frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 October 2017
doi: 10.3389/fpls.2017.01851

Identification of Substitutions and Small Insertion-Deletions Induced by Carbon-Ion Beam Irradiation in *Arabidopsis thaliana*

Yan Du[1†], Shanwei Luo[1,2†], Xin Li[1], Jiangyan Yang[1], Tao Cui[1,2], Wenjian Li[1], Lixia Yu[1], Hui Feng[1,2], Yuze Chen[3], Jinhu Mu[1,2], Xia Chen[1,2], Qingyao Shu[4], Tao Guo[5], Wenlong Luo[5] and Libin Zhou[1]*

[1] Biophysics Group, Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China, [2] College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, [3] College of Life Sciences and Technology, Gansu Agricultural University, Lanzhou, China, [4] National Key Laboratory of Rice Biology, Institute of Crop Sciences, Zhejiang University, Hangzhou, China, [5] National Engineering Research Center of Plant Space Breeding, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Michael Deyholos,
University of British Columbia, Canada

Reviewed by:
Chengdao Li,
Murdoch University, Australia
Mehboob-ur-Rahman,
NIBGE, Pakistan

***Correspondence:**
Libin Zhou
libinzhoulz@gmail.com

[†]These authors have contributed equally to this work.

Specialty section:
This article was submitted to Plant Genetics and Genomics, a section of the journal Frontiers in Plant Science

Received: 10 July 2017
Accepted: 11 October 2017
Published: 27 October 2017

Citation:
Du Y, Luo S, Li X, Yang J, Cui T, Li W, Yu L, Feng H, Chen Y, Mu J, Chen X, Shu Q, Guo T, Luo W and Zhou L (2017) Identification of Substitutions and Small Insertion-Deletions Induced by Carbon-Ion Beam Irradiation in Arabidopsis thaliana. Front. Plant Sci. 8:1851. doi: 10.3389/fpls.2017.01851

Heavy-ion beam irradiation is one of the principal methods used to create mutants in plants. Research on mutagenic effects and molecular mechanisms of radiation is an important subject that is multi-disciplinary. Here, we re-sequenced 11 mutagenesis progeny (M3) *Arabidopsis thaliana* lines derived from carbon-ion beam (CIB) irradiation, and subsequently focused on substitutions and small insertion-deletion (INDELs). We found that CIB induced more substitutions (320) than INDELs (124). Meanwhile, the single base INDELs were more prevalent than those in large size (≥2 bp). In details, the detected substitutions showed an obvious bias of C > T transitions, by activating the formation of covalent linkages between neighboring pyrimidine residues in the DNA sequence. An A and T bias was observed among the single base INDELs, in which most of these were induced by replication slippage at either the homopolymer or polynucleotide repeat regions. The mutation rate of 200-Gy CIB irradiation was estimated as 3.37×10^{-7} per site. Different from previous researches which mainly focused on the phenotype, chromosome aberration, genetic polymorphism, or sequencing analysis of specific genes only, our study revealed genome-wide molecular profile and rate of mutations induced by CIB irradiation. We hope our data could provide valuable clues for explaining the potential mechanism of plant mutation breeding by CIB irradiation.

Keywords: *Arabidopsis thaliana*, carbon-ion beam (CIB) irradiation, mutation, molecular spectrum, small INDELs, substitutions, whole genome-wide re-sequencing

INTRODUCTION

The heavy-ion beam is an effective and unique mutagen that can induce mutations at a high rate and on a broad spectrum. It has been widely used in the mutation breeding of plants and microbes because of their distinct physical and biological advantages. The most important physical aspect of the heavy-ion beam is the greater linear energy transfer (LET) than X-rays, gamma rays, and electrons. LET represents the energy deposition of ionizing radiations on their per unit track. This high concentration of deposited energy can cause more severe damage to the target biomolecules,

such as, proteins, membrane lipids, and nucleic acids. Hence, the heavy-ion beam provides a higher relative biological effectiveness (RBE) than the low LET radiation (Tanaka et al., 2010; Kazama et al., 2011; Nagata et al., 2016; Zhou et al., 2016). For plant mutation breeding, various mutant populations have been generated by heavy-ion beam irradiation: for instance, those of *Arabidopsis thaliana* (Tanaka et al., 1997, 2002), *Lotus japonicus* (Oka-Kira et al., 2005; Luo et al., 2016), Rice (*Oryza sativa* L.) (Ishikawa et al., 2012; Phanchaisri et al., 2012; Morita et al., 2017), Petunia (*Petunia hybrid*), Hase et al., 2010), *Tricyrtis hirta* (Nakano et al., 2010), Chrysanthemum (Matsumura et al., 2010), Wandering Jew (He et al., 2011), Geranium (Yu et al., 2016), and so on.

The creation of such mutations is an essential resource tool to better link phenotype screening, gene isolation, and function mining, in both forward and reverse genetics. DNA, carrying the genetic information, was originally thought as the main target of radiation. The radiation-induced DNA damage involved direct injury caused by thermal, dynamic, and ionization, as well as the indirect injury caused by the cytotoxic reaction to energy deposition (Ravanat et al., 2001; Kovacs and Keresztes, 2002; Alizadeh et al., 2013; Tokuyama et al., 2015). The mutagenic effects and molecular characteristics of mutations induced by heavy-ion beam are important in two respects: (1) For choosing the most suitable kinds of ionizing radiation; and (2) For elucidating the molecular profile and rate of mutations induced by heavy-ion beam at DNA level. Together, they form an important subject that is multi-disciplinary, such as radiobiology, molecular biology, genetics, biochemistry, etc. In the last century, the characterization of mutations induced by radiation was done by focusing on the phenotype, chromosome aberration, genetic polymorphism or sequencing analysis of specific genes only (Shikazono et al., 2005; Kazama et al., 2011; Hase et al., 2012; Hirano et al., 2012; Yan et al., 2014). However, whether restricted by detecting technology or by high operating costs, we still lack the investigation at whole genome level.

In response to the urgent demands of genome-wide studies, next generation sequencing (NGS) [also called the high-throughput sequencing (HTS)] techniques (Schuster, 2008; Shendure and Ji, 2008)—such as, Illummia/Solexa, Roche/454, and AB/SOLiD—were rapidly developed to enable the fast sequencing and mutation scanning for genomic mutation identification, species evolution, genetic diseases, etc. (Uchida et al., 2011; Belfield et al., 2012; Bolon et al., 2014; Xia et al., 2015; Zhou et al., 2015; Tiwari et al., 2016; Lehrbach et al., 2017; Tao et al., 2017). For instance, using NGS techniques, mutations induced by ethylmethane sulphonate (EMS) were found to be mainly consist of G/C–A/T transitions (Uchida et al., 2011). An EMS-induced causal mutation in *CTR1* required for boron-mediated root development by low-coverage genome re-sequencing in *A. thaliana* was successfully identified though association of re-sequencing and rough map-based cloning (Tabata et al., 2013). An efficient protocol, based on the NGS, to map and identify EMS-induced mutations in *Caenorhabditis elegans* were reported in 2017 (Lehrbach et al., 2017). Moreover, results obtained from NGS techniques have shown that the variations induced by radiation were, to some extent, different from the publications which have been previously reported. Taking the fast neutron exposure for instance: it was deemed to predominantly induce large deletions in size more than a kilo base in *A. thaliana* by the southern blot analysis of the restricted loci. However, a re-sequencing analysis of six mutants derived from fast neutron exposure indicated that the proportion of substitutions prevails over insertion-deletion (INDELs) mutations and that small deletions were more common than large fragment deletions (Bruggemann et al., 1996; Belfield et al., 2012; O'Rourke et al., 2013). Therefore, utilizing the NGS techniques would complement and complete

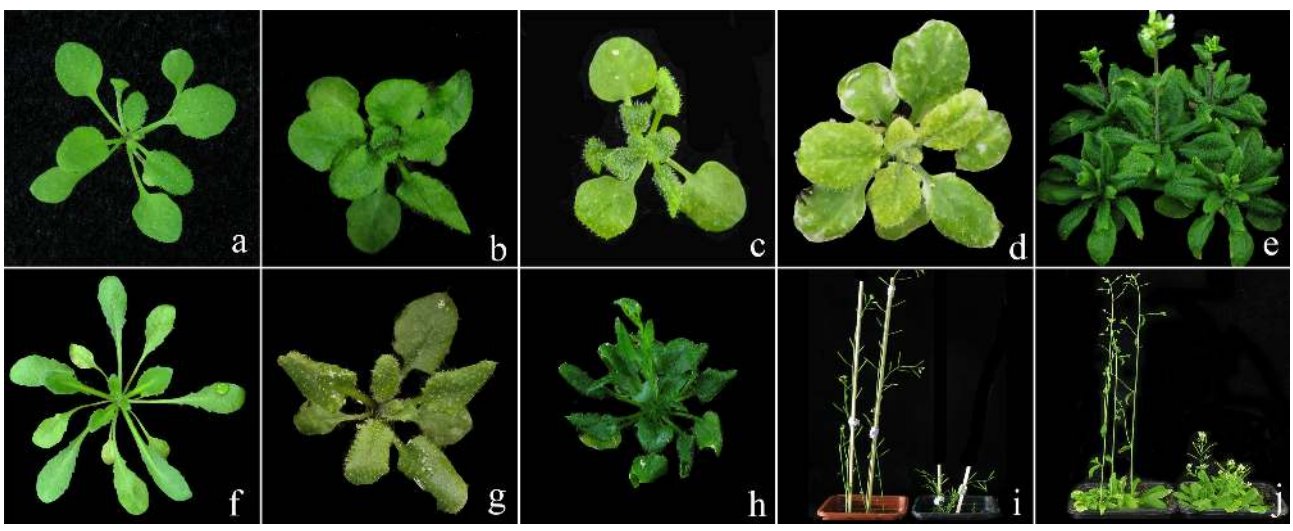

FIGURE 1 | Phenotypes of the nine re-sequenced lines that had stable mutation traits induced by CIB irradiation. **(a)**: wild type (ecotype Col); **(b–j)**: mutagenesis progeny lines in the corresponding order of C7, C197, C352, C357, C600, C828, C941, C116 and C541.

our knowledge of the actual nature of mutations induced by ionizing radiation. The corresponding algorithm tools, such as, the Sequence Alignment/Map (SAM) (Li et al., 2009), Bowtie (Langmead et al., 2009), the SHORE pipeline (Schneeberger et al., 2009), and VarScan (Koboldt et al., 2009, 2012), have been rapidly developed for processing big datasets obtained by NGS. Specifically, the flow of the Illummia/Solexa, one of the most mainstream NGS platforms, is as the following: (1) fragment the genomic DNA; (2) perform end-joining, followed by adding A to 3′ end; (3) ligation of the Solexa adapter; (4) recover and purify the target fragment, then amplify it via PCR; (5) cluster amplification and sequencing-by-synthesis reads obtained by sequencing, (6) data processing. For whole genome re-sequencing of model organisms, the clean reads are mapped against the reference genome (associated with bioinformatics analysis). The genomic mutations (i.e., the sequence and structure variants) are then called by using the corresponding algorithm tools.

