

## Research Article

# Whole-Genome Profiling of a Novel Mutagenesis Technique Using Proofreading-Deficient DNA Polymerase $\delta$

Yuh Shiwa,<sup>1</sup> Sanae Fukushima-Tanaka,<sup>2</sup> Ken Kasahara,<sup>2</sup>  
Takayuki Horiuchi,<sup>2</sup> and Hirofumi Yoshikawa<sup>1,3</sup>

<sup>1</sup>Genome Research Center, NODAI Research Institute, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan

<sup>2</sup>Neo-Morgan Laboratory Inc., 907 Nogawa, Miyamae-ku, Kawasaki, Kanagawa 216-0001, Japan

<sup>3</sup>Department of Bioscience, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan

Correspondence should be addressed to Hirofumi Yoshikawa, hiyoshik@nodai.ac.jp

Received 9 February 2012; Accepted 20 March 2012

Academic Editor: Hiromi Nishida

Copyright © 2012 Yuh Shiwa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel mutagenesis technique using error-prone DNA polymerase  $\delta$  (*pol* $\delta$ ), the disparity mutagenesis model of evolution, has been successfully employed to generate novel microorganism strains with desired traits. However, little else is known about the spectra of mutagenic effects caused by disparity mutagenesis. We evaluated and compared the performance of the *pol* $\delta$ *MKII* mutator, which expresses the proofreading-deficient and low-fidelity *pol* $\delta$ , in *Saccharomyces cerevisiae* haploid strain with that of the commonly used chemical mutagen ethyl methanesulfonate (EMS). This mutator strain possesses exogenous mutant *pol* $\delta$  supplied from a plasmid, thereby leaving the genomic one intact. We measured the mutation rate achieved by each mutagen and performed high-throughput next generation sequencing to analyze the genome-wide mutation spectra produced by the 2 mutagenesis methods. The mutation frequency of the mutator was approximately 7 times higher than that of EMS. Our analysis confirmed the strong G/C to A/T transition bias of EMS, whereas we found that the mutator mainly produces transversions, giving rise to more diverse amino acid substitution patterns. Our present study demonstrated that the *pol* $\delta$ *MKII* mutator is a useful and efficient method for rapid strain improvement based on *in vivo* mutagenesis.

## 1. Introduction

Random mutagenesis is a powerful tool for generating enzymes, proteins, metabolic pathways, or even entire genomes with desired or improved properties [1]. Due to the technical simplicity and applicability to almost any organism, chemical or radiation mutagenesis is frequently used for the generation of genetic variability in a microorganism. However, these methods tend to be inefficient because they can cause substantial cell damage when performed *in vivo* [2].

A novel mutagenesis technique using error-prone DNA polymerase  $\delta$  (*pol* $\delta$ ), based on the disparity mutagenesis model of evolution [3] has been successfully employed to generate novel microorganism strains with desired traits [4–11]. In the disparity model, mutations occur preferentially

on the lagging strand, due to the more complex, discontinuous DNA replication that takes place there. Computer simulation shows that the disparity model accumulates more mutations than the parity model, in which mutations occur stochastically and evenly in both strands [3]. In addition, the disparity model produces greater diversity because some offspring will have mutant DNA while some offspring will have nonmutated, wild-type DNA.

Several studies have shown that the disparity mutagenesis method often achieved more satisfactory results (i.e., higher mutation rate and quick attainment of the desired phenotype) than conventional methods such as the chemical mutagen, ethyl methanesulfonate (EMS) [5, 10], which is known to produce mainly G/C to A/T transitions [12]. However, little else is known about the spectra of mutagenic effects caused by disparity mutagenesis.

*polδ* is involved in the synthesis of the lagging strand of DNA [13]. Several mutants, including the proofreading-deficient *pol3-01* strain and several low-fidelity mutants, have been shown to elevate the mutation rate [14–18]. To generate the strains with the greatest mutagenicity, Neo-Morgan Laboratory (Kanagawa, Japan) has developed the plasmid YCplac33/*polδMKII*, expressing the *polδ* mutant allele with 2 mutations: one mutation to inactivate the proofreading activity (D321A and E323A) [15] and another mutation to decrease the fidelity of replication (L612M) [14, 17, 18].

With the recent advent of next-generation sequencing technologies, an accurate characterization of the mutant genome, relative to the parental reference strain, is now achievable. In fact, Flibotte et al. have analyzed the mutation spectra induced by various mutagens, such as EMS, ENU, and UV/TMP, in the whole genome of *Caenorhabditis elegans* [12]. Another group has also used these sequencing technologies to analyze the genetic variations between a parental and EMS-mutagenized strain of yeast [19].

