Theoretical and Applied Genetics (2020) 133:2323–2334
https://doi.org/10.1007/s00122-020-03600-5

ORIGINAL ARTICLE

Effective identification of CRISPR/Cas9-induced and naturally occurred mutations in rice using a multiplex ligation-dependent probe amplification-based method

Sukumar Biswas[1] · Rong Li[1] · Jun Hong[1] · Xiangxiang Zhao[2] · Zheng Yuan[1,2] · Dabing Zhang[1,2] · Jianxin Shi[1]

Received: 18 January 2020 / Accepted: 1 May 2020 / Published online: 14 May 2020

Abstract

Key message **A multiplex ligation-dependent probe amplification (MLPA)-based method was developed and successfully utilized to efficiently detect both CRISPR/Cas9-induced and naturally occurred mutations in rice.**

Abstract The site-specific nuclease-based CRISPR/Cas9 system has emerged as one of the most efficient genome editing tools to modify multiple genomic targets simultaneously in various organisms, including plants for both fundamental and applied researches. Screening for both on-target and off-target mutations in CRISPR/Cas9-generated mutants at the early stages is an indispensable step for functional analysis and subsequent application. Various methods have been developed to detect CRISPR/Cas9-induced mutations in plants. Still, very few have focused on the detection of both on- and off-targets simultaneously, let alone the detection of natural mutations. Here, we report a multiplex capable method that allows to detect CRISPR/Cas9 induced on- and off-target mutations as well as naturally occurred mutation based on a multiplex ligation-dependent probe amplification (MLPA) method. We demonstrated that unlike other methods, the modified target-specific MLPA method can accurately identify any INDELs generated naturally or by the CRISPR/Cas9 system and that it can detect natural variation and zygosity of the CRISPR/Cas9-generated mutants in rice as well. Furthermore, its high sensitivity allowed to define INDELs down to 1 bp and substitutions to a single nucleotide. Therefore, this sensitive, reliable, and cheap method would further accelerate functional analysis and marker-assisted breeding in plants, including rice.

Introduction

Genetic/allelic variations in plants, either occur naturally or artificially, may cause mutations. Artificial mutagenesis methods include chemical/physical treatments and novel plant breeding technologies, such as genome editing. Genome editing system utilizes site-specific nucleases to introduce precisely targeted double-strand breaks (DSB), and the desired modifications are obtained by subsequent endogenous DSB repair machinery robustly. Those site-specific nucleases include zinc finger nucleases (ZFNs), transcription-activator like effector nucleases (TALENs), and clustered regulatory interspaced short palindromic repeats associated protein Cas9 (CRISPR/Cas9) (Lusser et al. 2012; Zhu et al. 2017). Among them, the CRISPR/Cas9 system is more compelling, and particularly, it can modify multiple plant genes concurrently (Feng et al. 2014; Xie et al. 2015; Ishizaki 2016; Zong et al. 2017). CRISPR/Cas9 system introduces a single guide RNA (sgRNA) to guide the Cas9 protein to target genomic DNA consisting of the protospacer adjacent motif (PAM) to generate double-strand breaks

Communicated by Kan Wang.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-020-03600-5) contains supplementary material, which is available to authorized users.

✉ Jianxin Shi
jianxin.shi@sjtu.edu.cn

[1] Joint International Research Laboratory of Metabolic and Developmental Sciences, Shanghai Jiao Tong University–University of Adelaide Joint Centre for Agriculture and Health, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[2] Key Laboratory of Crop Marker-Assisted Breeding of Huaian Municipality, Jiangsu Collaborative Innovation Center of Regional Modern Agriculture and Environmental Protection, Huaian 223300, China

(DSB). Those DSBs in plants are repaired mainly through error-prone non-homologous end joining (NHEJ), which mostly generates insertion/deletion (INDEL) frame-shift mutations with only a few base pairs (bp) variation, leading to loss of function via premature translation termination (Ren et al. 2016; Pan et al. 2016; Zhu et al. 2017). However, Cas9 protein can also target homologous genomic sites of sgRNA concurrently, which may cause unintended off-target mutations with one to few bp variations (Jinek et al. 2012). These small mismatches in the genome cannot be identified through agarose-gel electrophoresis (Denbow et al. 2017), which averts the mutation screening and effects of subsequent functional analysis. Therefore, there is an urgent need to develop an effective, reliable, and inexpensive method for parallel analysis of CRISPR/Cas9-induced on- and off-target INDELs and to distinguish them from those that naturally occurred for plant science research.

Many different methods have been first developed in other model systems rather than plants to detect the mutations of target loci, all have their particular limitations (Zischewski et al. 2017). Most frequently used methods for INDELs identification include: (1) enzyme mismatch cleavage (EMC) assay; (2) fluorescence-based high-resolution melting analysis (HRMA) technique; and (3) modified migration-based heteroduplex mobility assay (HMA) (Thomas et al. 2014; Vouillot et al. 2015; Zischewski et al. 2017). EMC assay utilizes the most popular enzyme T7 endonuclease 1 (T7E1) or Surveyor nuclease to cleave heteroduplex DNA at mismatches with one or few nucleotides, and resulting mutation with these small mismatches can be analyzed by agarose gel electrophoresis (Vouillot et al. 2015). This method is easy to handle, cost-effective, suitable for large INDEL detection (Vouillot et al. 2015; Zischewski et al. 2017). But it is less sensitive; it cannot identify homozygous mutations, and it is not suitable for polymorphic locus analysis (Kim et al. 2011; Huang et al. 2012). HRMA method characterizes DNA samples based on their disassociation behavior and detects small sequence differences in PCR amplified sequences, just by direct melting. With the use of specific DNA dyes, high-end instrumentation and sophisticated analysis software, these differences are detected (Dahlem et al. 2012; Wang et al. 2015). HRMA is simple, rapid, sensitive, and compatible with high-throughput analysis, but it cannot detect comparatively large INDELs (> 100 bp) (Thomas et al. 2014; Zischewski et al. 2017). HMA takes advantage of the modified migration to separate re-hybridize PCR products in polyacrylamide gel electrophoresis (PAGE) (Ota et al. 2013). It is easy to operate, fast, cheap, suitable for detection of single nucleotide polymorphism (SNP) and small INDELs. But it cannot detect larger deletions (Ota et al. 2013; Zischewski et al. 2017). Other reported INDEL detection methods include PCR combined with ligation detection reaction (PCR-LDR) (Kc et al. 2016), restriction fragment length polymorphism (RFLP) (Kim et al. 2014), PCR based on two primer pairs (Yu et al. 2014), Tracking of Indels by Decomposition (TIDE) (Brinkman et al. 2014), CRISPR Genome Analyzer (CRISPR-GA) (Güell et al. 2014), and droplet digital PCR (ddPCR) (Findlay et al. 2016). Most of them are less sensitive, expensive, time-consuming, and not suitable for larger INDEL detection (Brinkman et al. 2014; Güell et al. 2014; Kim et al. 2014; Yu et al. 2014; Kc et al. 2016). In plant, similar methods have been developed recently. These methods include EMC (Nekrasov et al. 2013; Shan et al. 2014), HRMA (Denbow et al. 2017), annealing at critical temperature PCR (ACT-PCR) (Hua et al. 2017), PCR and Amplicon labeling-based method (Biswas et al. 2019), mutation sites-based specific primers PCR (MSBSP-PCR) (Guo et al. 2018), and cleaved amplified polymorphic sequence (CAPS) (Kohata et al. 2018). Although most of them are proved to be effective in certain cases, the applicability was adversely affected by intrinsic limitations, such as limited sensitivity and specificity, time/labor-consuming, and inapplicability for SNP detection (Nekrasov et al. 2013; Shan et al. 2014; Denbow et al. 2017; Hua et al. 2017; Guo et al. 2018; Kohata et al. 2018; Biswas et al. 2019). Most importantly, all of them lack multiplex capabilities; and none of them finds its application in natural variation discrimination.

