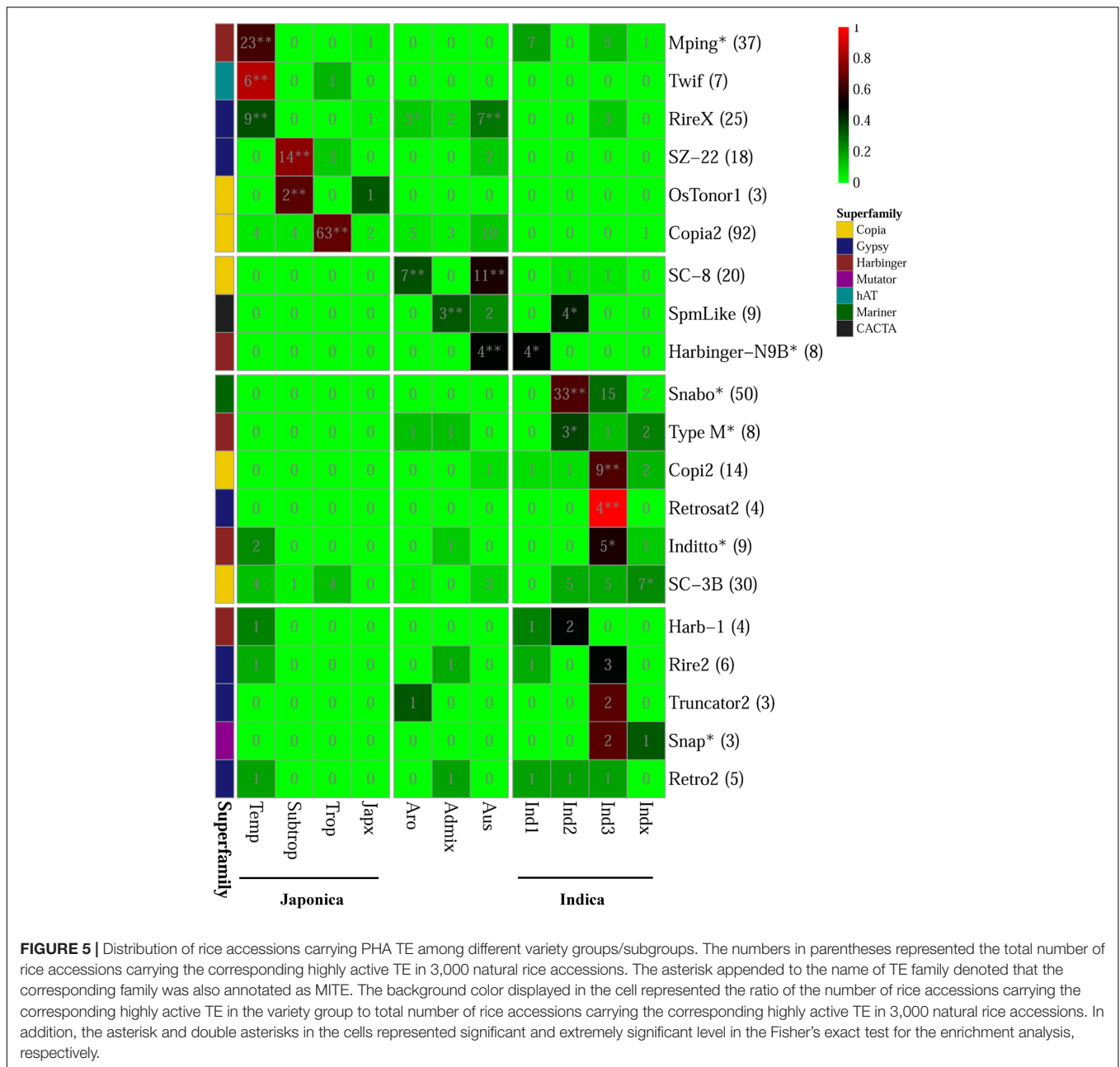


FIGURE 4 | Selected examples of potentially highly-active TEs identified from natural rice accessions. The histograms show the composition of singletons identified in each of the following rice accessions: **(A)** Trop/IRIS_313-7885; **(B)** Trop/IRIS_313-10703; **(C)** indx/IRIS_313-12300; **(D)** Subtrop/IRIS_313-11897; **(E)** aus/IRIS_313-11482; **(F)** Ind2/IRIS_313-9740; **(G)** aus1/IRIS_313-10930; **(H)** Trop/IRIS_313-10758; **(I)** Aro/IRIS_313-11068; **(J)** Temp/IRIS_313-11655; **(K)** ind3/IRIS_313-12142; **(L)** Subtrop/IRIS_313-8580; **(M)** ind2/IRIS_313-8703; **(N)** Indx/B090; **(O)** Temp/IRIS_313-11800; **(P)** Trop/IRIS_313-9550; **(Q)** Temp/IRIS_313-15904; **(R)** Temp/IRIS_313-8149. The three-letter prefix in the TE family name represents the superfamily to which it belongs. Its meaning is as follows: RLC: Copia; RLG: Gypsy; DTC: EnSpm/CACTA; DTH: PIF/Harbinger; DTA: hAT; DTT: Tc1/Mariner.

Mutator superfamily inserted specifically into AT-rich sequences (TAAATTATA), while the *MuDR2* family preferentially targeted CG-rich sequences (**Supplementary Table 4** and **Supplementary Figure 3**). The remaining TE families almost had no conserved sequences in TSDs. Nevertheless, the TSD length of TE families in the same superfamily showed exact consistency. For example, the *Copia* and *Gypsy* superfamilies had 5 bp TSDs, while *hAT* and *CACTA* superfamilies had 8 and 3 bp TSDs, respectively (**Supplementary Table 4** and **Supplementary Figure 3**).

In addition, we analyzed the association of singletons with the annotated genes to understand whether they potentially affect the function of genes. Of the 10,924 singletons, 926 (8.5%) and 1,993 (18.2%) were inserted into the coding regions and introns of genes (**Supplementary Table 5**), respectively, which may disrupt gene function through insertion mutagenesis or altering splicing patterns. The 238 (2.2%) and 409 (3.7%) singletons were located in the 5'-UTR and 3'-UTR of the genes, respectively (**Supplementary Table 5**). In addition, 479 (4.4%)



and 277 (2.5%) loci were found within the 200 bp upstream and downstream sequences adjacent to the annotated genes (Supplementary Table 5), respectively, which were likely to change gene expression by providing putative *cis*-regulatory elements or epigenetic regulation.

Extensively Current Activity of Transposable Elements and Potentially Highly Active Transposable Elements in Rice

The 10,924 singletons were distributed in 1,864 of 3,000 rice accessions (Supplementary Table 6), with an average of

5.86 insertions per accession, reflecting the extensive activity of TEs in 3,000 natural rice accessions. Among different accessions, the number of singletons displayed large differences, varying from 1 to 141 (Figure 3A and Supplementary Table 6). Particularly, nine rice accessions, including five japonicas (IRIS_313-15904; IRIS_313-7885; IRIS_313-10228; IRIS_313-10703; IRIS_313-10562) and one each of indica (B227), aus/boro (IRIS_313-11413), aromatic (IRIS_313-11217), admix (IRIS_313-9445), contributed more than 50 singletons (Figure 3A and Supplementary Table 6), suggesting that they may carry more active TEs. When categorizing rice accessions according to five distinct groups (japonica, indica, aus, aromatic, and admix), it was observed that the indica accessions had

the least number of singletons on average (4.75 singletons per accession), followed by japonica (6.58 singletons per accession), while aus, aromatic and admix accessions possessed relatively more singletons, with an average of 8.01, 8.43, and 8.71 insertions per accession, respectively (Figure 3B and Supplementary Table 6).

The identified singletons may be from different TE families in the individual rice accession. Therefore, they were further classified by families in each accession (Supplementary Table 7). Most TE families possessed only 1 to 4 singletons in all accessions (Figure 3C and Supplementary Table 7), while 35 families harbored at least 5 singletons in one or more rice accessions (Figure 4 and Supplementary Table 8), suggesting that they probably have highly current activity in the corresponding accession. For instance, *Copia2* family contributed 75 of 89 identified singletons in the tropical japonica accession IRIS_313-7885 (Figure 4A and Supplementary Table 8); *mPing* family created up to 91 singletons in the temperate japonica accession IRIS_313-15904, which was the only active family detected in this accession (Figure 4Q and Supplementary Table 8); *Olo24* contributed 43 of 53 identified singletons in the tropical japonica accession IRIS_313-10703 (Figure 4B and Supplementary Table 8). Of 35 TE families, 20 (11 retrotransposon and 9 DNA transposon families) displayed potentially high activity in multiple (≥ 3) rice accessions (Supplementary Table 8). Particularly, *Copia2* family maintained potentially high activity in up to 92 rice accessions, with the number of singletons per accession varying from 5 to 75.

