Molecular Therapy
Methods & Clinical Development
Original Article

Development of a new genetic reference material system based on *Saccharomyces cerevisiae* cells

Xin He,[1,3] Jiaqi Ding,[1,3] Zhenhua Xu,[1] Na Li,[1] Jingmin Yang,[2] Hongyan Chen,[1] and Daru Lu[1,2]

[1]State Key Laboratory of Genetic Engineering and MOE Engineering Research Center of Gene Technology, School of Life Sciences and Zhongshan Hospital, Fudan University, Shanghai 200438, China; [2]NHC Key Laboratory of Birth Defects and Reproductive Health, Chongqing Population and Family Planning Science and Technology Research Institute, Chongqing 401120, China

As an important quality control link of molecular diagnosis, genetic reference materials (RMs) are widely used in various gene detection platforms such as mutation detection, gene quantification, and second generation sequencing. However, contamination, construction, and storage of existing genetic RMs still remain challenges. Here, we established a new genetic RM system based on *Saccharomyces cerevisiae*. We chose the non-small cell lung cancer (NSCLC) mutation hotspots in Kirsten rat sarcoma viral oncogene (*KRAS*) and epidermal growth factor receptor (*EGFR*), using clustered regularly interspaced short palindromic repeats and CRISPR-associated protein (CRISPR-Cas9) system-mediated gene editing technology, combined with the high homologous recombination efficiency of *Saccharomyces cerevisiae*. A single copy of the target gene was inserted into the yeast genome, and the inserted target gene was stably inherited with the passage of yeast cells. The copy number calculation for the target gene can replays by cell counting. The RM system was evaluated by sequence, copy number, stability, and homogeneity. In summary, the recombinant yeast cell line has ease of construction and screening, stable genetic characteristics, accurate copy number calculation, and convenient culture and preservation. Our findings may provide new ideas and directions for the research and industrialization of genetic RMs.

INTRODUCTION

Genetic testing methods are widely used in many fields of human health, such as diagnosis of tumors and genetic diseases, detection of pathogenic microorganisms, and evaluation of genetically modified products. The certification of reference materials (RMs) for genetic molecular diagnosis are not comprehensive. The existing genetic diagnosis platform and the newly developed detection methods indicate an absence of a unified standard certification quality control, and the reported results are insufficiently reliable.[1] RMs with homogeneity, stability, and defined qualitative and quantitative characteristics are needed.[2,3] For RMs applied to genetic testing, the system should be homogeneous and stable, and the RMs should have clear DNA sequence and DNA molecular abundance and other basic characteristics.[4] Genetic RMs provide negative and positive controls for genetic detection to ensure the accuracy and reliability of experimental results.[5] The genetic RM plays an important role in the development of clinical test products, platform testing and calibration, quality control of the detection process, evaluation of test methods, determination of sample values, and reference of experimental repeatability.[6]

A series of RMs are used according to the experimental conditions:[7] commercial cell lines and DNA samples; remaining patient samples; sample sharing between laboratories; remaining samples of published research results;[8] cell lines containing genetic mutations;[9,10] genetically modified cell lines;[11] and plasmids, PCR products, and synthetic DNA.[12] Many types of genetic detection of RMs have their own advantages and limitations. Clinical patient samples have limitations such as inconvenience, low numbers, poor homogeneity, and inability to prepare RMs in large quantities;[13] the process of constructing cell lines with genetic mutations is cumbersome; and the double-copy characteristic of human cell genomes causes the possibility of heterozygous mutations. Compared with commercial cell lines with high cost and strict culture conditions, plasmid samples are simple and rapid to construct, but the plasmid copy number in the system is high, and dilution and simulation of genomic DNA and mixing of lower mutation percentage RMs will cause great errors and easy contamination, resulting in false positives of the experimental results.[14] The lack of suitable genetic reference standards has greatly hindered the development of genetic molecular diagnostic methods and the standardization of detection capability.

In this study, we selected yeast cells as a genetic background and established human cell mutations knock-in yeast cells by homologous recombination. Yeast cells have the following advantages as an RM carrier:[15] the genetic manipulation, cell culture, and DNA extraction methods of yeast cells are simple and easy and genetic modification in yeast cells can be stably inherited, amplified, and preserved. The transformation of haploid yeast cells can directly obtain homozygous mutant individuals that can be accurately mixed with yeast cells of different genetic types or wild-type human cells, as a quality control

Received 23 October 2020; accepted 13 January 2021;
https://doi.org/10.1016/j.omtm.2021.01.004.

[3]These authors contributed equally

Correspondence: Daru Lu, State Key Laboratory of Genetic Engineering and MOE Engineering Research Center of Gene Technology, School of Life Sciences and Zhongshan Hospital, Fudan University, Shanghai 200438, China.
E-mail: darulu@163.com

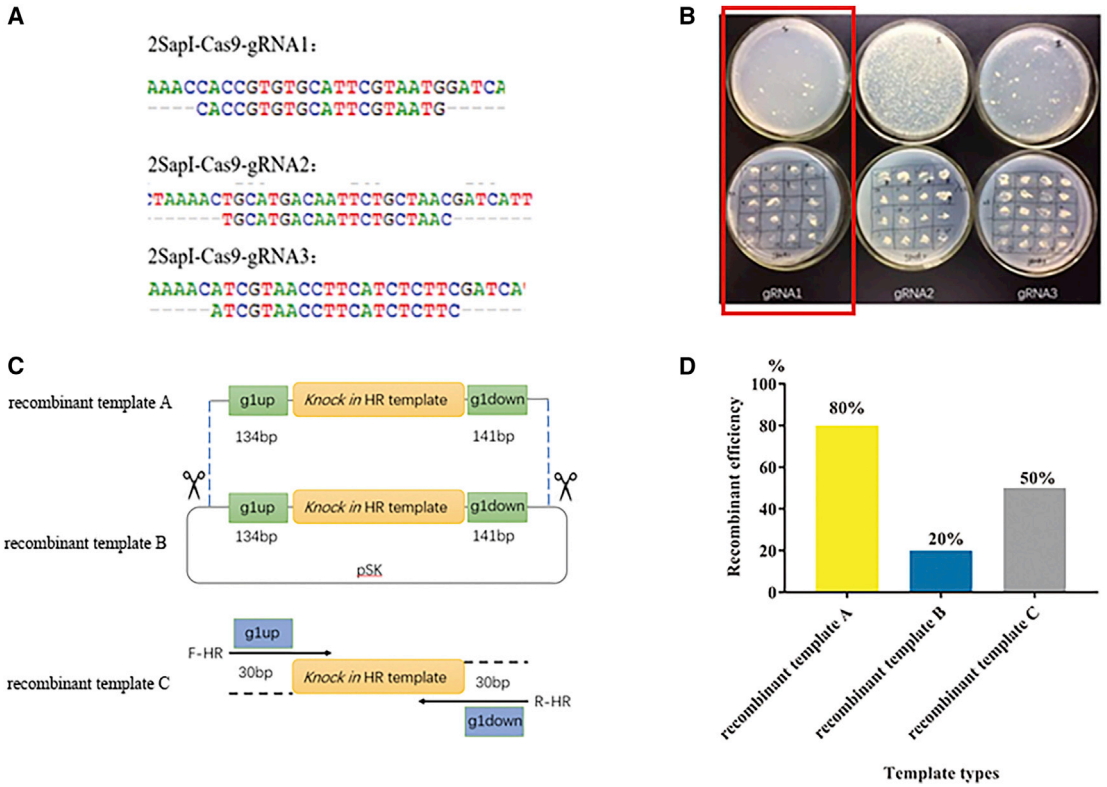


Figure 1. Construction and optimization of recombinant *S. cerevisiae* genetic RMs