In the past years, neither did basic functional analysis of candidate gene/sgRNA nor applied breeding use long-range PCR or long-read next-generation sequencing (NGS) methods, such as PacBio, to detect unintended genomic changes including off-targets in plants; therefore, several early reports claimed that the off-target events in CRISPR/Cas9 edited plants are rare; even it occurs, the frequency is very low (Feng et al. 2014; Zhang et al. 2014; Gao et al. 2015; Tang et al. 2018; Xu et al. 2015). However, increasing evidence showed unexpected high frequency of off-target mutagenesis in CRISPR/Cas9-induced Arabidopsis (Zhang et al. 2018) and rice (Endo et al. 2015; Li et al. 2016). Therefore, the detection of off-target is equally important to that of detection of on-target in plants, particularly for functional analysis. There are both bioinformatics compatible online-based tools (such as CRISPR.P, CCTop) and experimentation approaches (such as NGS, BLISS, BLESS, GUIDE-seq) to predict and identify putative off-target sites, respectively (Yan et al. 2017; Germini et al. 2018; Grohmann et al. 2019; Hahn and Nekrasov 2019). Online tools cannot stand alone without experimental validation, while most of proposed experimental approaches involve NGS, and generally are complex and time consuming. The most popular off-target identification approach is the amplification of silico predicted potential off-target sites and followed by Sanger sequencing. Nevertheless, it might overlook mutations at other alleles (Zischewski et al. 2017). EcoTILLING, a high-throughput-based method, has been reported to detect naturally occurred mutations, which, however,

is labor/time-consuming (Rigola et al. 2009). In addition, pan-genomes might help to identify natural mutations in the genome of interest (Zhao et al. 2018; Li et al. 2019). Notably, none of above-mentioned methods could fulfill the simultaneous detection of both on- and off-targets and natural variations in plants. A recently developed method qEva-CRISPR can detect all types of on-target mutations in human cells with high sensitivity regardless of mutation type and several off-targets (Dabrowska et al. 2018); however, no corresponding method has been established in plant. Development of competent, reliable and inexpensive multiplex capabilities-based method would help to screen on- and off-target as well as natural mutations concurrently, at the early stages in pooled samples, which would accelerate further functional analysis or breeding in plants.

In this study, taking advantage of the multiplex capability of a previously reported multiplex ligation-based probe amplification (MLPA) method, we developed a new MLPA-based method that allowed us to identify CRISPR/Cas9-induced on- and off-target INDELs, and natural occurred INDELs in rice. The sensitivity, reliability, and applicability were analyzed using CRISPR/Cas9-induced mutants targeting different genes and different rice cultivars harboring natural occurred SNPs in *semi-dwarf1* (*SD1*) loci.

Materials and methods

Plant materials

Several CRISPR/Cas9 edited mutant lines targeting on *semi-dwarf 1* (*SD1*) and other genes and several rice (*Oryza sativa*) varieties harboring natural mutation in *SD1* (including Kasalath, Xiushui, and Minghui63) were used in this experiment (Table S1). In addition, 9522 (*Oryza sativa* ssp. *japonica*) was used in this study as a reference. All above-mentioned rice lines, including mutants and wild types, were grown in the paddy field of Shanghai Jiao Tong University (30°N, 121°E), Shanghai, China, under natural rice growing conditions.

Plant genomic DNA extraction

Genomic DNA from leaf tissues was extracted as previously described with minor modifications (Murray and Thompson 1980). Leaf tissues were ground in the presence of liquid nitrogen and then incubated with lysis buffer (1.5 X CTAB) and RNase at 65 °C for 60 min. The liquid phase was collected after centrifugation, extracted again with phenol: chloroform and trichloromethane, and mixed with an equal volume of isopropyl alcohol to precipitate the genomic DNA. The pellet was washed twice with 70% ethanol, air-dried and then dissolved in ddH_2O. The quality and quantity of the extracted genomic DNA were evaluated using both the NanoDrop 1000 UV/vis Spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA) for OD_{260}/OD_{280} and OD_{260}/OD_{230} and the electrophoresis on 1% (w/v) agarose gel in 0.5 × TBE with Gel Red staining. All extracted genomic DNA was stored at − 20 °C until used in the experiments.

Probe design

All oligonucleotide probes for MLPA-based method were designed according to a previously adopted strategy (Kozlowski et al. 2007; Marcinkowska et al. 2010), and synthesized by Invitrogen Co., Ltd. (Shanghai, China). Each multiplex ligation-based probe was composed of two half-probes: 5′ half-probe and 3′ half-probe. Each half-probe consisted of the primer-specific binding sequence (PBS-red), the stuffer sequence (SS-gray), and the target-specific hybridization sequence (THS-black). In addition, the control probes were chosen outside of the genomic region of interest, dispersing ideally on different chromosomes. In general, paired half-probes for mutants were designed to directly adjacent to the putative mutated-regions, while control probes were designed to locate in the genomic regions free of repetitive elements, SNPs, and small INDELs (Fig. 1; Kozlowski et al. 2007; Marcinkowska et al. 2010). Sequences and detailed characteristics of all oligonucleotide probes used in this experiments are listed in Table S1.

MLPA program

MLPA reactions were carried out in a 50 µl of reaction mixtures as suggested (https://www.mlpa.com). Five µl of sample genomic DNA (approximately 100 ng) was denatured by heating at 98 °C for 5 min and subsequent cooling down; the resulting denatured DNA was mixed with 3 µl of hybridization master mix to a volume of 8 µl, heated at 95 °C for 2 min, and hybridized at 60 °C for 16 h. Hybridized probes were mixed with 32 µl of Ligase-65 master mix to reaction volume 40 µl, and ligated at 54 °C for 15 min, and followed by heating at 98 °C for 5 min to inactivate ligase enzyme. Ligated probes were cooled down to room temperature and mixed with 10 µl of polymerase master mix that contained 2 µl of SALSA PCR primer mix (universal primer pair, one of which was fluorescently labeled forward primer), 0.5 µl of polymerase, and 7.5 µl of ultrapure water. The final PCR reaction volume was 50 µl. PCR amplification reaction was carried out by 35 cycles at 95 °C for 30 s, 60 °C for 30 s and a final stage at 72 °C for 60 s. The amplified fluorophore-labeled MLPA PCR products were separated with LIZ GS500 size standard by capillary electrophoresis on an ABI Prism 3130XL apparatus (Applied Biosystem,