Variable Activity of Transposable Elements Among Different Rice Groups

For the 20 PHA TE families, we analyzed the distribution of their PHA accessions (referring to the rice accessions carrying the corresponding PHA TEs) among different groups and subgroups, where both indica and japonica were further subdivided into four subgroups (indica: indica1, indica2, indica3, indicaX; japonica: temperate japonica, tropical japonica, subtropical japonica, japonicaX) (Figure 5 and Supplementary Table 9). As a result, six TE families were significantly enriched in the japonica group (p -value < 0.05 , Fisher's exact test), including *mPing*, *Twif* and *RIREX* in the temperate japonica subgroup, with 62.2% (23/37), 85.7% (6/7) and 36.0% (9/25) of PHA accessions being in this subgroup, respectively; *SZ-22* and *OsTonor1* in the subtropical japonica subgroup, with 77.8% (14/18) and 66.7% (2/3) of PHA accessions being subtropical japonica, respectively; *Copia2* in the tropical japonica subgroup, with 68.5% (63/92) of PHA accessions being tropical japonica (Figure 5 and Supplementary Table 9). Similarly, eight TE families showed significant enrichment in the indica group (p -value < 0.05 , Fisher's exact test), including one (*Harbinger-N9B*) in the indica1 subgroup, three (*SpmLike*, *Snabo* and *Type M*) in the indica2 subgroup, three (*Copi2*, *Retrosat2*, *Inditto*) in the indica3 subgroup and one (*SC-3B*) in the indicaX subgroup. In addition, two (*RIREX* and *SC-8*), one

(*SpmLike*) and three (*RIREX*, *SC-8* and *Harbinger-N9B*) TE families presented significant enrichment in the aromatic, admix and aus group, respectively (p -value < 0.05 , Fisher's exact test; Figure 5 and Supplementary Table 9). For the remaining five TE families (*Harb-1*, *Rire2*, *Truncator2*, *Snap*, and *Retro2*), no significant enrichment could be detected in any group or subgroup.

To understand the evolutionary relationship among these PHA rice accessions, they were mapped to the phylogenetic tree of 3,000 rice accessions constructed from SNPs in the previous study (Mansueto et al., 2017). Most of the PHA accessions corresponding to the same TE family clearly gathered together to form one or several clusters, rather than scattered randomly throughout the tree (Supplementary Figure 4), suggesting that they may share a common origin of activity.

DISCUSSION

The Large-Scale Survey of Active Transposable Elements at the Population Scale

There was an eternal competition between TEs and hosts, which caused the birth and extinction of transposition activity to occur repeatedly as the genome evolved. It was important to determine which TEs were still active in the genome. Previous studies identified a few of active TEs from rice, but they only focused on a single TE family in one or few rice varieties (Hirochika et al., 1996; Jiang et al., 2003; Picault et al., 2009; Dong et al., 2012; Cheng et al., 2015; Tang et al., 2019). Considering the difference in transposition activity among different varieties (Hirochika et al., 1996; Jiang et al., 2003; Naito et al., 2006), it prevented us from comprehensively understanding the current activity of rice TEs. In this study, we systematically surveyed the activity of all TEs in 3,000 natural rice accessions by detecting UIS based on a computational approach. 10,924 high-confidence singletons from 251 TE families were identified from 1,864 of 3,000 rice accessions. It should be the first investigation for any organism to detect potential active TEs at the population scale.

Active Mobilization of Transposable Elements Under Normal Growth Conditions Rather Than Artificial Induction

The known active TEs identified from rice were mostly induced by artificial treatment up to now. *Tos17* (Hirochika et al., 1996), *mPing* (Jiang et al., 2003), *Lullaby* (Picault et al., 2009), *mGing* (Dong et al., 2012), *Karma* (Komatsu et al., 2003), *nDaiZ* (Huang et al., 2009) were activated under tissue culture; *dTok* (Moon et al., 2006), *nDart* (Tsugane et al., 2006), *mJing* (Tang et al., 2019), *Osr7*, *Osr23* (Wang et al., 2009) were isolated from the progeny of intraspecific or interspecific hybridization. By comparison, their active mobilization was rarely found under normal growth conditions. One exception was *mPing*, which

possessed very high activity in the temperate japonica cultivar Gimbozu EG4 and three related landraces (Aikoku A157, Aikoku A119, Aikoku A123) (Naito et al., 2006). These rice cultivars carrying naturally active *mPing* played a crucial role in biology research on TEs. They were recruited to successfully address two important questions concerning TE bursts, namely how TE achieved very high copy number without killing their host (Naito et al., 2006, 2009), and how the burst was sustained for generations without triggering host's silencing mechanisms (Lu et al., 2017).

