A prolific and robust whole-genome genotyping method using PCR amplification via primer-template mismatched annealing

Sheng Zhao[1,2†], Cuicui Zhang[2†], Liqun Wang[1†], Minxuan Luo[1, 2], Peng Zhang[2], Yue Wang[2], Waqar Afzal Malik[2], Yue Wang[2], Peng Chen[2], Xianjin Qiu[3], Chongrong Wang[4], Hong Lu[2], Yong Xiang[2], Yuwen Liu[2], Jue Ruan[2], Qian Qian[2*], Haijian Zhi[1*], Yuxiao Chang[2*]

SUPPORTING INFORMATION

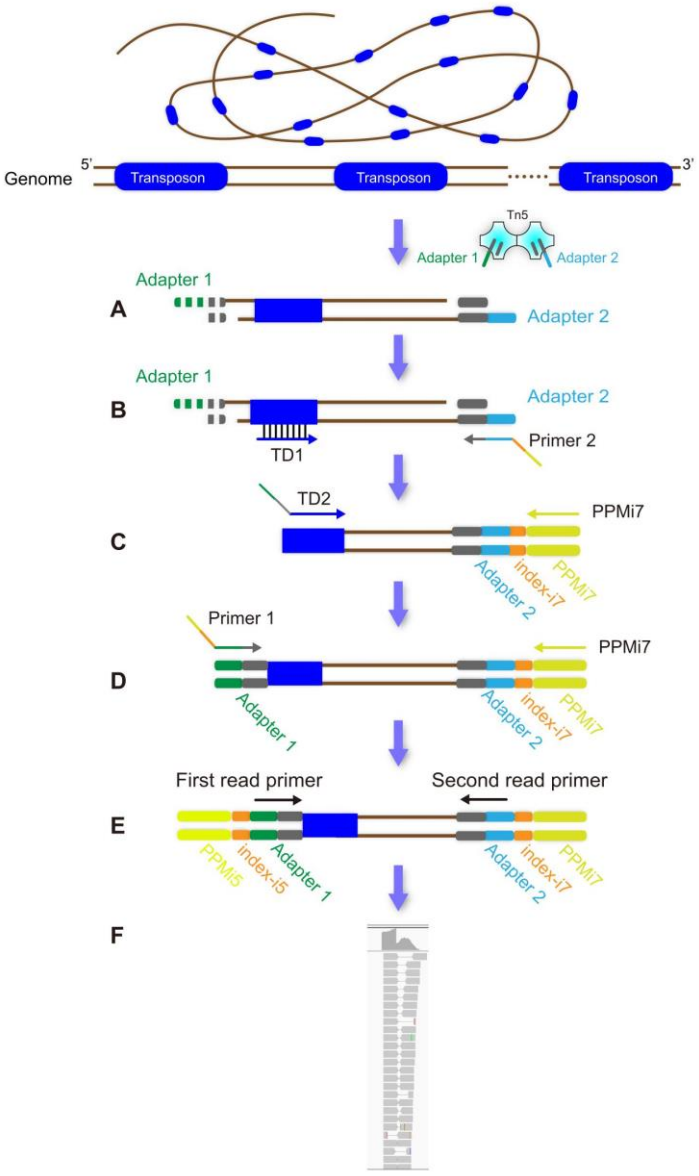


Figure S1. Initial design strategy of the whole-genome genotyping method based on transposon display using specific nested primers

(**A**) Fragmentation of genomic DNA and adapter ligation by the Tn5 transposase. (**B**) In the primary PCR, hundreds of flanking sequences of a certain transposon are amplified using the primer pair consisting of the specific primer TD1 annealing to the transposon sequence and Primer 2 targeting the ligated Adapter 2. (**C**) In the secondary PCR, the nested transposon primer TD2 with a tail compatible with Illumina sequencing is used to increase the PCR specificity and further enrich the target fragments. (**D**) In the tertiary PCR, all sequence elements required for Illumina sequencing are added. (**E**) The final library structure is prepared for Illumina-based sequencing. (**F**) An expected distribution of sequencing reads when aligning to a target transposon sequence.

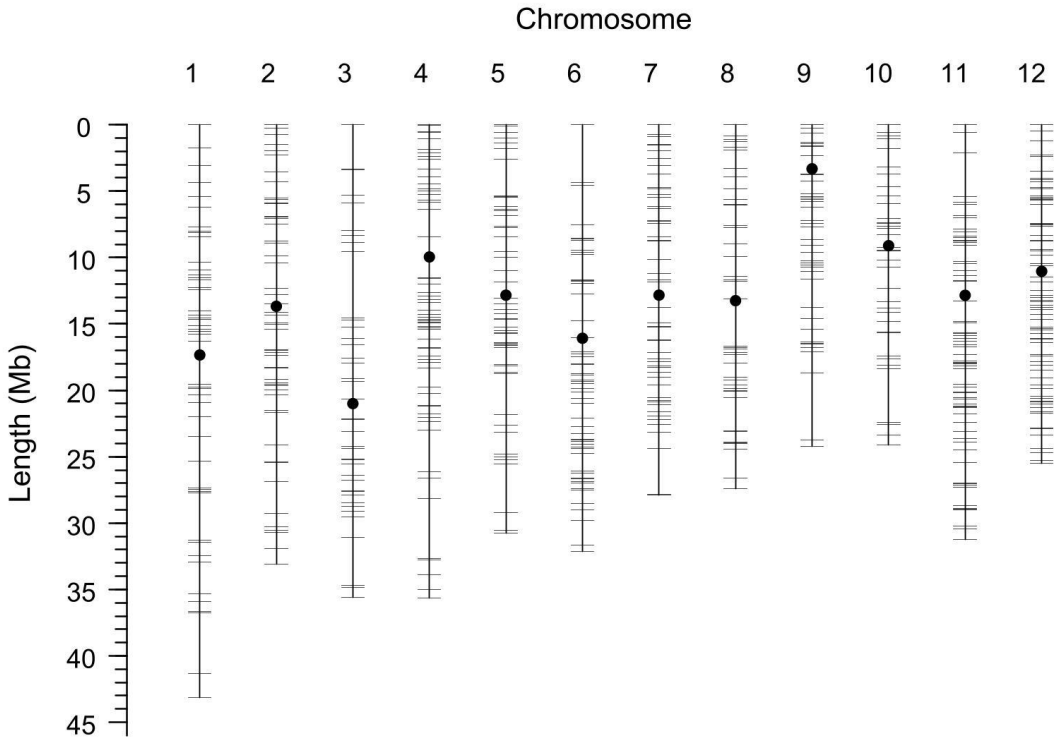


Figure S2. Distribution of 641 copies of the *Rim2/Hipa* CACTA transposon detected in R498

Black dots indicate the positions of centromeres.

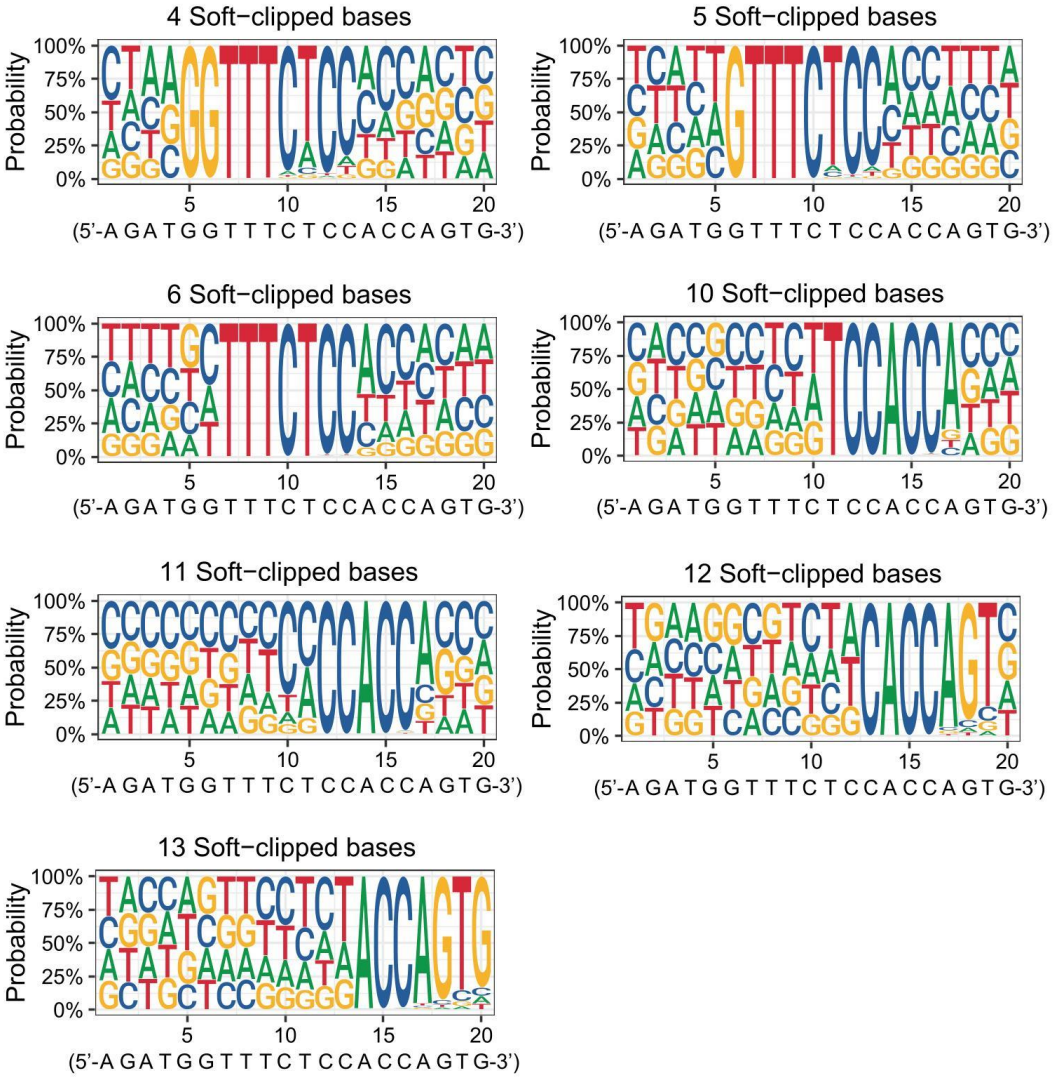


Figure S3. Stacked histograms displaying the genome sequences of FBI-*Rim2'* PTMA loci in the R498 genome

PTMA tags produced by different lengths of soft-clipped read1 are shown, with the *Rim2* primer sequence shown at the bottom.

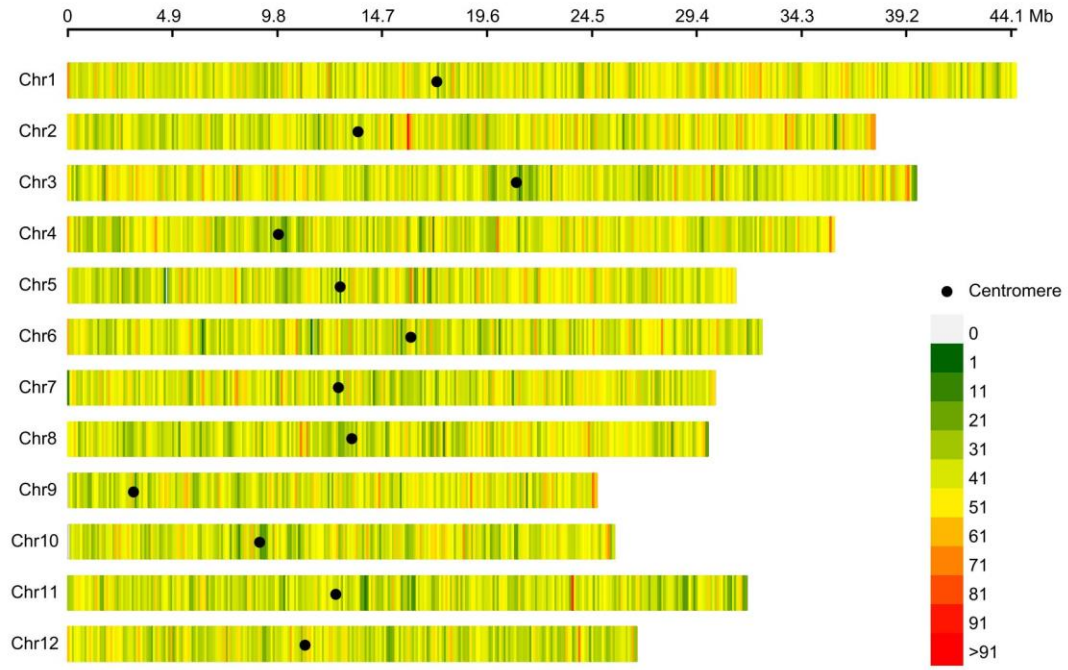


Figure S4. Density plot displaying the distribution of 162K PTMA tags in rice cultivar R498 Number of PTMA tags are calculated within a 0.1-Mb window size, and the coverage depth threshold of PTMA tags is ≥ 5×.