In the current study, we re-sequenced 11 mutagenesis progeny (M3) *A. thaliana* lines—nine mutants with visible and heritable traits, and two M3 mutagenesis progeny with inconspicuous phenotypes—derived from carbon-ion beam (CIB) irradiation by using the Illumina sequencing platform. Focusing on substitutions and small INDELs, we revealed the genome-wide molecular profile and rate of mutations induced by CIB irradiation in *A. thaliana*.

MATERIALS AND METHODS

Irradiation of CIB and Mutant Screening

The seeds of laboratory wild type *A. thaliana* (Lab-WT)—Columbia genetic background—were exposed to $^{12}C^{6+}$ ions at 200 Gy (energy, 43 MeV/nucleon; average LET within samples, 50 keV/μm) generated by the Heavy Ion Research Facility in Lanzhou (HIRFL) at the Institute of Modern Physics, Chinese Academy of Sciences (IMP-CAS). Methods of plant growth and mutation screening were described in a previous study (Yan et al., 2014). Nine mutant lines (M3) that displayed visible and heritable traits (C7, C116, C197, C352, C357, C541, C600, C828,

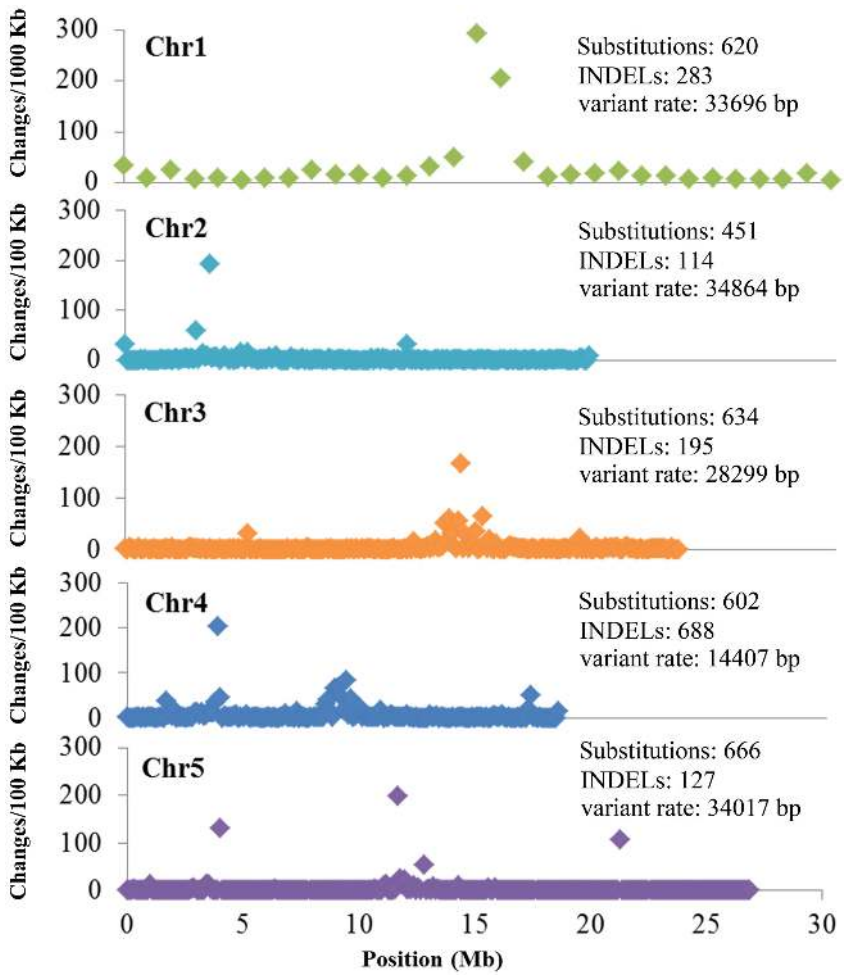

FIGURE 2 | Background mutations shared by the 12 re-sequenced *Arabidopsis thaliana* lines. Shown are the density and variant rate of the single base substitutions and small INDELs that were shared by 12 re-sequenced lines across each chromosome.

C941), as well as two mutagenesis progeny (M3) lines which with inconspicuous mutant phenotypes (C1001, C1322), along with the Lab-WT line were chosen for whole genome re-sequencing.

Whole Genome Re-sequencing

Young leaves were collected from the Lab-WT line and 11 M3 lines derived from CIB. Genomic DNA was extracted by using the routine protocol of CTAB (cetyltrimethylammonium bromide). The genomic re-sequencing was performed by the IIIumina HiSeq^{TM} 2500 system (Illumina, Inc.; San Diego, CA, USA) at the Biomarker Technologies Company (Beijing, China). Raw data were filtered following the standard process of Illumina, and the clean data obtained were then mapped onto the reference genome of Col-0 (TAIR10, http://www.ncbi.nlm.nih.gov/assembly/GCF_000001735.3) by using the Burrows-Wheeler Alignment tool (http://bio-bwa.sourceforge.net/bwa.shtml, v.0.7.15; Li and Durbin, 2009) and SAMtools (http://www.htslib.org/workflow/#mapping_to_variant, v.1.3.1; Li et al., 2009). The average depth of the re-sequenced lines had a 23–33-fold range, while those reads with a depth >10-fold accounted for up to 98.87% of the total reference genome, on average (Table S1).

Identification of the Variants

The VarScan 2 algorithms (v.3.9, http://varscan.sourceforge.net) were used to read the SAMtools mpileup output, and the substitutions and small INDELs were called with settings described below:

Min coverage: Minimum read depth at a position to make a call is eight;

Min reads2: Minimum supporting reads at a position to call variants is two;

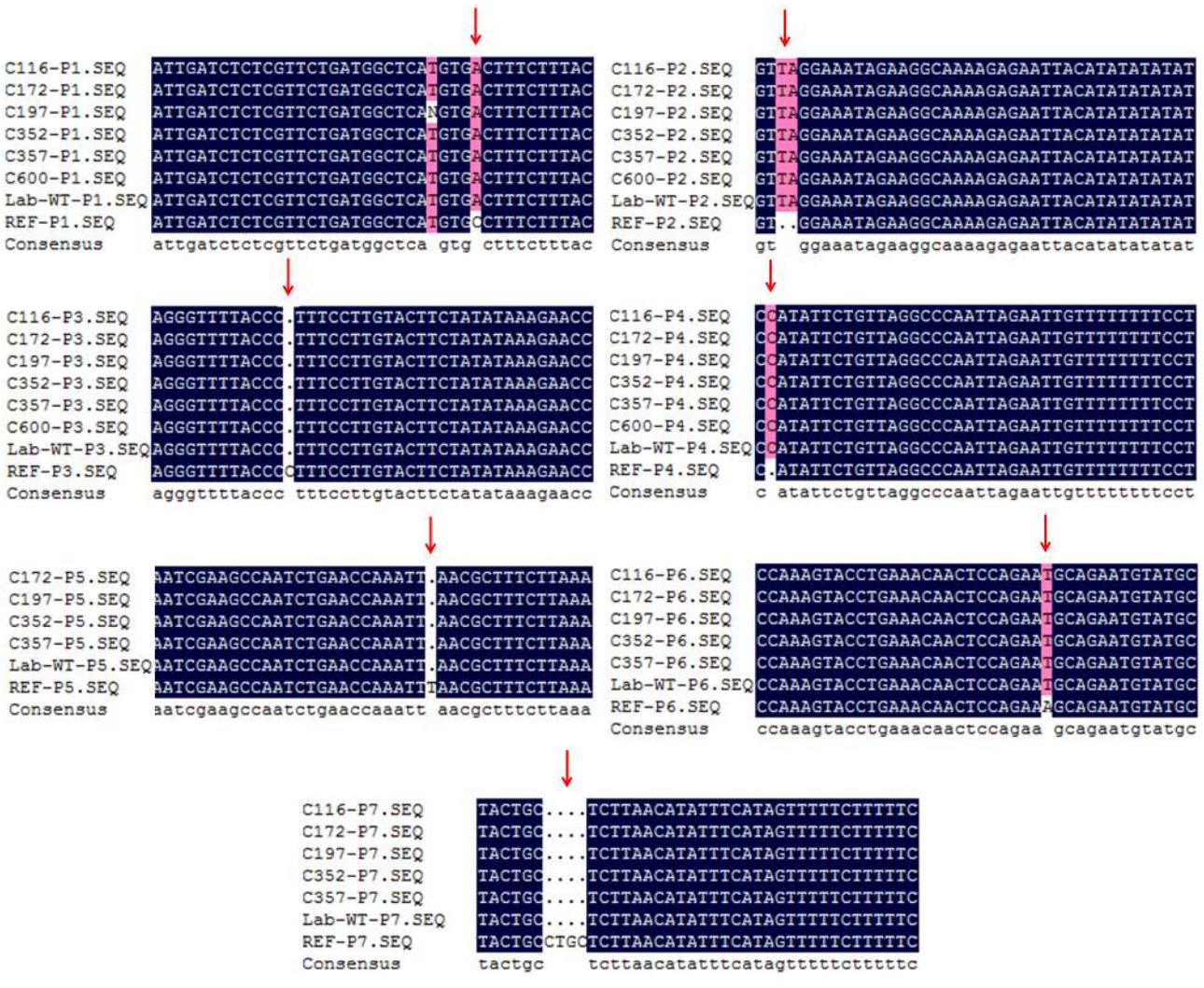

FIGURE 3 | Verification of background mutations by Sanger sequencing. The REF denotes the Col-0 reference genome. The sequence alignment of different primers (P1–P7) from C116, C172, C197, C352, C357, Lab-WT and reference were shown with different colors according to their sequence homology. Pink highlights the sequence which are different from the reference genome, the dot indicates deletions in the corresponding position.