In this study, we evaluate the performance of the *polδMKII* mutator, which expresses the proofreading-deficient and low-fidelity *polδ* in *S. cerevisiae* haploid strain, compared with the commonly used chemical mutagen EMS. This mutator strain possesses exogenous mutant *polδ* supplied from a plasmid, thereby leaving the genomic one intact. We measured the mutation rate of this mutator strain and found that the mutation frequency of *polδMKII* was approximately 7 times higher than that of EMS. We also performed high-throughput next generation sequencing with Illumina GAI to analyze the genome-wide mutation spectra produced by the 2 different mutagenesis methods and found that the mutator strain exhibited more pleiotropy and gave rise to more diverse amino acid substitution patterns. Our present study has demonstrated that a proofreading-deficient and low-fidelity *polδMKII* mutator is a useful and efficient method for rapid strain improvement based on *in vivo* mutagenesis. This mutator is also useful for studying the acceleration of evolution.

## 2. Materials and Methods

**2.1. Plasmid.** Plasmid YCplac33/*polδMKII* was constructed as follows: a 4.8 kb DNA fragment containing the *S. cerevisiae* BY2961 *pol3* gene, plus the UTR 1 kb upstream and 0.5 kb downstream, (*Matax ura3-52*, *his3-Δ300*, *trp1-Δ901*, *leu2-3*, *112 lys2-801*, *ade2-2*) was inserted into the *SalI-EcoRI* site of YCplac33, and 3 amino acid substitutions, D321A, E323A, and L612M, were introduced into the *pol3* gene using site-directed mutagenesis [20]. YCplac33 is low-copy number plasmid and is stably maintained in *S. cerevisiae* [20].

**2.2. Mutator Mutagenesis.** YCplac33/*polδMKII* vector (and YCplac33 empty vector as nonmutator control) was introduced into *S. cerevisiae* BY2961 strain cells using the LiCl method, and the transformants (mutator strains) were selected on synthetic complete (SC)-agar plates without uracil. Five mutator strains were picked and independently cultivated in 1 mL SC medium at 30°C for 24 h (about

30 generations) in order to introduce mutations into their chromosomes. To determine the mutation frequencies of the 5 mutator strains, aliquots were spread on SC-agar plates containing L-canavanine sulfate salt (0.06 mg/mL) (Sigma, St. Louis, MO, USA) to identify *CAN1* mutants, and incubated until resistant colonies were formed. The mutation frequencies were calculated as the number of drug-resistant colonies divided by the number of colonies on SC-agar plate without drug. Forward mutation rates at *CAN1* were determined by fluctuation analysis using these 5 independent cultures [21]. In order to fix mutations, another aliquot of the mutator culture was spread on SC-agar plates containing 5-fluoroorotic acid monohydrate (Wako) to obtain demutatorized cells curing from YCplac33/*polδMKII* vector. The genomic DNA was prepared from the demutatorized cells using the procedure described in the following section.

**2.3. EMS Mutagenesis.** *S. cerevisiae* BY2961 strain cells were suspended in 0.1 M phosphate-buffered saline (PBS) (pH 7.0) containing 1.5, 2.0, 2.5, or 3.0% ethyl methanesulfonate (EMS) and were incubated at 30°C for 1 h to introduce chromosomal mutations. The cells were washed 3 times with 5% sodium thiosulfate, suspended in sterilized water, and spread on SC-agar plates containing L-canavanine sulfate salt (0.06 mg/mL) (Sigma) to identify *CAN1* mutants. The mutation frequencies were calculated as described above. Another aliquot of the EMS-treated cell suspension was spread on a YPD-agar plate to isolate single clones. The genomic DNA was prepared from 5 single clones derived from the cells treated with 1.5% EMS using the procedure described in the following section.

**2.4. Library Preparation for Illumina Sequencing.** The genomic DNA from *S. cerevisiae* was extracted using the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA, USA). Each sequenced sample was prepared according to the Illumina protocols. Briefly, 3 μg of genomic DNA was fragmented to an average length of 200 bp by using the Covaris S2 system (Covaris, Woburn, MA, USA). The fragmented DNA was repaired, a single “A” nucleotide was ligated to the 3′ end, Illumina Index PE adapters (Illumina, San Diego, CA, USA) were ligated to the fragments, and the sample was size selected for a 300 bp product using E-Gel SizeSelect 2% (Invitrogen, Grand Island, NY, USA). The size-selected product was amplified by 18 cycles of PCR with the primers InPE1.0, InPE2.0, and the Index primer containing 6-nt barcodes (Illumina). The final product was validated using the Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA).