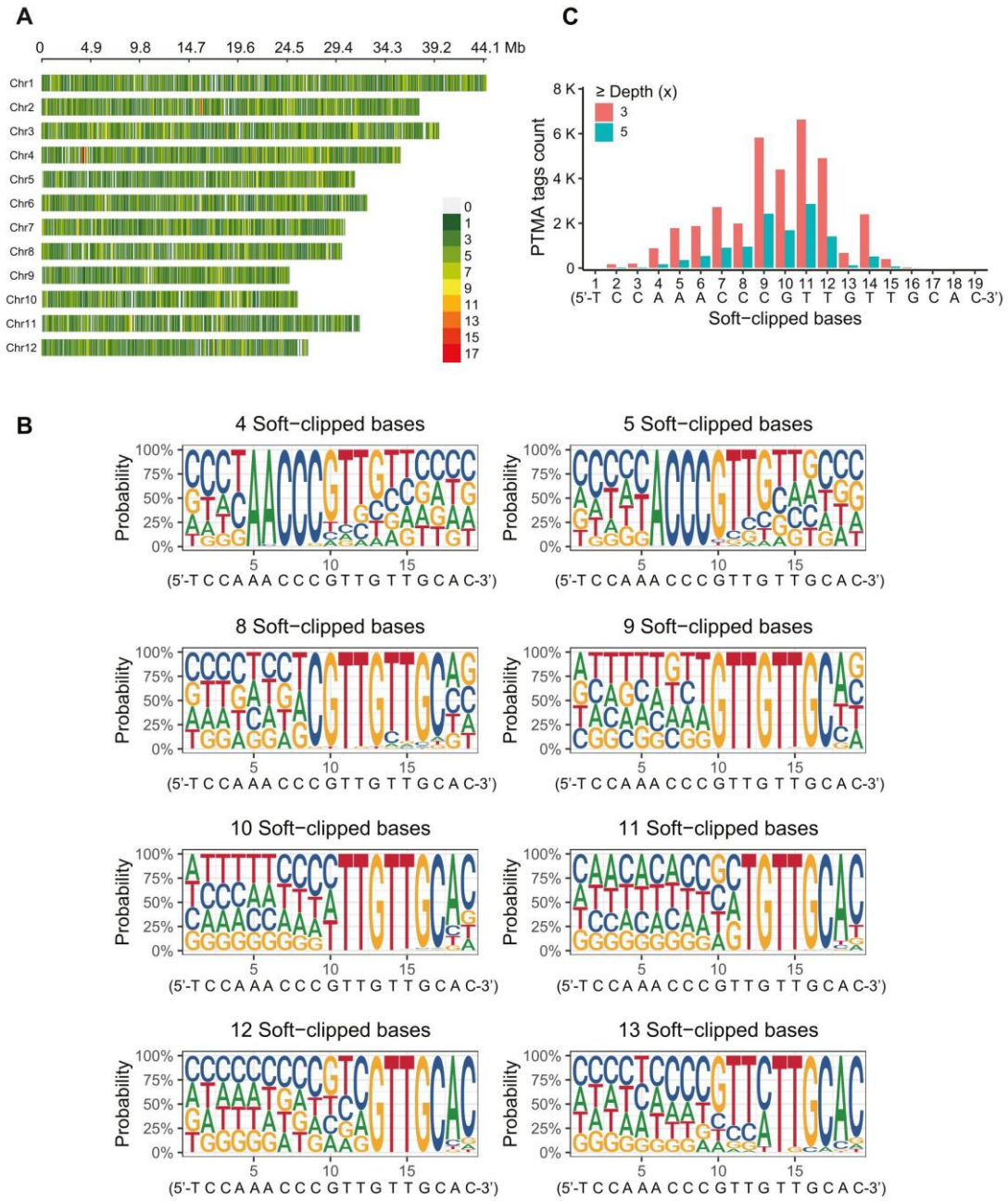


Figure S5. PTMA tags detected in rice R498 using the FBI-*Pi2*-1 primer

(**A**) Density plot showing the distribution of 12K PTMA tags (depth ≥ 5×) amplified by FBI-*Pi2*-1. The window size for tag density plot is 0.1 Mb. (**B**) Stacked histograms displaying the genome sequences of the FBI-*Pi2*-1′ PTMA loci in the R498 genome. PTMA tags produced by 4–5-nt and 8–13-nt soft-clipped read1 are shown, with the specific FBI-*Pi2*-1 primer sequence shown at the bottom. (**C**) Distribution of PTMA tag counts detected by different lengths of soft-clipped read1 in R498.

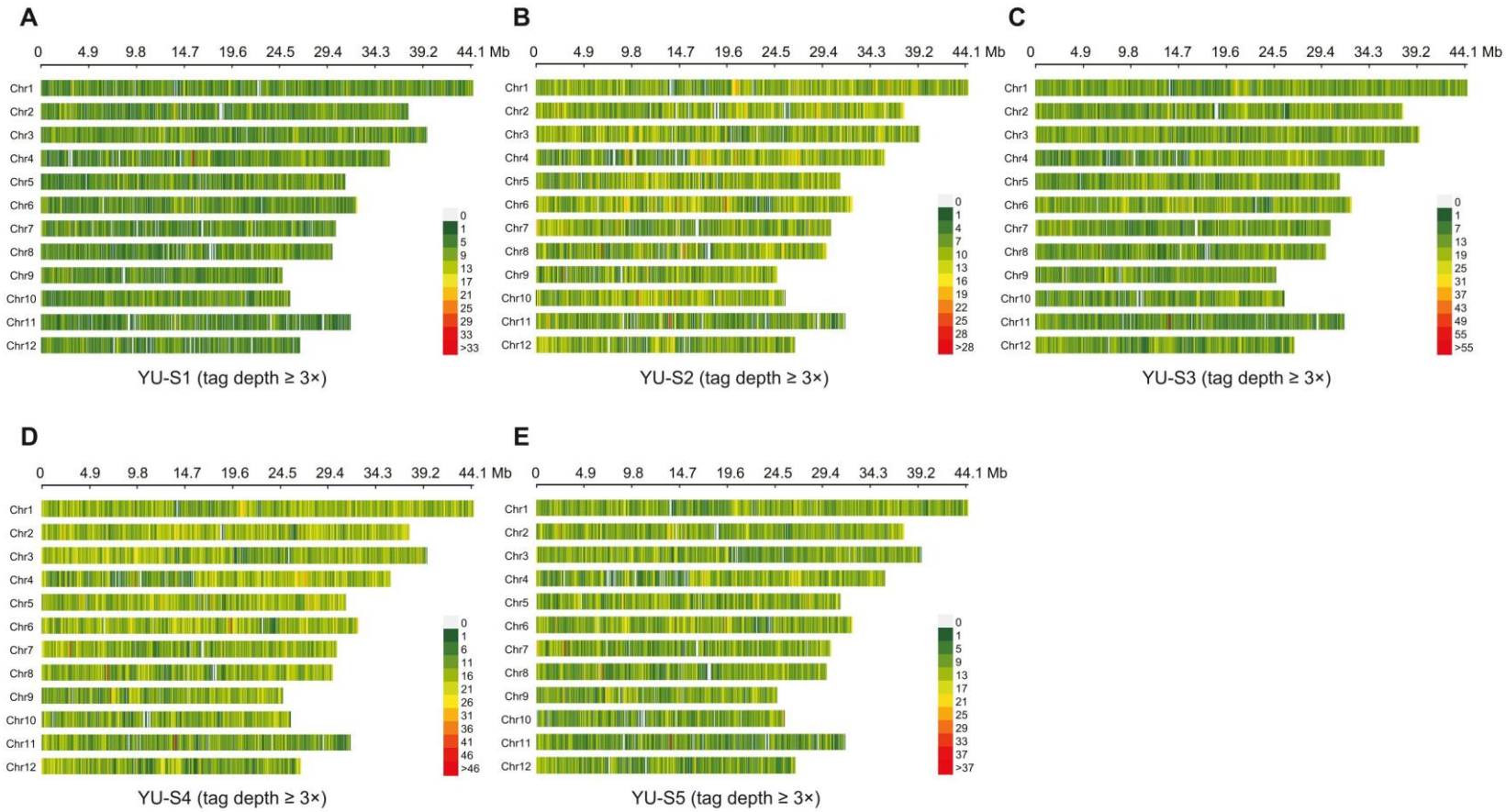
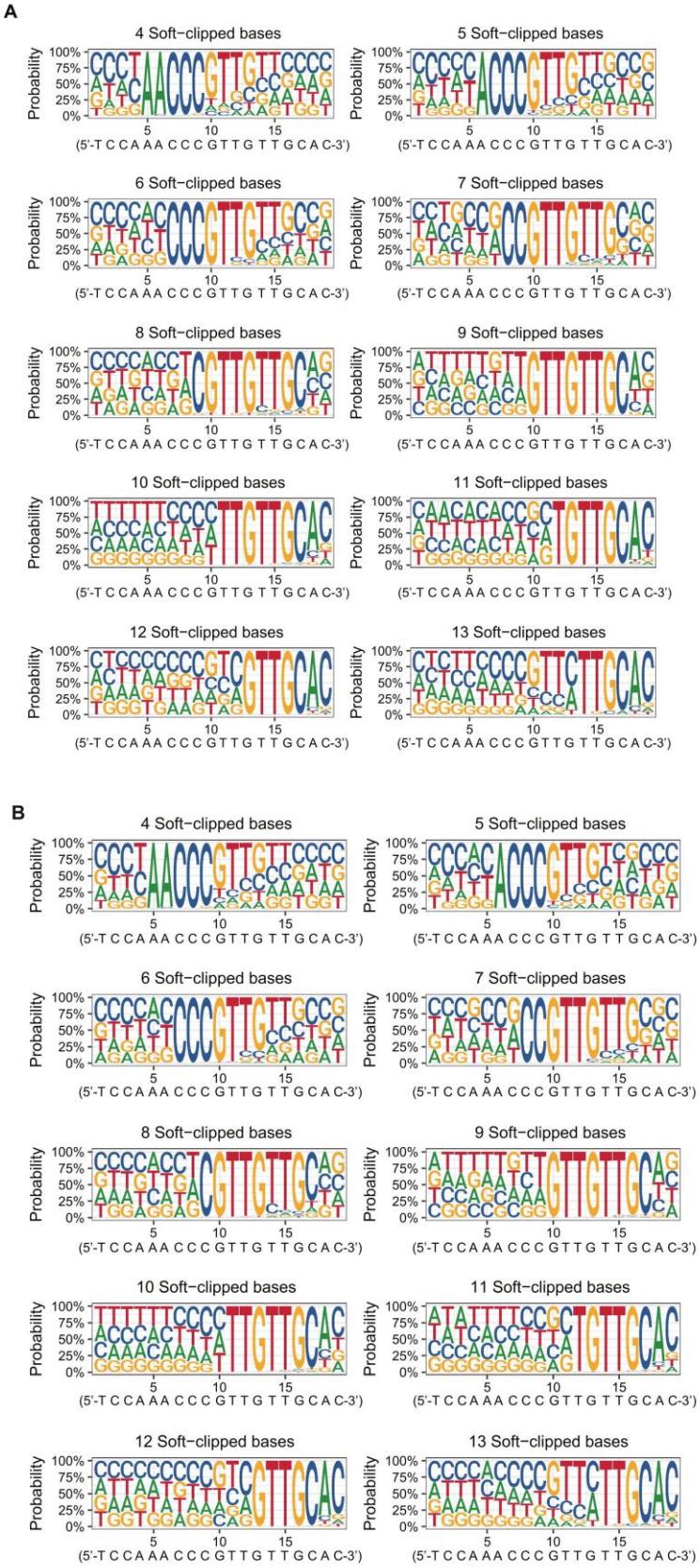


Figure S6. Density plot representing the distribution of total PTMA tags in five rice samples while using the FBI-*Pi2*-1 primer

The coverage depth threshold of PTMA tags is ≥ 3× and ≥ 5×, as indicated. The window size for tags density plot is 0.1 Mb.