(A) Three single guide RNAs (sgRNAs) were designed, and the results of 2SapI-Cas9-gRNA1, 2SapI-Cas9-gRNA2, and 2SapI-Cas9-gRNA3 plasmids were sequenced. (B) By transferring the single colonies grown on the plate to the new SD-Leu- plate, the number of colony growth was inversely proportional to the editing efficiency. (C) Schematic diagram of three homologous recombination templates. (D) Comparison of recombination efficiency of three homologous recombination templates.

in low-frequency mutation detection experiments. Yeast cells are preserved in hypertonic solution after protoplast formation, and they have a similar cell-wall-free morphology in the background RMs of human cells. They can be made into genomic DNA, a cell suspension, a formalin-fixed paraffin-embedded (FFPE) block, and other forms to meet the actual needs for a variety of RM forms. Yeast cells can accommodate large exogenous DNA fragments, including synthetic chromosomes.

The discovery and universal application of gene editing technology mediated by the CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats and CRISPR-associated protein) system provide a powerful tool for site-specific gene editing in different species and the establishment of human disease models.[16] The CRISPR-Cas9 system efficiently mediates site-directed DNA double-strand breaks (DSBs), usually activating non-homologous end joining (NHEJ) mechanisms of repair in the cell, followed by non-target insertions, deletions, and other mutations (indels).[17] This principle has been widely used for gene knockout.[18] DNA DSBs are also repaired by homologous recombination. When single- or double-stranded recombination templates are introduced, we can specifically knock in mutations.[19] *Saccharomyces cerevisiae* has efficient homologous recombination mechanisms and DSB repair efficiency, both of which are convenient for genetic modification.[20] Compared with the expression or replication of plasmid-transformed yeast cells, gene integration yeast has better strain stability, specific gene expression levels and copy number, and lower intercellular heterogeneity.[21] Gene editing of yeast cell genomic target sites can be performed by guide RNAs (gRNAs) that target specific sites. That is, by using CRISPR-Cas9 technology-mediated gene editing technology, combined with the higher homologous recombination efficiency of yeast cells, a single-copy target gene can be inserted into the yeast genome with inheritance.

Non-small cell lung cancer (NSCLC) accounts for ~80% of the total number of lung cancers. Most patients have lost the opportunity to undergo radical surgery by the time that they are diagnosed.[22,23] Epidermal growth factor receptor (*EGFR*) and Kirsten rat sarcoma viral oncogene (*KRAS*) are important biomarkers for NSCLC-targeted therapy.[24,25] *EGFR* mutations are mainly concentrated on 18–21 exons, with the most common type for 19 exon deletion and 21 exon L858R alteration, accounting for ~90% of the total mutation.[26] Ninety percent of *KRAS* mutations are point mutations in the 12/13 codons of exon 2. EGFR-tyrosine kinase inhibitors (TKIs) treatment

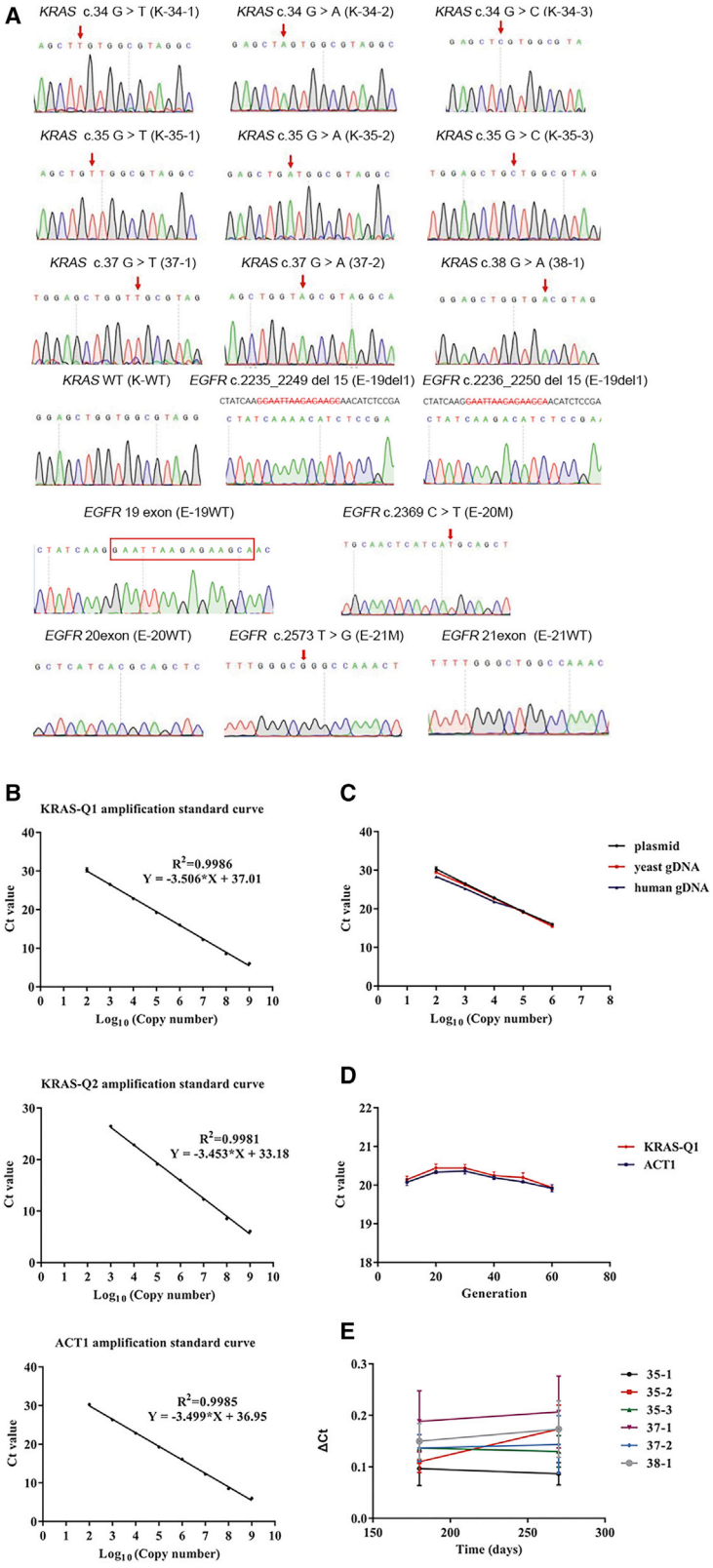


Figure 2. Verification of calibration ability of recombinant yeast cell genetic RMs