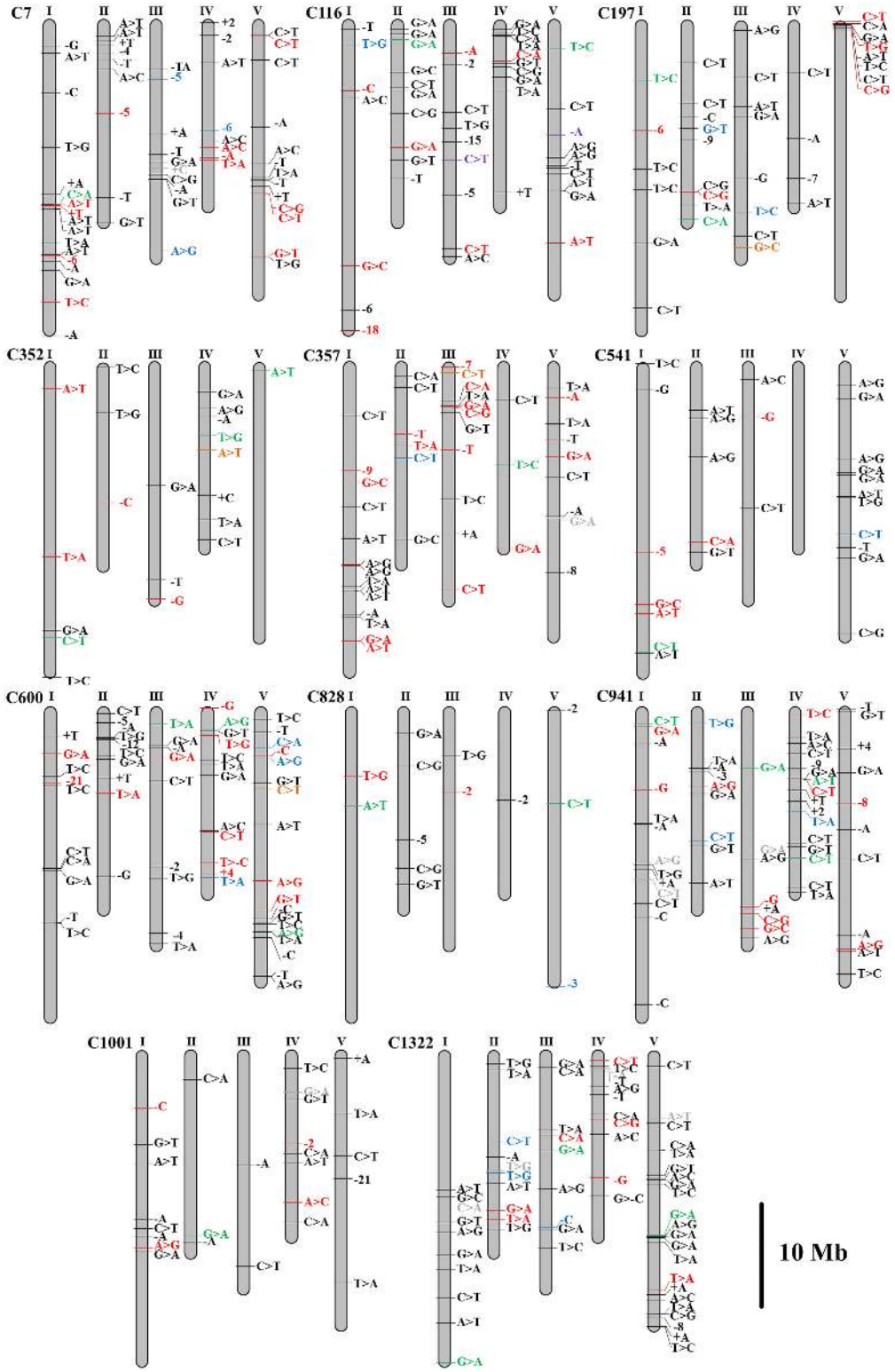

FIGURE 4 | Distributions of the substitutions and small INDELs across chromosomes in the genomes of 11 mutagenesis progeny lines (M3) of *Arabidopsis thaliana*. Single-base INDELs are indicated by base-designating letters, with a preceding minus sign (deletion) or plus sign (insertion). Multiple-base INDELs are indicated by a minus or plus sign with the number of deleted or inserted bases. Individual colors distinguish the possible mutation effects: missense, frame shift, in-frame deletion, stop gained/lost (red); synonymous (green); UTR (blue); splice site region, intron (orange); intergenic region (gray); up downstream regions (black); non-coding exon (purple).

Mmin-var-freq: Minimum variant allele frequency threshold is 0.01;
P-value: Default *p*-value threshold for calling variants is 0.99.
The true candidate mutations induced by CIB were screened as followed steps:

(1) Following the VarScan 2 algorithms, to filter the background mutations in the Lab-WT line—and to identify the particular mutations of each mutagenesis progeny—only those mutation sites with a variant allele frequency by read count were between 25–100% in a sample, while 0–10% in other samples were reserved; those variants with a variant allele frequency of ≥75% were called homozygous. Against this, the sites shared by the 12 sequenced lines were called as the background mutations.

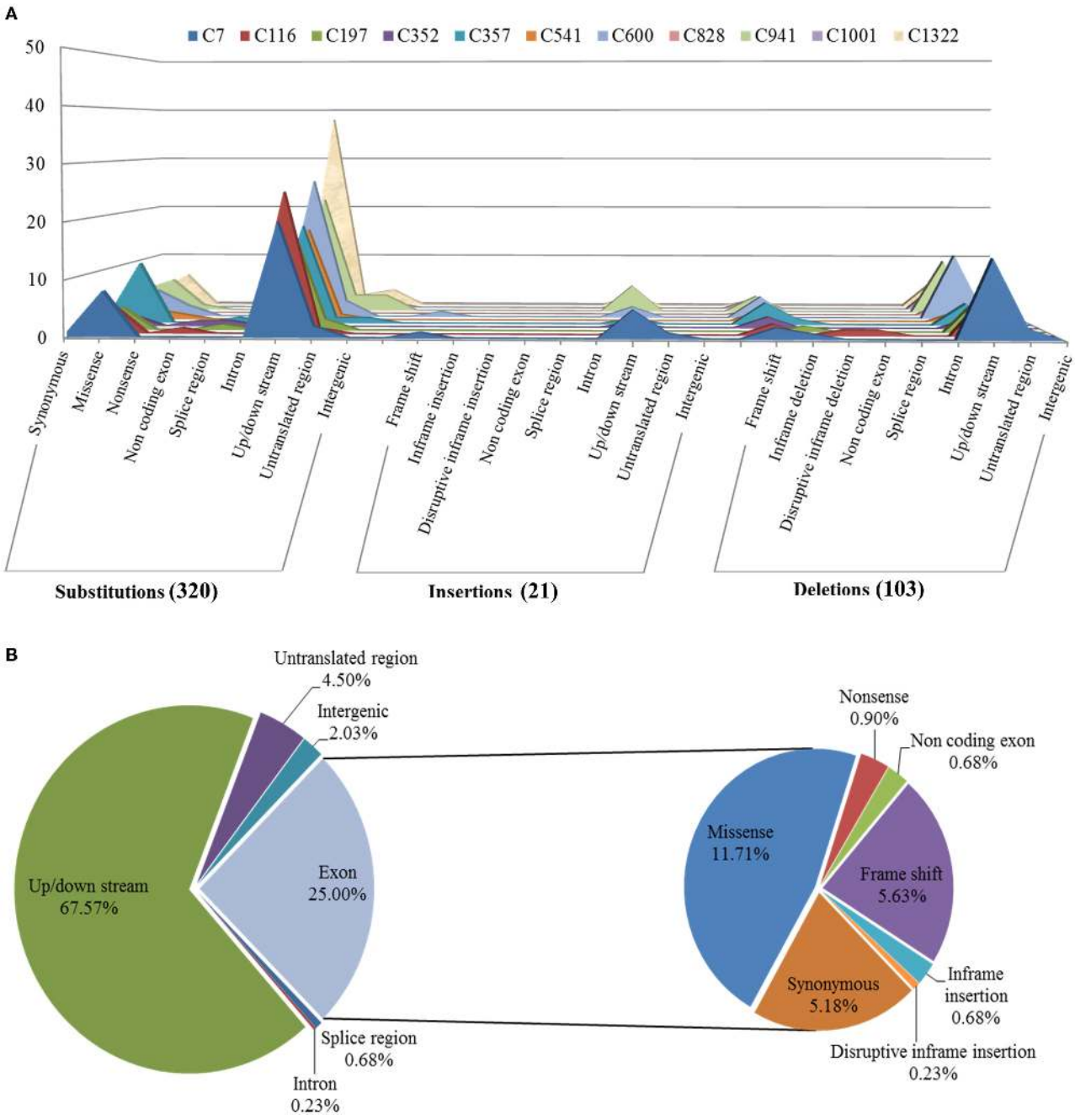

FIGURE 5 | Annotation of mutations induced by CIB irradiation in genomes of 11 mutagenesis progeny lines (M3) of *Arabidopsis thaliana*. **(A)** Mutations comprehensively inferred by BWA, SAMtools, VarScan 2, as well as their distributions among functional classes and mutation effects in each CIB-irradiated M3 plant line. **(B)** Overall distribution of mutations induced by CIB irradiation at the whole genomic level in *Arabidopsis thaliana*.

(2) Although the filtering standard of step (1) was effective at reducing the rate of false positives, to ensure the accuracy of mutation detection, the candidate mutations derived from step (1) of each sequenced line were visually confirmed in the Integrative Genomics Viewer (http://software.broadinstitute.org/software/igv/, v.2.3).

Mutation Annotation

Variant annotation and predicted effects were performed by the SnpEff toolbox (http://snpeff.sourceforge.net/index.html, v.4.2).

Verification of the Mutation Sites by Sanger Sequencing

The specific primers of mutation sites obtained by re-sequencing were designed by PRIMER3 (http://bioinfo.ut.ee/primer3-0.4.0/). The primers used for PCR are P1 (5′-TTTGCTGTTGGTAGTGTGCC-3′, 5′-TGGAAGTAGAGTGCAGCGAT-3′), P2 (5′-TATGCTTGCACGTTTGCTTC-3′, 5′-GGCTTGAGAGTTGGGTGAAC-3′), P3 (5′-CACTTCTTTTCCCCAACGGT-3′, 5′-TACATGCTTGCATCACCCAC-3′), P4 (5′-GACGACCTCCCAGATATGCC-3′, 5′-GGAGATTGTGGGGTCCCAAG-3′), P5 (5′-CAGGAGCAACTTACCAACGG-3′, 5′-GAACAATACAGGTGGCGTGG-3′), P6 (5′-TGTGATCAAGAGTGTCGGCT-3′, 5′-ACGTCAAAACTTCCTCCCCA-3′) and P7(5′-CCCCTGATAAGTTGCGTTTAAGT-3′, 5′-ACTTGATGACATGGGAGGCA-3′). Genomic DNA was extracted by using the routine protocol of CTAB. PCR amplification was performed with an initial denaturation step at 95°C for 5 min, followed by 38 cycles at 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s, with a final extension step at 72°C for 10 min. The PCR products were then detected by electrophoresis on a 1.5% agarose gel; only the qualified PCR products were used for Sanger sequencing.