The active TEs obtained in this study were identified from rice cultivated population. Therefore, their active transposition occurred under normal growth conditions. Compared with that derived from artificial induction, these active TEs could truly reflect the active status of TEs in the current rice genome. These naturally active TEs were likely to be activated by exogenous stresses. Previous studies have shown that TEs were capable of responding to various biotic and abiotic stresses (Grandbastien, 1998), such as wounding (Mhiri et al., 1997), radiation (Questa et al., 2010), cold (Ivashuta et al., 2002; Naito et al., 2009; Butelli et al., 2012), heat (Ito et al., 2011), salt (Naito et al., 2009), as well as pathogen infection (Grandbastien et al., 2005; Buchmann et al., 2009). The 3,000 diverse rice accessions involved in this study were grown worldwide and faced to different ecological and geographical environments. They were challenged by a variety of biotic and abiotic stresses during domestication, which provided abundant external conditions for TEs activation. The active transposition of these TEs enabled the host to rapidly acquire genetic diversity to cope with altering environments.

Another possibility was that the activation of these TEs was attributed to mutations in genes involved in the epigenetic pathway. Under normal conditions, the activity of most TEs was strictly repressed by epigenetic regulation. However, their active transposition could be triggered when the related genes were mutated during rice cultivation. For instance, disruption of rice SDG714, encoding a histone H3K9-specific methyltransferase, reduced the levels of CG and CHG methylation as well as H3K9 methylation and promoted frequent transposition of *Tos17* (Ding et al., 2007). The loss function of JMJ703 increased H3K4me3 and enhanced the transposition activity of *Karma* element (Cui et al., 2013). In this study, some rice accessions possessed several or even dozens of active TE families simultaneously (**Supplementary Table 7**), suggesting that their activities may be resulted from the relaxation of the epigenetic silencing.

The activation of TEs may also be due to alteration of TE sequences. It was reported that the burst of *mPing* in the domesticated rice accessions was associated with the acquisition of two variants of the transposase loci, *Ping16A* and *Ping16A_Stow* (Chen et al., 2019). Meanwhile, these naturally active TEs would be ideal candidates for elucidating the molecular mechanisms underlying the activation of TEs, and investigating the interaction between TEs and the host genome. In addition, they would also be potentially valuable resources for

tagging system of gene discovery in rice and other important crops.

A High-Throughput Strategy for Identifying Active Transposable Elements

Since active TEs are rare, it is very challenging to recognize them from numerous inactive elements. At present, the active TEs were mainly identified by the following methods. Firstly, inspecting the sequences between different copies to discover identical copies. It is based on the fact that new insertions arising from the current activity don't accumulate mutations. The active *mPing* was initially discovered by this method (Jiang et al., 2003). However, it is only applicable to the assembled genomes, since this method require full-length sequences of the TE. Secondly, characterizing spontaneous mutants. Most of the known active TEs in rice were isolated by this method up to now (Moon et al., 2006; Tsugane et al., 2006; Huang et al., 2009; Wang et al., 2009; Tang et al., 2019). It is undeniable that more active TEs will be revealed as more mutants are analyzed in the future. However, the acquisition of active TEs by this method is largely accidental, since the mutants caused by active transposition account for relatively small proportion. Thirdly, detecting transcriptional activity (Picault et al., 2009). Regardless of the retrotransposon or DNA transposon, the expression of genes encoding the enzymes required for transposition is a necessary prerequisite for successful transposition. Therefore, we can preliminarily speculate whether TEs are still active by detecting their transcripts or not. Nevertheless, the active transcription of TEs don't ensure that they are ultimately successfully transposed, because the hosts have evolved multiple mechanisms to regulate transposition activity in the post-transcription stages (Slotkin and Martienssen, 2007).