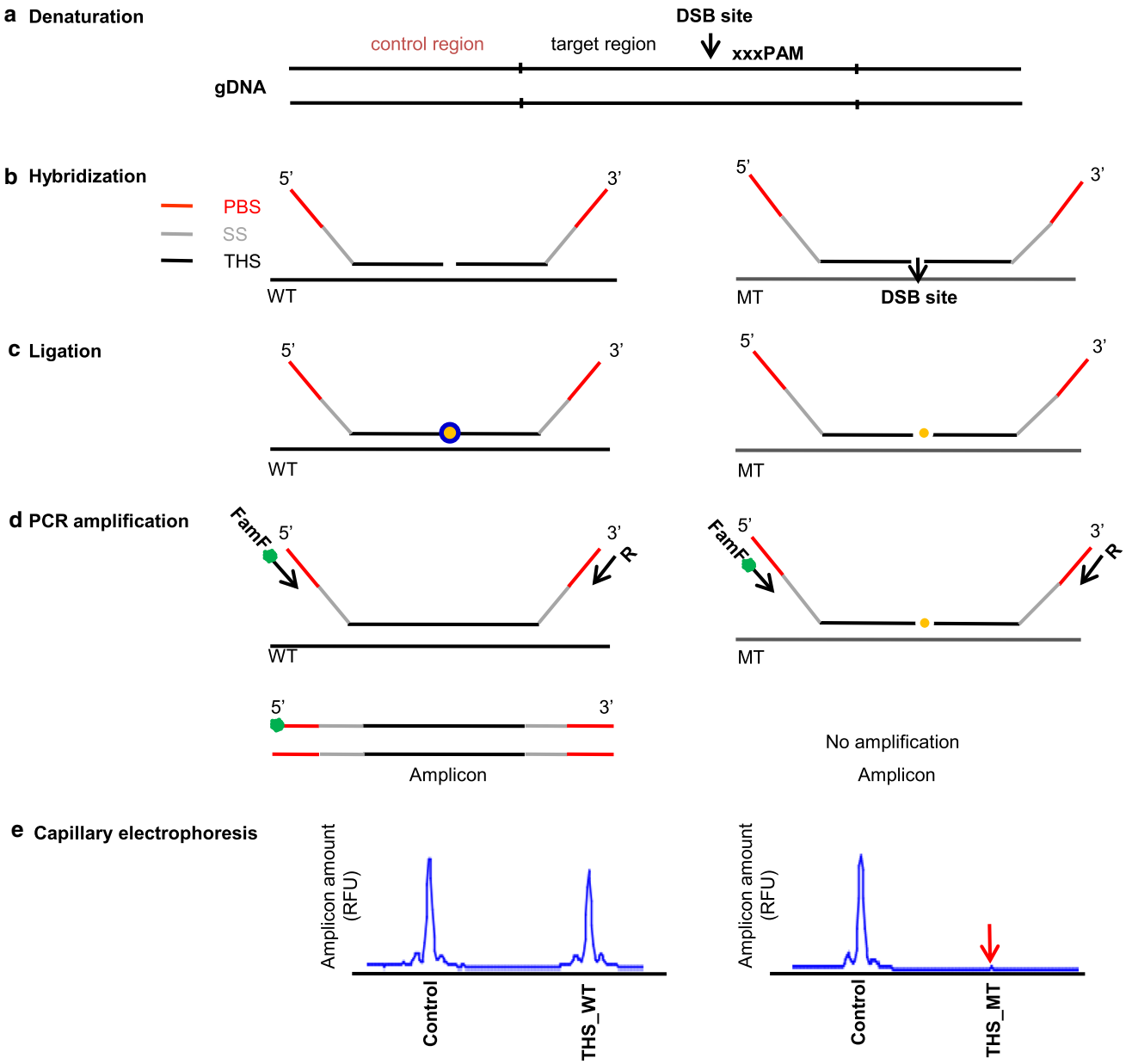

Fig. 1 Schematic presentation of the principle and steps for the identification of mutations by the developed multiplex ligation-dependent probe amplification (MLPA)-based method. MLPA-based method consists of five steps (from top to bottom). **a** Denaturation. Double stranded genomic DNA (gDNA) is denatured to two single strands by heating. In general, on-target, off-target and natural variation-specific MLPA paired half-probes are designed to directly adjacent to the putative mutated-sites (target region). Therefore, any mismatches will avert subsequent ligation of paired half-probes and further PCR amplification. DSB, double strand breaks; PAM, protospacer adjacent motif. **b** Hybridization. A probemix is added to the denatured gDNA sample for hybridization under stringent conditions. Each MLPA probe consists of two half-probes; 5′ half-probe and 3′ half-probe and each half-probe composed of the primer-specific binding sequence (PBS, red), stuffer sequence (SS, gray), and the target-specific hybridization sequence (THS, black) that correctly hybridize to the adjacent target of the sample gDNA. WT, wild-type; MT, mutant. **c** Ligation. Correctly hybridized 5′ half- and 3′ half-probes are ligated successfully into a single longer probe. **d** PCR amplification. All successfully ligated probes can be used as template for PCR amplification using universal primer pair that sits at the far 5′- and 3′- ends of the ligated probes. One of the universal primer (herein is the forward primer F) is fluorescent-labeled (FamF). R, reverse primer. **e** Capillary electrophoresis. All fluorophore-labeled amplicons are separated by capillary electrophoresis based on the lengths of their different length of SS, which show corresponding chromatograph peaks at the expected positions. RFU, relative fluorescence units (color figure online)

USA) and analyzed using Gene Marker software v2.6.4 (Soft Genetics, USA).

Results

Design of a MLPA-based system

The whole procedure of multiplex ligation-dependent probe amplification (MLPA)-based method included following five steps: DNA denaturation, probe hybridization, ligation, PCR amplification, and capillary electrophoresis (Fig. 1). Among them, probe design is crucial for the establishment of MLPA assay, because the PCR product from each DNA sequence must be independently detected and quantified in the capillary electrophoresis based on different length (Kozlowski et al. 2008). Each MLPA probe was composed of a 5′ half-probe and a 3′ half-probe, each half- probe consisted of one primer-specific binding sequence (PBS, red) for PCR amplification, one stuffer sequence (SS, gray) for determination of PCR product size, and one target-specific hybridization sequence (THS, black) for unique binding of probes to targets (Fig. 1). The 5′ and the 3′ THS probes in MLPA were generally adjacent to the predicted mutated regions, namely mutation hot spot region (MHS), and each sister THS sequence should be at least 21 nucleotides in length. The Phage M13 sequences (NCBI/GenBank ID V00604) between 3 and 119 bp were used as SS, which allowed to adjust the length of probes and to result in unique amplicon peak for each probe (Fig. 1a). Notably, the 5′ half-probe had a forward primer PBS, a SS, and a left THS sequence while the 3′ half-probe a right THS sequence, a right SS, and a reverse primer PBS (Fig. 1b). After hybridization, only were the two half-probes that hybridized next to each other on their target sequence under stringent conditions ligated to a single and longer probe (Fig. 1c), which could then be served as the template for the amplification using PBS primer pair (Fig. 1d). One of the PBS primers was labeled with a fluorescence dye, and the SS included in each probe endowed each amplification product a unique characteristic length, thus, the amplification products could be detected by capillary electrophoresis and fluorescence detection based on their sizes (Fig. 1e). Therefore, in the case of CRISPR/Cas9 induced or naturally occurred mutants, even a small mismatch at the ligation point (mutation site) would impair ligation and subsequent probe amplification. As a result, typical chromatogram peak would be utterly absent in the expected position for the affected target-specific probe, while the chromatogram peaks for the non-affected target-specific probes should appear as do for the control probes (Fig. 1e).

Based on this concept, probes for MLPA-based system were designed to identify CRIPSR/Ca9 induced *SD1* mutation in rice (Table S1). Those probes included four control probes that were randomly dispersed over rice genome, universal, and likely used in any MLPA assays in rice, and two target-specific probes (THS_*SD1*.1 and THS_*SD1*.2) that consisted of both 5′ half-probes and 3′ half-probes.