**2.5. Sequencing and Data Analysis.** The 11 barcoded libraries (the parental strain BY2961, 5 colonies from the mutator strain, and 5 colonies from the EMS-treated strain) were used for cluster generation in several multiplexed flow cell lanes in the Illumina Genome Analyzer II system. Ninety-one cycles of multiplexed paired-end sequencing was performed, running phi X 174 genomic DNA as a control in a separate lane of the flow cell. After the sequencing reactions were

complete, Illumina analysis pipeline (CASAVA 1.6.0) was used to carry out image analysis, base calling, and quality score calibration. Reads were sorted by barcode and exported in the FASTQ format. The quality of each sequencing library was assessed by evaluating the quality score chart and the nucleotide distribution plot using FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)).

Once the raw sequence data were curated, the reads of each sample were aligned to the S288c reference genome (<http://www.yeastgenome.org/>) using the BWA software (Ver. 0.5.1) with default parameters [22]. To avoid false positives and mutations from repetitive regions, we removed repetitive reads from the alignment files. We then used the SAMtools software (Ver. 0.1.9) [23] to produce the lists of mutations. To identify mutations that were produced by mutagenesis, we applied the following filtering criteria to the lists of mutations:

- (a) the coverage at the mismatch positions should be at least 10;
- (b) the variant is not present in the sequenced parental strain;
- (c) indels meet a SNP quality threshold of 50 and substitutions meet a SNP quality threshold of 20 (SAMtools assigns SNP quality, which is the Phred-scaled probability that the consensus is identical to the reference);
- (d) samples meet a mapping quality of 30 (SAMtools assigns Mapping quality, which is the Phred-scaled probability that the read alignment is wrong);
- (e) the percentage of reads showing the variant allele exceeds 90%.

A variant must pass this filter to be considered a mutation. Alignments of all mutations were inspected by Integrative Genomics Viewer (IGV) [24]. The lists of mutations were then annotated using COVA (comparison of variants and functional annotation) (<http://sourceforge.net/projects/cova>). COVA was specifically designed to annotate the large number of identified mutants using the Genbank annotation files.

### 3. Results

**3.1. Determination of Mutation Frequencies.** In this study, we evaluated the performance of the *polδMKII* mutator, compared with that of the commonly used chemical mutagen, EMS. To assess EMS efficiency, *S. cerevisiae* BY2961 cells were treated with different concentrations of EMS. The lethality and mutation frequencies of the canavanine resistant colonies are shown in Table 1. At an EMS concentration of 1.5%, the mutation frequency was approximately 18-fold higher than that in the control (untreated) strain. Above 2.0% EMS, the survival rate decreased with no increase in mutation frequency. Based on this result, we decided to use cells treated with 1.5% EMS for whole-genome sequencing.

To assess the effectiveness of the mutator, we transformed the haploid BY2961 strain with a yeast expression plasmid,

TABLE 1: Relationship between mutation frequency and survival after EMS treatment.

EMS concentration (%)	Mutation frequency of canavanine resistant ( $\times 10^{-7}$ )	Fold elevation*	Survival (%)
0.0	2	1	100
1.5	35	18	51
2.0	36	19	30
2.5	33	17	21
3.0	37	19	12

\* Fold elevation is relative to untreated cells.

TABLE 2: Frequency of drug-resistant mutants in the mutator strains.

Plasmids	Mutation frequency of canavanine resistant ( $\times 10^{-7}$ )	Fold elevation*
YCplac33	3.70	1
YCplac33/ <i>polδMKII</i>	486.7 $\pm$ 145.0 <sup>#</sup>	132

\* Fold elevation is relative to empty vector.

<sup>#</sup> Mean  $\pm$  standard deviation of 3 SC plates.

YCplac33/*polδMKII*, expressing the *polδ* mutant allele containing both the mutation to inactivate the proofreading activity (D321A and E323A) and the mutation to decrease the fidelity of replication (L612M). The mutator strain harboring the YCplac33/*polδMKII* plasmid will be referred to from here on as “mutator.” We determined the mutation frequency by resistance to canavanine. As summarized in Table 2, the mutation frequency of the mutator was approximately 132-fold higher than in the cells containing the empty vector. The forward mutation rate at the *CAN1* (arginine permease) locus was calculated to be  $7.9 \times 10^{-6}$ /cell division. These results show that the plasmid-generated mutated *polδ* protein effectively competes with the endogenous wild-type *polδ* protein that is produced from the chromosome, and the semidominant negative expression of mutated *polδ* was effective in introducing mutations. These results also demonstrate that the mutation frequency of the mutator was approximately 7 times higher than that of EMS.

**3.2. Whole-Genome Sequencing.** To analyze the genome-wide mutation spectra of the 2 different mutagenesis methods, we implemented a parallel sequencing approach with the Illumina Solexa technology (GAI instrument). We sequenced the parental haploid strain BY2961, each of the 5 clones from the mutator strains, and each of the 5 clones from the EMS-treated strains under nonselective conditions. Sequencing reads were aligned to the S288c reference genome using the BWA software [22]. To avoid false positives due to mutations from repetitive regions, reads mapped to multiple locations were discarded, and only uniquely mapped reads were used for subsequent analysis.