(A) The sequencing results of recombinant yeast genetic RMs. (B) Taking a recombinant yeast clone with *KRAS* mutation as an example, standard curves were drawn for the three pairs of primers, KRAS-Q1, KRAS-Q2, and ACT1, used in the experiment. (C) qPCR standard curve of recombinant template plasmid, human genomic DNA, and recombinant yeast genomic DNA. (D) The recombinant yeast cells were serially passaged, and the copy number of the target gene of 10th, 20th, 30th, 40th, 50th, and 60th generation yeast cells was detected by qPCR. (E) Verification of copy number stability of recombinant yeast cells stored in 20% glycerol at −80°C for 180 and 270 days.

Table 1. *KRAS* recombinant yeast genome qPCR amplification Ct values

Recombinant yeast type no.	KRAS-Q1	KRAS-Q2	ACT1
K-34-1	17.72 ± 0.04	–	17.64 ± 0.07
K-34-2	18.13 ± 0.04	–	18.10 ± 0.03
K-34-3	18.29 ± 0.06	–	18.14 ± 0.06
K-35-1	18.01 ± 0.09	–	17.93 ± 0.09
K-35-2	18.28 ± 0.07	–	18.21 ± 0.06
K-35-3	15.91 ± 0.04	15.96 ± 0.05	15.78 ± 0.06
K-37-1	15.18 ± 0.02	15.27 ± 0.10	15.22 ± 0.06
K-37-2	18.15 ± 0.06	–	18.01 ± 0.10
K-38-1	17.98 ± 0.07	–	17.84 ± 0.07
K-WT	17.36 ± 0.07	–	17.30 ± 0.03

Values are means ± SD.

is effective in patients with EGFR-sensitive mutations in NSCLC, but *KRAS* mutations can cause resistance to EGFR-TKIs in NSCLC patients.[27,28] The detection of *KRAS* and *EGFR* mutations has important value for the individual treatment of cancer. In the diagnosis and targeted therapy of cancer patients, the cells to be detected are mainly circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA).[29–31] In addition to ctDNA, circulating peripheral blood contains normal tissue DNA, and the proportion of ctDNA is usually less than 10%, but sometimes as low as one-thousandth to one-ten thousandth.[32] High-resolution, low-frequency mutation detection methods and genetic RMs that mimic low-concentration mutants are required for gene detection.

This study explored and optimized the CRISPR-Cas9-mediated method for site-directed knocked-in *KRAS* and *EGFR* mutations in yeast. Single-copy target gene knock-in recombinant yeast cells can be an RM for low-frequency mutation detection. The qualitative and quantitative characteristics, stability, homogeneity, and other important indicators of genetic RMs were evaluated.

RESULTS

Establishment and optimization of recombinant *S. cerevisiae* RMs mediated by CRISPR-Cas9 system

To select gRNA with high editing efficiency, three gRNA annealing products targeting the *URA3* gene were screened and ligated into the p425-Sap-TEF1p-Cas9-CYC1t-2xSap vector to obtain three *URA3*-target plasmids. The sequencing results are shown in Figure 1A. The three plasmids were transformed into *Saccharomyces cerevisiae* W303-1A cells, and the editing efficiency of three gRNAs was URA3-gRNA1, 30%; URA3-gRNA2, 0%; and URA3-gRNA3, 13% (Figure 1B). The number of colonies was inversely related to the editing efficiency. URA3-gRNA1 with high editing efficiency was selected for subsequent study (Figure 1B).

According to the targeted cleavage position of gRNA1, we designed the upstream and downstream homologous arms and constructed

Table 2. Stability of insert fragment copy number of recombinant yeast in glycerol for short-term storage

Recombinant yeast type (temperature and no.)	20 day KRAS-Q1	20 day ACT1	ΔCt_{20}	0 day ΔCt_0
−20°C K-34-1	21.09 ± 0.10	20.90 ± 0.08	0.20	0.17
−20°C K-35-1	20.92 ± 0.15	20.63 ± 0.09	0.29	0.07
4°C K-34-2	20.93 ± 0.02	20.73 ± 0.03	0.20	0.27
4°C K-35-3	21.21 ± 0.14	20.93 ± 0.10	0.27	0.17
Room temperature K-34-3	20.71 ± 0.04	20.46 ± 0.05	0.25	0.24

Values are means ± SD.

the three homologous recombinant templates. The three types of templates are as follows: recombinant template A, a linearized plasmid containing 134 bp upstream homologous arm and 141 bp downstream homologous arm and knock-in fragments of *EGFR* and *KRAS* mutations; recombinant template B, a circular plasmid knock-in fragment containing 134 bp upstream homologous arm and 141 bp downstream homologous arm and *EGFR* and *KRAS* mutations; and recombinant template C, a PCR product containing 30 bp upstream homologous arm and 30 bp downstream homologous arm and knock-in fragments of *EGFR* and *KRAS* mutations (Figure 1C). Meanwhile, we compared the recombination efficiency of three different types of templates. For all three of the recombinant templates, we selected 1 μg as the concentration of integration and co-transformed *Saccharomyces cerevisiae* with Cas9 plasmid. The results showed that the recombination efficiency of the recombinant template A, recombinant template B, and recombinant template C was 80%, 20%, and 50%, respectively (Figure 1D). Therefore, 1 μg of the recombinant template A was selected to co-transform *S. cerevisiae* with Cas9 plasmid, and 17 recombinant yeast monoclones containing *KRAS*/*EGFR* wild-type and mutant fragments were rapidly obtained.

Verification of calibration ability of recombinant yeast cell genetic RMs

To verify the reliability of recombinant yeast cell sequences, both the *URA3* gene amplification primers and the primers in the middle of the knock-in sequence were selected for PCR amplification of the positive products for Sanger sequencing. As shown in Figure 2A, the results showed that the knock-in sequence was inserted into the gRNA1 targeting site of the *URA3* gene.

