Studies on the adaptational strategies to the heat stress in *Saccharomyces cerevisiae* and the reconstruction of thermotolerance

Atsushi SATOMURA

2017

Contents

Abbreviations

ABC	ATP binding cassette
ATP	Adenosine triphosphate
cAMP	Cyclic adenosine monophosphate
Cas	CRISPR associated protein
CRISPR	Clustered regulatory interspaced short palindromic repeat
DSB	Double strand brake
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
gRNA	Guide RNA
GTP	Guanosine triphosphate
HDR	Homology directed repair
HSE	Heat shock element
HSP	Heat shock protein
NHEJ	Non-homologous end joining
OD_{600}	Optical density at 600 nm
PAM	Protospacer-adjacent motif
PKA	Protein kinase A
SD	Synthetic dextrose
TALEN	Transcription activator like effector nuclease
YPD	Yeast extract- peptone-dextrose
ZFN	Zinc finger nuclease

Introduction

Microbial fermentation has attracted attention as an eco-friendly strategy for the production of useful materials. In bioproduction, multiple reactions can occur at lower pressures and temperatures that use much less energy than chemosynthetic production. However, the fermentation condition is limited to environments where cells are able to survive and grow, as bioproduction is usually associated with cellular growth. For instance, heat produced during fermentation diminishes or abolishes the cellular growth and fermentation in bioproduction. It is, however, not cost-efficient to cool down the fermentation apparatus by using a lot of water resources and cooling energy. Therefore, thermotolerance is essential for cost-efficient bioproduction, both to save energy consumption and costs associated with the cooling apparatus. In addition, fermentation at high temperatures potentially diminishes contamination risks and costs to distill the product, leading to more stable and inexpensive bioproduction.

Stress tolerance of *Saccharomyces cerevisiae*

In bioindustrial processes, organisms are constantly subjected to stresses, such as high or low pH, organic solvents, and high temperatures [1]. In the case of bioethanol production, fermentation heat and organic acids, such as acetic acid produced during pretreatment of the biomass, can reduce fermentation and survival rates. To save the costs to relieve those stresses, stress tolerant organisms have been in demand. To this end, stress-tolerant organisms, (e.g., thermophilic *Kluyveromyces* species for ethanol production) have been studied for bioproduction [2, 3]. Despite these efforts, the budding yeast *Saccharomyces cerevisiae* continues to be primarily used in bioproduction because it has several advantages over other strains [4]. First, *S. cerevisiae* grows quickly in inexpensive media. Second, it is able to be grown both in aerobic and anaerobic conditions. Third, since the genetic manipulation for *S. cerevisiae* has been well developed, genetic engineering and molecular analysis are much easier than in other species.

To confer stress tolerance to *S. cerevisiae*, several approaches have been utilized, such as the modification of cell walls. Since cell walls are in direct contact with the external environment, artificial cell wall modification by cell surface engineering [5, 6] has successfully improved stress tolerances against acidic stress and organic solvents [7-9]. These cell wall modifications play a buffering role between the external environment and internal cytoplasm.

However, cell wall modification cannot be used to treat with heat stress because heat directly affects cellular proteins and membrane fluidity. Although there are various advantages to thermotolerant yeast, effective strategies to enhance thermotolerance have not been demonstrated so far. To isolate thermotolerant *S. cerevisiae*, several breeding systems, such as adaptation [10], genome shuffling [11], protoplast fusion [12], and mutagenesis techniques [13] have been utilized. However, the genetic characteristics for thermotolerance have not been uncovered and the reconstruction of thermotolerance on an ordinary yeast strain has not been carried out.

Genomic mutations that confer the yeast cells with thermotolerance will be required for cost-effective bioproduction. In addition, analysis of the signaling pathways of reconstructed thermotolerant yeasts will expand the knowledge in this area.