A
4 Soft−clipped bases
5 Soft−clipped bases
6 Soft−clipped bases
7 Soft−clipped bases
8 Soft−clipped bases
9 Soft−clipped bases
10 Soft−clipped bases
11 Soft−clipped bases
12 Soft−clipped bases
13 Soft−clipped bases
Probability
100%
75%
50%
25%
0%
5
10
15
(5'-T C C A A A C C C G T T G T T G C A C-3')
B
4 Soft−clipped bases
5 Soft−clipped bases
6 Soft−clipped bases
7 Soft−clipped bases
8 Soft−clipped bases
9 Soft−clipped bases
10 Soft−clipped bases
11 Soft−clipped bases
12 Soft−clipped bases
13 Soft−clipped bases
Probability
100%
75%
50%
25%
0%
5
10
15
(5'-T C C A A A C C C G T T G T T G C A C-3')

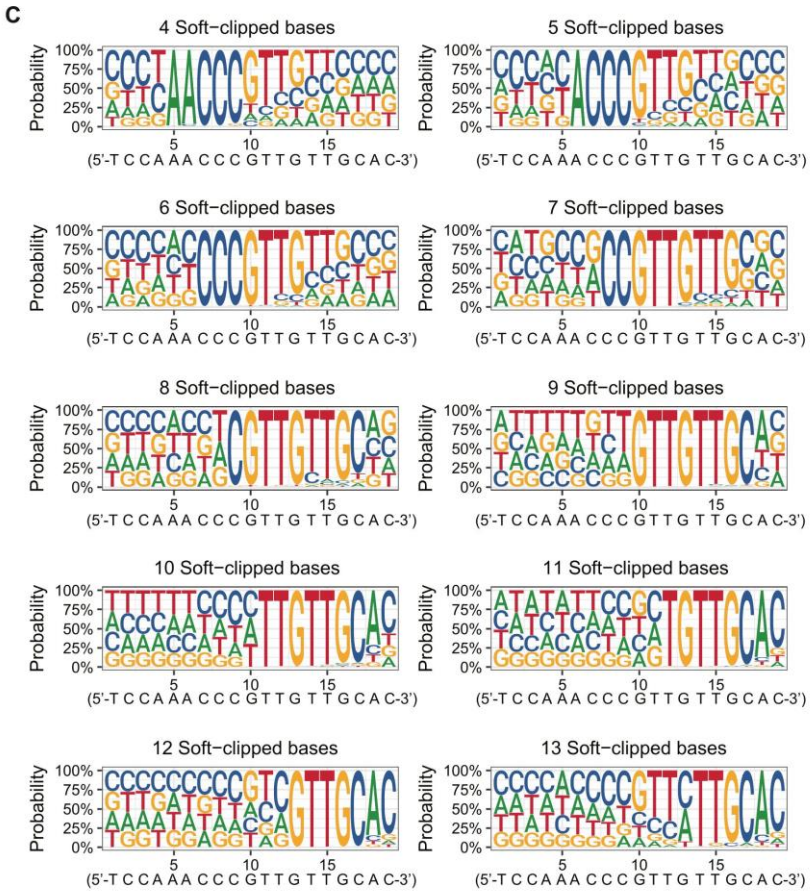


Figure S7. Stacked histograms displaying the genome sequences of the FBI-*Pi2*-1′ PTMA loci in three rice breeding lines

(**A**) YU-S1; (**B**) YU-S3; and (**C**) YU-S5. For each sample, PTMA tags produced by 4–13-nt soft-clipped read1 are shown, with the specific FBI-*Pi2*-1 primer sequence shown at the bottom.

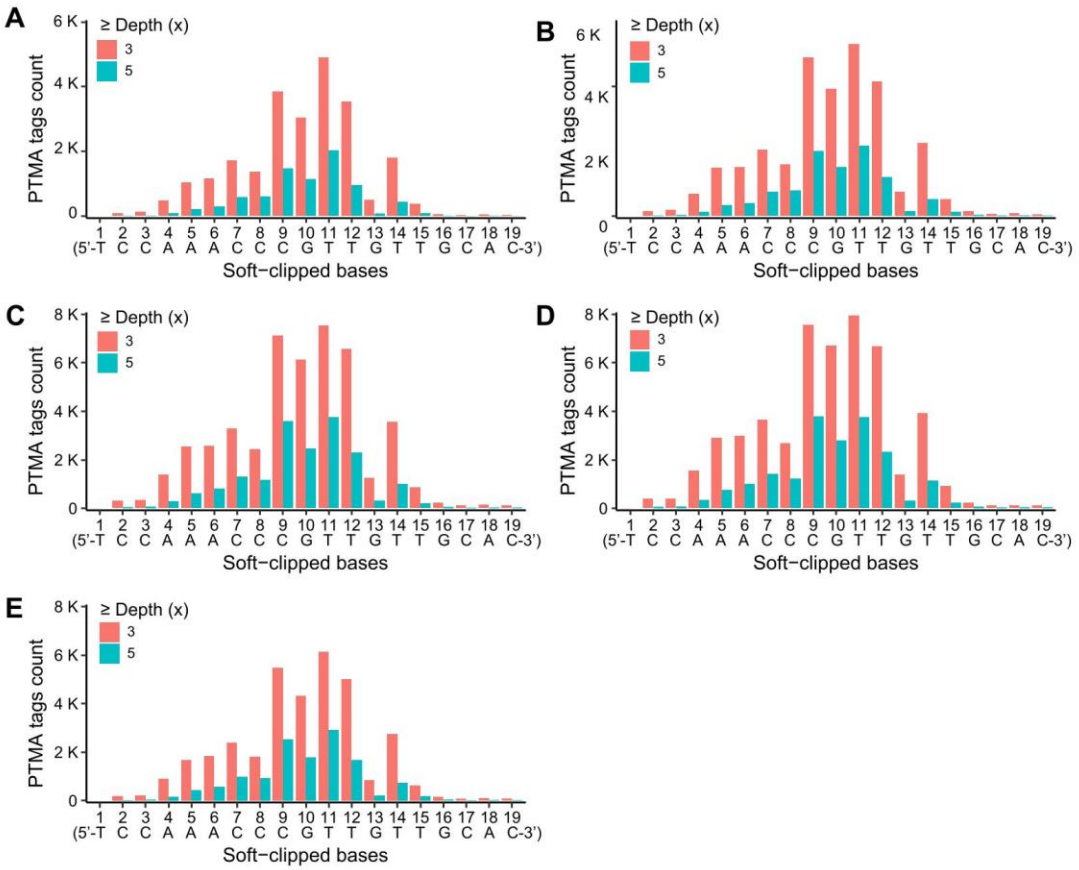


Figure S8. Distribution of PTMA tag counts detected by different lengths of soft-clipped read1 in five rice breeding lines

(**A**) YU-S1; (**B**) YU-S2; (**C**) YU-S3; (**D**) YU-S4; and (**E**) YU-S5. The specific FBI-*Pi2*-1 primer sequence is shown at the bottom. The coverage depth threshold of PTMA tags is ≥ 3× and ≥ 5×, as indicated.

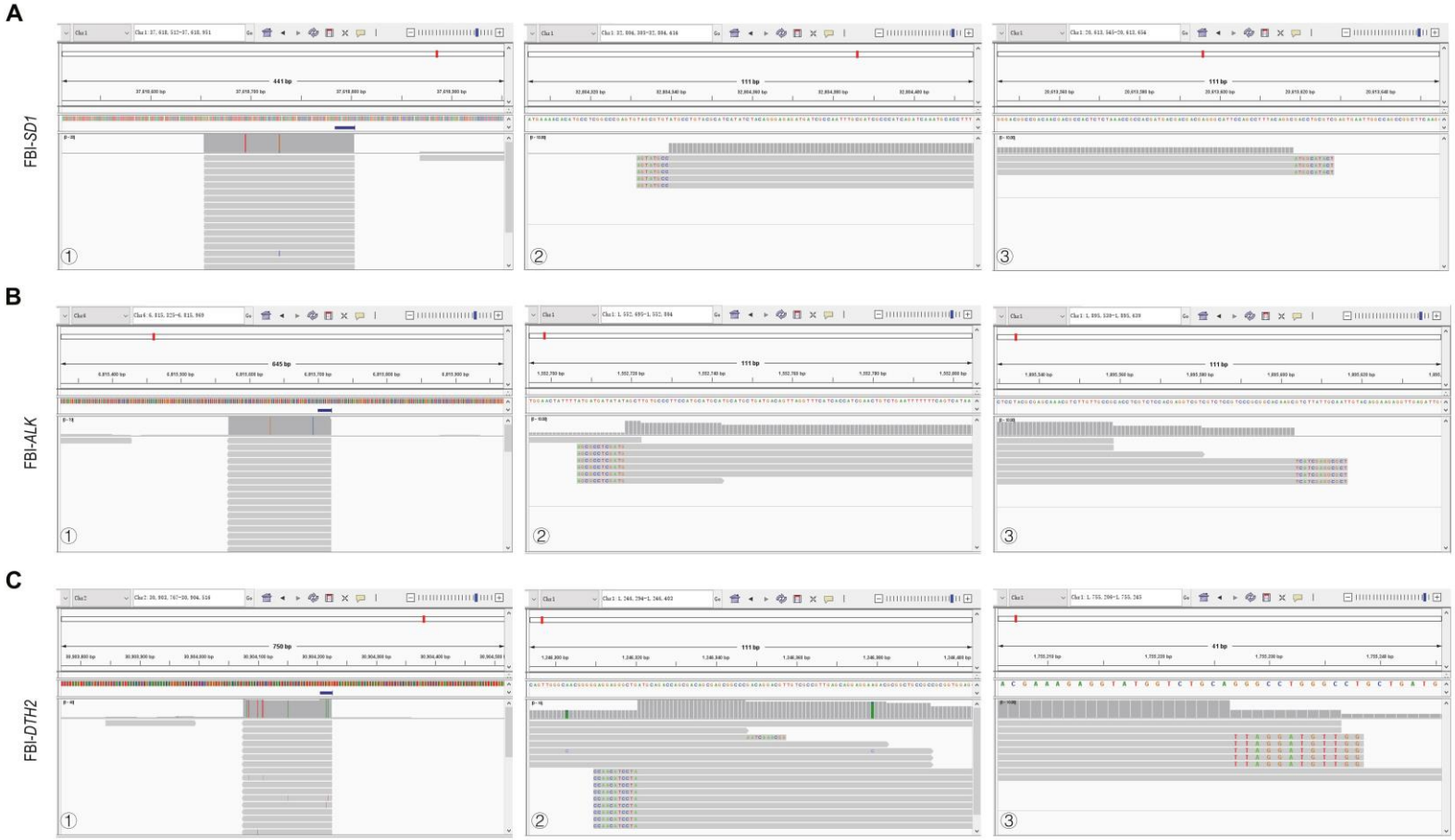


Figure S9. Amplification of PTPA and PTMA reads with three FBI-seq primers on the R498 genome

(**A**) FBI-*SD1*; (**B**) FBI-*ALK*; and (**C**) FBI-*DTH2*. ① Alignment of PTPA reads and the corresponding target genes (*SD1*, *ALK*, and *DTH2*) were deeply covered; ② and ③ PTMA tags formed by read1 aligning to the positive and negative strands of the R498 genome with few soft-clipped bases (colored bases), respectively.