Specificity and sensitivity of the developed MLPA-based system

The MLPA-based system assay was first optimized and validated using genomic DNA (gDNA) from wild type (WT) and known *SD1* deletion mutants. The results showed that control and WT probes generated correct chromatogram peak at the expected positions in both WT and mutant genomes, while probes of CRISPR/Cas9-induced mutants (1–2 bp deletions) (THS_*SD1*.1 and THS_*SD1*.2), did not generate corresponding chromatogram peaks in the expected positions (Fig. S1). This result indicated that MLPA-based system can be used for the identification of CRISPR/Cas9 edited mutants in rice. Its specificity was confirmed in different *SD1* mutants with a single nucleotide insertion or deletion, or a single nucleotide substitution. In all cases, probes of CRISPR/Cas9-induced *SD1* mutants, either 1 bp deletion, 1 bp insertion, or 1 bp replacement, failed to generate correct chromatogram peaks in the expected positions (Fig. S2).

To demonstrate the applicability and accuracy of the developed MLPA-based method for different gene targets, MLPA probes for mutants of three CRISPR/Cas9 targeted genes including *Os06g0135460, Os07g0445800,* and *SD1* were designed (Table S1). MLPA-based analyses were separately performed for each target-specific probe using mixed genomic DNA samples of different mutations targeting the same corresponding sgRNA. Similarly, all MLPA probes generated correct chromatogram peaks in the expected positions when WT DNA sample was used (Fig. 2a). In contrast, no MLPA probes, THS_Os06g0135460.1 for *Os06g0135460*, THS_Os07g0445800.1 for *Os07g0445800,* or THS_SD1.1 and THS_SD1.2 for *SD1,* generated correct chromatogram peaks in the expected positions when mixed DNA samples for corresponding mutants for *Os06g0135460* (Fig. 2b), *Os07g0445800* (Fig. 2c), and *SD1* (Fig. 2d, e) were used, respectively.

The sensitivity of the developed MLPA-based method was tested using mixed genomic DNA samples from both WT with mutant, in which the genomic DNA of the *SD1* mutation (1 bp deletion) was mixed with the WT genomic DNA to corresponding mutant/WT ratios of 100% (100% mutant), 50% (50% mutant/50% WT), and 0% (100% WT), respectively. The results showed that, as compared to 100% WT, 50% mixed DNA samples generated significantly reduced chromatogram peak, while 100% mutant failed to generate chromatogram peak for the target-specific probe (THS_*SD1*.2) (Fig. S3). This result was also confirmed by the quantifying the relative peak of the target-specific

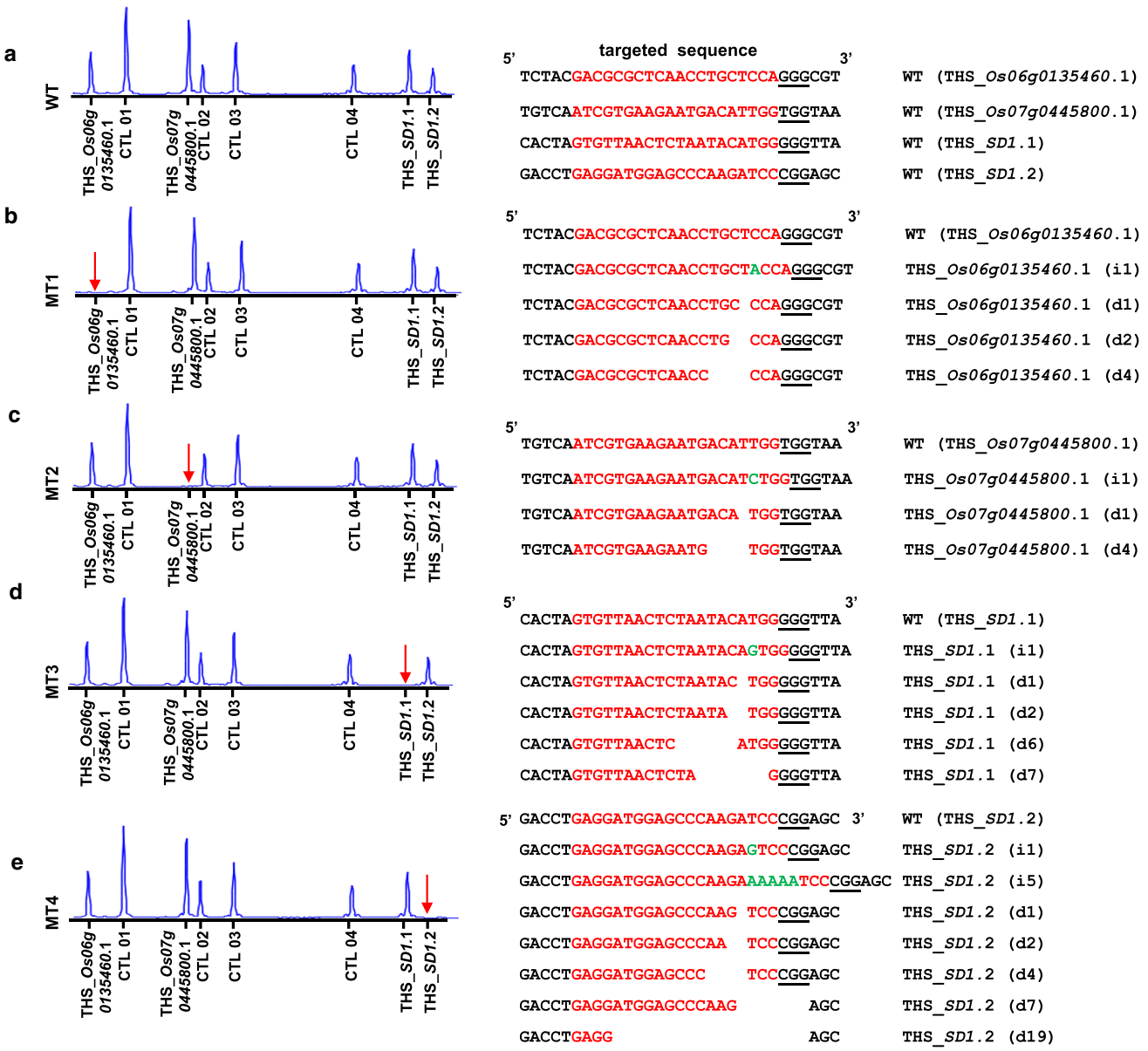

Fig. 2 Detection of on-target mutations by the developed MLPA-based method. The resulting chromatogram peaks separated by capillary electrophoresis of MLPA-based method are presented in the left panels, while Sanger sequencing results of the PCR products of WT and mutants used for the analysis are presented in the right panels. **a** Wild-type (WT) genomic DNA. **b** Mixed mutant genomic DNA samples containing different mutations targeting *Os06g0135460*.1 (MT1). **c** Mixed mutant genomic DNA samples containing different mutations targeting *Os07g0445800*.1 (MT2). **d** Mixed genomic DNA samples containing different mutations targeting *SD1*.1 (MT3). **e** Mixed mutant genomic DNA samples containing different mutations targeting *SD1*.2 (MT4). Red arrowhead in each left panel represents the expected chromatogram peak site of the corresponding target-specific probe. Red, green, and empty letters in the right panel represent target sequence, inserted and deleted base in the target, respectively. d#, deletion with #bp; i#, insertion with #bp; CTL, control probe; THS, target-specific hybridization sequence; WT, wild-type; MT, mutant (color figure online)

probe (THS_*SD1*.2) using Coffalyser.Net software (MRC-Holland) (Fig. S3). Since the 100% mutant and 50% mixed DNA samples represented the homozygous and heterozygous mutant status, respectively, the developed MLPA-based method could thus be useful for zygosity analysis in CRISPR/Cas9-induced rice mutants.