Signaling pathway for thermotolerance

Hsf1p, a transcription factor related to thermotolerance in *S. cerevisiae*, regulates the expression of more than 5% of genes involved in protein folding, energy generation, carbohydrate metabolism, and cell wall organization [14, 15]. Hsf1p binds in both constitutive and inducible manners to a heat shock element (HSE) in the promoter of target genes [16]. Hsf1p only poorly activates transcription of HSP-encoding genes at low temperatures, but strongly activates transcription in response to a shift to high temperatures [16, 17]. Although Hsf1p regulates the expressions of several stress-responsive genes, the genetically modified Hsp1p, which hyperactivates the downstream genes, failed to induce thermotolerance sufficiently [18]. This indicates that these Hsf1p mutants did not fully activate all genes essential for thermotolerance and there should be an additional signaling pathway to induce those genes. A recent study demonstrated that genes involved in metabolism, signal transduction, and chromatin remodeling, as well as HSP-encoding genes, played a crucial role in thermotolerance [19].

In addition to Hsf1p, thermotolerance in yeast is reported to be controlled by the transcription activators, Msn2p and Msn4p [20]. They activate more than 200 genes, including genes encoding heat shock proteins (HSPs) (*HSP12*, *HSP104*, and *SSA4*) and trehalose synthesis genes (*TPS1*, *TPS2*, and *TSL1*) [20-22]. Msn2p and Msn4p are regulated by a cAMP-dependent protein kinase (PKA) signaling pathway. The cAMP/PKA signaling pathway

is a well-conserved pathway, operated by intracellular cAMP as a second messenger [23]. This pathway controls a variety of processes including cell-cycle progression [24], life span [25], diauxic shift transition [26], and stress response [27]. Intracellular cAMP level is regulated by adenylate cyclase (Cyr1p), which converts ATP to cAMP [28]. The monomeric G proteins (Ras1p and Ras2) control the activity of Cyr1p depending on the activity of Cdc25p [29]. Cdc25p is a membrane-bound guanine nucleotide exchange factor (GEF) that activates Ras1p and Ras2p by stimulating the release of GDP and the binding of GTP (Fig. 1) [30]. Increasing the intracellular level of cAMP is known to antagonize the activities of Msn2p and Msn4p through the inhibition of Bcy1p and the activation of cAMP-dependent protein kinase (PKA) composed of Tpk1p, Tpk2p, and Tpk3p [27]. The kinase is responsible for the inactivation of various transcriptional regulators including Msn2p and Msn4p [31]. The lack of the cAMP/PKA pathway by disruptions of *CDC25* or *RAS2* has been reported to enhance thermotolerance. However, the excessive downregulation of this pathway leads to severe growth defects, because the cell cycle is under the control of this pathway [32]. The fine tuning of this pathway, which enhances thermotolerance but does not inhibit growth rates, will be important in constructing thermotolerant strains for bioproduction.

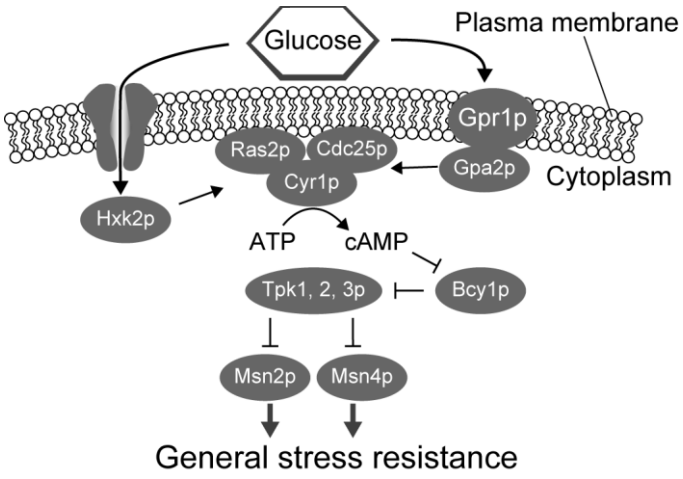

Fig. 1 cAMP/PKA signaling pathway Hxk2p, Hexokinase2; Ras2p, Ras family GTPase; Cdc25p, Ras family guanine nucleotide exchange factor; Cyr1p, Adenylate cyclase; Gpr1p, Plasma membrane G protein coupled receptor; Gpa2p, Guanine nucleotide-binding protein subunit alpha; Bcy1p, cAMP-dependent protein kinase; Tpk1, 2, 3p, Thiamin pyrophosphokinase 1, 2, 3; Msn2p and Msn4p; Stress-responsive transcriptional activator.