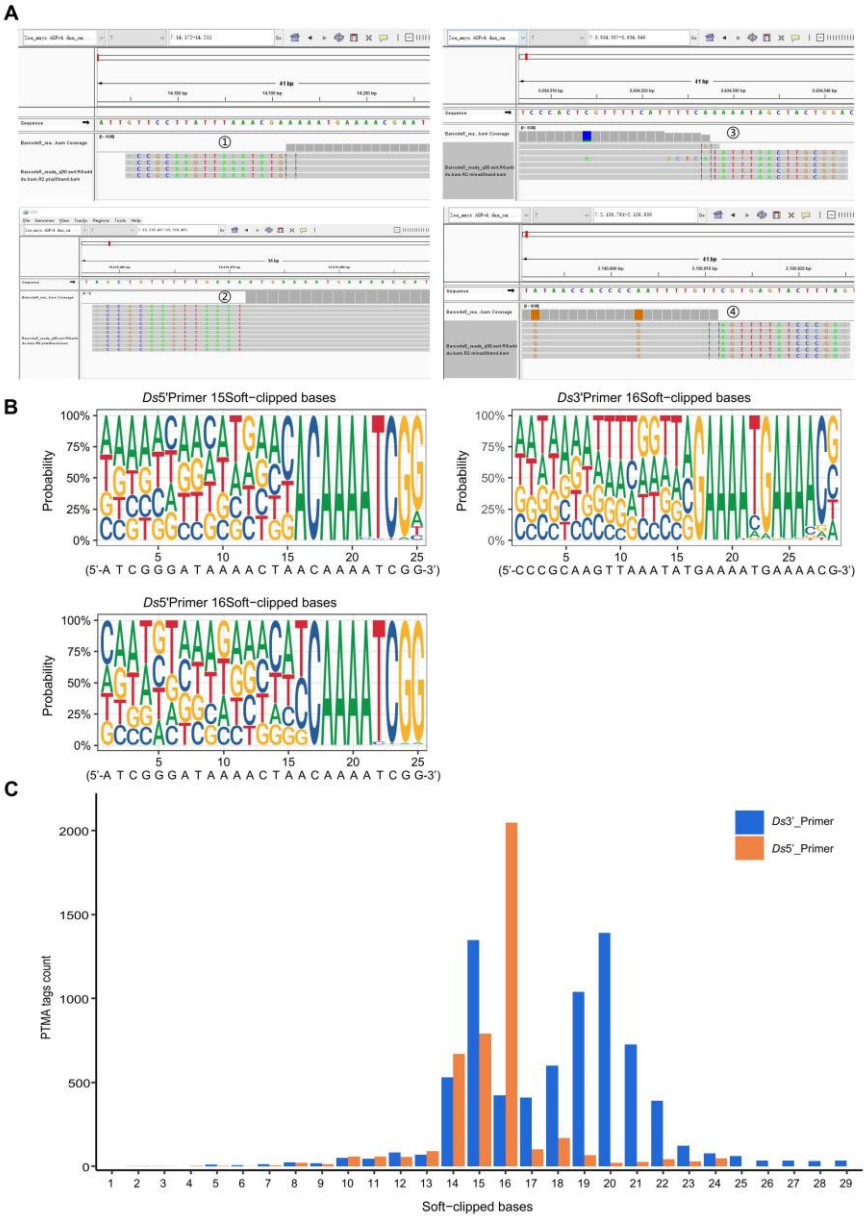


Figure S10. PTMA tags detected in randomly selected sample_1 of maize *Ds* mutants from TEAseq

(**A**), Screenshots of the IGV browser showing PTMA tags. Left [① ②] and right [③ ④] panels indicate PTMA tags formed by read2 aligning to the positive and negative strands of the maize B73_V4.0 reference genome with few soft-clipped bases (colored bases). (**B**), Stacked histograms displaying the genome sequences of *Ds*3′ and *Ds*5′ PTMA loci in the B73 genome. PTMA tags produced by 15- and 16-nt soft-clipped read2 are shown, and corresponding *Ds*3′ and *Ds*5′ primer sequences are shown at the bottom. (**C**), Distribution of 12K PTMA tags detected by different lengths of soft-clipped read2 in sample_1. The coverage depth threshold of PTMA tags is ≥ 3×.

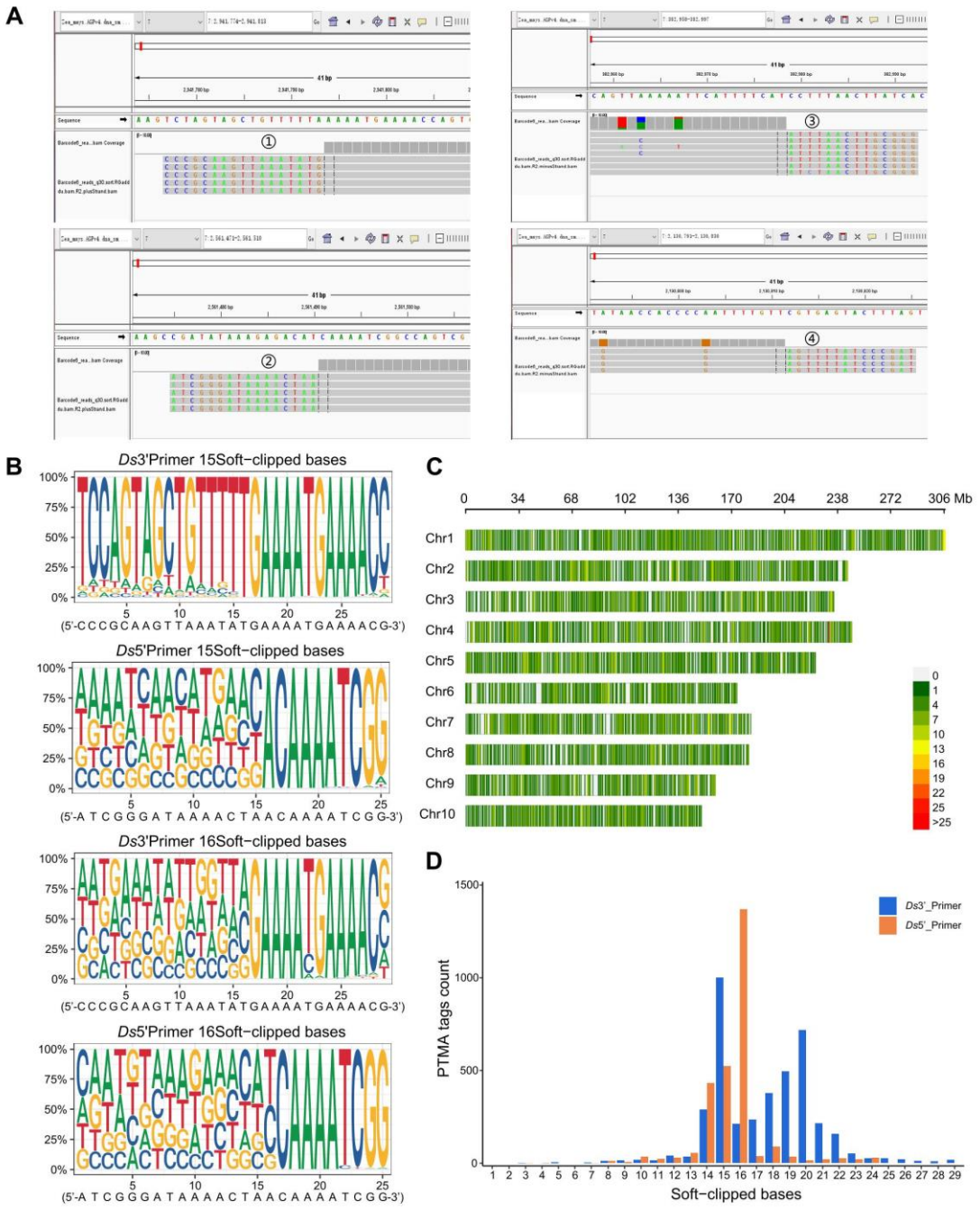


Figure S11. PTMA tags detected in randomly selected sample_2 of maize *Ds* mutants from TEAseq

(**A**) Screenshots of the IGV browser showing PTMA tags. Left [① ②] and right [③ ④] panels indicate PTMA tags formed by read2 aligning to the positive and negative strands of the maize B73_V4.0 reference genome with few soft-clipped bases (colored bases). (**B**) Stacked histograms displaying the genome sequences of *Ds*3′ and *Ds*5′ PTMA loci in the B73 genome. PTMA tags produced by 15- and 16-nt soft-clipped read2 are shown, and

corresponding *Ds*3′ and *Ds*5′ primer sequences are shown at the bottom. (**C**) Density distribution of 7K PTMA tags (depth ≥ 3×) detected in sample_2 mutant line. Number of PTMA tags are calculated within a 1.0-Mb window size. (**D**) Distribution of PTMA tag counts detected by different lengths of soft-clipped read2 in sample_2.

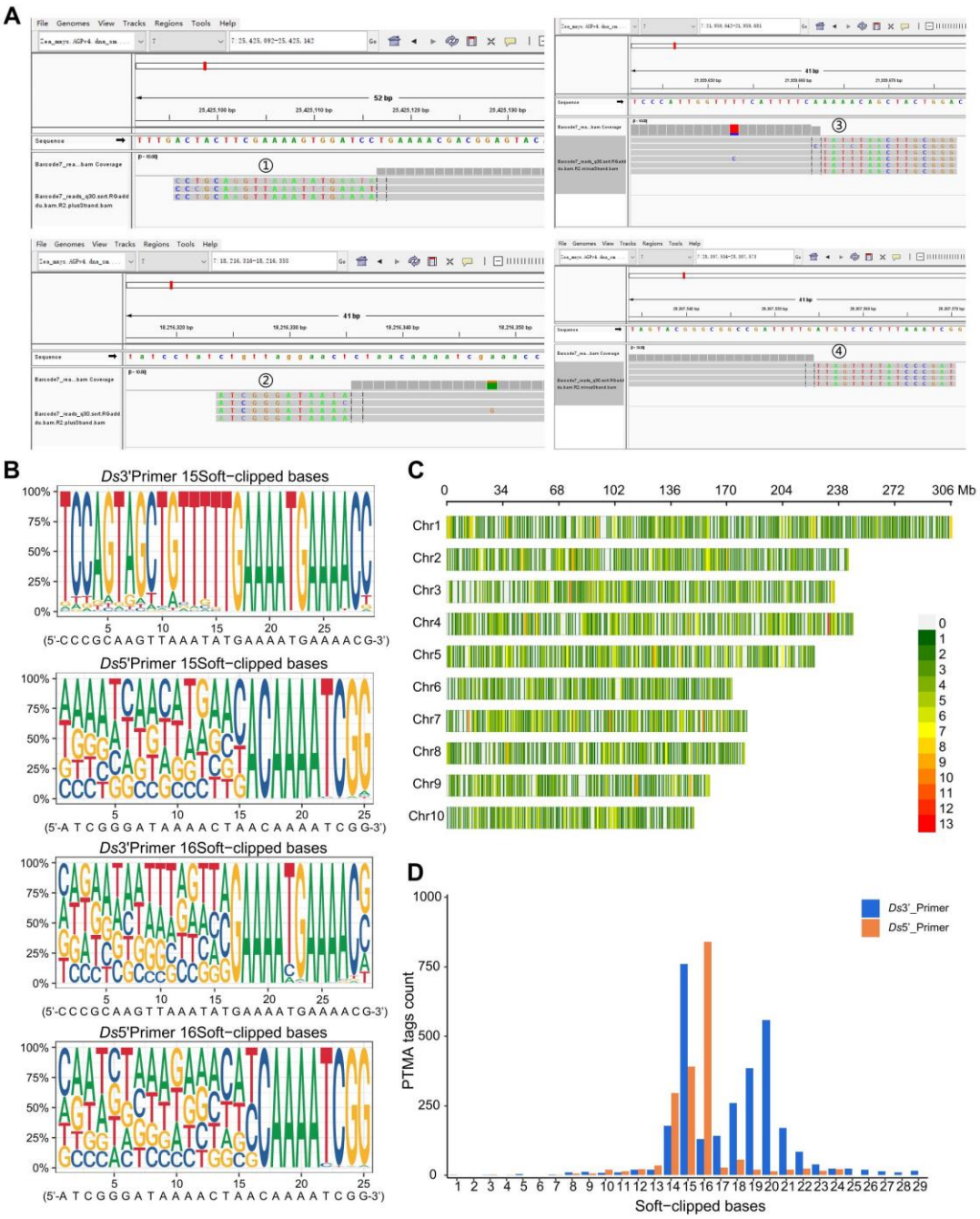


Figure S12. PTMA tags detected in randomly selected sample_3 of maize *Ds* mutants from TEAseq

(**A**) Screenshots of the IGV browser showing PTMA tags. Left [① ②] and right [③ ④] panels indicate PTMA tags formed by read2 aligning to the positive and negative strands of the maize B73_V4.0 reference genome with few soft-clipped bases (colored bases), respectively. (**B**) Stacked histograms displaying the genome sequences of *Ds*3′ and *Ds*5′ PTMA loci in the B73 genome. PTMA tags produced by 15- and 16-nt soft-clipped read2 are shown,

and corresponding *Ds*3′ and *Ds*5′ primer sequence are shown at the bottom. (**C**) Density distribution of 5K PTMA tags (depth ≥ 3×) detected in sample_3 mutant line. Number of PTMA tags are calculated within a 1.0-Mb window size. (**D**) Distribution of PTMA tag counts detected by different lengths of soft-clipped read2 in sample_3.