Data Availability

The whole-genome sequencing data reported in this study have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics, 2017) in BIG Data Center (Nucleic Acids Res, 2017), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession number CRA000464 that is publicly accessible at http://bigd.big.ac.cn/gsa.

RESULTS

CIB Irradiation and Mutagenesis Progeny

The Lab-WT *A. thaliana* seeds were exposed to CIB irradiation with the aim (though not the focus of the present study) of constructing a comprehensive carbon-ion-induced mutation resource collection for *A. thaliana*. The M2 seedlings were screened for abnormal phenotypes by comparison with the Lab-WT; meanwhile, several lines lacking obvious mutant traits were also reserved. In total, more than 1,000 independent plants were isolated. Our subsequent analyses were based on nine mutant lines (M3) that displayed visible and heritable traits (i.e., C7, C116, C197, C352, C357, C541, C600, C828, C941), as well as two mutagenesis progeny (M3) lines which with inconspicuous mutant phenotypes (i.e., C1001, C1322), together with the Lab-WT line were chosen for whole genome re-sequencing. The phenotypes of the nine stable lines are shown in **Figure 1**.

Mutations Shared by the 12 Sequenced Lines

DNA sequence data from a single Lab-WT line and other 11 mutagenesis progeny lines were aligned to the *A. thaliana* Information Resource (TAIR 10) (http://www.arabidopsis.org) Col-0 reference genome by using BWA, Samtools, and VarScan 2. The mapping results for the mutant sequencing reads for each line are shown in Table S1. Before the variants analysis, we surmised that there were a number of substitutions and INDELs which might be the background discrepancies that originally existed between the Lab-WT and Col-0 reference genomes. To identify those variations truly induced by CIB irradiation, it was reasonable to rule out these background discrepancies firstly.

TABLE 1 | Variant rate of the single base substitutions and small INDELs across each chromosome in the 11 re-sequenced lines.

Line	Chromosome 1		Chromosome 2		Chromosome 3		Chromosome 4		Chromosome 5	
	Variants[a]	Rate[b]	Variants	Rate	Variants	Rate	Variants	Rate	Variants	Rate
C7	17	1789863	9	2188699	10	2345983	8	2323132	13	2075039
C116	7	4346810	10	1969829	9	2606648	10	1858506	10	2697550
C197	6	5071279	9	2188699	8	2932479	4	4646264	8	3371938
C352	5	6085534	3	6566096	3	7819943	8	2323132	1	26975502
C357	14	2173405	6	3283048	11	2132712	3	6195019	9	2997278
C541	7	4346810	5	3939658	3	7819943	0	/	11	2452318
C600	10	3042767	10	1969829	9	2606648	12	1548755	18	1498639
C828	2	15213836	5	3939658	2	11729915	1	18585056	3	8991834
C941	13	2340590	9	2188699	8	2932479	16	1161566	11	2452318
C1001	8	3803459	3	6566096	2	11729915	8	2323132	5	5395100
C1322	10	3042767	10	1969829	9	2606648	11	1689551	22	1226159

[a] *Variants indicate the number of mutations in each chromosome.*

[b] *Rate equals to length of chromosome/variants. Length (bp) of chromosome 1 to 5 is 30427671, 19698289, 23459830, 18585056, and 26975502, respectively.*

Analyzed by group, since the 11 mutagenesis progeny lines were in the same background (Columbia), the variants shared by the 12 sequenced lines were considered as the background mutations. On this premise, a total of 2,973 substitutions and 1,407 INDELs were identified (**Figure 2**).

To verify whether the background mutations were indeed correct, seven sites shared by the 12 sequenced lines were selected for Sanger sequencing. To increase the credibility, another M3 line (C172), one not re-sequenced by NGS, was also included. The Sanger sequencing results indicated that those sites obtained by NGS were indeed common background mutations in our Lab-WT (**Figure 3**). These background mutations were unevenly distributed in the Lab-WT genome: the highest mutation number was observed on chromosome 4, with a change rate of one per 14,407 nucleotide base pairs along the genome. Interestingly, substitutions and INDELs tended to occur in the peri-centromeric regions of each chromosome as clusters, especially in chromosome 1. This result is basically consistent with the spontaneous mutations in *A. thaliana* in previous report (Ossowski et al., 2010), although the mechanism of such a mutational bias remains unexplored.

Distribution and Rates of Mutations Induced by CIB Irradiation

To identify the exclusive mutations of each mutagenesis progeny line, combining the mutation calling by VarScan 2 and the visual confirmation by the Integrative Genomics Viewer (IGV), a total of 320 substitutions and 124 INDELs were detected, and the ratio of substitutions to INDELs was calculated as 2.58:1 (**Figures 4**, **5A**). The number of variations per line ranged from 20 to 62, and their distribution across the chromosomes in each of the 11 lineages is shown in **Figure 4**. However, no remarkable regularity in variant rate of single base substitutions and small INDELs across each chromosome were found in the 11 re-sequenced lines (**Table 1**). At the genome-wide scale, 67.57% of the 444 detected CIB irradiation-induced mutations occurred in the upstream and downstream regions, 4.50% in the 3′/5′-untranslated regions (UTR), 25.00% in exon, 0.68% in the splice region, 0.23% in intron, and 2.03% in the intergenic region (**Figure 5B**).

Mutations involved in missense, stop gained/lost, frame shift, in-frame deletion, and 3′/5′-UTR, are usually predicted to affect gene function more probably. Classifying all the above sites including both heterozygous and homozygous, 103 mutations located in 97 genes were sorted, and the number of mutated genes affected by irradiation in the M3 lines ranged from 3 to 15 (**Table 2** and Table S2). The homozygous mutations in M3 were used to estimate the M1 heterozygous mutation events caused by CIB irradiation followed Mendelian principles (Belfield et al., 2012). The total number of single base substitutions, single base insertions, and single base deletions in M1 of the 11 sequenced lines were 362.67, 18.67, and 69.33, respectively (Table S3). After correcting the spontaneous mutations by using the Col-0 MA

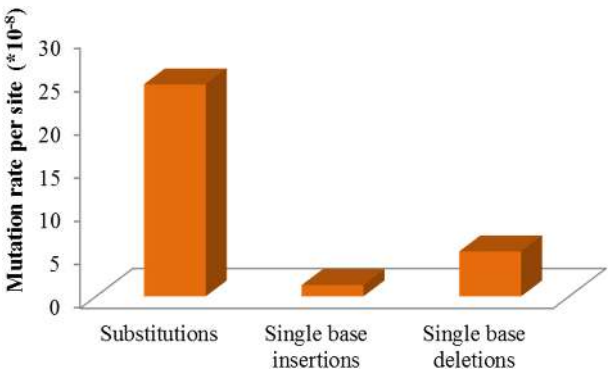

FIGURE 6 | Single nucleotide mutation rates of CIB irradiation on *Arabidopsis thaliana*.

TABLE 2 | Numbers of genes predicted to occur function changes in the 11 re-sequenced lines.

	Homozygous				Heterozygous				Total mutated genes[c]
Line	3′/5′UTR	Non-synonymous[a]	Frame shift	Disruptive inframe_del[b]	3′/5′UTR	Non-synonymous	Frame shift	Disruptive inframe_del	
C7	2	4	1	0	2	3	3	1	13
C116	0	5	2	1	1	0	0	0	9
C197	0	0	0	0	2	4	0	1	7
C352	0	2	2	0	0	0	0	0	4
C357	1	6	3	0	0	3	1	1	15
C541	0	1	1	0	1	2	1	0	6
C600	1	1	0	0	2	6	5	0	15
C828	0	0	0	0	1	1	1	0	3
C941	2	3	1	0	1	4	2	0	13
C1001	0	1	1	0	0	1	1	0	4
C1322	0	0	0	0	2	6	1	0	9

[a] *Non-synonymous mutations include those mutations with the effect of missense, stop gained, stop lost, and initiator codon variant.*

[b] *Disruptive inframe_del indicates that the mutation lead one codon changed, and one or more codons are deleted.*

[c] *When multiple mutations are located in the same gene in one re-sequenced line, and these mutations are viewed as a single gene-mutated.*

line mutation rates (Ossowski et al., 2010), on average the single base mutation rate was estimated to be 3.37×10^{-7} per site per genome (**Figure 6**). The mutation rate of the 200-Gy CIB irradiation was nearly 47-fold that of the spontaneous rate (7.1×10^{-9} per site). Since the phenotypes of C7, C357, C116, and C541 were heritable, the homozygous mutations in M3 were used to detect whether there was any common mutations at the genomic level. However, no shared genetic factors were found between C116 and C541, or between C7 and C357. To some extent, this indicated that the CIB irradiation was able to induce mutation in multiple genes which could give rise to similar traits (**Figure 7**).

Single Base Substitutions Induced by CIB Irradiation

Exposure to CIB irradiation induced plenty of substitutions. Their type may be classified into two categories: transition (mutations that happen among the same type of bases; e.g., purine > purine or pyrimidine > pyrimidine) and transversion (mutations that happen among the different types of bases; e.g., purine > pyrimidine or pyrimidine > purine). A total of 320 CIB irradiation-induced substitutions were identified and classed, yielding a transition to transversion (Ti/Tv) ratio of 0.99. This indicated that the CIB irradiation was able to induce transitions and transversions at nearly the same frequency (**Figure 8A**). The G:C > A:T were the most frequently observed substitutions in the re-sequenced mutagenesis progeny lines: 51 of the 320 detected CIB-induced mutations were G > A transitions, 55 were C > T mutations (Table S4). The C > T transitions hold the biggest ratio of substitutions induced by the CIB. By analyzing the flanking DNA sequences of the most prominent substitutions of C > T variations, it was found that 37 of these (67.2%) occurred at the pyrimidine dinucleotide sites (**Table 3**). In addition, the substitutions seemed more likely to happen at the C base site (87) rather than at other base sites (A, G, and T) (**Figure 8D**).