The emergence of new insertions in the genome was the most powerful and direct evidence to prove the active transposition. These loci were expected to be only present in one accession of the whole population, as their transposition occurred after divergence between different accessions. Therefore, it was a feasible strategy to identify active TEs by detecting singleton insertions from a large population. However, it was almost impossible to perform such research by traditional experimental methods because of the enormous workload. Recently, the development of next-generation sequencing technology has provided an excellent opportunity for efficient researches on genetic variation. Here we provided a high-throughput strategy for capturing active TEs from the resequenced population. The main content of this method was to construct a comprehensive profile of TIPs by analyzing the resequenced data from the population using our previously developed pipeline (Liu et al., 2020b). It consisted of the identification of TE loci and the genotyping of each locus in the population. Then, singleton insertions were screened out from the profile. The advantage of this method was that it was applicable to all types of TEs and was capable of detecting a large number of active TEs from the population in a single analysis. It could also be applied directly to other

species when the whole-genome resequenced data for their population were available.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <http://ibi.zju.edu.cn/Rtrip/index.html>.

AUTHOR CONTRIBUTIONS

J-HX and ZL contributed to the conception and design of the study and wrote the manuscript. ZL, HZ, YY, M-XW, Y-CZ, E-KY, MA, KS, M-HD, and J-HX analyzed the data. HZ and Y-CZ provided the computer resource. All authors read and contributed to the manuscript.

FUNDING

We would like to thank Minghua Duan and Zihua Wang from Zhejiang Zhengjingyuan Pharmacy Chain Co., Ltd., for their funds supporting (H20151699 and H20151788).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.745526/full#supplementary-material>

Supplementary Figure 1 | The outline of the procedure for the identification of high-confidence singleton insertions. In the TIPs profile, 27,531 of 60,743 TE loci were detected simultaneously from both forward and reverse directions flanking the TE insertion. Therefore, each of them has two records corresponding to the

forward and reverse directions in the profile. In total, there are 88,274 records corresponding to 60,743 TE loci in the profile. We used these data as the starting point for screening singleton insertions. As a result, 23,036 of 88,274 insertions were found to be singletons. Among them, 2,924 singletons were heterozygous, 2,343 singletons had missing calls in more than 10% of rice accessions, 972 singletons had no sufficient supporting reads, 5,873 singletons had two records corresponding to the forward and reverse directions. Finally, a filtered dataset consisting of 10,924 high-confidence non-redundant TE singletons was obtained.

Supplementary Figure 2 | The general schematic view of the method to detect TSDs for singleton loci that were identified from both forward and reverse directions around the insertion sites.

Supplementary Figure 3 | Analysis of sequence conservation of TSD for different TE families. The sequence logos were generated by WebLogo3 (Schneider and Stephens, 1990; Crooks et al., 2004).

Supplementary Figure 4 | Distribution of rice accessions carrying PHA TE in the phylogenetic tree of 3,000 natural rice accessions, which was constructed from SNPs in the previous study (Mansueto et al., 2017).

Supplementary Table 1 | The information of singletons identified from 3,000 natural rice accessions.

Supplementary Table 2 | The number of singletons in the 251 TE families.

Supplementary Table 3 | The density of singletons by each family on the 12 chromosomes of rice.

Supplementary Table 4 | TSD of 5,849 singletons detected from both ends of the insertions.

Supplementary Table 5 | The information of the singletons associated with genes.

Supplementary Table 6 | The number of singletons in each rice accession and variety group.

Supplementary Table 7 | The classification of singletons by TE family in each rice accession.

Supplementary Table 8 | The PHA TEs identified from 3,000 natural rice accessions.

Supplementary Table 9 | The distribution of rice accessions carrying PHA TEs among different variety groups and subgroups.