TABLE 3: Sequencing and mapping statistics.

Sample name	Number of mapped unique reads	% mapped reads	% genome covered* by unique reads	Average coverage by unique reads
BY2961	11,155,487	96.13	94.97	87.9×
EMS1	5,406,681	96.94	94.81	42.2×
EMS2	6,240,554	97.26	94.85	48.7×
EMS3	5,275,583	98.12	94.81	41.2×
EMS4	4,502,271	97.17	94.80	35.2×
EMS5	4,113,345	96.27	94.83	32.1×
Mutator 1	9,612,541	93.93	94.90	75.8×
Mutator 2	5,111,531	92.39	94.79	39.9×
Mutator 3	5,649,822	96.11	94.95	44.1×
Mutator 4	4,226,405	98.79	94.85	33.0×
Mutator 5	9,855,938	97.36	95.10	77.6×

\* Coverage is defined as the percentage of bases in the genome that have at least 1 uniquely mapped read at that position.

In the current study, the average genomic coverage ranged from 32× to 87× (Table 3). On average, 94.18% of the S288c reference genome was covered with at least 1 uniquely mapped read at each base. Subsequently, we analyzed the data for 2 kinds of mutational events: single nucleotide variants (SNVs) and small insertions and deletions (Indels). Illumina sequencing found 6,766 genetic differences between our parental strain BY2961 and the S288c. Mutations induced by these mutagens were identified by subtracting the parental mutations. Sequence-processing details can be found in Section 2.

**3.3. The Mutation Spectra of Mutator and EMS.** We compared the average number of mutations between mutator strains and EMS-mutagenized strains (Figure 1). Mutator produced fewer SNVs than EMS (7.2 versus 55.8 per strain, resp.,  $P < 0.05$ ). Mutator and EMS produced few deletions (1.6 versus 2.8 per strain, resp.), as well as few insertions (0.2 versus 0.6 per strain, resp.). An average of  $1.14 \times 10^7$  nucleotide sites fulfilled our criteria of read depth ( $\geq 10$ ), with an average base-substitutional mutation rate estimate of EMS:  $4.87$  (SE =  $1.34$ )  $\times 10^{-6}$  per site, Mutator:  $2.09$  ( $0.55$ )  $\times 10^{-8}$  per site per cell division (about 30 generations). The rate we calculated for the mutator is 100-fold higher than the previously reported spontaneous mutation rate,  $3.3$  ( $0.8$ )  $\times 10^{-10}$ , based on 454 analyses of 4 mutation-accumulation (MA)-lines [26]. The 2 mutagens generate mutations that are distributed similarly across the various gene features although the mutator did produce more SNVs within exons than did EMS (Figure 2).

The mutation spectra are shown in Figure 3(a). In the genome-wide profile, we found that the mutator primarily induced transversions (72%) while EMS primarily induced transitions (97%), well in accord with the known mutagenic specificity of EMS [12]. Similarly, the mutator primarily

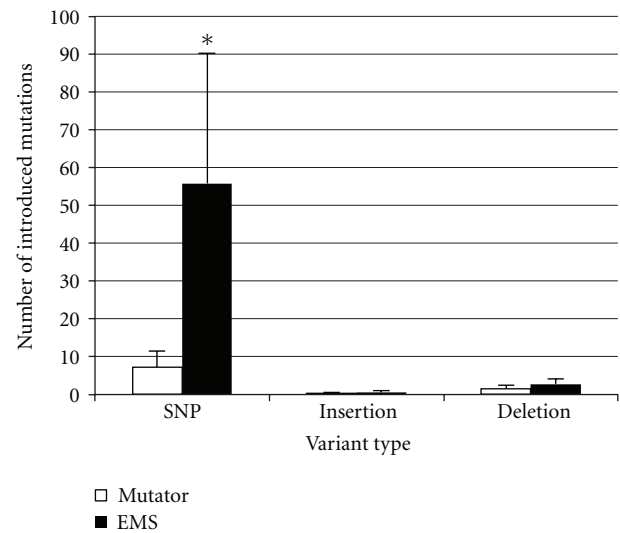


FIGURE 1: Average number of introduced mutations. By subtracting parental mutations from each mutagenized strain, we determined the number of mutations that were introduced by each mutagen. Bars represent mean  $\pm$  standard error for 5 clones. \*  $P < 0.05$  versus mutator in a two-sample  $t$ -test.

induced transversions (69%) in the nonsynonymous substitutions in exons (Figure 3(b)), similar to what has been seen in *pol3-01* study using *URA3* reporter gene [16]. EMS treatment was also in agreement with the genome-wide spectra, induced transitions with a prevalence of 98%.