Small INDELs Induced by CIB Irradiation

For a long time, deletions were considered the main mutation type induced by heavy-ion beam irradiation, or at least had a similar frequency to substitutions, judging from the sequencing analysis of specific genes only. In present study, at a genome-wide scale, 124 small INDELs were detected, of which 103 (83.06%) were deletion mutations, whereas only 21 (16.94%) were insertions (**Figure 8B**). Among the 103 small deletions, 35 (33.98%) multiple bases deletions (≥2 bp) ranging in size from 2–21 bp were detected (**Figures 8B,C**). In addition, 68 (66.02%) single base deletions were detected as well, which seemed the most common class of deletion mutation. Meanwhile, 17 (80.95%) of the 21 insertions were single base mutations, of which only four (19.05%) has a size ≥2 bp. Interestingly, for the single-base INDELs, we noticed that there was a bias to the A and T bases: 45 of the 68 single base deletions and 15 of the 17 single base insertions were A and T (**Figure 8D**). To explore the rule of small INDELs occurring, we investigated the flanking sequence of 63 small INDELs in six randomly selected sequenced lines (**Table 4**). Eight (88.89%) of the nine single base insertions and 29 (82.86%) of the 35 single base deletions, along with 16 (84.21%) of the 19 INDELs that had a size ≥2 bp occurred within or near the homopolymer or polynucleotide repeats.

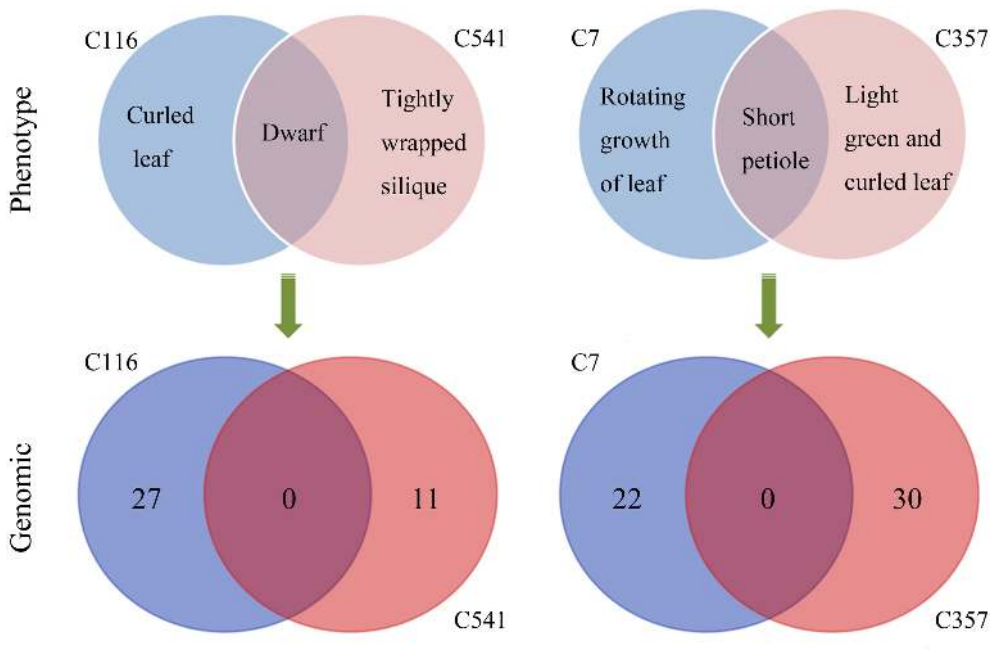

FIGURE 7 | Similarity between the mutant lines with similar phenotypes at the genomic level. The homozygous mutations in M3 were used to detect whether there were any common mutations at the genomic level between the re-sequenced lines sharing similar phenotypes.

DISCUSSION

In this study, we re-sequenced 11 mutagenesis progeny lines (M3) of *A. thaliana* lines derived from CIB irradiation. Based on the obtained data, we could reveal the mutation effects of CIB irradiation on *A. thaliana* at whole genome level, as well as its related molecular mutation spectrum and mutation rates. In contrast to previous studies, we preferred to filter the background mutations shared by multiple re-sequenced lines, rather than rely on the published reference genome. This step was an essential premise for detecting the mutations caused by CIB irradiation. To reduce false positives, we verified the detected mutations by using the IGV, a high-performance visualization tool for the interactive exploration of large, integrated genomic datasets.

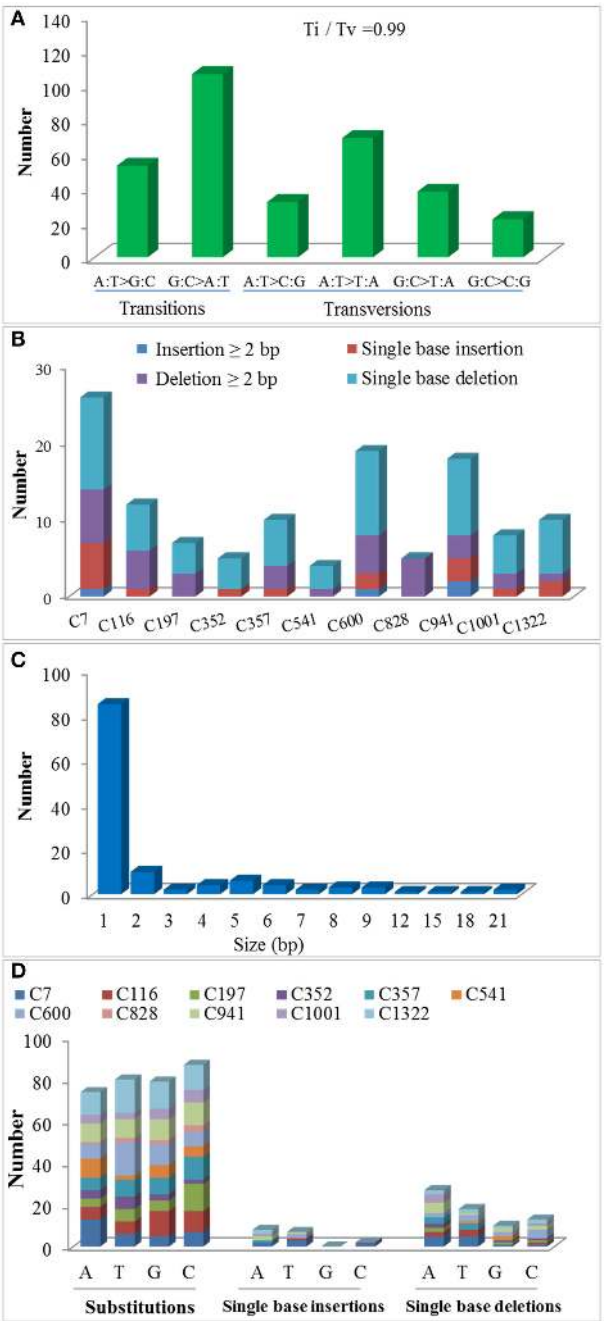

FIGURE 8 | Molecular spectrum of CIB irradiation induced substitutions and small INDELs in the genomes of 11 M3 lines. **(A)** The ratio of transition to transversion (Ti/Tv). **(B,C)** Size distributions of the small INDELs. **(D)** Base bias of CIB irradiation-induced single base mutations.

Heavy-ion beam irradiation is thought to generate mutations in the form of substitutions, small INDELs and structure variants (i.e., large fragment deletion, inversions, intra-chromosomal translocations, and inter-chromosomal translocations; Tanaka et al., 2010). In this study, we found that the CIB irradiation-induced substitutions prevailed over the INDELs, with the ratio of substitutions to INDELs of 2.58:1 (**Figure 5A**). *Brachypodium distachyon* mutant line which was induced by chronic gamma radiation, its ratio was 11.90:1 (Lee et al., 2017). As for the fast neutron irradiation, this ratio was 1.45:1 in *A. thaliana*, and 1.26:1 in rice mutant line (Belfield et al., 2012; Li et al., 2016). Among the 124 small INDELs detected here, the number of deletions was more than those of insertions, and the single base INDELs were more prevalent than those in size equal to or greater than 2 bp (**Figures 8B,C**). Although our current result showed a similar tendency to gamma and fast neutron irradiation (induced more substitution mutations), the proportion of INDELs varied with the quality of radiations. The high LET particles irradiation (i.e., CIB and fast neutron) induced more INDELs than the low LET irradiation. (Belfield et al., 2012; Li et al., 2016; Lee et al., 2017). In fact, we have tried to detect the large INDELs, by simultaneously using both the Pindel (https://trac.nbic.nl/pindel/; Ye et al., 2009) and Break Dancer Max (http://breakdancer.sourceforge.net/breakdancermax.html; Chen et al., 2009). Although hundreds of large deletions were eventually detected (Table S5), most of them were turned out to be false positives after validating them by IGV. This outcome may due to that most commonly used re-sequencing strategy in current is based on the paired-end libraries with an insertion size of 350 bp, for which the read length is 2 × 125 bp rather than the mate pair sequencing libraries whose inserts size can reach up to 2–5 kb. Actually, the long-insert paired-end libraries may be much more suitable for structural variant detection. Secondly, the LET value of the CIB irradiation used in this study was only 50 keV/μm. Previous studies showed that the deletion size and complexity of DNA damage is closely related to the LET value of irradiation (Hada and Georgakilas, 2008; Sage and Harrison, 2011; Hirano et al., 2015). For example, by using a high resolution melting (HRM) technique, CIB irradiation around 30 keV/μm mainly induced substitutions or deletions/insertions, ranging in size from 1 to 53 bp, which was on par with the mutagenesis efficiency of classical chemical mutagen EMS (Kazama et al., 2011). Hirano et al. have reported that heavy-ion beam of 290 keV/μm led to an increased proportion of large deletions (>1 kb) and chromosomal rearrangements (Hirano et al., 2012). Thirdly, the majority of severe DNA damage, including large deletions (from kb to Mb) induced by radiation, cannot be inherited to the offspring, as these non-heritable mutations may be involved in genes that are vital for gamete development or viability. In general, progeny could inherit mutations of 1- or 4-bp deletions (Naito et al., 2005). Considering all these reasons above, to ensure the accuracy and reliability of results in the present study, it was

TABLE 3 | Pyrimidine dinucleotide analyses at C > T substitutions sites in 11 CIB irradiated M3 lines.