We then detected whether a single copy of the target gene is inserted into the yeast cell. Taking a recombinant yeast clone inserted with a *KRAS* mutation as an example, three pairs of primers, KRAS-Q1, KRAS-Q2, and ACT1, were used for PCR amplification, and standard curves were drawn for the three pairs of primers (Figure 2B). The primer amplification standard curve R^2 value of KRAS-Q1, KRAS-Q2, and ACT1 was 0.9986, 0.9981, and 0.9985, respectively (Figure 2B), and the amplification efficiency of the primers was 92.75%, 94.83%, and 93.1%, respectively. The standard curve has a high degree of fit, and the primer amplification efficiency is high

Table 3. Stability of insert fragment copy number of recombinant yeast protoplasts stored in hypertonic PB for short-term storage

Recombinant yeast type (temperature and no.)	20 day KRAS-Q1	20 day ACT1	ΔCt_{20}	0 day ΔCt_0
4°C K-34-2	20.25 ± 0.05	20.20 ± 0.08	0.05	0.27
4°C K-35-3	19.91 ± 0.07	19.80 ± 0.04	0.27	0.17
Room temperature K-34-3	21.17 ± 0.07	19.75 ± 0.04	1.42	0.24
Room temperature K-35-2	22.63 ± 0.13	20.17 ± 0.05	2.46	0.22

Values are means ± SD.

and similar. To determine whether the target fragment has a random insertion other than a specific site, we compared the cycle threshold (Ct) values amplified by the KRAS-Q1, KRAS-Q2, and ACT1 primers. The amplification values of the target sequence primers and the internal reference primers were both $\Delta Ct < 0.2$ (Table 1), demonstrating that the CRISPR-Cas9-mediated homologous recombination method can knock a single copy of the exogenous gene in the *Saccharomyces cerevisiae* W303-1A cell.

The plasmid and human genome are the most common RMs for genetic testing. The qPCR standard curves were used to determine whether the recombinant yeast genomic DNA can be used as an RM for DNA quantification. The results demonstrated that recombinant yeast genomic DNA can be used as an RM for DNA quantification (Figure 2C).

To verify the stability of recombinant yeast RMs, we tested the following aspects: passaging; long-term storage at −80°C in 20% glycerol; short-term storage in 20% glycerol at −20°C, 4°C, and room temperature; and recombinant yeast protoplasts stored at 4°C and room temperature for the short term. As shown in the Table S3 and Figure 2D, with the increase of growth generation, the inserted target gene was consistent with the internal reference gene copy number on the yeast genome, and the $\Delta Ct < 0.3$ of the qPCR of the two indicated that the inserted target gene can be stably inherited with the passage of yeast cells. We can see in Figure 2E that under long-term cryopreservation conditions, the ΔCt value tends to be stable or slightly increased with an increase in storage time. On the whole, when the recombinant yeast was stored for a long time in 20% glycerol at −80°C, the genome and insert fragment copy number of the recombinant yeast were stable. For the recombinant yeast strains preserved with 20% glycerol under different temperature conditions, the ΔCt values of the target gene and the internal reference gene amplification were all lower than 0.3, and the ΔCt value did not change significantly after 20 days of storage (Table 2). Therefore, when the recombinant yeast was stored in 20% glycerol at −20°C, 4°C, and room temperature for a short period of time, the genome and insert fragment copy number of the recombinant yeast were stable. As shown in Table 3, there was no significant change in the ΔCt value of yeast protoplasts stored at 4°C for 20 days, whereas the ΔCt value of protoplasts stored at room temperature for 20 days was increased, and the Ct value of the target gene became larger. Therefore, when the recombinant yeast protoplasts were stored for a short period of time at 4°C for 20 days, the copy number of the recombinant yeast genome and insert was stable, whereas the copy number of the target gene was reduced at room temperature for 20 days. In summary, after continuous passage of recombinant yeast genetic RMs, long-term frozen storage at −80°C, storage in glycerol at −20°C, 4°C, and room temperature for the short term, and recombinant yeast protoplasts stored at 4°C for 20 days, the number of copies of the target gene all remained stable, i.e., $\Delta Ct < 0.3$ with single-copy reference gene.

Table 4. Universal buffer formulation and cell state of different yeast protoplasts and human cells

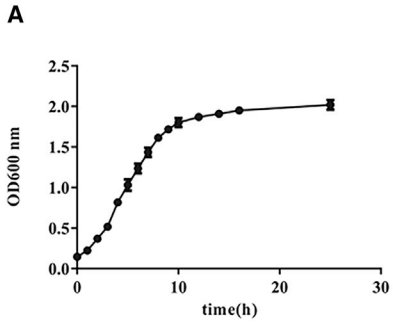
Cell fixative type	Cell state and morphology (293T)	Cell state and morphology (recombinant yeast protoplast)
1× PBS	normal	shrink
10× PBS	shrink	swelling is not broken
0.8 M hypertonic PB solution	severe shrinkage	normal
4% paraformaldehyde	shrink	cells become smaller and the edges become darker
4% neutral formaldehyde	slightly shrink and maintain cell morphology	complete cell morphology and round
5% glacial acetic acid	normal	normal
Ethanol/glacial acetic acid (3:1)	good shape, great fluidity	good shape, great fluidity
Methanol/glacial acetic acid (3:1)	good shape, great fluidity	good shape, great fluidity
95% ethanol	slight dehydration, large fluidity	slight dehydration, large fluidity
70% ethanol	good shape, great fluidity	good shape, great fluidity

Homogenization of recombinant yeast cells and human cells

The genetic RM should be homogeneous with the biological sample for genetic testing, which can be manipulated by the same processes and DNA extraction methods. Yeast cells are eukaryotes with cell walls that can be subjected to lytic enzymatic digestion to obtain spherical protoplasts that are closer to the morphology and preservation of human cells. Thus, it is necessary to develop a universal buffer solution common to yeast protoplasts and human cells.

Yeast cell growth is mainly divided into four phases: growth lag phase, logarithmic growth phase, plateau phase, and senescence death phase. The growth curve of *S. cerevisiae* W303-1A cells is plotted in Figure 3A and shows an optical density 600 (OD600) value of 0.5–1.5 in the logarithmic growth phase.