REFERENCES

- Bennetzen, J. L., and Wang, H. (2014). The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annu. Rev. Plant Biol.* 65, 505–530. doi: 10.1146/annurev-arplant-050213-035811
- Buchmann, R. C., Asad, S., Wolf, J. N., Mohannath, G., and Bisaro, D. M. (2009). Geminivirus AL2 and L2 proteins suppress transcriptional gene silencing and cause genome-wide reductions in cytosine methylation. *J. Virol.* 83, 5005–5013. doi: 10.1128/JVI.01771-08
- Butelli, E., Licciardello, C., Zhang, Y., Liu, J., Mackay, S., Bailey, P., et al. (2012). Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24, 1242–1255. doi: 10.1105/tpc.111.095232
- Carpentier, M. C., Manfroi, E., Wei, F. J., Wu, H. P., Lasserre, E., Llauro, C., et al. (2019). Retrotranspositional landscape of Asian rice revealed by 3000 genomes. *Nat. Commun.* 10:24. doi: 10.1038/s41467-018-07974-5
- Chen, J. F., Lu, L., Benjamin, J., Diaz, S., Hancock, C. N., Stajich, J. E., et al. (2019). Tracking the origin of two genetic components associated with transposable element bursts in domesticated rice. *Nat. Commun.* 10:641. doi: 10.1038/s41467-019-08451-3
- Cheng, C., Tarutani, Y., Miyao, A., Ito, T., Yamazaki, M., Sakai, H., et al. (2015). Loss of function mutations in the rice chromomethylase OsCMT3a cause a burst of transposition. *Plant J.* 83, 1069–1081. doi: 10.1111/tpj.12952
- Chuang, E. B., Elde, N. C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86. doi: 10.1038/nrg.2016.139
- Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Cui, X., and Cao, X. (2014). Epigenetic regulation and functional exaptation of transposable elements in higher plants. *Curr. Opin. Plant Biol.* 21, 83–88. doi: 10.1016/j.pbi.2014.07.001
- Cui, X., Jin, P., Cui, X., Gu, L., Lu, Z., Xue, Y., et al. (2013). Control of transposon activity by a histone H3K4 demethylase in rice. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1953–1958. doi: 10.1073/pnas.1217020110
- Ding, Y., Wang, X., Su, L., Zhai, J., Cao, S., Zhang, D., et al. (2007). SDG714, a histone H3K9 methyltransferase, is involved in *Tos17* DNA methylation and transposition in rice. *Plant Cell* 19, 9–22. doi: 10.1105/tpc.106.048124
- Dominguez, M., Dugas, E., Benchouaia, M., Leduque, B., Jimenez-Gomez, J. M., Colot, V., et al. (2020). The impact of transposable elements on tomato diversity. *Nat. Commun.* 11:4058. doi: 10.1038/s41467-020-17874-2
- Dong, H. T., Zhang, L., Zheng, K. L., Yao, H. G., Chen, J., Yu, F. C., et al. (2012). A *Gaijin*-like miniature inverted repeat transposable element is mobilized in rice