Line	Chromosome	Loci	Mutation	Flanking sequence	Pyr-Pyr
C7	5	1496355	C > T	AATT**C**TTGA	T**C** **C**T
C7	5	1454369	C > T	CCTT**C**ACAG	T**C**
C7	5	3829172	C > T	TCAA**C**ATCA	
C7	5	16650819	C > T	GACT**C**TCCA	T**C** **C**T
C116	2	6339827	C > T	GGCT**C**TCAA	T**C** **C**T
C116	3	8857172	C > T	AGTA**C**GCGT	
C116	3	13445052	C > T	CAGT**C**GTTG	T**C**
C116	3	21948981	C > T	TTCT**C**CACT	T**C** **C**C
C116	5	8571804	C > T	CCTT**C**AGGT	T**C**
C116	5	14749784	C > T	ACAA**C**TGTC	**C**T
C197	1	27664800	C > T	TTTG**C**TTTT	**C**T
C197	2	7777170	C > T	CTCT**C**AGCT	T**C**
C197	2	3980109	C > T	CGGG**C**GATG	
C197	3	5423074	C > T	TTAA**C**AGGC	
C197	3	20612671	C > T	AATA**C**ACGT	
C197	4	4989792	C > T	ACGC**C**CAAT	C**C** **C**C
C197	5	12293490	C > T	CTCT**C**ATGT	T**C**
C197	5	5133036	C > T	ATTG**C**AACA	
C352	1	26491174	C > T	TTTT**C**AGTA	T**C**
C352	4	17139849	C > T	TCTA**C**ATAT	
C357	1	5147165	C > T	ATAT**C**AGAT	T**C**
C357	1	13896252	C > T	TGTT**C**TCAC	T**C** **C**T
C357	2	2382455	C > T	ACAC**C**AGGA	C**C**
C357	2	9045207	C > T	TGCA**C**AACA	
C357	3	959295	C > T	AGTT**C**TTGC	T**C** **C**T
C357	3	21754406	C > T	GAAG**C**ATTT	
C357	4	3608407	C > T	ACTA**C**TATC	**C**T
C357	5	11079083	C > T	TGTT**C**AACG	T**C**
C541	1	27821135	C > T	TAGG**C**AAGA	
C541	3	13945744	C > T	TGTT**C**CGCG	T**C** **C**C
C541	5	16565496	C > T	GTTT**C**TTCT	T**C** **C**T
C600	1	15506551	C > T	ACTG**C**AGAG	
C600	2	629920	C > T	CACT**C**TGAT	T**C** **C**T
C600	3	7132293	C > T	GAAA**C**GAGA	
C600	4	12005191	C > T	AACA**C**ATTT	
C600	5	7896135	C > T	CTGA**C**ATTC	
C828	5	9262872	C > T	AGCA**C**ATTT	
C941	1	1654865	C > T	AAAT**C**AGCA	T**C**
C941	1	16595744	C > T	AACT**C**TAAA	T**C** **C**T
C941	1	18884415	C > T	GAAA**C**AAGG	
C941	2	12645936	C > T	TCTT**C**GCTC	T**C**
C941	4	4481201	C > T	TTAA**C**TTTT	**C**T
C941	4	6924288	C > T	AGCA**C**CAGC	**C**C
C941	4	13051872	C > T	ATTA**C**TTCA	**C**T
C941	4	14525988	C > T	GATT**C**TACT	T**C** **C**T
C941	4	17299998	C > T	TAAA**C**AGTT	
C941	5	14679245	C > T	CCAA**C**TCCT	**C**T
C1001	1	17132595	C > T	TTAT**C**ATGT	T**C**
C1001	3	20688685	C > T	AAAA**C**CTAA	**C**C
C1001	5	10126078	C > T	ATAC**C**ACCT	C**C**
C1322	1	23839536	C > T	AATG**C**GTTT	
C1322	2	8613063	C > T	CTCT**C**AATA	T**C**
C1322	4	980039	C > T	GAAT**C**CATC	T**C** **C**C
C1322	5	1431634	C > T	TTTC**C**TTTC	C**C** **C**T
C1322	5	6887544	C > T	GGTT**C**TGTT	T**C** **C**T

The C base occurring in the substitution sites are in bold and underlined.

reasonable to focus on substitutions and small INDELs. On the other hand, the mutation rate of the 200-Gy CIB (50 keV/μm) irradiation might be underestimated in this study, because it omitted the detection for the non-transmissible mutations. Our sequencing analyses were based on nine mutagenesis progeny lines (M3) with visible and heritable mutation phenotypes (C7, C116, C197, C352, C357, C541, C600, C828, and C941), as well as two M3 lines with inconspicuous mutation phenotypes (C1001, C1322). In fact, there were still several mutations been detected in C1001, C1322 genomes. This suggests that it is too simple and imprudent to correlate the visible phenotypes to the molecular mutations. Besides, it will underestimate or neglect the potential mutants without eyeable mutation traits. Yan et al. also reported that the effects of the CIB irradiation were underestimated by counting the plants that only displayed abnormal and visible phenotypes (Yan et al., 2014). Otherwise, we previously thought that there might be some mutation hotspot in genome, for example, whether CIB irradiation was prone to induce mutations that were located on certain specific chromosomes? However, based on the variant rate statistics of mutations across each chromosome in the 11 re-sequenced lines, our answer is that so-called hotpots were not observed.

Among the detected substitutions, the ratio of transitions to transversions induced by CIB irradiation was 0.99 in this study. This value differs greatly from the 2.73 of spontaneous substitutions reported in the mutation accumulation line (Ossowski et al., 2010), but closes to the 0.86 reported in a fast neutron-induced mutation line (Belfield et al., 2012). It suggests that artificial mutagenesis can balance the ratio of transitions and transversions. Although an obvious bias of G:C > A:T transitions was observed in this study, compared to mutations induced by EMS which mostly were G:C > A:T substitutions (about 88%), the proportion of this kind of mutations is much lower when treated by CIB irradiation (Henry et al., 2014). For the single base INDELs, however, we found a different bias of A and T bases from the substitutions, and the majority of small INDELs happened at, or adjacent to, homopolymeric or polynucleotide repeats (especially those of A or T bases). Actually, such biases were also prominent in other radiation, such as UV (Daya-Grosjean and Sarasin, 2005) and fast neutron (Belfield et al., 2012). Radiation exposure could induce C > T transitions by causing the formation of covalent linkages between neighboring pyrimidine residues (for instance, CC, CT, TC, and TT) in the DNA sequence, thus resulting in a predominance of UV-induced

TABLE 4 | Flanking sequences analysis of the small INDELs in six randomly selected re-sequenced lines.

Line	Chromosome	Loci	INDEL size (bp)	Sequence
C7	1	16803045	+1	AAAAATAATT**A**AAAAACGAAA
C7	1	17871566	+1	GATGAGTCTC**T**TTGACAATGA
C7	2	2330971	+1	CTAATCTCTG**T**TTTTTTTTT
C7	3	14218519	+1	AAAAACCGAT**C**AGAAGAATTC
C7	3	10963641	+1	AACATGTGGC**A**AAAATAAATT
C7	5	15938548	+1	TAAAGTTAGA**T**TTTTTATTTT
C116	4	16566783	+1	TTTTTTTTC**T**TTTTTTTTT
C352	4	12840961	+1	TTTAAAAAAA**C**TATTCACAAT
C357	3	16396219	+1	CAAACTTTCT**A**AAAACTCAGC
C7	1	2500882	−1	AAACCTAATA**G**GAAAAGGGAC
C7	1	7060076	−1	TGCGGCCTTG**C**GGGAGCAATC
C7	1	23165210	−1	ATCAATGGCT**A**AAAAACCATC
C7	1	30364445	−1	AATGCAAGAG**A**AAGCATTTCA
C7	2	3736597	−1	GTGTATGACC**T**TTATATTTTT
C7	2	16776720	−1	ATCCATAACA**T**TTTTTTTTG
C7	3	12880551	−1	AGAATGCTCA**T**TTATCTCATT
C7	3	15247466	−1	AAATCATACG**A**CGAACACTAC
C7	4	13256395	−1	TGCTCTTGCC**A**AGGTTAGTTC
C7	5	10287067	−1	CCAAGATCCG**A**ACCTAGAAAT
C7	5	15120725	−1	AATAGATTTC**T**TTTATCGAAA
C7	5	15424664	−1	TGTTTTTTTG**T**GTGTTTTCTT
C116	1	6787418	−1	CAAGAGCGT**C**CACGACGAGTT
C116	1	899668	−1	TGGAGCTGC**T**TCATAAGTTC
C116	2	14954774	−1	AAGCGAAAC**T**TTTATTGCTA
C116	3	3201387	−1	AGCGGTTAT**A**TACATCATAA
C116	5	14227661	−1	TTTTTTTTC**T**TTTTTTTTAA
C116	5	11032451	−1	CAATTACAGG**A**ATGTCGATTT
C197	2	9064276	−1	GAGAACCAAT**C**AAGCCCTAAG
C197	2	17412340	−1	GGAAAGTAAT**A**AGAGCGTTTT
C197	3	15116009	−1	TAGATAAGTA**G**GTTTTTGCGC
C197	4	11372899	−1	AAGAAAATTG**A**ATAGAAAAAA
C352	2	13175805	−1	GTTGCTGCAA**C**CAATAGAGCA
C352	3	20903968	−1	GCGCTTGGAA**T**TTTTTTAATT
C352	3	22734373	−1	AGTGGTTAAA**G**GGTTCCAGCC
C352	4	5256307	−1	CTAAATTCAG**A**AAAAAAAACA
C357	1	24487663	−1	ATGTTACAGT**A**AAAAAAAAAA
C357	2	6789742	−1	ACTAAAGTTG**T**TGCTGCTGAT
C357	3	8430932	−1	ATTCTACTTG**T**CCTCTGAAAT
C357	5	3418139	−1	CACCTCTAAC**A**AAAGTCTTAG
C357	5	7522732	−1	TGAGAATCTC**T**TTTTCTCATT
C357	5	14755177	−1	AAACAATTTC**A**AAGTCCAACC
C541	1	2611420	−1	CCACGTATAT**G**GTAATCACAA
C541	3	5203766	−1	GAAAACTGGT**G**GGAGGTGATC
C541	5	17816447	−1	GACTTGAGAG**T**TTTAACAGAA
C7	3	4723240	−2	GTAGTCATTT**TA**TATATATAGA
C7	4	1435510	−2	AAAACCCCAA**AT**ATAATACTAC
C116	3	4315154	−2	TTTATAACTC**TG**TGCAGTGCTA
C7	2	3072553	−4	AAACGCTCGG**CGAT**GGTGATGATT
C7	3	5701019	−5	TATTAAAAAG**CCAAC**TTGGTAAAAA
C7	2	8793345	−5	AGCCCATGGA**ATGCT**AATGATTTTG
C116	3	16824813	−5	GGGTGTGAAAA**CTGGT**CAGTTAATAG
C541	1	18202184	−5	TCCTCTCAGT**TCTCA**TCTGTGGCAT
C7	4	10587368	−6	AGAATTAATC**CACTCT**TTTTCTTTTT

(Continued)

TABLE 4 | Continued

Line	Chromosome	Loci	INDEL size (bp)	Sequence
C7	1	22655118	−6	ACTCCGAAGC**TGTAGA**TGTCTGATTT
C116	1	27885672	−6	CATTTTTATA**TCGTTT**TGATGGTGAT
C197	1	10625807	−6	TCGGTTGATG**AAGACA**ATGGTAACAT
C197	4	15198376	−7	TTTATCTCCA**AGTTTTC**CATCTTTCTC
C357	3	442084	−7	CTGAACCTGTG**AACCA**GGAACCCAGTC
C357	5	20249679	−8	ACTTTAAGTT**ACTTTCTC**ACCAAAAAAA
C197	2	11198490	−9	TTTACTTATG**AATCCTGCT**AATTGAATGAT
C357	1	10405785	−9	TACAGTACAT**AATAAGTAG**ATAAGTGTAA
C116	3	11700393	−15	CTGCCTTCCCC**CATTTGCAGGACTTT**CATTGTTATA
C116	1	29852923	−18	GGCCTAGGTA**GATTAAGAGGCTTAAGCT**GCTGTTGAAT

Homopolymeric and polynucleotide stretches are underlined; the inserted or deleted base is in bold.