We hope to find a suitable enzymatic hydrolysis condition to maintain protoplast morphology. The rate of production of yeast protoplasts is



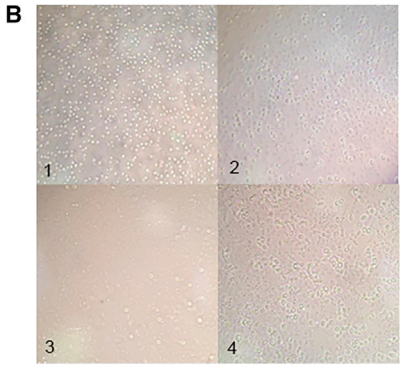


Figure 3. Homogeneous treatment of recombinant yeast cells and human cells

(A) The growth curve of *Saccharomyces cerevisiae* W303-1A strain. (B) The protoplast morphology of *Saccharomyces cerevisiae* W303-1A cells: (B-1) the yeast protoplast in the osmotic buffer (10 × 40 times magnification by optical microscopy); (B-2) with the addition of pure water to the yeast protoplast suspension, the cell swells; (B-3) protoplasts rose and broke after absorbing water for a period of time; and (B-4) the enzymatic hydrolysis time is too long and the protoplast cell atypicality increased.

related to cell generation, cell growth stage, lytic enzyme concentration, and the type of osmotic stabilizer. We explored different lyticase concentrations and enzymolysis time gradients to observe the rate of protoplast formation under different conditions (Table S4). Yeast protoplast cells are transparent spheres that rise and break after a certain degree of water absorption, and cell debris can be seen in the visual field (Figure 3B, 1–3), whereas yeast cells, which do not form protoplasts, retain their original cell morphology. A lyticase concentration of 0.08 U/mL and enzymatic hydrolysis for 40 min was the optimal combination. When the enzymatic hydrolysis time was insufficient or the enzyme concentration was too low, cell wall digestion was incomplete, and the protoplast release rate was low. When the enzymatic hydrolysis time was too long, enzymatic toxicity resulted in protoplast release too early, which caused the protoplast to rupture. Protoplast formation rate decreased after 60 min of enzymatic hydrolysis, and the protoplast cell atypicality increased (Figure 3B, 4), thereby being easily damaged and unconducive to cell count and resolution of cell morphology. The results showed that the protoplast effect was the best when OD600 = 1.4, the production rate was high (>95%), and the cell permeability was regular (Figure 3B, 1).

To obtain a humanized yeast-human cell mixed RM, we explored yeast protoplast and human cell universal buffer formulations. Recombinant yeast protoplasts and human embryonic kidney 293T cells proliferated *in vitro* were obtained. The two cells were separately resuspended in 10 commonly used cell fixatives and buffers (Table 4). The results, summarized in Table 4, show that the preferred cell morphology retention in several buffers, from high to low, was as follows: 5% glacial acetic acid, ethanol/glacial acetic acid mixture (3:1), methanol/glacial acetic acid mixture (3:1), 4% neutral formaldehyde, and 70% ethanol. Therefore, 5% glacial acetic acid was the optimal immobilization buffer (Table 4), and the cells preserved in this fixative were positive for DNA extraction and amplification of the target fragment.

Recombinant yeast cells mixed to mimic low-frequency mutations in tumor tissues

In this study, we used an automatic cell counter to count yeast cells, and a mixture of different genotypes of recombinant yeast cells yielded 10%, 1%, 0.5%, and 0.1% samples, simulating the proportion of low-frequency mutations present in the circulating DNA of the tumor. To count more accurately, we performed protoplast treatment on yeast cells mixed in a certain proportion, DAPI nuclear staining of yeast protoplasts, and observation of the presence of nuclei that were not detached from the mother by laser confocal microscopy. As can be seen from the Figure 4A, the yeast that is budding does not leave the mother's progeny and also has a blue nucleus after DAPI staining. Therefore, when the yeast is counted, the yeast that is budding is also recorded as one copy.

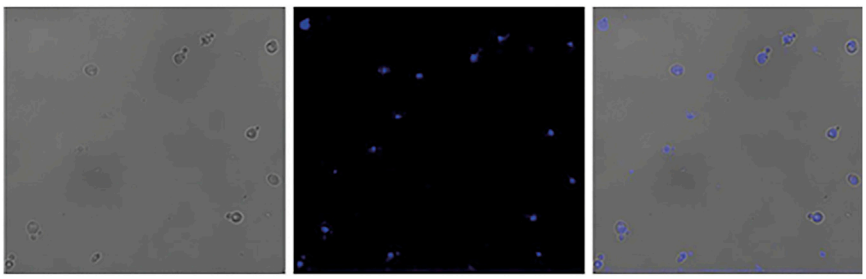
Next generation sequencing and micro-drop digital PCR were used to detect the mutation frequency of the mixed sample, and the difference between the mutation frequency of the mixed cell system and the expected value was examined. The detection results show the gradient of the mutation frequency of the sample, and the existence mode of the low-frequency mutation in the simulated tumor clinical samples was double verified. By the mixing of different genotypes of recombinant yeast cells, the theoretical value of *EGFR* 19exon c. 2235_2249del15 mutation frequency was 10%, 1%, and 0.1% and the theoretical value of *EGFR* exon21 p. L858R_c. 2573T>G mutation frequency was 10%, 1.1%, 1%, 0.5%, and 0.1%. The verification results of digital PCR and second generation sequencing are shown in Table S5. For samples with different mutation frequencies of L858R, the value of the next generation sequencing mutation frequency was higher than that obtained via digital PCR. As shown in Figure 4B, the detection results of the mixed sample mutation frequency were close to the expected values, and the two detection results had a good linear relationship, indicating that this genetic RM with different mutation frequencies is reliable.

Application form of genetic RMs

For development of multiple genetic RMs for different needs, recombinant yeast cells with specific mutations or mutation frequencies were used to form protoplast suspensions, FFPE cell masses, and slices (Figure S1). Also, the genomic DNA of these forms can be extracted and used by enzymatic hydrolysis of the yeast cell wall by using lytic enzyme.

DISCUSSION

This study has made bold explorations and attempts in the field of genetic RMs. Yeast cells were used as the genetic background, and the target gene was stably integrated into the genome of yeast cells by



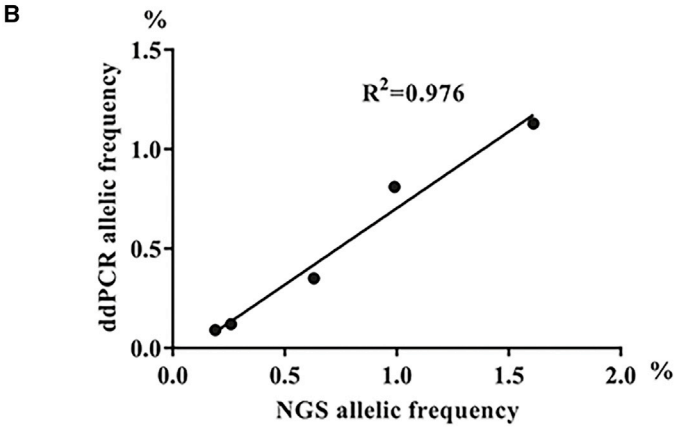


Figure 4. Recombinant yeast cells mixed to mimic low-frequency mutations in tumor tissues

(A) Representative images of DAPI nuclear staining of yeast protoplasts. (B) Detection of target mutation frequency of the samples by the next generation sequencing and micro-drop digital PCR.