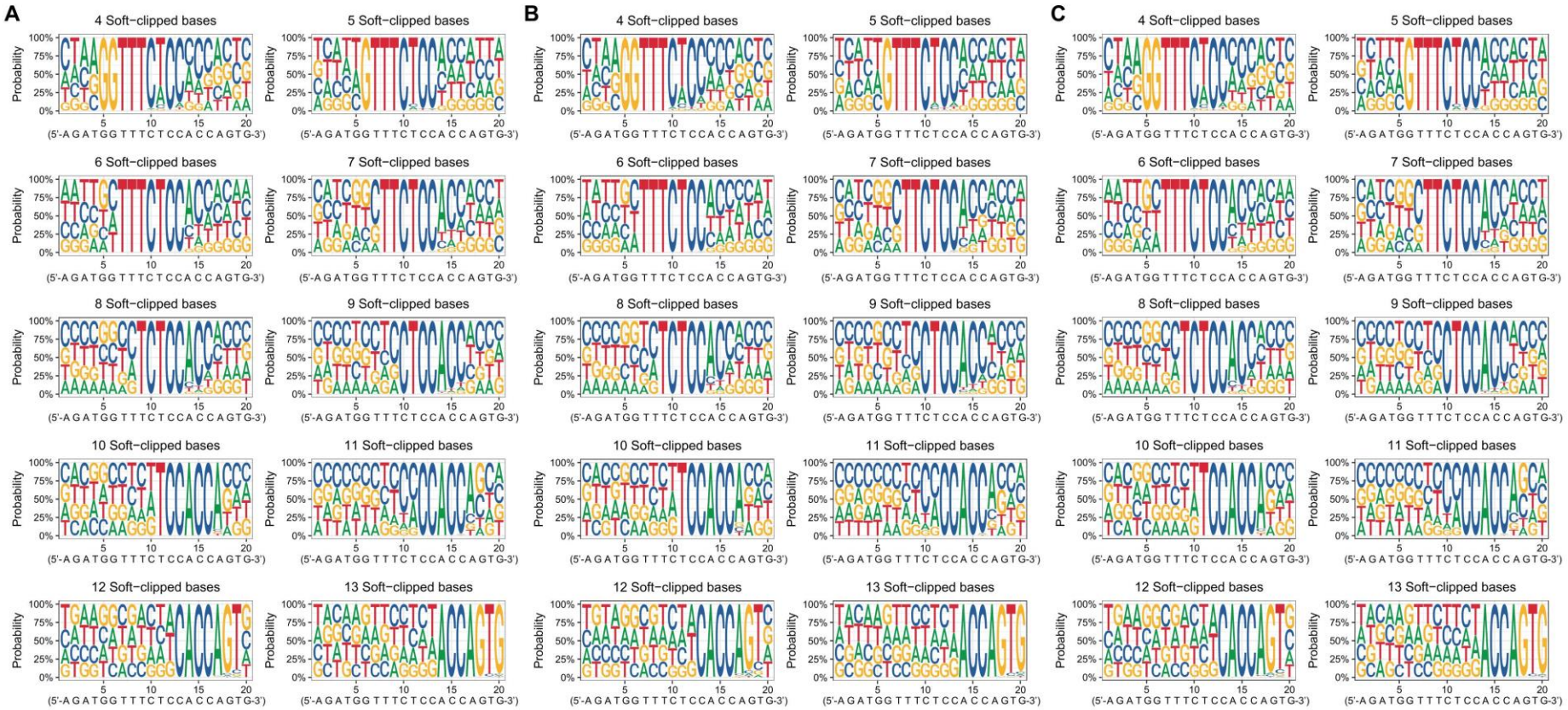


Figure S13. Stacked histograms of three replicates displaying the genome sequences of FBI-*Rim2'* PTMA loci in R498

PTMA tags produced by 4–13-nt soft-clipped read1 are shown in replicate 1 (**A**), replicate 2 (**B**), and replicate 3 (**C**). The specific FBI-*Rim2* primer sequence is shown at the bottom.

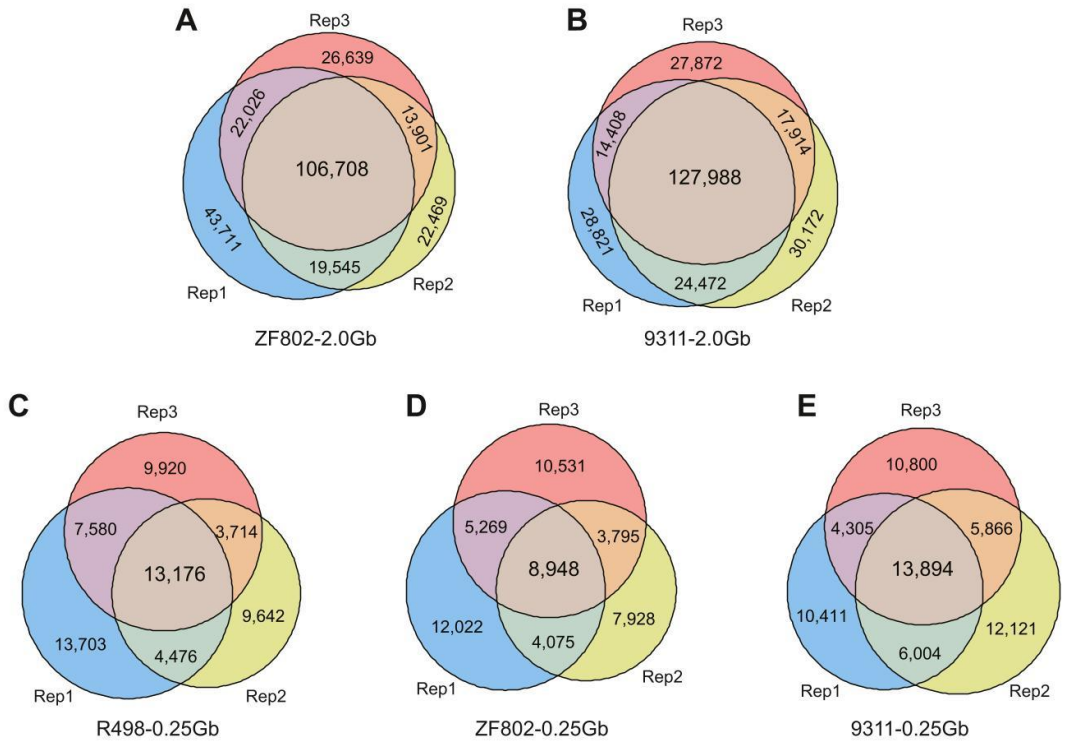


Figure S14. Number of shared PTMA tags detected in three replicates for R498, ZF802, and 9311 with sequencing data of 2.0 or 0.25 Gb

The coverage depth threshold of PTMA tags is ≥ 3×.

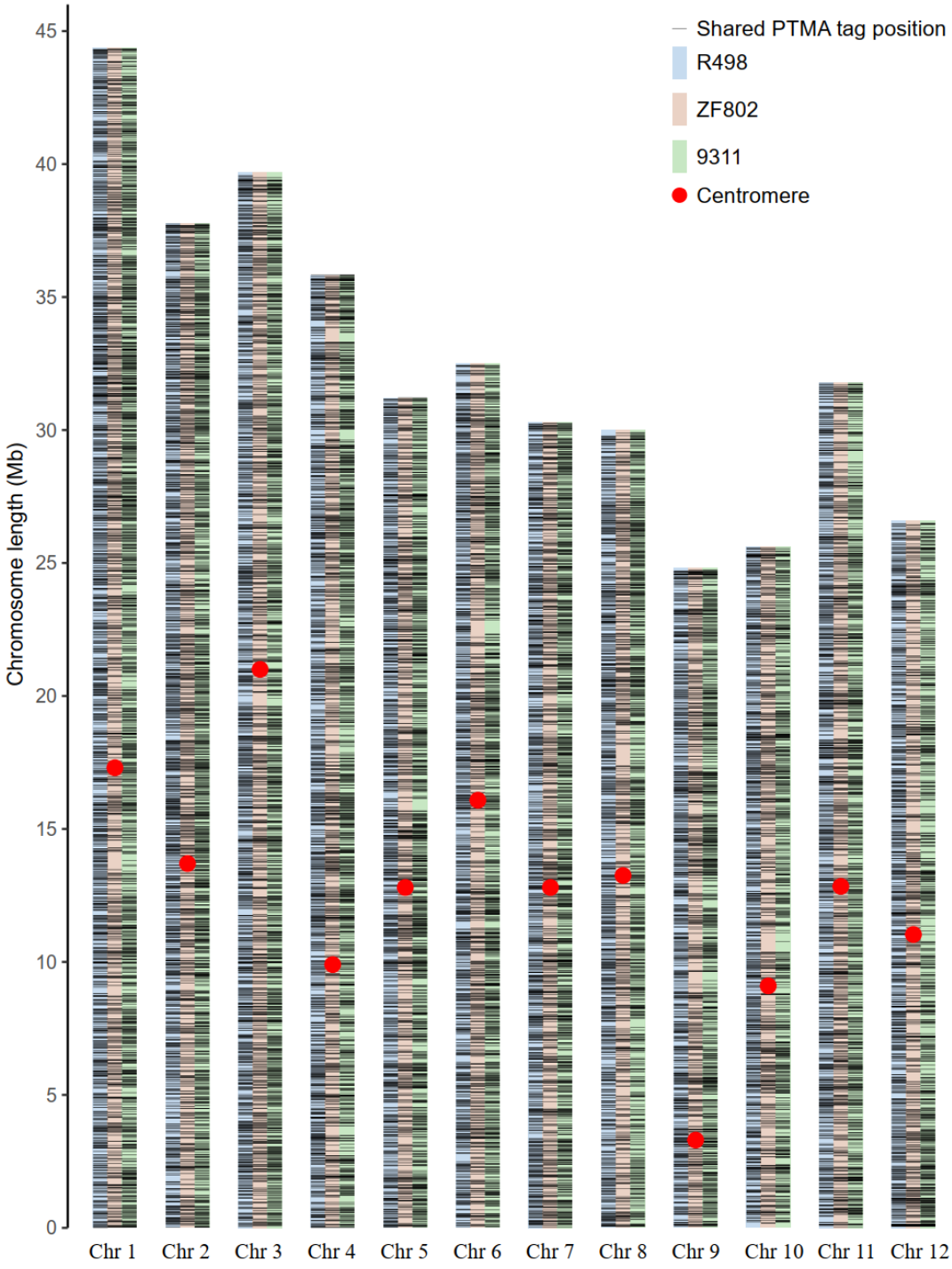


Figure S15. Density plot representing the distribution of shared PTMA tags among three technical replicates in three rice cultivars with sequencing data of 0.25 Gb

The coverage depth threshold of PTMA tags shown here is ≥ 3×.