- during cell differentiation. *BMC Genomics* 13:135. doi: 10.1186/1471-2164-13-135
- Fuentes, R. R., Chebotarov, D., Duitama, J., Smith, S., De la Hoz, J. F., Mohiyuddin, M., et al. (2019). Structural variants in 3000 rice genomes. *Genome Res.* 29, 870–880. doi: 10.1101/gr.241240.118
- Fultz, D., Choudury, S. G., and Slotkin, R. K. (2015). Silencing of active transposable elements in plants. *Curr. Opin. Plant Biol.* 27, 67–76. doi: 10.1016/j.pbi.2015.05.027
- Gonzalez, J., Lenkov, K., Lipatov, M., Macpherson, J. M., and Petrov, D. A. (2008). High rate of recent transposable element-induced adaptation in *Drosophila melanogaster*. *PLoS Biol.* 6:e251. doi: 10.1371/journal.pbio.006break0251 doi: 10.1371/journal.pbio.0060251
- Grandbastien, M. A. (1998). Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.* 3, 181–187. doi: 10.1016/S1360-1385(98)01232-1
- Grandbastien, M. A., Audeon, C., Bonnivard, E., Casacuberta, J. M., Chalhou, B., Costa, A. P., et al. (2005). Stress activation and genomic impact of *Tnt1* retrotransposons in Solanaceae. *Cytogenet. Genome Res.* 110, 229–241. doi: 10.1159/000084957
- Grandbastien, M. A., Spielmann, A., and Caboche, M. (1989). *Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. *Nature* 337, 376–380. doi: 10.1038/337376a0
- Hirochika, H. (1993). Activation of tobacco retrotransposons during tissue culture. *EMBO J.* 12, 2521–2528. doi: 10.1002/j.1460-2075.1993.tb05907.x
- Hirochika, H. (2001). Contribution of the *Tos17* retrotransposon to rice functional genomics. *Curr. Opin. Plant Biol.* 4, 118–122. doi: 10.1016/S1369-5266(00)00146-1
- Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H., and Kanda, M. (1996). Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7783–7788. doi: 10.1073/pnas.93.15.7783
- Huang, C. R., Burns, K. H., and Boeke, J. D. (2012). Active transposition in genomes. *Annu. Rev. Genet.* 46, 651–675. doi: 10.1146/annurev-genet-110711-155616
- Huang, J., Zhang, K., Shen, Y., Huang, Z., Li, M., Tang, D., et al. (2009). Identification of a high frequency transposon induced by tissue culture, *nDaiZ*, a member of the *hAT* family in rice. *Genomics* 93, 274–281. doi: 10.1016/j.ygeno.2008.11.007
- Iskow, R. C., McCabe, M. T., Mills, R. E., Torene, S., Pittard, W. S., Neuwald, A. F., et al. (2010). Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141, 1253–1261. doi: 10.1016/j.cell.2010.05.020
- Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472, 115–119. doi: 10.1038/nature09861
- Ivashuta, S., Naumkina, M., Gau, M., Uchiyama, K., Isobe, S., Mizukami, Y., et al. (2002). Genotype-dependent transcriptional activation of novel repetitive elements during cold acclimation of alfalfa (*Medicago sativa*). *Plant J.* 31, 615–627. doi: 10.1046/j.1365-313X.2002.01383.x
- Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S. R., McCouch, S. R., et al. (2003). An active DNA transposon family in rice. *Nature* 421, 163–167. doi: 10.1038/nature01214
- Kikuchi, K., Terauchi, K., Wada, M., and Hirano, H. Y. (2003). The plant MITE *mPing* is mobilized in anther culture. *Nature* 421, 167–170. doi: 10.1038/nature01218
- Komatsu, M., Shimamoto, K., and Kyoizuka, J. (2003). Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon *Karma*. *Plant Cell* 15, 1934–1944. doi: 10.1105/tpc.011809
- Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J. III., et al. (2012). Landscape of somatic retrotransposition in human cancers. *Science* 337, 967–971. doi: 10.1126/science.1222077
- Lisch, D. (2009). Epigenetic regulation of transposable elements in plants. *Annu. Rev. Plant Biol.* 60, 43–66. doi: 10.1146/annurev.arplant.59.032607.092744
- Lisch, D. (2013). How important are transposons for plant evolution? *Nat. Rev. Genet.* 14, 49–61. doi: 10.1038/nrg3374
- Liu, Z., Wang, T., Wang, L., Zhao, H., Yue, E., Yan, Y., et al. (2020b). RTRIP: a comprehensive profile of transposon insertion polymorphisms in rice. *Plant Biotechnol. J.* 18, 2379–2381. doi: 10.1111/pbi.13425