Analysis on CRISPR/Cas9-induced off-targets in rice

CRISPR/Cas9 may also generate off-target mutation because it tolerates up to three mismatches between the sgRNA and the target sequence (Zhu et al. 2017). To explore if the developed MLPA-based method can be used to screen and

detect CRISPR/Cas9-induced off-target mutations, we first applied it to *Os06g0135460* and *Os07g0445800* mutants using two corresponding probes OTS_*Os06g0135460*.1 and OTS_*Os07g0445800*.1, respectively. Our previous screening for the top five potential off-targets of each target identified by CRISPR-P software (https://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR) using PCR found that there was off-target in mutants targeting *Os07g0445800*. In MLPA-based method, when WT DNA was used, probe OTS_*Os07g0445800*.1–1, together with other two probes *OTS_SD1.2–1* and *OTS_SD1.2–2*, generated chromatogram peaks at the expected positions (Fig. 3a). In contrast, when mixed mutant genomic DNA samples were used, probe OTS_*Os07g0445800*.1–1 did not generate corresponding chromatogram peak, while other two probes *OTS_SD1.2–1* and *OTS_SD1.2–2* did (Fig. 3b), indicating the occurrence of off-target mutation in *Os07g0445800.1*.

The effectiveness and usefulness of this MLPA-based method for screening and detection of CRISPR/Cas9-induced off-targets were validated with other *SD1* mutants targeting two different sgRNA (THS_*SD1*.1 and THS_*SD1*.2) using corresponding off-target specific probes OTS_ *SD1*.1 and OTS_*SD1*.2, respectively. Previously screening for the top five potential off-targets of each target identified by CRISPR-P software using PCR found off-target mutations only in mutants targeting THS_*SD1*.2.

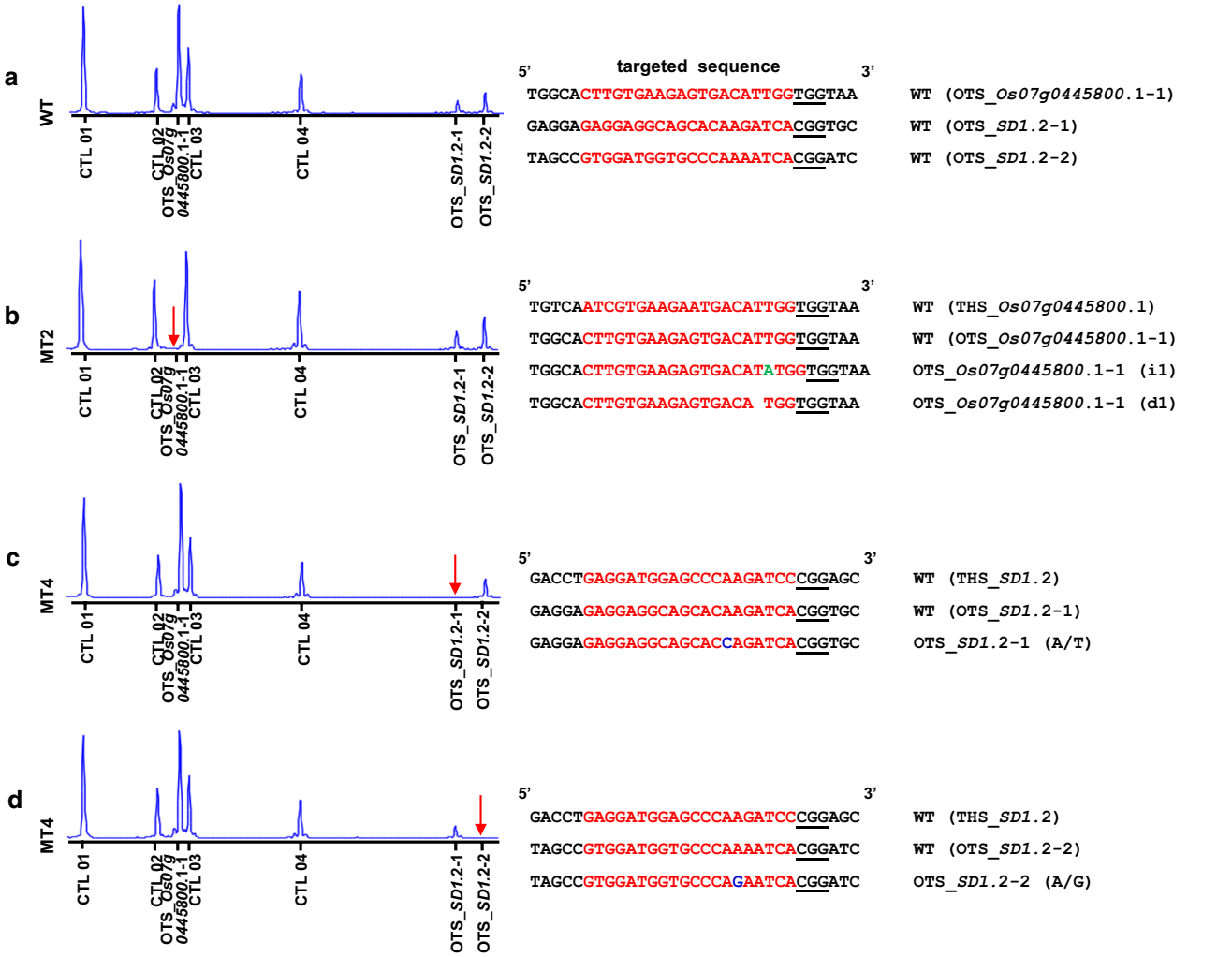

Fig. 3 Detection of off-target mutations by the developed MLPA-based method. The resulting chromatogram peaks separated by capillary electrophoresis of MLPA-based method are presented in the left panels, while Sanger sequencing results of the PCR products of WT and mutants used for the analysis are presented in the right panels. **a** Wild-type (WT) genomic DNA. **b** Mixed mutant genomic DNA samples containing different mutations targeting *Os07g0445800*.1 (MT2). **c** Genomic DNA sample of *SD1*.2 mutant (MT4). **d** gDNA sample of *SD1*.2 mutant (MT4). Red arrowhead in each left panel represents the expected chromatogram peak site of the corresponding target-specific probe. Red, green, blue and empty letters in the right panel represent target sequence, inserted, replaced, and deleted base in the target, respectively. d#, deletion with #bp; i#, insertion with #bp; CTL, control probe; OTS, off-target-specific; WT, wild-type; MT, mutant (color figure online)

MLPA-based method verified PCR results, in which different from those in WT genomic DNA samples (Fig. 3a), no chromatogram peaks appeared at the expected sites for the off-target-specific probes *OTS_ SD1*.2–1 and *OTS_SD1*.2–2 in mixed mutant genomic DNA samples (Fig. 3c–d). Sanger sequencing confirmed that the off-target mutation was caused by a single nucleotide substitution in both off-target sites (Fig. 3c–d). These results indicated that this MLPA-based method is quite effective and highly sensitive for identification of CRISPR/Cas9-induced off-target mutations as well.