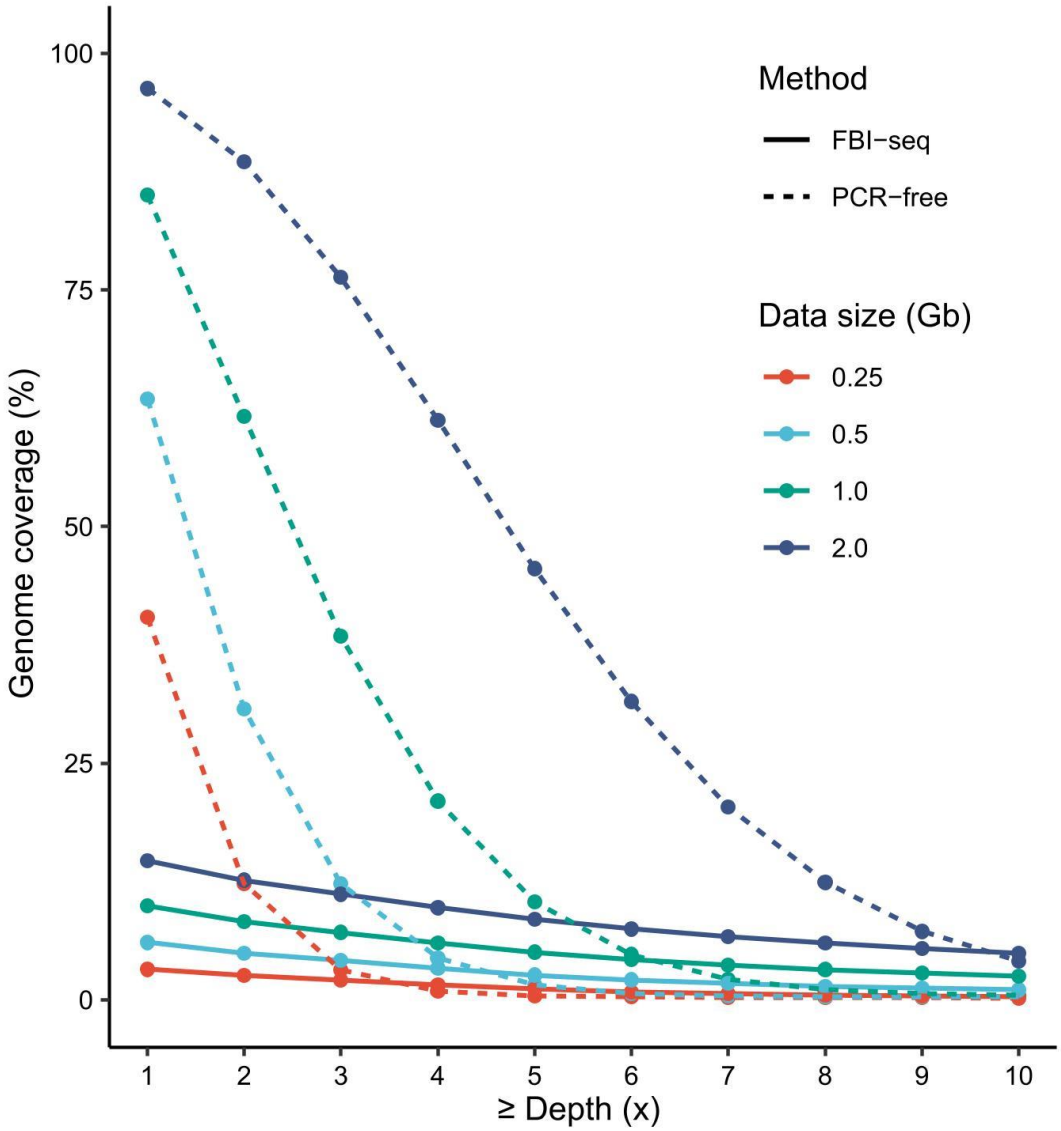


Figure S16. Genome coverage of different sequencing data (0.25–2.0 Gb) measured with FBI-seq and a PCR-free approach in R498

For FBI-seq data, low-coverage random reads are excluded.

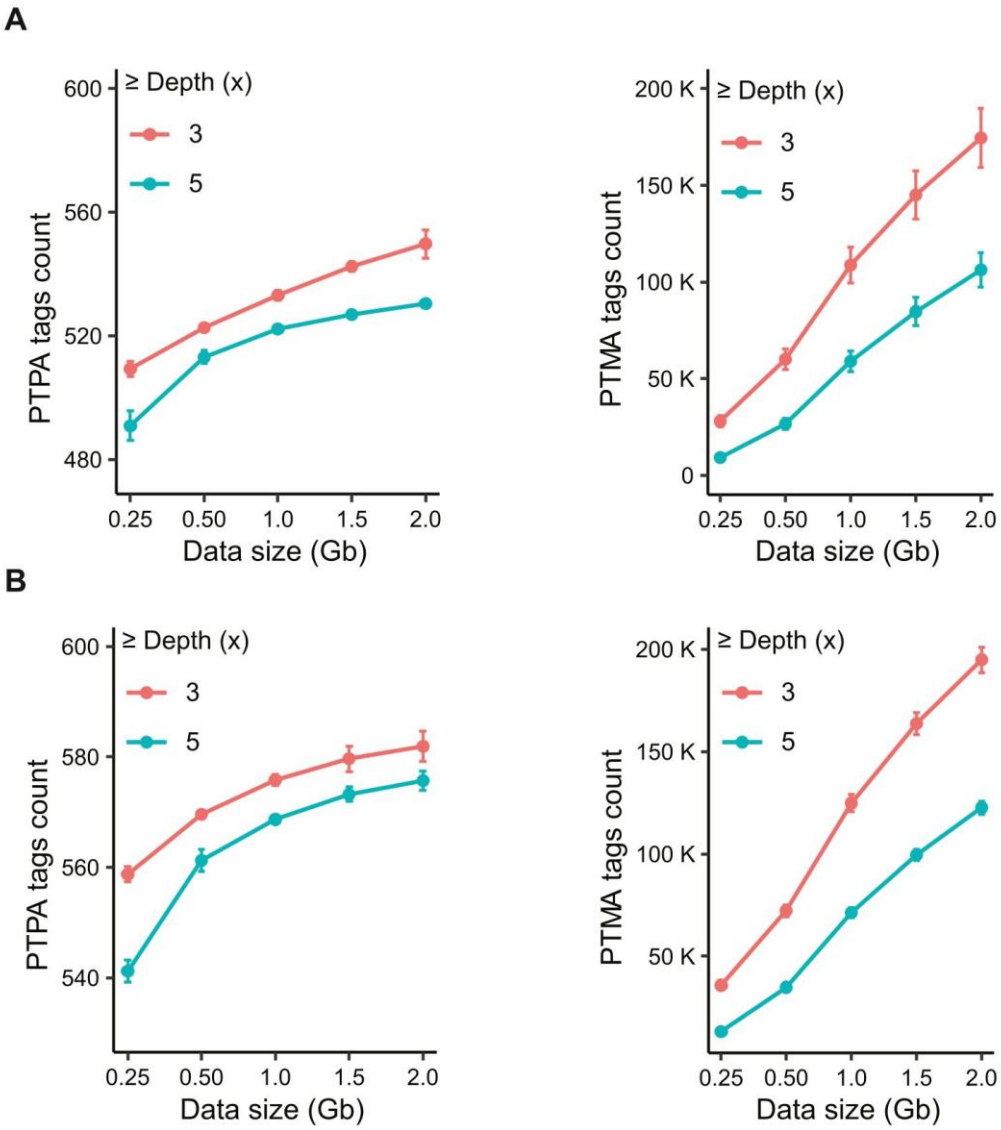
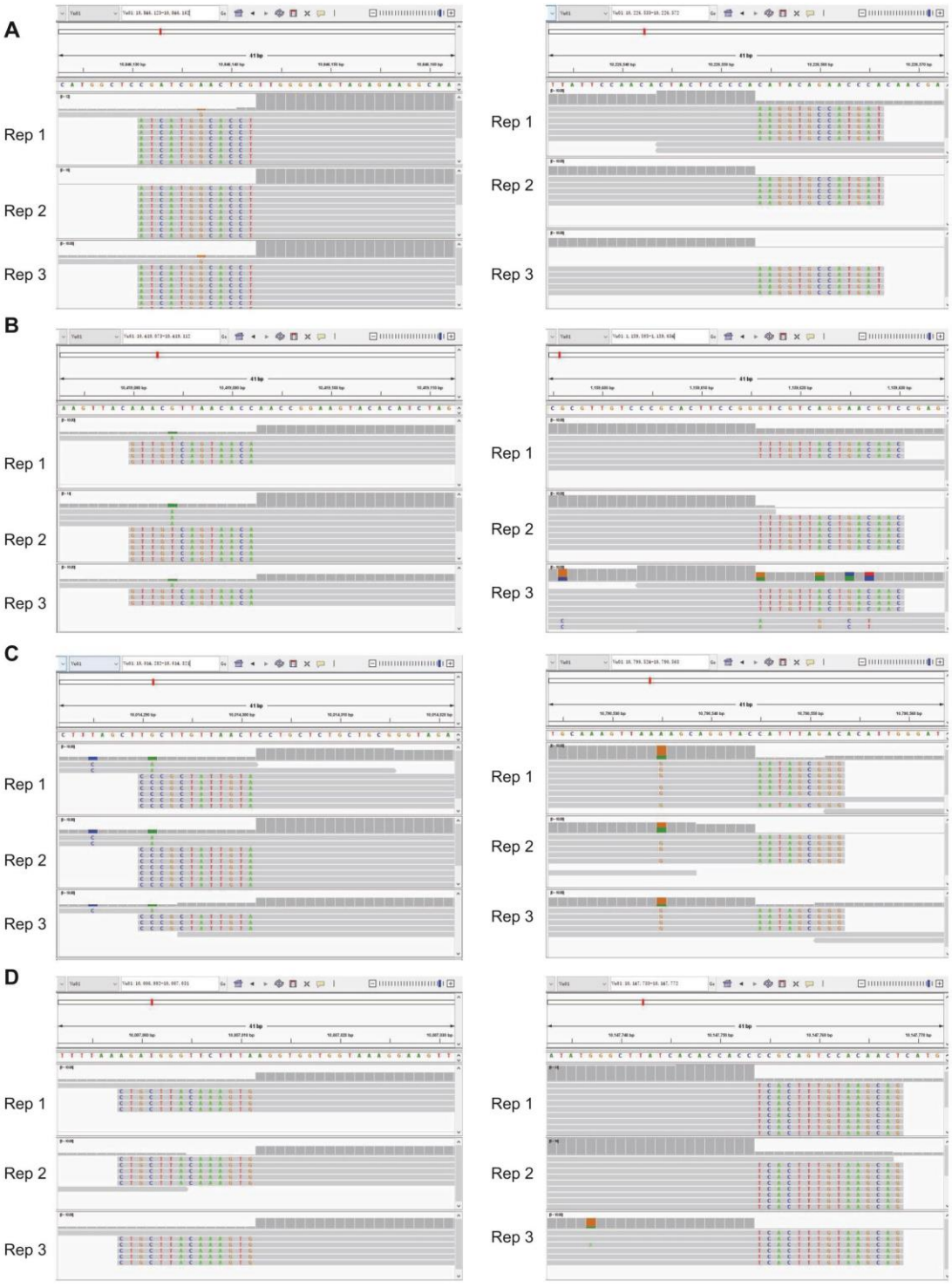


Figure S17. Number of PTPA and PTMA tags detected in rice cultivars ZF802 (A) and 9311 (B) with different sequencing coverage

The coverage depth threshold of PTPA and PTMA tags is ≥ 3× and ≥ 5×, as indicated.