- Liu, Z., Fan, M., Yue, E. K., Li, Y., Tao, R. F., Xu, H. M., et al. (2020a). Natural variation and evolutionary dynamics of transposable elements in *Brassica oleracea* based on next-generation sequencing data. *Hortic. Res.* 7:145. doi: 10.1038/s41438-020-00367-0
- Lu, L., Chen, J. F., Robb, S. M. C., Okumoto, Y., Stajich, J. E., and Wessler, S. R. (2017). Tracking the genome-wide outcomes of a transposable element burst over decades of amplification. *Proc. Natl. Acad. Sci. U. S. A.* 114, E10550–E10559. doi: 10.1073/pnas.1716459114
- Ma, J., Devos, K. M., and Bennetzen, J. L. (2004). Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res.* 14, 860–869. doi: 10.1101/gr.1466204
- Macko-Podgorni, A., Stelmach, K., Kwolek, K., and Grzebelus, D. (2019). Stowaway miniature inverted repeat transposable elements are important agents driving recent genomic diversity in wild and cultivated carrot. *Mob. DNA* 10:47. doi: 10.1186/s13100-019-0190-3
- Mansueti, L., Fuentes, R. R., Borja, F. N., Detras, J., Abriol-Santos, J. M., Chebotarov, D., et al. (2017). Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Res.* 45, D1075–D1081. doi: 10.1093/nar/gkw1135
- McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U. S. A.* 36, 344–355. doi: 10.1073/pnas.36.6.344
- Mhiri, C., Morel, J. B., Vernhettes, S., Casacuberta, J. M., Lucas, H., and Grandbastien, M. A. (1997). The promoter of the tobacco *Tnt1* retrotransposon is induced by wounding and by abiotic stress. *Plant Mol. Biol.* 33, 257–266. doi: 10.1023/A:1005727132202
- Mills, R. E., Bennett, E. A., Iskow, R. C., and Devine, S. E. (2007). Which transposable elements are active in the human genome? *Trends Genet.* 23, 183–191. doi: 10.1016/j.tig.2007.02.006
- Moon, S., Jung, K. H., Lee, D. E., Jiang, W. Z., Koh, H. J., Heu, M. H., et al. (2006). Identification of active transposon *dTok*, a member of the *hAT* family, in rice. *Plant Cell Physiol.* 47, 1473–1483. doi: 10.1093/pcp/plc012
- Naito, K., Cho, E., Yang, G., Campbell, M. A., Yano, K., Okumoto, Y., et al. (2006). Dramatic amplification of a rice transposable element during recent domestication. *Proc. Natl. Acad. Sci. U. S. A.* 103, 17620–17625. doi: 10.1073/pnas.0605421103
- Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C. N., Richardson, A. O., et al. (2009). Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* 461, 1130–1134. doi: 10.1038/nature08479
- Nakazaki, T., Okumoto, Y., Horibata, A., Yamahira, S., Teraishi, M., Nishida, H., et al. (2003). Mobilization of a transposon in the rice genome. *Nature* 421, 170–172. doi: 10.1038/nature01219
- Picault, N., Chaparro, C., Piegu, B., Stenger, W., Formey, D., Llauro, C., et al. (2009). Identification of an active LTR retrotransposon in rice. *Plant J.* 58, 754–765. doi: 10.1111/j.1365-313X.2009.03813.x
- Piegu, B., Guyot, R., Picault, N., Roulin, A., Sanyal, A., Kim, H., et al. (2006). Doubling genome size without polyploidization: dynamics of retrotransposon-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res.* 16, 1262–1269. doi: 10.1101/gr.5290206
- Quadana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddeloh, J. A., et al. (2016). The *Arabidopsis thaliana* mobilome and its impact at the species level. *Elife* 5:e15716. doi: 10.7554/eLife.15716.046
- Questa, J. I., Walbot, V., and Casati, P. (2010). Mutator transposon activation after UV-B involves chromatin remodeling. *Epigenetics* 5, 352–363. doi: 10.4161/epi.5.4.11751
- Rubin, G. M., Kidwell, M. G., and Bingham, P. M. (1982). The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. *Cell* 29, 987–994. doi: 10.1016/0092-8674(82)90462-7
- Schneider, T. D., and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 18, 6097–6100. doi: 10.1093/nar/18.20.6097
- Schrader, L., and Schmitz, J. (2019). The impact of transposable elements in adaptive evolution. *Mol. Ecol.* 28, 1537–1549. doi: 10.1111/mec.14794