Analysis on natural variations in rice

Since the developed MLPA-based method can effectively detect CRISPR/Cas9-induced INDEL mutations, we assume that it might work for the detection of natural variation mutations in rice. To explore this possibility, we designed new set of natural variation (NV) probes for this purpose based on the SNPs of the *SD1* gene in the pan-genome identified with SNP-Seek database (https://snp-seek.irri.org/_snp.zul). The designed target specific probes NV_ *SD1*.1, NV_ *SD1*.2, NV_ *SD1*.3, and NV_ *SD1*.4 were specific for one 1 bp and one 2 bp substitution mutations in rice varieties Minghui63, Kasalath and Xiushui, respectively. A set of controlled probes designed previously were also used. As shown in Fig. 4, no mutant genomic DNA samples generated chromatogram peaks at the expected positions for corresponding probe NV_ *SD1*.1, NV_ *SD1*.2, NV_ *SD1*.3, and NV_ *SD1*.4, respectively, indicating the applicable of this method for natural variation detection.

Discussion

Although the CRISPR/Cas9 system is still being improved for better targeted genome editing, it has been widely utilized both in basic and applied sciences. The predominant DSB repair pathway of CRISPR/Cas9 system in plants tends to generate small INDELs down to 1 bp or single nucleotide substitutions, and such small INDELs might occur naturally in plants (Grohmann et al. 2019). Therefore, a practical, accurate, and economic screening method for INDELs or single nucleotide substitutions caused by targeted genome editing or natural mutation is essential not only for applied breeding but also for fundamental functional research in plants. Previously developed methods can identify CRISPR/Cas9-induced on-target INDELs (Nekrasov et al. 2013; Shan et al. 2014; Denbow et al. 2017; Hua et al. 2017; Guo et al. 2018; Kohata et al. 2018; Biswas et al. 2019), off-target INDELs (Zischewski et al. 2017), or natural variation (Rigola et al. 2009); separately, very few of them can have multiplex capacity to detect all above-mentioned mutations and variations simultaneously. In this study, we modified an established MLPA method, developed a new approach to detect CRISPR/Cas9-induced on-target and off-target mutants and natural variations, and proved its sensitivity and applicability in different lines in rice.

Various molecular approaches have been developed previously to detect CRISPR/Cas9-induced mutations. These methods included HRMA (Denbow et al. 2017), EMC (Nekrasov et al. 2013; Shan et al. 2014), ACT-PCR (Hua et al. 2017), MSBSP-PCR (Guo et al. 2018), CAPS (Kohata et al. 2018), WGS (Zhang et al. 2014) and others (Biswas et al. 2019). All except WGS proved not to be effective in the identification of them simultaneously. In addition, each method's intrinsic limitations restrain its fully application (Biswas et al. 2019). Compared with HRMA, EMC, ACT-PCR, MSBSP-PCR, and CAPS, the developed MLPA-based method is sensitive (down to 1 bp INDELs and single nucleotide substitutions) (Fig. 2), consistent (suitable for different targets) (Figs. 2, 3), accurate (consistent with Sanger sequencing results) (Figs. 2, 3, 4), and quantitative (also applicable in zygosity analysis) (Fig. S3). Compared with WGS, one of the most potential approaches to identify CRISPR/Cas9-induced mutations including both on and off-target ones and natural variations in the genomes of interest (Zhang et al. 2014), MLPA-based method is cheap (< 8$ per sample, including probe synthesis cost), time-saving (within maximum 2 days), and no bioinformatics knowledge demanding (no assembly). Furthermore, MLPA-based method is very useful in the identification of CRISPR/Cas9-induced off-targets using non-WGS approaches, particularly at the early screening stage, due mainly to its multiplicity. Amplification of in silico predicted putative off-target sites and subsequent Sanger sequencing are the easiest ways to identify off-target mutation in plants (Zischewski et al. 2017). However, costs of Sanger sequencing and difficulties in management limit its application for a high number of off-target sites and samples. MLPA-based method becomes the good alternative in such cases because of its ability to detect about 60 target (off-target) sites in a single assay (www.mlpa.com). Sequencing of MLPA verified positive mutants will save time and money at this stage. However, for off-target detection, MLPA-based method can only be used for simultaneously detection of previously identified off-target, but not unknown off-targets. In this case, NGS is super advantageous over MLPA. As for the multiplicity of established MLPA for simultaneous detection of on- and off-targets, so far, only at most four target sites plus four controls were simultaneously assessed in this study, more targets need to be included for simultaneous detection of both on- and off-target mutations in the future studies.

What's more, compared with high-throughput sequence-based EcoTILLING technology, which identifies naturally induced mutations in plants (Rigola et al.