A
Rep 1
Rep 2
Rep 3
Rep 1
Rep 2
Rep 3
B
Rep 1
Rep 2
Rep 3
Rep 1
Rep 2
Rep 3
C
Rep 1
Rep 2
Rep 3
Rep 1
Rep 2
Rep 3
D
Rep 1
Rep 2
Rep 3
Rep 1
Rep 2
Rep 3

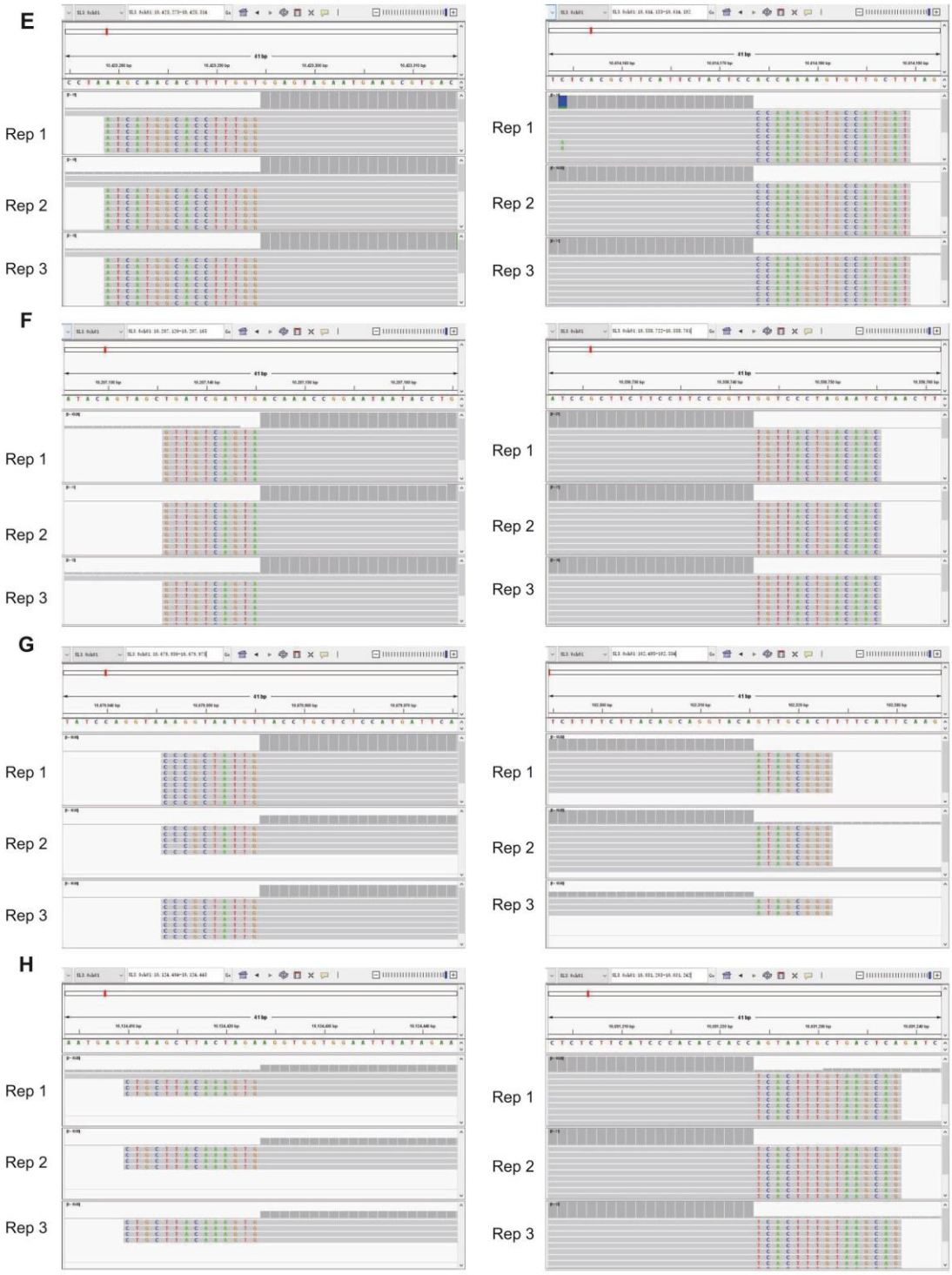


Figure S18. Screenshots of shared PTMA tags among three replicates when four FBI-seq primers are used in cowpea or tomato

For each species, four specific FBI-seq primers are designed according to rice gene sequences (*BADH2*, *Waxy*, *Rf3*, and *Rf4*) with known functions and given here as follows: (**A**) and (**E**), FBI-*BADH2*; (**B**) and (**F**), FBI-*Waxy*; (**C**) and (**G**), FBI-*Rf3*; and (**D**) and (**H**), FBI-*Rf4*. In (**A**)-(**H**), three technical replicates are performed, and shared PTMA tags are identified. The left and right

ideograms represent PTMA tags formed by read1 aligning to the positive and negative strands of the corresponding reference genome with few soft-clipped bases (colored).

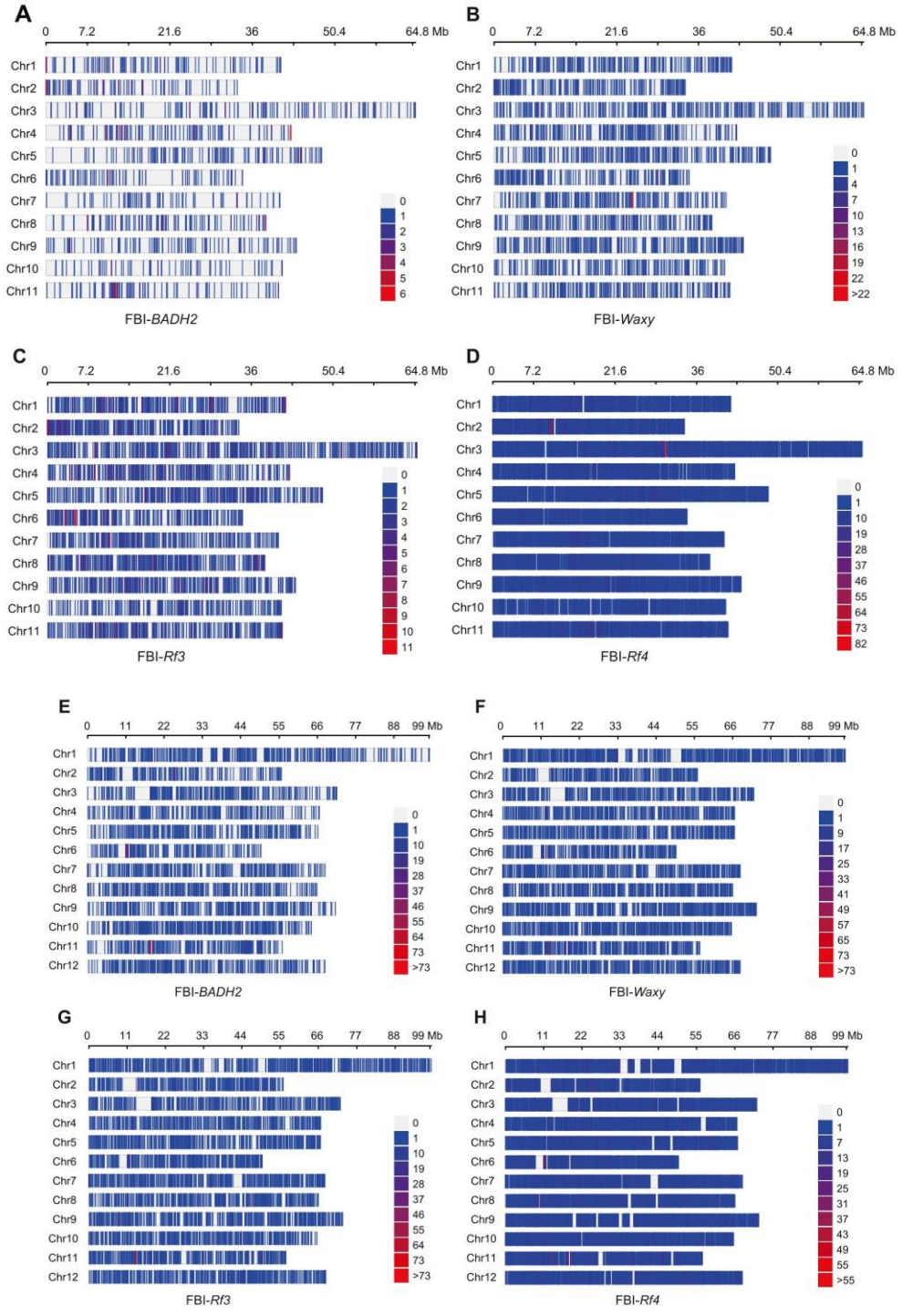


Figure S19. Density plots displaying the distribution of shared PTMA tags among three replicates when four FBI-seq primers are used in cowpea (A-D) or tomato (E-H)

The coverage depth threshold of PTMA tags is ≥ 3×. The window size for the density plot is 0.1 Mb. All four FBI-seq primers used here are designed according to rice gene sequences with known functions.

Table S1. Summary of tags detected in rice cultivar R498 in the pilot test

Cultivar	Raw reads	Raw bases (bp)	Coverage depth threshold ≥ 3×					Coverage depth threshold ≥ 5×					Note
			No. of PTPA tags†	Ratio of PTPA reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	Ratio of random reads (%)	No. of PTPA tags†	Ratio of PTPA reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	Ratio of random reads (%)	
R498	7,727,242	2,318,172,600	641	1.13	305,500	43.80	55.07	640	1.13	162,442	33.63	65.24	Reference genome size of R498 is ~ 391.0 Mb

†Theoretically, there were a total of 651 FBI-*Rim2* primer binding regions existed in R498 reference genome.

Table S2. Summary of PTMA tags detected in six rice samples using FBI-*Pi2*-1 primer

Rice Sample†	Raw reads	Raw bases (bp)	Reads duplication rate (%)	Coverage depth threshold ≥ 3×				Coverage depth threshold ≥ 5×				Note
				No. of PTMA tags	Ratio of PTMA reads (%)	Ratio of PTPA reads (%)	Ratio of random reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	Ratio of PTPA reads (%)	Ratio of random reads (%)	
R498	1,878,959	563,687,700	10.50	36,511	10.13	0.0020	89.87	12,291	5.33	0.0020	94.67	Reference genome size of R498 is ~ 391.0 Mb
YU-S1	1,778,970	533,691,000	11.54	25,614	7.57	0.0032	92.43	8,305	3.87	0.0032	96.13	
YU-S2	1,986,238	595,871,400	11.51	32,272	8.53	0.0025	91.47	10,699	4.48	0.0025	95.51	
YU-S3	2,822,486	846,745,800	12.94	50,807	10.26	0.0024	89.74	19,026	6.01	0.0024	93.99	
YU-S4	2,550,714	765,214,200	13.63	55,122	12.18	0.0025	87.82	20,426	7.04	0.0025	92.96	
YU-S5	2,366,220	709,866,000	12.08	37,258	8.82	0.0035	91.18	13,827	5.11	0.0035	94.89	

†R498 is a rice cultivar, and YU-S1~YU-S5 are five rice breeding materials.

Table S3. Number of PTPA and PTMA tags with coverage depth ≥ 3× detected in TEAseq data

Maize mutants†	Raw reads	Raw bases (bp)	No. of PTPA tags‡	Ratio of PTPA reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	Note
sample_1	10,864,418	3,259,325,400	31	5.66	12,052	23.44%	Reference genome size of B73_v4 is ~2.1 Gb
sample_2	3,614,455	1,084,336,500	29	5.64	6,835	22.09%	
sample_3	2,119,094	635,728,200	24	2.97	4,733	17.05%	

†the sequenced data of three randomly selected maize *Dissociation* (*Ds*) insertional mutants were obtained from Lyu et al. (2021).

‡Theoretically, there are a total of 32 intact specific primers (*Ds* 3' primer and *Ds* 5' primer) binding regions existed in B73_v4 reference genome.