- Slotkin, R. K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* 8, 272–285. doi: 10.1038/nrg2072
- Song, X., and Cao, X. (2017). Transposon-mediated epigenetic regulation contributes to phenotypic diversity and environmental adaptation in rice. *Curr. Opin. Plant Biol.* 36, 111–118. doi: 10.1016/j.pbi.2017.02.004
- Tang, Y., Ma, X., Zhao, S., Xue, W., Zheng, X., Sun, H., et al. (2019). Identification of an active miniature inverted-repeat transposable element *mJing* in rice. *Plant J.* 98, 639–653. doi: 10.1111/tpj.14260
- Tsugane, K., Maekawa, M., Takagi, K., Takahara, H., Qian, Q., Eun, C. H., et al. (2006). An active DNA transposon *nDart* causing leaf variegation and mutable dwarfism and its related elements in rice. *Plant J.* 45, 46–57. doi: 10.1111/j.1365-3113X.2005.02600.x
- Vitte, C., Fustier, M. A., Alix, K., and Tenaillon, M. I. (2014). The bright side of transposons in crop evolution. *Brief Funct. Genomics* 13, 276–295. doi: 10.1093/bfpg/elu002
- Wang, H., Chai, Y., Chu, X., Zhao, Y., Wu, Y., Zhao, J., et al. (2009). Molecular characterization of a rice mutator-phenotype derived from an incompatible cross-pollination reveals transgenerational mobilization of multiple transposable elements and extensive epigenetic instability. *BMC Plant Biol.* 9:63. doi: 10.1186/1471-2229-9-63
- Wessler, S. R. (1988). Phenotypic diversity mediated by the maize transposable elements *Ac* and *Spm*. *Science* 242, 399–405. doi: 10.1126/science.2845581
- Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., et al. (2007). A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982. doi: 10.1038/nrg2165

Conflict of Interest: M-HD was employed by Zhejiang Zhengjingyuan Pharmacy Chain Co., Ltd., and Hangzhou Zhengcaiyan Pharmaceutical Co., Ltd., Hangzhou, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Zhao, Yan, Wei, Zheng, Yue, Alam, Smartt, Duan and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.