CRISPR-Cas9-mediated homologous recombination. A new genetic RM system based on *Saccharomyces cerevisiae* was established to evaluate sequence, copy number, stability, and homogeneity. Recombinant yeast cells can be used as RMs for positive control of target gene detection and quality control materials for mutation abundance detection. The rapid development of new scientific disciplines such as proteomics, metabolomics, and genomics also requires the development of new RMs. The complexity of the preparation of new RMs and the guarantee of their homogeneity and stability are huge challenges.[33] These issues are addressed in our study.

Compared with the genetic RMs that are commonly used, the genetic RM system established in this study can to some extent address the disadvantages of easy contamination, short preservation time, and inaccurate quantification of plasmid or synthetic DNA.[14,34] Compared with commercial cell lines,[8] this system eliminated the complicated cell line construction process and cell immortalization operation. In addition, the system obtained a genetic reference cell line with the same clear mutation frequency and copy number and convenient DNA extraction, amplification, and detection. Compared with the genetically modified cell RM, the system has the advantages of high homologous recombination efficiency, controllable insertion copy number, and simple and rapid culture.[15,35] Compared with previous controls for mutation frequency detection, this system does not need to be calculated by mixing the molar ratios of different mutation types.[36] The direct cell number calculation reduces the uncertainty of the genetic standard system establishment compared with calculated number of genomic moles indirectly.

In addition, this study explored the protoplastization of recombinant yeast RMs and the universal fixed buffer formulation with human cells, enabling the newly established genetic RM system to better simulate the clinical samples. By simultaneously simulating the characteristics of cell genetic heterogeneity in tumor tissue samples, yeast cells with specific mutation/wild-type sequences were counted and mixed, and the mutant frequency of mixed samples was verified by second generation sequencing and micro-drop digital PCR. The results show that this method can obtain mixtures of cells with specific gene mutation frequencies both quickly and accurately. In addition, cells can be stored for a long time, thereby serving as a quality control for various genetic detection methods such as second generation sequencing. Yeast cells can accommodate larger fragments of exogenous DNA, and large-fragment deletions, duplications, and even chromosomal number variations can be modified and used with this RM system as a genetic background.[37]

This new RM system still has some shortcomings. First, the cell diameters of the two systems are 5–10 times different, which is not conducive to the mixing of RMs with specific mutation frequencies. Second, when recombinant yeast cells are used as genetic RMs, protoplast treatment is required, and the treatment process and the DNA extraction process increase the loss of sample. Finally, in the genetic RMs based on *KRAS* and *EGFR* mutation hotspots developed in this study, the actual clinical samples are mostly ctDNA of patients, and subsequent studies will obtain genetic RM forms that mimic the length of circulating DNA fragments of the tumor by means of restriction endonuclease treatment, sonication, and micrococcal nuclease digestion.[38] In this study, the system of integrating exogenous genes into yeast is suitable for most genetic mutations such as point mutations, deletions, insertions, and copy number repeats. However, for large fragment deletions and chromosome number variation, further exploration by yeast artificial chromosomes is needed. Next, we plan to further explore the application of genetic RMs in genetic diseases, such as using RMs to simulate the form of cell-free DNA in the maternal plasma[39] and developing multiple forms of genetic RMs to facilitate scientific research and clinical applications.

MATERIALS AND METHODS

Preparation of experimental materials

The following materials were obtained from Lv Hong's lab (School of Life Sciences, Fudan University): W303-1A *ura3Δ*, a uracil-deficient strain of *Saccharomyces cerevisiae*; pBlueScript-SK (pSK) plasmid, for ligation of homologous recombination templates and homology arm; pSK-URA3-up-down plasmid, a targeted yeast *URA3* homologous recombination template framework; and p425-SapI-TEF1p-Cas9-CYC1t-2xSapI, a Cas9-mediated gene editing system in *Saccharomyces cerevisiae*. Human embryonic kidney 293T cells, *Escherichia coli* DH5α cells, Top10 competent cells.

Primers and probes design

Primer-optimized design was performed by Primer 3 online software, Primer Premier 5 software, and NCBI Primer BLAST website. Probe annealing temperature and secondary structure prediction were performed on Mfold web server website (https://www.rna.albany.edu/unafold-web-server-discontinued). All primers were synthesized by Sango Biotech, and the probes were synthesized by Ningbo Kangbei Biotechnology. The primer and probe sequences used in this study are listed in Table S1.

Establishment of CRISPR-Cas9 gene editing system based on *Saccharomyces cerevisiae* genome safety site *URA3*

p425-Sap-TEF1p-Cas9-CYC1t-2xSap is a yeast-*E. coli* shuttle-type vector that can edit yeast cell genomic DNA at gRNAs targeting-specific sites. Based on the *URA3* gene-coding region sequence, three gRNAs were designed on the CRISPRdirect website (http://crispr.dbcls.jp/) with low off-target rate in different positions of the *URA3* gene. The complementary oligonucleotide (oligo) with the gRNA sequence was annealed and ligated into the BspQI single-stranded linearized p425-SapI-TEF1p-Cas9-CYC1t-2xSapI vector and transformed into *E. coli*. The positive clone was identified and acquired to obtain *URA3* gene targeting the CRISPR-Cas9 plasmid. Plasmids were extracted by using the Axygen Plasmid Mini Kit. The *URA3-target* Cas9 plasmid was transformed into *S. cerevisiae* W303-1A cells, and the product was coated on an SD-Leu- plate and cultured for 2–3 days at 30°C. We found that the colony number on the plate can serve as a preliminary indicator the gRNA editing efficiency. Five to 10 single colonies were picked from each plate and lysed in 0.2% SDS. The *URA3* gene was amplified, and a small amount of PCR product was electrophoresed. The positive clones were selected to verify the editing efficiency by sequencing, and the gene-edited clones were used to perform deletion, insertion, and frameshift in the gRNA targeting and recognition regions. We selected one of the most efficient gRNAs for the next step.

Construction of homologous recombination templates of *Saccharomyces cerevisiae*

According to the preferred gRNA targeting position, primers were designed to extend ~150 bp upstream and downstream of the cleavage site. The upstream and downstream homologous arms were amplified. Next, we performed electrophoresis detection and gel extraction of the PCR products. The upstream homologous arm and the downstream homologous arm were, respectively, ligated into the *Eco*RV and *Not*I restriction sites of pBlueScript vector. After identification, the recombinant universal vector(pSK-*URA3*-up-down) with yeast homology arm was obtained. In this study, we selected the 19/20/21 exon of the *EGFR* gene and the 12/13 codon common mutation hotspot of exon 2 of the *KRAS* gene to establish and verify a novel gene detection standard material system. The information related to the mutation site is shown in Table S2. The fragments with the wild-type sequence/mutation site were amplified or obtained by overlap extension PCR. The sequence length is 0.4–1.5 kb, which can be ligated into any restriction site between the upstream and downstream homologous arms of pSK-*URA3*-up-down. The successfully constructed plasmid can be double digested by selecting a suitable restriction site on the outer side of the upstream and downstream homologous arms, and the purified fragment can be a homologous recombination template for co-transformation with the Cas9 plasmid. Using NEB restriction enzymes and enzyme digestion buffer, DNA was purified using the AxyPrep-96 DNA Gel Extraction Kit.