**3.4. Amino Acid Substitution Patterns.** The mutation spectra of a given mutagenesis method influences the repertoire of changed amino acids at the protein level, and we were able to evaluate the amino acid substitution patterns generated by our 2 protocols (Table 4). Initially, we classified mutations into those that preserved the corresponding amino acid,

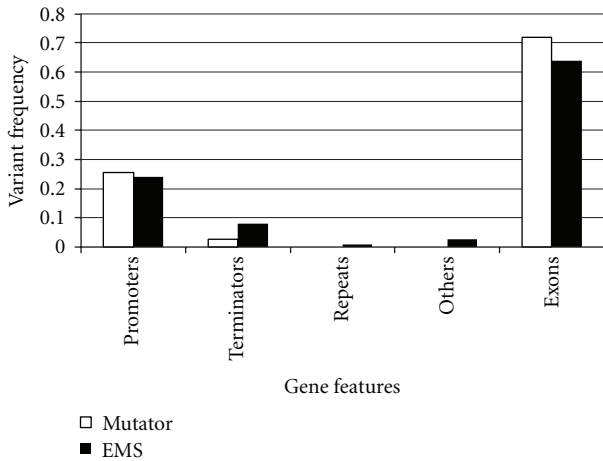


FIGURE 2: Relative frequency of SNVs affecting various gene features. The mutator and EMS generated mutations that were distributed similarly across the various gene features. The data for individual strains were combined according to the mutagen used. Promoters indicate the region 1 kb upstream of each gene. Terminators indicate the region 200 bp downstream of each gene.

changed the amino acid, or generated a stop codon. A clear difference was seen between mutator and EMS. Of the total mutations, the mutator changed the amino acid in approximately 85%, whereas EMS changed the amino acid in approximately 61%. The mutator also generated more stop codons than EMS (7% versus 2%, resp.). While mutator generated more changes to the first or second nucleotide of the codon, EMS generated changes in all 3 positions in approximately equal proportions.

Amino acid changes were classified into conservative and nonconservative substitutions, where a conservative substitution changed the encoded amino acid to a similar amino acid according to the criteria of the BLOSUM62 matrix [25]. Of the amino acid changes, mutator produced more nonconservative substitutions than EMS (83% and 53%). For the comparison of random mutagenesis methods, Wong et al. [27] proposed a useful structure indicator that takes into account Gly and Pro substitutions as well as stop codons. In our study, the mutator produced an equivalent number of Gly/Pro and stop codon substitutions, whereas EMS generated only stop codon substitutions.

#### 4. Discussion

In this study, we evaluated the performance of a novel mutagenesis technique using error-prone proofreading-deficient and low-fidelity DNA polymerase  $\delta$  by determining the mutation rate of the strain harboring the enzyme. We also analyzed the spectra of mutations across the entire *S. cerevisiae* genome and then assessed the diversity of mutation types at the amino acid level.

Proofreading-deficient *pol* $\delta$  mutants, such as *pol3-01* strain, and several low-fidelity *pol* $\delta$  mutants, such as L612M, have been shown to present a mutator phenotype and to elevate the mutation rate [14–18]. We generated a BY2961

strain expressing a *pol* $\delta$  MKII mutator, *pol* $\delta$  mutant allele containing a combination of mutations to inactivate the proofreading activity (D321A and E323A) and to decrease the fidelity of replication (L612M). This mutant allele acts as a strong mutator, as evidenced by the high frequency of spontaneous mutations (131-fold over control, compared to 18-fold for EMS strains). Vencatesan et al. reported the forward *CAN1* mutation rates of *pol* $\delta$  mutants as  $1.5 \times 10^{-6}$  in L612M, and  $5.6 \times 10^{-6}$  in *pol3-01* [18]. These mutant strains were constructed by integrating the *pol3-01* or *pol3-L612M* allele into the chromosomal *POL3* gene by targeted integration, thereby disrupting the endogenous *POL3* gene. In contrast, our mutator plasmid expressing the *pol* $\delta$  mutant allele produced a mutation rate of  $7.9 \times 10^{-6}$ , which shows a high mutation rate as well as chromosomal integration. The use of the *pol* $\delta$  MKII mutator plasmid allows the continued expression of the endogenous wild-type *POL3* and provides for an efficient restoration of the wild-type mutation rate by curing the yeast strains of the mutator plasmid. Once the desired trait(s) has been selected, curing the cells from the mutator plasmid can stabilize the newly obtained phenotype.