C > T mutations at the dipyrimidine sequences (Daya-Grosjean and Sarasin, 2005). We had speculated that CIB irradiation-induced C > T mutations likewise followed the mechanisms of preferentially anchoring to pyrimidine dinucleotide sites. By investigating the flanking sequencing of all the 55 C > T mutation sites, 67.27% did occur at the pyrimidine dinucleotide. Given the associations among the UV irradiation, fast neutron, and CIB analyses to date, we infer that radiation of different qualities share a common characteristic of pyrimidine dinucleotide-related C > T transitions. It has been reported that the single base INDELs may be caused by replication slippage at homopolymer or polynucleotide repeat regions (Viguera et al., 2001). To verify this view, the flanking sequence of small INDELs was investigated in the present study. It was found that most of the single base insertions and deletions and the 19 INDELs having sizes ≥2 bp occurred at or close to the homopolymer or polynucleotide repeats. This indicates that CIB irradiation might induce the occurrence of DNA replication slippage.

According to genetic variant annotation and the effect prediction analysis of the 444 detected mutations in the 11 re-sequenced lines, a total of 97 genes—that is to say less than nine genes on average in each genome—incurred relatively high-impact mutations (such as, missense, nonsense, frame shift, or 3′/5′-UTR), which were capable of disrupting the corresponding gene functions. Hence, only a minor proportion of genes would be affected by CIB. Similar results were also observed for Ar-ion (290 keV/μm) and Fe-ion (640 keV/μm) induced mutations (Hirano et al., 2015). This feature possessing by heavy-ions might be a promising tool for plant breeding, because the heavy-ion beam irradiation could alter some characteristics of interest without interfering with other key traits.

Although the main objective of present study is to investigate the mutation spectrum and rate of single base substitutions and small INDELs induced by CIB, it is an essential issue to excavate the mutation resource genes that are associated with plant traits at post-genomics era. Among the 11 re-sequenced lines, C7, C116, C197, C352, C357, C541, C600, C828, and C941 were mutants with stable phenotypes, therefore according to the VarScan 2, homozygous variants were sorted out and performed Gene Ontology (GO) annotation. The detail information of gene ID, family, and biological process involved in were listed in Table S6, the pseudogenes and hypothetic protein genes were filtered out. The rough map based cloning of C197, C352, and C357 were completed in our previous study (Yan et al., 2014). C197 displayed frostbite-like, and uneven leaves. Its mutation sites was located on chromosome 1 (26305380) and chromosome 4 (14740942) respectively, based on the re-sequencing results. AT4G31320, the only gene encoding an SAUR-like auxin-responsive protein family known to be involved in auxin response, was associated with the rough mapping region. C352 displayed an analogous phenotype to the *var2* mutant characterized by a variegated stem, leaves, sepals, and siliques (Takechi et al., 2000). Targeting this trait, we previously performed rough map-based cloning, and the mutated genes were closely linked to the marker T20P8 (Chr 2, 11595846). Fortunately, among the four candidate mutation sites (**Figure 4**) of C352, one site (chromosome 2, 13175805) with a deletion of a single cytosine which led to the frame shift variant of *var2* gene, was identified. The mutation site of C357 displaying short petiole and compact growth pattern was positioned on the chromosome 3 (989952). Six candidate genes (AT3G02260, AT3G03780, AT3G11540, AT3G12830, AT3G13235, and AT3G44910) were provided by re-sequencing. To ensure the accuracy of results, strict data filtering criterion can minimize the false positive variants. However, it may lead to miss detection. But to a certain extent, the association analysis of rough mapping and whole-genome re-sequencing can provide crucial clues for identifying the corresponding mutations. In fact, to identify gene or locus which are responsible for the interested trait, whole genome re-sequencing analysis, bulked-segregant analysis (BSA), as well as MutMap sequencing strategies have been successfully used for rapidly mapping mutant genes in plant species (Song et al., 2017; Win et al., 2017).

CONCLUSION

This study revealed molecular profile and rate of substitutions and small INDELs mutations induced by CIB irradiation in *A. thaliana* at the whole genome level. As heavy-ion beam mutagenesis is an effective and unique mutagen, we will continue our efforts to develop genome sequencing and mutation detection strategies that are more suitable for and more targeted

at mutation breeding induced by heavy-ion beam irradiation in the near future. We hope our data could provide valuable clues for explaining the potential mechanism of plant mutation breeding by CIB irradiation.

AUTHOR CONTRIBUTIONS

LZ coordinated and supervised the project. YD designed the experiments and wrote the manuscript. YD and SL analyzed the data. YD, XL, JY, TC, LY, HF, YC, JM, and XC performed experiments. WJL, QS, TG, and WLL corrected the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by National Key Research and Development Program (2016YFD0102106), the Strategic Priority Research Program of the Chinese Academy of Sciences (CAS) (No. XDA08040111), the National Natural Science Foundation of China (Nos. 11205218, 11275171, U1432126), the Light of West China Program (Nos. Y506020, Y606020), and the Youth Innovation Promotion Association (No. 2015337) of CAS.

ACKNOWLEDGMENTS

We are very grateful to members of the biophysics group at our institute for the *A. thaliana* mutation screening and planting. We appreciate Dr. P Li from Medical Physics Group at IMP-CAS for helping us modify the English grammar and academic writing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2017.01851/full#supplementary-material

REFERENCES

Alizadeh, E., Sanz, A. G., García, G., and Sanche, L. (2013). Radiation damage to DNA: the indirect effect of low-energy electrons. *J. Phys. Chem. Lett.* 4, 820–825. doi: 10.1021/jz4000998

Belfield, E. J., Gan, X., Mithani, A., Brown, C., Jiang, C., Franklin, K., et al. (2012). Genome-wide analysis of mutations in mutant lineages selected following fast-neutron irradiation mutagenesis of *Arabidopsis thaliana*. *Genome Res.* 22, 1306–1315. doi: 10.1101/gr.131474.111

Bolon, Y. T., Stec, A. O., Michno, J. M., Roessler, J., Bhaskar, P. B., Ries, L., et al. (2014). Genome resilience and prevalence of segmental duplications following fast neutron irradiation of soybean. *Genetics* 198, 967–981. doi: 10.1534/genetics.114.170340

Bruggemann, E., Handwerger, K., Essex, C., and Storz, G. (1996). Analysis of fast neutron-generated mutants at the *Arabidopsis thaliana* HY4 locus. *Plant J.* 10, 755–760. doi: 10.1046/j.1365-313X.1996.10040755.x

Chen, K., Wallis, J. W., McLellan, M. D., Larson, D. E., Kalicki, J. M., Pohl, C. S., et al. (2009). Breakdancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* 6, 677–681. doi: 10.1038/nmeth.1363

Daya-Grosjean, L., and Sarasin, A. (2005). The role of UV induced lesions in skin carcinogenesis: an overview of oncogene and tumor suppressor gene modifications in xeroderma pigmentosum skin tumors. *Mutat. Res.* 571, 43–56. doi: 10.1016/j.mrfmmm.2004.11.013

Du, Y., Li, W., Yu, L., Chen, G., Liu, Q., and Luo, S. et al. (2014). Mutagenic effects of carbon-ion irradiation on dry *Arabidopsis thaliana* seeds. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 759, 28–36. doi: 10.1016/j.mrgentox.2013.07.018

Hada, M., and Georgakilas, A. G. (2008). Formation of clustered DNA damage after high-LET irradiation: a review. *J. Radiat. Res.* 49, 203–210. doi: 10.1269/jrr.07123

Hase, Y., Okamura, M., Takeshita, D., Narumi, I., and Tanaka, A. (2010). Efficient induction of flower-color mutants by ion beam irradiation in petunia seedlings treated with high sucrose concentration. *Plant Biotechnol.* 27, 99–103. doi: 10.5511/plantbiotechnology.27.99

Hase, Y., Yoshihara, R., Nozawa, S., and Narumi, I. (2012). Mutagenic effects of carbon ions near the range end in plants. *Mutat. Res.* 731, 41–47. doi: 10.1016/j.mrfmmm.2011.10.004

He, J. Y., Lu, D., Yu, L. X., and Li, W. J. (2011). Pigment analysis of a color-leaf mutant in wandering jew (*Tradescantia fluminensis*) irradiated by carbon ions. *Nuclei Sci. Tech.* 22, 77–83. doi: 10.13538/j.1001-8042/nst.22.77-83

Henry, I. M., Nagalakshmi, U., Lieberman, M. C., Ngo, K. J., Krasileva, K. V., Vasquez-Gross, H., et al. (2014). Efficient genome-wide detection and cataloging of EMS-Induced mutations using exome capture and next-generation sequencing. *Plant Cell* 26, 1382–1397. doi: 10.1105/tpc.113.121590

Hirano, T., Kazama, Y., Ishii, K., Ohbu, S., Shirakawa, Y., and Abe, T. (2015). Comprehensive identification of mutations induced by heavy-ion beam irradiation in *Arabidopsis thaliana*. *Plant J.* 82, 93–104. doi: 10.1111/tpj.12793

Hirano, T., Kazama, Y., Ohbu, S., Shirakawa, Y., Liu, Y., Kambara, T., et al. (2012). Molecular nature of mutations induced by high-LET irradiation with argon and carbon ions in *Arabidopsis thaliana*. *Mutat. Res.* 735, 19–31. doi: 10.1016/j.mrfmmm.2012.04.010

Ishikawa, S., Ishimaru, Y., Igura, M., Kuramata, M., Abe, T., Senoura, T., et al. (2012). Ion-beam irradiation, gene identification, and marker-assisted breeding in the development of low-cadmium rice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19166–19171. doi: 10.1073/pnas.1211132109

Kazama, Y., Hirano, T., Saito, H., Liu, Y., Ohbu, S., Hayashi, Y., et al. (2011). Characterization of highly efficient heavy-ion mutagenesis in *Arabidopsis thaliana*. *BMC Plant Biol.* 11:161. doi: 10.1186/1471-2229-11-161