Engineering of master regulator

Transcription factors that function at the top of a regulatory hierarchy are called “master regulators” [33]. Introductions of mutations into master regulators are able to comprehensively alter cellular phenotypes and enhance cellular stress tolerance by directly and indirectly modifying the expression of a wide variety of genes. For example, mutations in Pdr1p and Pdr3p, master regulators of pleiotropic drug resistance [34], have been reported to induce organic solvent tolerance through the comprehensive upregulation of downstream genes, such as those encoding ATP-binding cassette (ABC) transporters [35, 36]. Engineering master regulators is a very easy and simple technique to confer stress tolerance, such as organic solvent tolerance, because it only requires the introduction of one or a few mutations in regulatory elements. Since thermotolerance is regulated by various kinds of proteins, such as those involved in metabolism, signal transduction, and heat shock response [19], overexpression of only some genes will not be enough to induce thermotolerance. The regulation of a master regulator, which can comprehensively regulate genes related to thermotolerance, is considered to be an ideal approach to enhance it.

There have been, however, no reports of point mutations involved in thermotolerance. In addition, techniques to introduce mutations into *S. cerevisiae* genome are inefficient and time-consuming. Even if mutations related to thermotolerance are discovered, the lack of fast and efficient techniques will become a barrier to the construction of thermotolerant yeast.

Genome editing

To introduce mutations into the *S. cerevisiae* genome, a vector that includes a marker removing system has been used so far. The pAUR135 vector is commercially available (Takara Bio, Shiga, Japan), and is the most common vector system to edit the yeast genome (Fig. 2). This vector contains *AUR1-C* gene that confers aureobasidin tolerance with yeast cells and *GIN11M86* gene under the control of *GAL10* promoter [37]. Integration of vector backbone including mutant sequence is achieved using aureobasidin selection. When transferring those colonies onto galactose medium, only pAUR135-vector-removed clones as a result of homologous recombination are able to grow due to the expression of lethal *GIN11M86* [38]. The recombinant strains are free of vector sequences and marker genes. It takes approximately two weeks to construct mutants with this technique and the efficiencies to obtain the mutants are

theoretically less than 50%. Although the gene manipulation techniques have been well developed for yeasts, the genome-design tools remained low efficiency and take relatively long time. For the reconstruction of thermotolerant yeasts by point mutations in a regulatory element, not only the exploration of mutations involved in thermotolerance but also the development of a fast, efficient, and versatile tool for genome design is necessary.