In general, all random mutagenesis methods developed to date are biased toward transition mutations, although efforts have been made to overcome this [28]. While transition bias was observed in EMS, we actually observed transversion bias with the mutator (Figure 3(a)). Because of this, the mutator yielded a broader spectrum of nucleotide changes across the entire genome. The mutator was also biased toward transversions in the nonsynonymous substitutions (Figure 3(b)). For EMS, the spectrum of mutation events we observed is similar to what has been reported by others [12].

At the protein level, the amino acid substitution pattern differed between the mutator and EMS (Table 4). Mutations generated by the mutator resulted in amino acid substitutions more often than did mutations generated by EMS (85% versus 61%, resp.). Most of the substitutions made by the mutator were nonconservative, whereas only half of the substitutions made by EMS were nonconservative. In addition, the mutator generated more structure-disturbing amino acid changes (Gly/Pro). The transversion bias of non-synonymous substitutions by the mutator generates more diverse amino acid substitution patterns than does the transition bias of EMS.

Although the average base-substitution mutation rate of EMS was approximately 100 times higher than that of the mutator, the mutation frequency of the mutator was approximately 7 times higher than that of EMS. This gap between a higher apparent mutation frequency and fewer mutations may be explained by the higher proportion of amino acid changes and the diversity of amino acid substitutions by the mutator. This suggests one plausible explanation for the effectiveness of the disparity mutagenesis.

The disparity mutagenesis technique has been successfully applied to not only eukaryotic microorganisms such as *S. cerevisiae* [5, 7–9], *S. pombe* [9], and *Ashbya gossypii* [10], but also to prokaryotic microorganisms such as *Escherichia coli* [4] and *Bradyrhizobium japonicum* [6]. We believe that

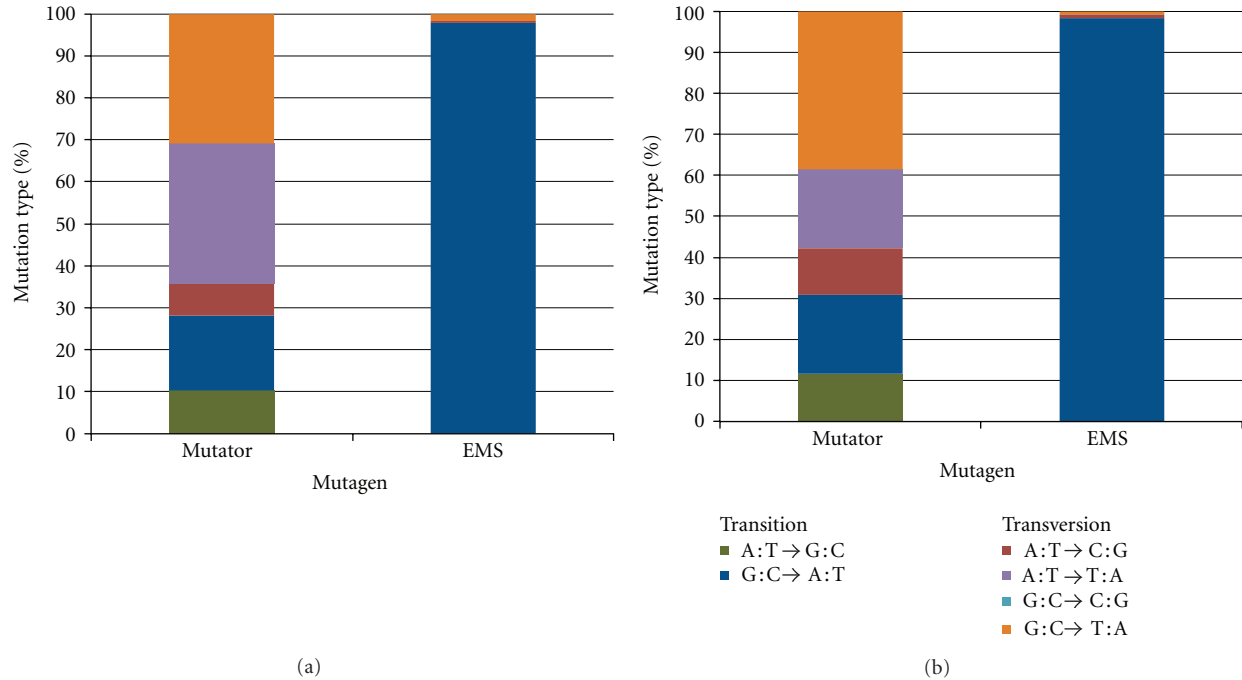


FIGURE 3: Relative frequency of transitions and transversions induced by *polδ* and EMS. The mutations spectra show the frequency of transitions and transversions generated by the mutator and EMS. The data for the individual strains were combined according to the mutagen used. The color key identifying the type of mutation is provided in the inset. Complementary mutations, such as A → C and T → G, are pooled. (a) Genome-wide profile; (b) Non-synonymous substitutions only.