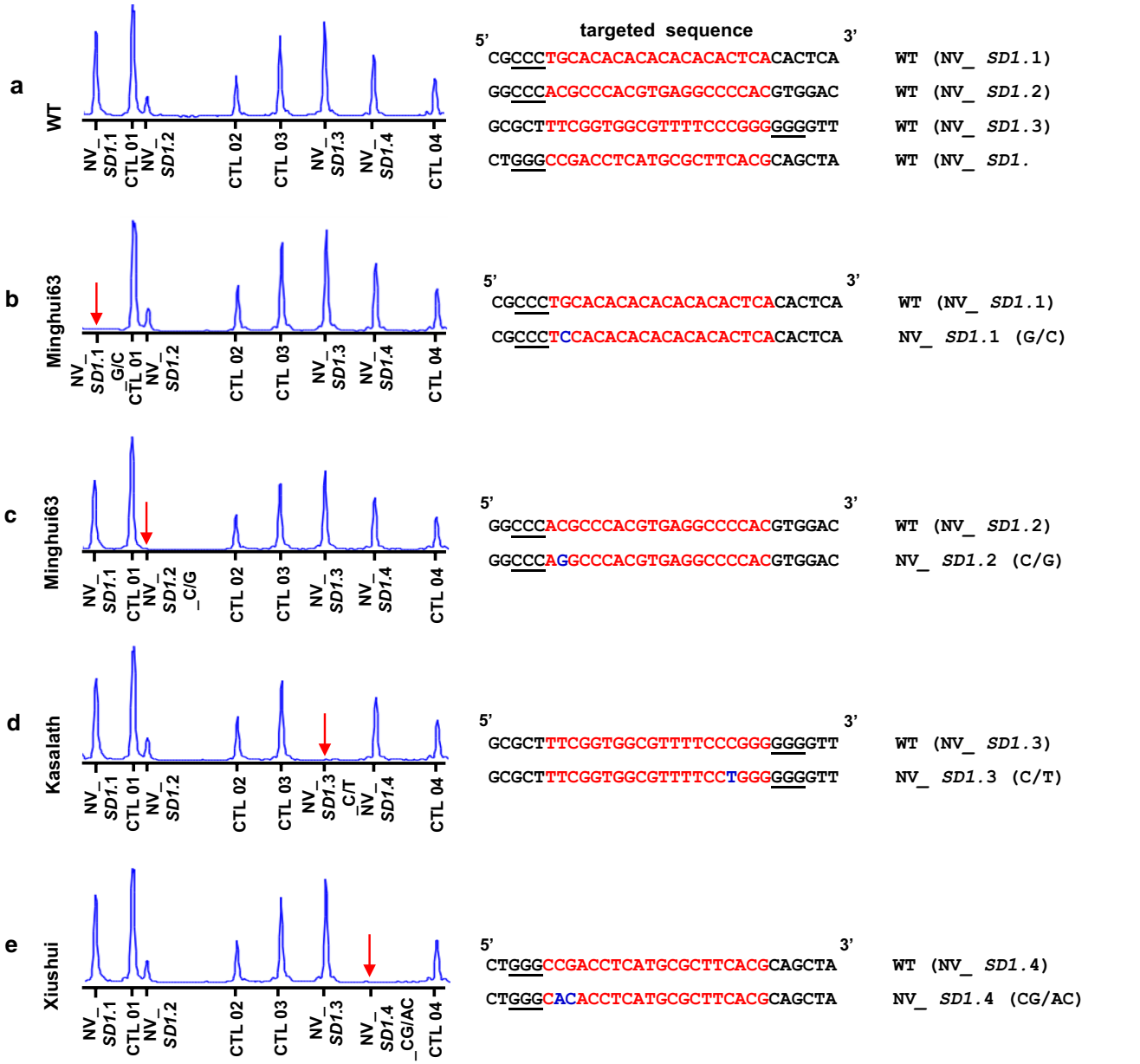

Fig. 4 Detection of natural occurred mutations by the developed MLPA-based method. The resulting chromatogram peaks separated by capillary electrophoresis of MLPA-based method are presented in the left panels, while Sanger sequencing results of the PCR products of corresponding lines with different natural variations used for the analysis are presented in the right panels. **a** Wild-type (WT) genomic DNA. **b** Minghui63 genomic DNA. **c** Minghui63 genomic DNA. **d** Kasalath genomic DNA. **e** Xiushui genomic DNA. Red arrowhead in each left panel represents the expected chromatogram peak site of the corresponding target-specific probe. Red and blue letters in the right panel represent target sequence and replaced base in the target, respectively. CTL, control probe; NV, natural variation (color figure online)

2009), MLPA-based method is labor/time-saving (no need for library preparation), cheap (no library preparation and high-throughput sequence), and sensitive (particularly for the detection of SNPs from pooled of samples). In sum, this MLPA-based method is effective in the simultaneous identification of CRISPR/Cas9 or naturally induced INDELs in plants (Fig. S4).

Different from other previously developed methods such as PCR and Amplicon labeling-based method (Biswas et al. 2019), MLPA-based method designed target-specific paired half-probes to directly adjacent to the predicted mutated-sites overlapping the target sequence. Therefore, any mismatches in the ligation point avert ligation and subsequent PCR amplification, which renders

MLPA-based method more sensitive to detect any mutations in the targeted sites. In rice, small (1–2 bp) nucleotide mismatches caused mutations are often observed in CRISPR/Cas9 targets, potential off-targets, and natural variation regions (Grohmann et al. 2019). These molecular features made MLPA-based method the most suitable tool in rice to identify CRISPR/Cas9 or natural induced INDELs in a single-tube assay (Kozlowski et al. 2008), as evidenced by the observation in this study, in which 1 bp INDELs and single nucleotide substitutions successfully prevented ligation and subsequent probe amplification (Figs. S2b–d). In addition, MLPA-based method uses mixed mutant genomic DNA, which endows it an added value for its application in rice research community, specifically for functional analysis of the targeted gene to identify mutations in a pool of mutants. Because genome edited mutants in many other cereal crops have similar characteristics (Zhu et al. 2017), this MLPA-based method developed in rice is plausible to be applicable in other cereal crops for similar purpose.

Similar to other methods, this MLPA-based method also has some drawbacks. For example, it cannot tell the exact genotype of the tested sample, which can only be resolved by Sanger sequencing; it cannot cover the whole-genome analysis and does not allow to detect mutations outside of the targeted regions. Regarding probe design, factors, such as GC contents, Tm value and probe length, can adversely affect the efficiency of the MPLA-based assay, which need careful design following proposed formula (Samelak-Czajka et al. 2017). For target sequences, those are unspecific or those with repeat elements or various SNPs cannot be readily detected by this method (Kozlowski et al. 2008). Nevertheless, combined with Sanger sequencing, the MLPA-based method is one of the most economic alternatives for effective detection or screening for CRISPR/Cas9 or natural induced INDELs in plants. To make it more efficient, it is strongly recommended to sequence the target regions in the intended germ plasm before designing CRISPR/Cas9 guide RNAs.

In summary, an effective, accurate, economic and multiplex capable MLPA-based method was developed and utilized to simultaneously detect CRISPR/Cas9-induced and naturally occurred INDELs in rice (Fig. S4), which would facilitate both rice breeding and rice functional analysis using tools such as genome editing and allelic diversification. In the future, the applicability of this MLPA-based method would be tested for more targets in rice and other plant species.

Acknowledgments This work was supported by grants from the China National Transgenic Plant Special Fund (2016ZX08012-002, 2016ZX08009-003-007 and 2017ZX08013001-001), and the Programme of Introducing Talents of Discipline to Universities (111 Project, B14016).

Author contribution statement Biswas S carried out most of the experiments and drafted the manuscript. Li R and Hong J assisted in experimentations, Zhao X, Yuan Z, and Zhang D helped in the data analysis and discussion, Shi J supervised the experiment and revised the manuscript. All authors have read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

Biswas S, Li R, Yuan Z et al (2019) Development of methods for effective identification of CRISPR/Cas9-induced indels in rice. Plant Cell Rep 38:503–510. https://doi.org/10.1007/s00299-019-02392-3

Brinkman EK, Chen T, Amendola M, Van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 42:e168. https://doi.org/10.1093/nar/gku936

Dabrowska M, Czubak K, Juzwa W, Krzyzosiak WJ, Olejniczak M, Kozlowski P (2018) QEva-CRISPR: a method for quantitative evaluation of CRISPR/Cas-mediated genome editing in target and off-target sites. Nucleic Acids Res 46(10):e101. https://doi.org/10.1093/nar/gky505

Dahlem TJ, Hoshijima K, Jurynec MJ et al (2012) Simple Methods for Generating and Detecting Locus-Specific Mutations Induced with TALENs in the Zebrafish Genome. PLoS Genet 8(8):e1002861. https://doi.org/10.1371/journal.pgen.1002861

Denbow CJ, Lapins S, Dietz N et al (2017) Gateway-compatible CRISPR-Cas9 vectors and a rapid detection by high-resolution melting curve analysis. Front Plant Sci 8:1171. https://doi.org/10.3389/fpls.2017.01171

Endo M, Mikami M, Toki S (2015) Multigene knockout utilizing off-target mutations of the CRISPR/cas9 system in rice. Plant Cell Physiol 56(1):41–47. https://doi.org/10.1093/pcp/pcu154

Feng Z, Mao Y, Xu N et al (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proc Natl Acad Sci U S A 111:4632–4637. https://doi.org/10.1073/pnas.1400822111

Findlay SD, Vincent KM, Berman JR, Postovit LM (2016) A digital pcr-based method for efficient and highly specific screening of genome edited cells. PLoS ONE 11:e0153901. https://doi.org/10.1371/journal.pone.0153901

Gao J, Wang G, Ma S et al (2015) CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. Plant Mol Biol 87(1–2):99–110. https://doi.org/10.1007/s11103-014-0263-0

Germini D, Tsfasman T, Zakharova VV et al (2018) A comparison of techniques to evaluate the effectiveness of genome editing. Trends Biotechnol 36(2):147–159. https://doi.org/10.1016/j.tibtech.2017.10.008

Grohmann L, Keilwagen J, Duensing N et al (2019) Detection and identification of genome editing in plants: challenges and opportunities. Front Plant Sci 10:1–8. https://doi.org/10.3389/fpls.2019.00236

Güell M, Yang L, Church GM (2014) Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). Bioinformatics 30:2968–2970. https://doi.org/10.1093/bioinformatics/btu427