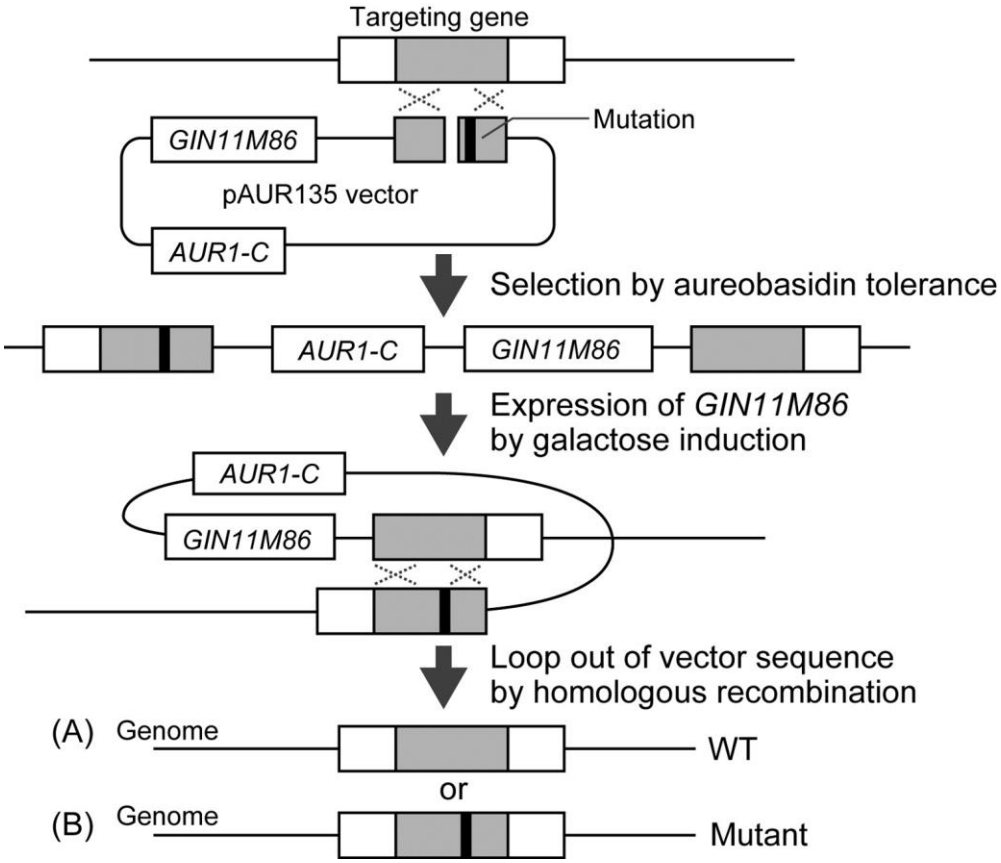

Fig. 2 pAUR135 system for the design of *S. cerevisiae* genome *AUR1-C*, Aureobasidin A resistance gene; *GIN11M86*, modified version of a gene encoding a galactose-inducible growth inhibitory protein.

In the past decades, new technologies for genome design have enabled the introduction of mutations into the genomes of a range of organisms including *S. cerevisiae*. This technology, which is referred to as "genome editing," is based on target-specific nucleases that cause a double strand DNA break (DSB) in a specific site [39]. DSBs can induce two independent repairing pathways, non-homologous end jointing (NHEJ) and homology directed repair (HDR)

[40, 41] (Fig 3). NHEJ is an error-prone repairing pathway, which leaves an insertion or deletion at the cleavage site. HDR is a high-fidelity repairing pathway depending on donor templates. Thus, genome cleavage can stimulate gene knock-out through NHEJ and flexible gene designs through HDR at the cleaved site.

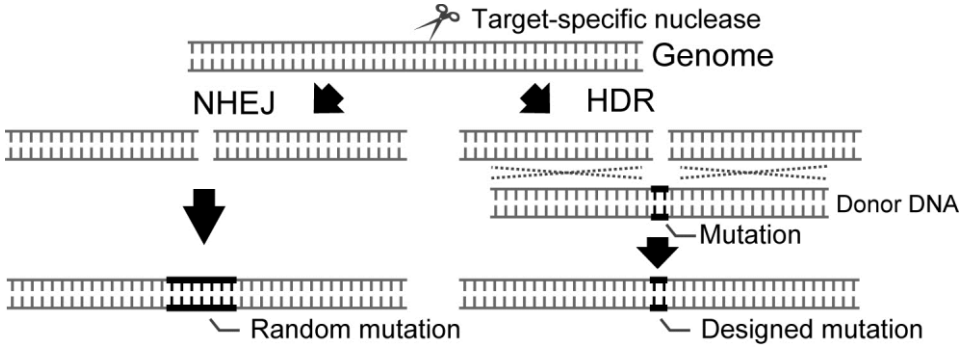

Fig. 3 DSB-repairing pathways NHEJ, non-homologous end jointing; HDR, homology directed repair.

1. ZFN

A flexible genome editing technology was first established in 2003 by Porteus and Baltimore [42]. The new technology, called zing finger nuclease (ZFN), couples a zinc finger DNA-binding domain with a non-specific FokI nuclease. Each zinc finer module recognizes a unique three base pair DNA. Thus, the combination of each zinc finger module generates a target specific DNA binding domain. The recognition of the DNA binding domains enables the FokI nucleases to cleave a specific DNA site. Although ZFN can theoretically cleave any specific DNA sequences by modulating the zinc finger units, ZFN has not been well used thus far. This is because the specificities of individual zinc fingers depend on the contexts of the adjacent zinc finger modules and optimization of the zinc figure modules is often required to target specific sites [43].

2. TALEN

In 2009, transcription activator-like effector nuclease (TALEN) emerged as an alternative technology to ZFN [44, 45]. TALEN is composed of customizable DNA binding domains derived from *Xanthomonas* bacteria fused to a non-specific FokI nuclease. DNA recognition is mediated by the conserved 33–35 amino acid sequences of each DNA binding module [44, 45].