TABLE 4: Mutations at protein level.

	Mutator		EMS	
	<i>n</i>	%	<i>n</i>	%
Total mutations	28	100	201	100
Preserved amino acids	2	7.1	74	36.8
Amino acid changes	24	85.7	123	61.2
Stop	2	7.1	4	2.0
Changes in codon letter	28	100	201	100
1st	11	39.3	64	31.8
2nd	13	46.4	65	32.3
3rd	4	14.3	72	35.8
Impact of amino acid change	24	100	123	100
Conservative <sup>a</sup>	4	16.7	57	46.3
Nonconservative	20	83.3	66	53.7
Stop and Gly/Pro codons	4	15.4	4	3.1
Stop	2	50.0	4	100.0
Gly/Pro	2	50.0	0	0.0

<sup>a</sup> Conservative and nonconservative amino acid substitutions were defined according to the BLOSUM62 matrix [25].

this novel mutagenesis technique has the potential to be applied to a wide variety of microorganisms.

Our present study has demonstrated that a proofreading-deficient and low-fidelity *polδ**MKII* mutator is a useful and efficient method for rapid strain improvement based on

*in vivo* mutagenesis. It has been suggested that organisms may accelerate evolution by decreasing the fidelity of the proofreading activity of *polδ* in nature [29]; therefore, this mutator may also be useful for studying the acceleration of evolution.

## Abbreviations

EMS: Ethyl methanesulfonate  
 SNV: Single nucleotide variant  
 Indel: Insertions and deletions  
 SC: Synthetic complete.

## Data Access

The raw reads used in this study are available on the DDBJ Sequence Read Archive (DRA) under accession DRA000522.

## Acknowledgments

Authors thank Dr. M. Furusawa for critical reading of the manuscript and for useful suggestions. This study was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (S0801025).