Guo J, Li K, Jin L et al (2018) A simple and cost-effective method for screening of CRISPR/Cas9-induced homozygous/biallelic mutants. Plant Methods 14:40. https://doi.org/10.1186/s13007-018-0305-8

Hahn F, Nekrasov V (2019) CRISPR/Cas precision: do we need to worry about off-targeting in plants? Plant Cell Rep 38(4):437–441. https://doi.org/10.1007/s00299-018-2355-9

Hua Y, Wang C, Huang J, Wang K (2017) A simple and ef fi cient method for CRISPR / Cas9-induced mutant screening. J Genet Genomics 44:207–213. https://doi.org/10.1016/j.jgg.2017.03.005

Huang MC, Cheong WC, Lim LS, Li MH (2012) A simple, high sensitivity mutation screening using Ampligase mediated T7 endonuclease I and Surveyor nuclease with microfluidic capillary electrophoresis. Electrophoresis 33:788–796. https://doi.org/10.1002/elps.201100460

Ishizaki T (2016) CRISPR/Cas9 in rice can induce new mutations in later generations, leading to chimerism and unpredicted segregation of the targeted mutation. Mol Breed 36(12):165. https://doi.org/10.1007/s11032-016-0591-7

Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821. https://doi.org/10.1126/science.1225829

Kc R, Srivastava A, Wilkowski JM et al (2016) Detection of nucleotide-specific CRISPR/Cas9 modified alleles using multiplex ligation detection. Sci Rep 6:32048. https://doi.org/10.1038/srep32048

Kim H, Um E, Cho SR et al (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods 8:941–948. https://doi.org/10.1038/nmeth.1733

Kim JM, Kim D, Kim S, Kim JS (2014) Genotyping with CRISPR-Cas-derived RNA-guided endonucleases. Nat Commun 5:3157. https://doi.org/10.1038/ncomms4157

Kohata R, Koitabashi K, Kitashiba H, Nishio T (2018) Sensitive mutant detection by concentrating mutant DNA with allele-specific capture and its application to analysis of contaminated grains in rice. Plant Cell Rep 37:865–872. https://doi.org/10.1007/s00299-018-2274-9

Kozlowski P, Jasinska AJ, Kwiatkowski DJ (2008) New applications and developments in the use of multiplex ligation-dependent probe amplification. Electrophoresis 29(23):4627–4636. https://doi.org/10.1002/elps.200800126

Kozlowski P, Roberts P, Dabora S et al (2007) Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. Hum Genet 121:389–400. https://doi.org/10.1007/s00439-006-0308-9

Li J, Manghwar H, Sun L et al (2019) Whole genome sequencing reveals rare off-target mutations and considerable inherent genetic or/and somaclonal variations in CRISPR/Cas9-edited cotton plants. Plant Biotech J 17:858–868. https://doi.org/10.1111/pbi.13020

Li M, Li X, Zhou Z et al (2016) Reassessment of the four yield-related genes Gn1a, DEP1, GS3, and IPA1 in rice using a CRISPR/Cas9 system. Front Plant Sci 7:377. https://doi.org/10.3389/fpls.2016.00377

Lusser M, Parisi C, Plan D, Rodríguez-Cerezo E (2012) Deployment of new biotechnologies in plant breeding. Nat Biotechnol 30(3):231–239

Marcinkowska M, Wong KK, Kwiatkowski DJ, Kozlowski P (2010) Design and generation of MLPA probe sets for combined copy number and small-mutation analysis of human genes: EGFR as an example. Sci World J 10:2003–2018. https://doi.org/10.1100/tsw.2010.195

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4325. https://doi.org/10.1093/nar/8.19.4321

Nekrasov V, Staskawicz B, Weigel D et al (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol 31:691–693

Ota S, Hisano Y, Muraki M et al (2013) Efficient identification of TALEN-mediated genome modifications using heteroduplex mobility assays. Genes Cells 18:450–458. https://doi.org/10.1111/gtc.12050

Pan C, Ye L, Qin L et al (2016) CRISPR/Cas9-mediated efficient and heritable targeted mutagenesis in tomato plants in the first and later generations. Sci Rep 6:24765. https://doi.org/10.1038/srep24765

Ren C, Liu X, Zhang Z et al (2016) CRISPR/Cas9-mediated efficient targeted mutagenesis in Chardonnay (*Vitis vinifera* L.). Sci Rep 6:32266. https://doi.org/10.1038/srep32289

Rigola D, van Oeveren J, Janssen A et al (2009) High-throughput detection of induced mutations and natural variation using KeyPoint™ technology. PLoS ONE 4(3):e4761. https://doi.org/10.1371/journal.pone.0004761

Shan Q, Wang Y, Li J, Gao C (2014) Genome editing in rice and wheat using the CRISPR/Cas system. Nat Protoc 9(10):2395–2410. https://doi.org/10.1038/nprot.2014.157

Samelak-Czajka A, Marszalek-Zenczak M, Marcinkowska-Swojak M, Kozlowski P, Figlerowicz M, Zmienko A (2017) MLPA-based analysis of copy number variation in plant populations. Front Plant Sci 8:222. https://doi.org/10.3389/fpls.2017.00222

Tang X, Liu G, Zhou J et al (2018) A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. Genome Biol 19(1):84. https://doi.org/10.1186/s13059-018-1458-5

Thomas HR, Percival SM, Yoder BK, Parant JM (2014) High-throughput genome editing and phenotyping facilitated by high resolution melting curve analysis. PLoS ONE 9(12):e114632. https://doi.org/10.1371/journal.pone.0114632

Vouillot L, Thélie A, Pollet N (2015) Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. G3 Genes. Genomes Genet 5:407–415. https://doi.org/10.1534/g3.114.015834

Wang K, Mei DY, Liu QN et al (2015) Research of methods to detect genomic mutations induced by CRISPR/Cas systems. J Biotechnol 214:128–132. https://doi.org/10.1016/j.jbiotec.2015.09.029

Xie K, Minkenberg B, Yang Y (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci U S A 112(11):3570–3575. https://doi.org/10.1073/pnas.1420294112

Xu RF, Li H, Qin RY et al (2015) Generation of inheritable and "transgene clean" targeted genome-modified rice in later generations using the CRISPR/Cas9 system. Sci Rep 5:11491. https://doi.org/10.1038/srep11491

Yan WX, Mirzazadeh R, Garnerone S et al (2017) BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. Nat Commun 8:15058. https://doi.org/10.1038/ncomms15058

Yu C, Zhang Y, Yao S, Wei Y (2014) A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish. PLoS ONE 9(6):e98282. https://doi.org/10.1371/journal.pone.0098282

Zhang H, Zhang J, Wei P et al (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol J 12(6):797–807. https://doi.org/10.1111/pbi.12200

Zhang Q, Xing HL, Wang ZP et al (2018) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. Plant Mol Biol 96(4–5):445–456. https://doi.org/10.1007/s11103-018-0709-x

Zhao Q, Feng Q, Lu H et al (2018) Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. Nat Genet 50:278–284. https://doi.org/10.1038/s41588-018-0041-z

Zhu C, Bortesi L, Baysal C et al (2017) Characteristics of genome editing mutations in cereal crops. Trends Plant Sci 22(1):38–52. https://doi.org/10.1016/j.tplants.2016.08.009

Zischewski J, Fischer R, Bortesi L (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. Biotechnol Adv 35:95–104. https://doi.org/10.1016/j.biotechadv.2016.12.003

Zong Y, Wang Y, Li C et al (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat Biotechnol 35(5):438–440. https://doi.org/10.1038/nbt.3811