The target base of each unit is determined by the divergent amino acids at the 12th and 13th amino acids in the modules. Combining each protein unit corresponding to the target sequences allows delivery of FokI nucleases to the specific site, resulting in the target cleavage. Unlike ZFN, the DNA recognition of TALEN is not dependent on the contexts of surrounding DNA binding motifs; thus, TALEN recognizes target sites more robustly than ZFN. However, both ZFN and TALEN need a correct alignment of each protein module, which requires time-consuming DNA manipulations.

3. CRISPR/Cas9 system

A novel genome editing technology emerged three years after the development of TALEN. The technology uses an adaptive immune system in prokaryotic cells called a clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system [46]. The CRISPR/Cas system cleaves foreign DNA sequences in bacteria, such as genetic elements of phages and alien plasmids [47, 48]. The most established CRIPSR system derived from *Streptococcus pyogenes* consists of the following three components: Cas9 nuclease, CRISPR RNA (crRNA), and trans-activating crRNA (tracrRNA) [49]. The artificial tethering of the 3' end of crRNA to the 5' end of tracrRNA can also generate an active RNA form (gRNA) [46]. This system recognizes the target DNA sequence by the function of a gRNA, which binds to the complementary 20 nucleotide sequence immediately followed by a protospacer adjacent motif (PAM), typically "NGG." In contrast to ZFN and TALEN, the target recognition is carried out by RNA; therefore, laborious efforts to align the recognition modules are unnecessary in this system, which makes it much more simple and versatile to design the genome. Because of this advantage, the CRISPR/Cas9 system has been applied to various species including humans [50], zebrafish [51], *Caenorhabditis elegans* [52], *Escherichia coli* [53], and *S. cerevisiae* [54].

Despite the high efficiency and robustness of the CRISPR/Cas9 system, there is a major limitation. When mutations are introduced into the outside areas of PAM and gRNA-targeting sequences by HDR, these sequences remain intact after editing, causing re-cleavage by the CRISPR/Cas9 system (Fig. 4). The re-cleavage will continuously induce HDR until the PAM or gRNA-targeting sequence is disrupted by error-prone NHEJ. Therefore, the CRISPR-Cas9 system is unable to precisely edit the outside areas of PAM and gRNA-targeting sequences by HDR, while it can precisely edit the inside of these sequences with nearly 100% efficiencies in

yeast [54].

Genome-wide base editing system that can be applied to all regions of the genome is necessary to construct the desired yeast strains by designing master regulators.