A simplified homologous recombination template generation method is as follows: according to the preferred gRNA targeting position, the sequence extending ~30 bp upstream and downstream of the cleavage site was added to the 5′ end of the knock-in sequence amplification primer, and the homologous recombination template with short arm can be obtained by PCR amplification, and the PCR product can be purified as homologous recombination template and Cas9 plasmid co-transformed *Saccharomyces cerevisiae*.

Recombinant yeast RMs stability verification

In experiments to verify the stability of recombinant yeast genome passages, the recombinant yeast was transferred to 50 mL of YPD medium in a ratio of 0.2‰ after resuscitation, transferred to the same ratio after 20 h as the first 10 generations, and so on. We collected and extracted genomic DNA from 10th, 20th, 30th, 40th, 50th, and 60th generation yeast cells, and 5 ng of yeast genomic DNA served as the qPCR template.

In experiments to verify the long-term storage stability of recombinant yeast at −80°C in 20% glycerol for 180 days and 270 days, the yeast cells were resuscitated on YPD plates. The colonies were picked after resuscitation and cultured overnight in test tubes. The yeast solution was collected for yeast genomic DNA extraction and calibrated to a concentration of 5–10 ng/μL. qPCR was performed using KRAS-Q1 and ACT1 primers. Three replicates were taken for each sample, and the average Ct value was taken. Differences were made between the average Ct values of the two pairs of primers to obtain a ΔCt value.

In experiments to verify the short-term storage stability of recombinant yeast in glycerol under different temperature conditions, we obtained a small amount of colonies suspended in a 20% glycerol cryotube. The samples were stored at −20°C, 4°C, and room temperature for 20 days. Next, the yeast liquid was coated on the plate, and the colonies that grew were subjected to genomic DNA extraction; the

concentration was adjusted to 5–10 ng/μL. qPCR was performed using KRAS-Q1 and ACT1 primers. Three replicates were taken for each sample, and the average Ct value was obtained. Differences were made between the average Ct values of the two pairs of primers to obtain a ΔCt value.

In experiments to verify the short-term storage stability of recombinant yeast protoplasts, protoplasts were prepared from overnight cultured recombinant yeasts that were stored at 4°C or room temperature for 20 days. qPCR was then used to detect Ct values amplified by KRAS-Q1 and ACT1 primers, by comparing the change in the ΔCt between the two pairs of primers.

Optimization of protoplast formation conditions in yeast cells

For the *Saccharomyces cerevisiae* W303-1A strain, lyticase was selected to hydrolyze the cell wall and hypertonic phosphate buffer (PB) solution was used as the osmotic pressure stabilizer. The mycelium growth state in the middle and late stages of logarithmic growth is consistent, metabolism is strong, and the enzyme has strong sensitivity. The yeast solution was cultured overnight for 13 h, and the measured value of OD600 was 1.37. We took 1 mL of yeast solution as an experimental group, and set the concentration of lyticase (10 U/μL) in hypertonic buffer to five gradients: 0.01, 0.02, 0.04, 0.08, and 0.1 U/mL, with enzymatic hydrolysis conducted for 60 min. Selection of the 0.08 U/mL enzyme concentration was determined via enzymatic hydrolysis for the following five set times: 30, 40, 50, 60, and 90 min. After treatment, the protoplast suspension was treated with pure water under an optical microscope, and the protoplast formation rate was estimated according to the water absorption and rupture of the protoplasts.

Exploration of universal buffer formulation of yeast protoplast and human cells

To obtain a humanized yeast-human cell mixed RM, we simulated the state of cellular genetic heterogeneity in tumor tissues through mixing the recombinant yeast cell RMs with human embryonic kidney 293T cells in a specific ratio. Therefore, it is vital to develop a buffer solution formulation common to yeast protoplasts and human cells for establishing a humanized yeast-human cell mixed RM. Recombinant yeast protoplasts and human embryonic kidney 293T cells that proliferated *in vitro* were obtained, and the two cell types were separately placed in the following 10 commonly used cell fixatives and buffers: 0.8 M hypertonic PB solution, 1× PBS, 10× PBS, 4% paraformaldehyde, 4% neutral formaldehyde, 5% glacial acetic acid, ethanol/glacial acetic acid mixture (3:1), methanol/glacial acetic acid mixture (3:1), 95% ethanol, and 70% ethanol. We observed the cell morphology under a light microscope (Table 4).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtm.2021.01.004.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (81630087) and Shanghai Science and Technology Commission (17JC1400900). This research was also supported by Chongqing Science and Technology Commission (cstc2020jcyj-zdxm0180 and cstc2019jxjl130001). We thank the members of our own laboratories for helpful experimental instruction.

AUTHOR CONTRIBUTIONS

D.L., X.H., and J.D. are responsible for the conception of the work. X.H., J.D., and Z.X. participated in the acquisition and analysis of the experimental data. X.H. and Z.X. contributed to the preparation of the manuscript. All authors discussed and interpreted the results.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Hawkins, J.R., Hawkins, M., Boyle, J., Gray, E., Matejtschuk, P., and Metcalfe, P. (2010). Genetic reference materials and their application to haematology. Biologicals *38*, 467–473.

2. Scharf, H., and Bremser, W. (2015). Certification of a reference material for determination of total cyanide in soil to support implementation of the International Standard ISO 11262:2011. Anal. Bioanal. Chem. *407*, 3219–3223.

3. Siekmann, L. (2007). Requirements for reference (calibration) laboratories in laboratory medicine. Clin. Biochem. Rev. *28*, 149–154.

4. Grimalt, S., Harbeck, S., Shegunova, P., Seghers, J., Sejerøe-Olsen, B., Emteborg, H., and Dabrio, M. (2015). Development of a new cucumber reference material for pesticide residue analysis: feasibility study for material processing, homogeneity and stability assessment. Anal. Bioanal. Chem. *407*, 3083–3091.

5. Levin, B.C., Richie, K.L., and Jakupciak, J.P. (2006). Advances in Huntington's disease diagnostics: development of a standard reference material. Expert Rev. Mol. Diagn. *6*, 587–596.