## References

- [1] N. E. Labrou, "Random mutagenesis methods for in vitro directed enzyme evolution," *Current Protein and Peptide Science*, vol. 11, no. 1, pp. 91–100, 2010.
- [2] O. Selifonova, F. Valle, and V. Schellenberger, "Rapid evolution of novel traits in microorganisms," *Applied and Environmental Microbiology*, vol. 67, no. 8, pp. 3645–3649, 2001.
- [3] M. Furusawa and H. Doi, "Promotion of evolution: disparity in the frequency of strand-specific misreading between the lagging and leading DNA strands enhances disproportionate accumulation of mutations," *Journal of Theoretical Biology*, vol. 157, no. 1, pp. 127–133, 1992.
- [4] K. Tanabe, T. Kondo, Y. Onodera, and M. Furusawa, "A conspicuous adaptability to antibiotics in the *Escherichia coli* mutator strain, dnaQ49," *FEMS Microbiology Letters*, vol. 176, no. 1, pp. 191–196, 1999.
- [5] C. Shimoda, A. Itadani, A. Sugino, and M. Furusawa, "Isolation of thermotolerant mutants by using proofreading-deficient DNA polymerase  $\delta$  as an effective mutator in *Saccharomyces cerevisiae*," *Genes and Genetic Systems*, vol. 81, no. 6, pp. 391–397, 2006.
- [6] M. Itakura, K. Tabata, S. Eda et al., "Generation of *Bradyrhizobium japonicum* mutants with increased N<sub>2</sub>O reductase activity by selection after introduction of a mutated dnaQ gene," *Applied and Environmental Microbiology*, vol. 74, no. 23, pp. 7258–7264, 2008.
- [7] H. Abe, Y. Fujita, Y. Chiba, Y. Jigami, and K. I. Nakayama, "Upregulation of genes involved in gluconeogenesis and the glyoxylate cycle suppressed the drug sensitivity of an N-glycan-deficient *Saccharomyces cerevisiae* mutant," *Bioscience, Biotechnology and Biochemistry*, vol. 73, no. 6, pp. 1398–1403, 2009.
- [8] H. Abe, Y. Fujita, Y. Takaoka et al., "Ethanol-tolerant *Saccharomyces cerevisiae* strains isolated under selective conditions by over-expression of a proofreading-deficient DNA polymerase  $\delta$ ," *Journal of Bioscience and Bioengineering*, vol. 108, no. 3, pp. 199–204, 2009.
- [9] H. Abe, Y. Takaoka, Y. Chiba et al., "Development of valuable yeast strains using a novel mutagenesis technique for the effective production of therapeutic glycoproteins," *Glycobiology*, vol. 19, no. 4, pp. 428–436, 2009.
- [10] E. Y. Park, Y. Ito, M. Nariyama et al., "The improvement of riboflavin production in *Ashbya gossypii* via disparity mutagenesis and DNA microarray analysis," *Applied Microbiology and Biotechnology*, vol. 91, no. 5, pp. 1315–1326, 2011.
- [11] T. Matsuzawa, Y. Fujita, N. Tanaka, H. Tohda, A. Itadani, and K. Takegawa, "New insights into galactose metabolism by *Schizosaccharomyces pombe*: isolation and characterization of a galactose-assimilating mutant," *Journal of Bioscience and Bioengineering*, vol. 111, no. 2, pp. 158–166, 2011.
- [12] S. Flibotte, M. L. Edgley, I. Chaudhry et al., "Whole-genome profiling of mutagenesis in *Caenorhabditis elegans*," *Genetics*, vol. 185, no. 2, pp. 431–441, 2010.
- [13] T. Fukui, K. Yamauchi, T. Muroya et al., "Distinct roles of DNA polymerases delta and epsilon at the replication fork in *Xenopus* egg extracts," *Genes to Cells*, vol. 9, no. 3, pp. 179–191, 2004.
- [14] L. Li, K. M. Murphy, U. Kanevets, and L. J. Reha-Krantz, "Sensitivity to phosphonoacetic acid: a new phenotype to probe DNA polymerase  $\delta$  in *Saccharomyces cerevisiae*," *Genetics*, vol. 170, no. 2, pp. 569–580, 2005.
- [15] A. Morrison, A. L. Johnson, L. H. Johnston, and A. Sugino, "Pathway correcting DNA replication errors in *Saccharomyces cerevisiae*," *The EMBO Journal*, vol. 12, no. 4, pp. 1467–1473, 1993.
- [16] A. Morrison and A. Sugino, "The 3'  $\rightarrow$  5' exonucleases of both DNA polymerases  $\delta$  and  $\epsilon$  participate in correcting errors of DNA replication in *Saccharomyces cerevisiae*," *Molecular and General Genetics*, vol. 242, no. 3, pp. 289–296, 1994.
- [17] S. A. N. McElhinny, C. M. Stith, P. M. J. Burgers, and T. A. Kunkel, "Inefficient proofreading and biased error rates during inaccurate DNA synthesis by a mutant derivative of *Saccharomyces cerevisiae* DNA polymerase," *The Journal of Biological Chemistry*, vol. 282, no. 4, pp. 2324–2332, 2007.
- [18] R. N. Venkatesan, J. J. Hsu, N. A. Lawrence, B. D. Preston, and L. A. Loeb, "Mutator phenotypes caused by substitution at a conserved motif A residue in eukaryotic DNA polymerase  $\delta$ ," *The Journal of Biological Chemistry*, vol. 281, no. 7, pp. 4486–4494, 2006.
- [19] B. Timmermann, S. Jarolim, H. Russmayer et al., "A new dominant peroxiredoxin allele identified by whole-genome re-sequencing of random mutagenized yeast causes oxidant-resistance and premature aging," *Aging*, vol. 2, no. 8, pp. 475–486, 2010.
- [20] R. D. Gietz and A. Sugino, "New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites," *Gene*, vol. 74, no. 2, pp. 527–534, 1988.
- [21] D. E. Lea and C. A. Coulson, "The distribution of the numbers of mutants in bacterial populations," *Journal of Genetics*, vol. 49, no. 3, pp. 264–285, 1949.
- [22] H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, 2009.
- [23] H. Li, B. Handsaker, A. Wysoker et al., "The sequence alignment/map format and SAMtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, 2009.
- [24] J. T. Robinson, H. Thorvaldsdóttir, W. Winckler et al., "Integrative genomics viewer," *Nature Biotechnology*, vol. 29, no. 1, pp. 24–26, 2011.
- [25] S. Henikoff and J. G. Henikoff, "Amino acid substitution matrices from protein blocks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 22, pp. 10915–10919, 1992.

- [26] M. Lynch, W. Sung, K. Morris et al., "A genome-wide view of the spectrum of spontaneous mutations in yeast," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 27, pp. 9272–9277, 2008.
- [27] T. S. Wong, D. Roccatano, M. Zacharias, and U. Schwaneberg, "A statistical analysis of random mutagenesis methods used for directed protein evolution," *Journal of Molecular Biology*, vol. 355, no. 4, pp. 858–871, 2006.
- [28] T. S. Rasila, M. I. Pajunen, and H. Savilahti, "Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, *Escherichia coli* mutator strain, and hydroxylamine treatment," *Analytical Biochemistry*, vol. 388, no. 1, pp. 71–80, 2009.
- [29] M. Furusawa, "Implications of double-stranded DNA structure for development, cancer and evolution," *Open Journal of Genetics*, vol. 01, no. 03, pp. 78–87, 2011.