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., et al. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25, 2283–2285. doi: 10.1093/bioinformatics/btp373

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. doi: 10.1101/gr.129684.111

Kovacs, E., and Keresztes, A. (2002). Effect of gamma and UV-B/C radiation on plant cells. *Micron* 33, 199–210. doi: 10.1016/S0968-4328(01)00012-9

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Lee, M. B., Kim, D. Y., and Seo, Y. W. (2017). Identification of candidate genes for the seed coat colour change in a *Brachypodium distachyon* mutant induced by gamma radiation using whole-genome re-sequencing. *Genome* 60, 581–587. doi: 10.1139/gen-2016-0145

Lehrbach, N. J., Ji, F., and Sadreyev, R. (2017). Next-generation sequencing for identification of EMS-induced mutations in *Caenorhabditis elegans*. *Curr. Protoc. Mol. Biol.* 117, 7.29.21–7.29.12. doi: 10.1002/cpmb.27

Li, G., Chern, M., Jain, R., Martin, J. A., Schackwitz, W. S., Jiang, L., et al. (2016). Genome-wide sequencing of 41 rice (*Oryza sativa* L.) mutated lines reveals diverse mutations induced by fast-neutron irradiation. *Mol. Plant* 9, 1078–1081. doi: 10.1016/j.molp.2016.03.009

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Luo, S., Zhou, L., Li, W., Du, Y., Yu, L., Feng, H., et al. (2016). Mutagenic effects of carbon ion beam irradiations on dry *Lotus japonicus* seeds. *Nucleic. Inst. Methods Phys. Res. B Beam Interact. Materials Atoms* 383, 123–128. doi: 10.1016/j.nimb.2016.06.021

Matsumura, A., Nomizu, T., Furutani, N., Hayashi, K., Minamiyama, Y., and Hase, Y. (2010). Ray florets color and shape mutants induced by 12C5? ion beam irradiation in chrysanthemum. *Sci. Hortic.* 123, 558–561. doi: 10.1016/j.scienta.2009.11.004

Morita, R., Nakagawa, M., Takehisa, H., Hayashi, Y., Ichida, H., Usuda, S., et al. (2017). Heavy-ion beam mutagenesis identified an essential gene for chloroplast development under cold stress conditions during both early growth and tillering stages in rice. *Biosci. Biotechnol Biochem.* 81, 271–282. doi: 10.1080/09168451.2016.1249452

Nagata, K., Hashimoto, C., Watanabe-Asaka, T., Itoh, K., Yasuda, T., Ohta, K., et al. (2016). *In vivo* 3D analysis of systemic effects after local heavy-ion beam irradiation in an animal model. *Sci. Rep.* 6:28691 doi: 10.1038/srep 28691

Naito, K., Kusaba, M., Shikazono, N., Takano, T., Tanaka, A., Tanisaka, T., et al. (2005). Transmissible and nontransmissible mutations induced by irradiating *Arabidopsis thaliana* pollen with gamma-rays and carbon ions. *Genetics* 169, 881–889. doi: 10.1534/genetics.104. 033654

Nakano, M., Amano, J., Watanabe, Y., Nomizu, T., Suzuki, M., Mizunashi, K., et al. (2010). Morphological variation in *Tricyrtis hirta* plants regenerated from heavy ion beam-irradiated embryogenic calluses. *Plant Biotechnol.* 27, 155–160. doi: 10.5511/plantbiotechnology.27.155

Oka-Kira, E., Tateno, K., Miura, K., Haga, T., Hayashi, M., Harada, K., et al. (2005). klavier (klv), a novel hypernodulation mutant of *Lotus japonicus* affected in vascular tissue organization and floral induction. *Plant J.* 44, 505–515. doi: 10.1111/j.1365-313X.2005.02543.x

O'Rourke, J. A., Iniguez, L. P., Bucciarelli, B., Roessler, J., Schmutz, J., McClean, P. E., et al. (2013). A re-sequencing based assessment of genomic heterogeneity and fast neutron-induced deletions in a common be an cultivar. *Front. Plant Sci.* 4:210. doi: 10.3389/fpls.2013.00210

Ossowski, S., Schneeberger, K., Lucas-Lledo, J. I., Warthmann, N., Clark, R. M., Shaw, R. G., et al. (2010). The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* 327, 92–94. doi: 10.1126/science.1180677

Phanchaisri, B., Samsang, N., Yu, L. D., Singkarat, S., and Anuntalabhochai, S. (2012). Expression of OsSPY and 14-3-3 genes involved in plant height variations of ion-beam-induced KDML 105 rice mutants. *Mutat. Res.* 734, 56–61. doi: 10.1016/j.mrfmmm.2012.03.002

Ravanat, J. L., Douki, T., and Cadet, J. (2001). Direct and indirect effects of UV radiation on DNA and its components. *J. Photochem. Photobiol. B* 63, 88–102. doi: 10.1016/S1011-1344(01)00206-8

Sage, E., and Harrison, L. (2011). Clustered DNA lesion repair in eukaryotes: relevance to mutagenesis and cell survival. *Mutat. Res.* 711, 123–133. doi: 10.1016/j.mrfmmm.2010.12.010

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A. H., Nielsen, K. L., et al. (2009). SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat. Methods* 6, 550–551. doi: 10.1038/nmeth0809-550

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18. doi: 10.1038/nmeth1156

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145. doi: 10.1038/nbt1486

Shikazono, N., Suzuki, C., Kitamura, S., Watanabe, H., Tano, S., and Tanaka, A. (2005). Analysis of mutations induced by carbon ions in *Arabidopsis thaliana*. *J. Exp. Bot.* 56, 587–596. doi: 10.1093/jxb/eri047

Song, J., Li, Z., Liu, Z., Guo, Y., and Qiu, L.-J. (2017). Next-Generation sequencing from bulked-segregant analysis accelerates the simultaneous identification of two qualitative genes in soybean. *Front. Plant Sci.* 8:919. doi: 10.3389/fpls.2017.00919

Tabata, R., Kamiya, T., Shigenobu, S., Yamaguchi, K., Yamada, M., Hasebe, M., et al. (2013). Identification of an EMS-induced causal mutation in a gene required for boron-mediated root development by low-coverage genome re-sequencing in Arabidopsis. *Plant Signal. Behav.* 8:e22534. doi: 10.4161/psb.22534

Takechi, K., Sodmergen, Murata, M., Motoyoshi, F., and Sakamoto, W. (2000). The YELLOW VARIEGATED (VAR2) locus encodes a homologue of FtsH, an ATP-dependent protease in Arabidopsis. *Plant Cell Physiol.* 41, 1334–1346. doi: 10.1093/pcp/pcd067

Tanaka, A., Sakamoto, A., Ishigaki, Y., Nikaido, O., Sun, G., Hase, Y., et al. (2002). An ultraviolet-B-resistant mutant with enhanced DNA repair in Arabidopsis. *Plant Physiol.* 129, 64–71. doi: 10.1104/pp.010894

Tanaka, A., Shikazono, N., and Hase, Y. (2010). Studies on biological effects of ion beams on lethality, molecular nature of mutation, mutation rate, and spectrum of mutation phenotype for mutation breeding in higher plants. *J. Radiat. Res.* 51, 223–233. doi: 10.1269/jrr.09143

Tanaka, A., Tano, S., Chantes, T., Yokota, Y., Shikazono, N., and Watanabe, H. (1997). A new Arabidopsis mutant induced by ion beams affects flavonoid synthesis with spotted pigmentation in testa. *Genes Genet. Syst.* 72, 141–148. doi: 10.1266/ggs.72.141

Tao, Y., Mace, E. S., Tai, S., Cruickshank, A., Campbell, B. C., Zhao, X., et al. (2017). Whole-genome analysis of candidate genes sssociated with seed size and weight in sorghum bicolor reveals signatures of artificial selection and insights into parallel domestication in cereal crops. *Front. Plant Sci.* 8:1237. doi: 10.3389/fpls.2017.01237

Tiwari, A., Bahr, A., Bahr, L., Fleischhauer, J., Zinkernagel, M. S., Winkler, N., et al. (2016). Next generation sequencing based identification of disease-associated mutations in Swiss patients with retinal dystrophies. *Sci. Rep.* 6:28755 doi: 10.1038/srep28755

Tokuyama, Y., Furusawa, Y., Ide, H., Yasui, A., and Terato, H. (2015). Role of isolated and clustered DNA damage and the post-irradiating repair process in the effects of heavy ion beam irradiation. *J. Radiat. Res.* 56, 446–455. doi: 10.1093/jrr/rru122

Uchida, N., Sakamoto, T., Kurata, T., and Tasaka, M. (2011). Identification of EMS-induced causal mutations in a non-reference *Arabidopsis thaliana* accession by whole genome sequencing. *Plant Cell Physiol.* 52, 716–722. doi: 10.1093/pcp/pcr029

Viguera, E., Canceill, D., and Ehrlich, S. D. (2001). Replication slippage involves DNA polymerase pausing and dissociation. *Embo J.* 20, 2587–2595. doi: 10.1093/emboj/20.10.2587

Win, K. T., Vegas, J., Zhang, C., Song, K., and Lee, S. (2017). QTL mapping for downy mildew resistance in cucumber via bulked segregant analysis using next-generation sequencing and conventional methods. *Theor. Appl. Genet.* 130, 199–211. doi: 10.1007/s00122-016-2806-z

Xia J. H., Bai, Z., Meng, Z., Zhang, Y., Wang, L., Liu, F., et al. (2015). Signatures of selection in tilapia revealed by whole genome resequencing. *Sci. Rep.* 5:14168. doi: 10.1038/srep14168

Ye, K., Schulz, M. H., Long, Q., Apweiler, R., and Ning, Z. (2009). Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25, 2865–2871. doi: 10.1093/bioinformatics/btp394

Yu, L.-X., Li, W.-J., Du, Y., Chen, G., Luo, S.-W., Liu, R.-Y., et al. (2016). Flower color mutants induced by carbon ion beam irradiation of geranium (*Pelargonium x hortorum*, Bailey). *Nucleic. Sci. Tech.* 27, 112–119. doi: 10.1007/s41365-016-0117-3

Zhou, X., Yang, Z., Jiang, T. T., Wang, S. Y., Liang, J. P., Lu, X. H., et al. (2016). The acquisition of *Clostridium tyrobutyricum* mutants with improved bioproduction under acidic conditions after two rounds of heavy-ion beam irradiation. *Sci. Rep.* 6:29968. doi: 10.1038/srep29968

Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–441. doi: 10.1038/nbt.3096

Conflict of Interest Statement: The 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.