6. Corbisier, P., Pinheiro, L., Mazoua, S., Kortekaas, A.M., Chung, P.Y., Gerganova, T., Roebben, G., Emons, H., and Emslie, K. (2015). DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials. Anal. Bioanal. Chem. *407*, 1831–1840.

7. Kalman, L.V., Datta, V., Williams, M., Zook, J.M., Salit, M.L., and Han, J.Y. (2016). Development and Characterization of Reference Materials for Genetic Testing: Focus on Public Partnerships. Ann. Lab. Med. *36*, 513–520.

8. Chen, B., O'Connell, C.D., Boone, D.J., Amos, J.A., Beck, J.C., Chan, M.M., Farkas, D.H., Lebo, R.V., Richards, C.S., Roa, B.B., et al. (2005). Developing a sustainable process to provide quality control materials for genetic testing. Genet. Med. *7*, 534–549.

9. Hawkins, M., Boyle, J., Wright, K.E., Elles, R., Ramsden, S.C., O'Grady, A., Sweeney, M., Barton, D.E., Burgess, T., Moore, M., et al. (2011). Preparation and validation of the first WHO international genetic reference panel for Fragile X syndrome. Eur. J. Hum. Genet. *19*, 10–17.

10. Boyle, J., Hawkins, M., Barton, D.E., Meaney, K., Guitart, M., O'Grady, A., Tobi, S., Ramsden, S.C., Elles, R., Gray, E., et al. (2011). Establishment of the first WHO international genetic reference panel for Prader Willi and Angelman syndromes. Eur. J. Hum. Genet. *19*, 857–864.

11. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51–59.

12. Beranek, M., Drastikova, M., Dulicek, P., and Palicka, V. (2014). Preparing compound heterozygous reference material using gene synthesis technology: a model of thrombophilic mutations. Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech Repub. *158*, 539–543.

13. Pratt, V.M., Everts, R.E., Aggarwal, P., Beyer, B.N., Broeckel, U., Epstein-Baak, R., Hujsak, P., Kornreich, R., Liao, J., Lorier, R., et al. (2016). Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. J. Mol. Diagn. *18*, 109–123.

14. Baoutina, A., Bhat, S., Zheng, M., Partis, L., Dobeson, M., Alexander, I.E., and Emslie, K.R. (2016). Synthetic certified DNA reference material for analysis of human erythropoietin transgene and transcript in gene doping and gene therapy. Gene Ther. *23*, 708–717.

15. Jessop-Fabre, M.M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M.K., Keasling, J.D., and Borodina, I. (2016). EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. Biotechnol. J. *11*, 1110–1117.

16. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

17. Inui, M., Miyado, M., Igarashi, M., Tamano, M., Kubo, A., Yamashita, S., Asahara, H., Fukami, M., and Takada, S. (2014). Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. Sci. Rep. *4*, 5396.

18. Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

19. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

20. Schiestl, R.H., and Petes, T.D. (1991). Integration of DNA fragments by illegitimate recombination in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *88*, 7585–7589.

21. Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J., and Borodina, I. (2014). EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. FEMS Yeast Res. *14*, 238–248.

22. Paez, J.G., Jänne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497–1500.

23. Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA Cancer J. Clin. *65*, 87–108.

24. Soh, J., Toyooka, S., Matsuo, K., Yamamoto, H., Wistuba, I.I., Lam, S., Fong, K.M., Gazdar, A.F., and Miyoshi, S. (2015). Ethnicity affects *EGFR* and *KRAS* gene alterations of lung adenocarcinoma. Oncol. Lett. *10*, 1775–1782.

25. Wang, L., Hu, H., Pan, Y., Wang, R., Li, Y., Shen, L., Yu, Y., Li, H., Cai, D., Sun, Y., and Chen, H. (2014). PIK3CA mutations frequently coexist with EGFR/KRAS mutations in non-small cell lung cancer and suggest poor prognosis in EGFR/KRAS wildtype subgroup. PLoS ONE *9*, e88291.

26. Mitsudomi, T., and Yatabe, Y. (2010). Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer. FEBS J. *277*, 301–308.

27. Pao, W., Wang, T.Y., Riely, G.J., Miller, V.A., Pan, Q., Ladanyi, M., Zakowski, M.F., Heelan, R.T., Kris, M.G., and Varmus, H.E. (2005). KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS Med. *2*, e17.

28. Do, H., Krypuy, M., Mitchell, P.L., Fox, S.B., and Dobrovic, A. (2008). High resolution melting analysis for rapid and sensitive EGFR and KRAS mutation detection in formalin fixed paraffin embedded biopsies. BMC Cancer *8*, 142.

29. Schwarzenbach, H., Hoon, D.S., and Pantel, K. (2011). Cell-free nucleic acids as biomarkers in cancer patients. Nat. Rev. Cancer *11*, 426–437.

30. Heitzer, E., Ulz, P., and Geigl, J.B. (2015). Circulating tumor DNA as a liquid biopsy for cancer. Clin. Chem. *61*, 112–123.

31. Diaz, L.A., Jr., and Bardelli, A. (2014). Liquid biopsies: genotyping circulating tumor DNA. J. Clin. Oncol. *32*, 579–586.

32. Fleischhacker, M., and Schmidt, B. (2007). Circulating nucleic acids (CNAs) and cancer—A survey. Biochim. Biophys. Acta *1775*, 181–232.

33. Kiełbasa, A., Gadzała-Kopciuch, R., and Buszewski, B. (2016). Reference Materials: Significance, General Requirements, and Demand. Crit. Rev. Anal. Chem. *46*, 224–235.

34. Zhang, L., Wu, Y., Wu, G., Cao, Y., and Lu, C. (2014). Correction of the lack of commutability between plasmid DNA and genomic DNA for quantification of genetically modified organisms using pBSTopas as a model. Anal. Bioanal. Chem. *406*, 6385–6397.

35. DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343.

36. Singh, R.R., Patel, K.P., Routbort, M.J., Reddy, N.G., Barkoh, B.A., Handal, B., Kanagal-Shamanna, R., Greaves, W.O., Medeiros, L.J., Aldape, K.D., and Luthra, R. (2013). Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes. J. Mol. Diagn. *15*, 607–622.

37. Foote, S., and Denny, C. (2002). Construction of YAC libraries with large inserts. Curr. Protoc. Hum. Genet. *Chapter 5*, Unit5.2.

38. Zhang, R., Peng, R., Li, Z., Gao, P., Jia, S., Yang, X., Ding, J., Han, Y., Xie, J., and Li, J. (2017). Synthetic Circulating Cell-free DNA as Quality Control Materials for Somatic Mutation Detection in Liquid Biopsy for Cancer. Clin. Chem. *63*, 1465–1475.

39. Drury, S., Hill, M., and Chitty, L.S. (2016). Cell-Free Fetal DNA Testing for Prenatal Diagnosis. Adv. Clin. Chem. *76*, 1–35.