Table S4. Summary of PTPA and PTMA tags detected in three rice cultivars using the FBI-*Rim2* primer

Cultivar	Technical replicates	Raw reads	Raw bases (bp)	Reads duplication rate (%)	Coverage depth threshold ≥ 3×							Coverage depth threshold ≥ 5×						Note
					No. of PTPA tags†	Ratio of PTPA reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	No. of shared PTMA tags	Ratio of shared PTMA tags (%)	Ratio of random reads (%)	No. of PTPA tags†	Ratio of PTPA reads (%)	No. of PTMA tags	Ratio of PTMA reads (%)	No. of shared PTMA tags	Ratio of shared PTMA tags (%)	Ratio of random reads (%)	
R498	1	15,361,546	4,608,463,800	33.34	642	5.67	295,520	54.60		79.31	39.73	642	5.67	204,445	51.80		83.16	42.53	Reference genome size of R498 is ~ 391.0 Mb
	2	20,970,546	6,291,163,800	35.41	643	3.40	485,202	53.49	234,391	48.31	43.11	642	3.40	336,284	50.43	170,019	50.56	46.17	
	3	19,427,302	5,828,190,600	36.75	642	6.01	328,148	54.07		71.43	39.92	642	6.01	236,978	52.01		71.74	41.98	
ZF802	1	15,741,797	4,722,539,100	32.25	624	1.98	396,666	51.96		62.84	46.06	611	1.98	273,636	48.69		66.26	49.33	
	2	17,803,851	5,341,155,300	32.46	628	2.32	379,726	50.15	249,269	65.64	47.53	616	2.32	268,125	47.49	181,306	67.62	50.19	
	3	13,125,466	3,937,639,800	28.88	626	1.98	312,558	48.15		79.75	49.87	609	1.98	216,322	45.09		83.81	52.93	
9311	1	21,115,400	6,334,620,000	31.40	631	3.13	413,764	53.73		67.18	43.14	623	3.13	290,817	51.12		70.40	45.75	
	2	19,533,364	5,860,009,200	30.64	625	2.57	386,918	53.74	277,957	71.84	43.68	617	2.57	269,554	50.99	204,731	75.95	46.43	
	3	18,649,191	5,594,757,300	30.86	626	2.85	367,200	52.65		75.70	44.50	617	2.85	259,373	50.13		78.93	47.02	

†Theoretically, there are a total of 651 FBI-*Rim2* primer binding regions existed in R498 reference genome.

Table S5. Summary of detected PTMA tags number in three rice cultivars when randomly sampling 2.0 or 0.25 Gb of sequencing data

Sampled data size (Gb)	Cultivar	Technical replicates	No. of PTMA tags[†]	No. of shared PTMA tags	Ratio of shared PTMA tags (%)
2.0	R498	1	164,408	103,964	63.24
		2	202,897		51.24
		3	158,481		65.60
	ZF802	1	191,990	106,708	55.58
		2	162,623		65.62
		3	169,274		63.04
	9311	1	195,689	127,988	65.40
		2	200,546		63.82
		3	188,182		68.01
0.25	R498	1	38,935	13,176	33.84
		2	31,008		42.49
		3	34,390		38.31
	ZF802	1	30,314	8,948	29.52
		2	24,746		36.16
		3	28,543		31.35
	9311	1	34,614	13,894	40.14
		2	37,885		36.67
		3	34,865		39.85

[†]The coverage depth threshold of PTMA tags being ≥ 3× was showed here.

Table S6. Comparison of R498 genome coverage rate between FBI-seq and PCR-free libraries with different sequencing data sizes

Sampled data size (Gb)	FBI-seq				PCR-free				*P*_value
	Replicate 1	Replicate 2	Replicate 3	Mean	Replicate 1	Replicate 2	Replicate 3	Mean	
0.25	14.43%	14.42%	14.42%	14.42%	40.44%	40.42%	40.43%	40.43%	2.59E-14
0.5	21.30%	21.33%	21.27%	21.30%	63.49%	63.54%	63.53%	63.52%	5.37E-13
1.0	30.37%	30.44%	30.36%	30.39%	85.06%	85.09%	85.12%	85.09%	5.84E-13
2.0	42.08%	42.09%	42.09%	42.09%	96.27%	96.29%	96.29%	96.28%	2.11E-15

Table S7. Number of PTPA or PTMA tags with different sequencing data sizes in three rice cultivars

Cultivar	Tag types	Technical replicates	Coverage depth threshold ≥ 3×					Coverage depth threshold ≥ 5×					Note
			0.25 Gb	0.50 Gb	1.0 Gb	1.5 Gb	2.0 Gb	0.25 Gb	0.50 Gb	1.0 Gb	1.5 Gb	2.0 Gb	
R498	PTPA	1	628	632	633	634	635	622	629	625	633	635	Reference genome size of R498 is ~ 391.0 Mb
		2	628	636	638	639	639	614	630	636	638	639	
		3	632	635	636	640	640	628	632	636	637	638	
	PTMA	1	38,935	69,623	109,244	140,579	164,299	16,905	37,018	64,877	87,546	104,735	
		2	31,008	67,584	123,830	167,351	202,780	10,190	29,650	65,981	95,718	121,108	
		3	34,390	64,024	105,170	134,887	158,481	14,797	33,768	63,357	85,286	103,180	
ZF802	PTPA	1	513	524	533	543	551	498	515	524	525	529	
		2	512	522	529	547	558	495	515	524	526	530	
		3	504	522	535	542	545	482	510	523	529	529	
	PTMA	1	30,314	65,324	118,675	158,865	191,787	9,836	29,006	64,041	92,358	116,083	
		2	24,746	54,595	100,284	135,121	162,882	7,676	23,435	53,514	77,948	98,245	
		3	28,543	60,197	107,312	141,461	169,370	9,586	27,396	59,209	83,934	104,444	
9311	PTPA	1	556	569	577	582	585	538	564	572	574	577	
		2	556	571	573	578	579	538	561	570	571	574	
		3	562	570	575	578	582	543	564	569	571	574	
	PTMA	1	34,614	71,041	121,953	163,985	196,023	12,339	33,349	69,866	98,459	122,049	
		2	37,885	75,498	129,577	169,033	200,424	14,362	36,858	73,978	102,674	126,148	
		3	34,865	70,106	121,101	157,941	187,944	12,982	34,344	69,706	97,161	119,401	

Table S8. Number of SNPs and genetic background recovering ratio calculated for each individual of a rice BC_2F_4 population

BC_2F_4 population	Raw reads	Raw bases (bp)	No. of shared SNPs†	No. of shared SNPs detected	No. of high depth SNPs‡	No. of low depth SNPs§	Total SNP count	Ratio of recurrent parent (%)
8GWB5-54	1,059,471	317,841,300	37,520	35,613	65,096	240,581	341,290	93.0
8GWB5-67	921,110	276,333,000		36,400	72,129	350,018	458,547	83.8
8GWB5-83	849,040	254,712,000		36,467	63,988	341,365	441,820	88.1
8GWB6-61	861,163	258,348,900		36,153	61,958	352,845	450,956	87.4
8GWB6-69	1,290,821	387,246,300		35,952	93,931	265,087	394,970	87.4
VE6219 (Non-recurrent Parent)	17,210,118	5,163,035,400	1,921,989	-	-	-	-	-
CNG1 (Recurrent Parent)	17,895,222	5,368,566,600		-	-	-	-	-

†Shared SNP means theses SNPs were detected in at least 80% individuals, and coverage depth of each shared SNP should be at 3× at least. For two parents, only homozygotic and polymorphic SNPs were considered.

‡High depth SNP means coverage depth of each snp (exluding shared SNPs) should be at 3× at least.

§Low depth SNP means coverage depth of each snp should be less than 3×.

Table S9. Number of PTMA tags detected in cowpea and tomato using different FBI-seq primers

Species	Primer name†	Technical replicates	Raw reads	Raw bases (bp)	No. of PTMA tags‡	No. of shared tags	Note
Cowpea (*Vigna unguiculata* ssp. *sesquipedalis*)	FBI-*BADH2*	1	5,154,686	1,546,405,800	8,691	1,416	Reference genome size of Vunguiculata_540_v1.0 is ~ 519.4 Mb.
		2	4,742,251	1,422,675,300	8,810		
		3	5,452,451	1,635,735,300	7,707		
	FBI-*Waxy*	1	4,934,506	1,480,351,800	12,153	3,049	
		2	5,157,300	1,547,190,000	9,603		
		3	5,476,705	1,643,011,500	14,349		
	FBI-*Rf3*	1	5,570,692	1,671,207,600	16,329	4,190	
		2	5,455,172	1,636,551,600	17,389		
		3	3,101,351	930,405,300	6,509		
	FBI-*Rf4*	1	4,959,834	1,487,950,200	51,117	23,471	
		2	4,961,905	1,488,571,500	58,844		
		3	4,960,194	1,488,058,200	42,542		
Tomato (*Solanum lycopersicum*) cv. Heinz 1706	FBI-*BADH2*	1	5,551,989	1,665,596,700	17,844	7,067	Reference genome size of Slycopersicum_514_SL3.0 is ~ 828.1 Mb.
		2	4,919,674	1,475,902,200	19,999		
		3	5,965,883	1,789,764,900	21,394		
	FBI-*Waxy*	1	6,974,974	2,092,492,200	25,895	7,657	
		2	4,587,743	1,376,322,900	16,300		
		3	5,212,465	1,563,739,500	20,054		
	FBI-*Rf3*	1	6,086,863	1,826,058,900	21,757	7,352	
		2	5,669,357	1,700,807,100	17,582		
		3	6,173,661	1,852,098,300	22,248		

		1	4,709,309	1,412,792,700	57,823		
	FBI-*Rf4*	2	4,484,586	1,345,375,800	47,936	28,110	
		3	4,688,003	1,406,400,900	70,474		

†All four specific FBI-seq primers used here were designed according to rice gene sequences with known function.

‡The coverage depth threshold of PTMA tags being ≥ 3× was showed here.

Table S10. Comparison of FBI-seq with commonly used genotyping methods

Methods	Technical keys	Advantages	Pitfalls	Costs	Time
FBI-seq	PTPA and PTMA amplification	only one specific primer needed for simultaneous foreground and background genotyping and almost no upfront investment for any species	PTMA loci were not predictable	low	short
Whole genome sequencing	whole genome shotgun sequencing	detection of whole profile of genomic variation	high cost spending on sequencing step and over sequenced for breeding projects	high	moderate
GBS, RAD-seq and its derivatives	reduce-representation of whole genome by selecting DNA fragment near restriction enzyme sites	only sequenced ~ 10% of the whole genome and hence reduced the cost spending on sequencing step	contained missing data arising from uneven restriction enzyme digestion and adapter ligation efficiency	moderate	moderate
Multiplex PCR	PCR amplification of up to thousands of loci in one tube	low cost for library preparation and sequencing	high upfront investment for primer design and synthesis; low marker density	low	moderate
Target enrichment	hybrid selection of partial genome by biotinylated DNA/RNA probes in solution	ability to target any locus and low cost for sequencing step	high upfront investment for probe design and manufacture; complex protocol for library preparation	high	long
DNA arrays	DNA hybridization with probes on solid surface	high stability and reliability; convenient for data comparing and sharing across different laboratories and experiments	very low flexibility and high upfront investment for array design and manufacture	high	long

Table S11. Summary of primer sequences information used in FBI-seq

Primer name†	Primer sequence (5'-3')	Primer structure†
FBI-*ALK*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+AGCGCCTCGATGAGCTTGTGC	Adapter1 + *ALK* gene primer
FBI-*BADH2*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+ATCATGGCACCTTTGGGGAGTAG	Adapter1 + *BADH2* gene primer
FBI-*DTH2*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCAACATCCTAATGCAGACCA	Adapter1 + *DTH2* gene primer
FBI-*Pi2*-1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+TCCAAACCCGTTGTTGCAC	Adapter1 + *Pi2*-1 gene primer
FBI-*Pi2*-2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCTCCGAACAACGCCAACTG	Adapter1 + *Pi2*-2 gene primer
FBI-*qGL3*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCGTTTGATTGGGCCGCAGGAAAA	Adapter1 + *qGL3* primer
FBI-*qSH1*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+TTCTGGAGGTCCGCGAACGAG	Adapter1 + *qSH* primer
FBI-*Rf3*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CCCGCTATTGTACCTGCTCTTAA	Adapter1 + *Rf3* gene primer
FBI-*Rf4*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+CTGCTTACAAAGTGAGGTGGTGT	Adapter1 + *Rf4* gene primer
FBI-*SD1*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+AGTATGCCATGCCTGTAAAAT	Adapter1 + *SD1* gene primer
FBI-*Rim2*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+AGATGGTTTCTCCACCAGTG	Adapter1 + *Rim2* transposon primer
FBI-*Waxy*	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG+GTTGTCAGTAACAAACCGGAAGG	Adapter1 + *Waxy* primer
PPMi7	CAAGCAGAAGACGGCATACGAGAT	
Primer 1	AATGATACGGCGACCACCGAGATCTACAC+[i5]+TCGTCGGCAGCGTC	PPMi5 + index-i5 + Adapter1 (partial)
Primer 2	CAAGCAGAAGACGGCATACGAGAT+[i7]+GTCTCGTGGGCTCGG	PPMi7 + index-i7 + Adapter2 (partial)

†Italic fonts represent genes/loci from rice genome.