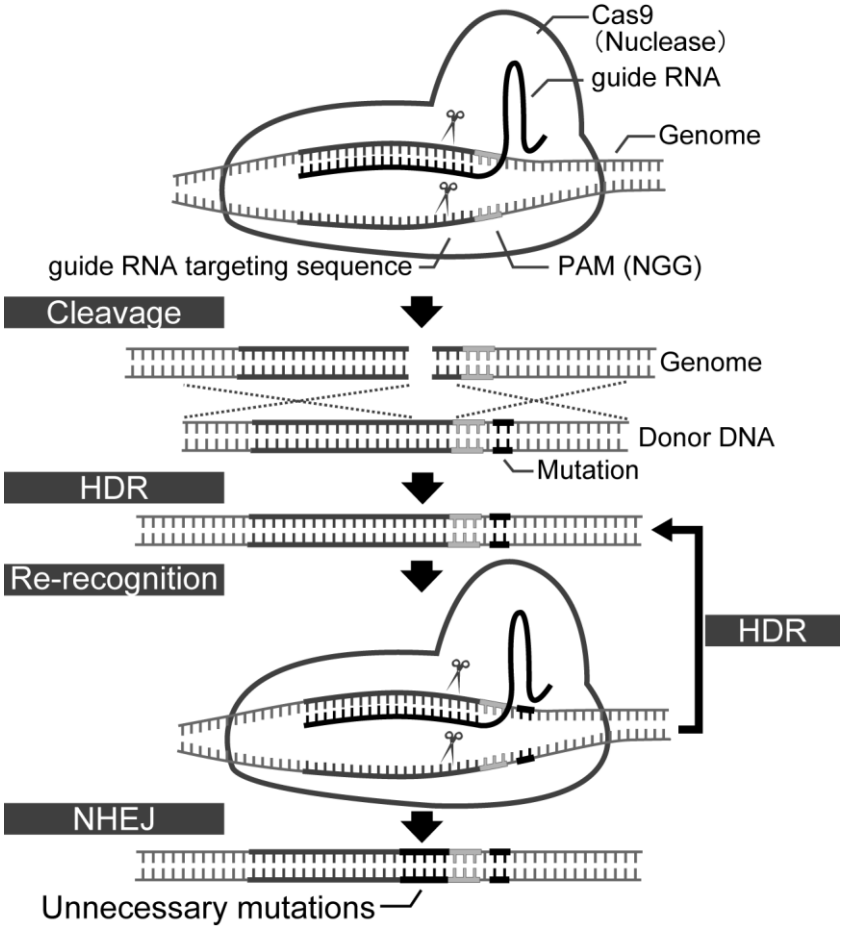

Fig. 4 Limitation of CRISPR/Cas9 system

To identify mutations involved in thermotolerance, we started from a breeding of thermotolerant *S. cerevisiae*. A comparative genome analysis between bred thermotolerant strains and the parental strain allowed us to reveal mutations involved in thermotolerance. These mutations successfully enhanced thermotolerance without impairing growth. However, the conventional technique to introduce mutations into the yeast genome is time-consuming (approximately two weeks) and inefficient. To address these problems, I developed a new genome editing tool that can genome-widely introduce mutations within only five days and with approximately 50% efficiency.

References

1. Ruis H, Schüller C. Stress signaling in yeast. Bioessays. 1995; 17: 959-965.
2. Banat I, Singh D, Marchant R. The use of a thermotolerant fermentative *Kluyveromyces marxianus* IMB3 yeast strain for ethanol production. Acta Biotechnol. 2004; 16: 215-223.
3. Singh D, Banat I, Nigam P, Marchant R. Industrial scale ethanol production using the thermotolerant yeast *Kluyveromyces marxianus* IMB3 in an Indian distillery. Biotechnol Lett. 1998; 20: 753-755.
4. Satomura A, Kuroda K, Ueda M. Environmental stress tolerance engineering by modification of cell surface and transcription factor in *Saccharomyces cerevisiae*. Curr Envron Eng. 2014; 1: 149-156.
5. Ueda M, Tanaka A. Genetic immobilization of proteins on the yeast cell surface. Biotechnol Adv. 2000; 18: 121-140.
6. Kuroda K, Ueda M. Arming technology in yeast-novel strategy for whole-cell biocatalyst and protein engineering. Biomolecules. 2013; 3: 632-650.
7. Matsui K, Kuroda K, Ueda M. Creation of a novel peptide endowing yeasts with acid tolerance using yeast cell-surface engineering. Appl Microbiol Biotechnol. 2009; 82: 105-113.
8. Zou W, Ueda M, Yamanaka H, Tanaka A. Construction of a combinatorial protein library displayed on yeast cell surface using DNA random priming method. J Biosci Bioeng. 2001; 92: 393-396.
9. Zou W, Ueda M, Tanaka A. Screening of a molecule endowing *Saccharomyces cerevisiae* with *n*-nonane-tolerance from a combinatorial random protein library. Appl Microbiol Biotechnol. 2002; 58: 806-812.
10. Balakumar S, Arasaratnam V, Balasubramaniam K. Isolation and improvement of a thermotolerant *Saccharomyces cerevisiae* strain. World J Microbiol Biotechnol. 2001; 17: 739-746.
11. Shi D, Wang C, Wang K. Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae*. J Ind Microbiol Biotechnol. 2009; 36: 139-147.
12. Kida K, Kume K, Morimura S, Sonoda Y. Repeated-batch fermentation process using a thermotolerant flocculating yeast constructed by protoplast fusion. J Ferment Bioeng. 1992; 74: 169-173.
13. Sridhar M, Kiran Sree N, Venkateswar Rao L. Effect of UV radiation on thermotolerance, ethanol tolerance and osmotolerance of *Saccharomyces cerevisiae* VS

$_1$ and VS_3 strains. Bioresour Technol. 2002; 83: 199-202.

14. Hahn J-S, Hu Z, Thiele DJ, Iyer VR. Genome-wide analysis of the biology of stress responses through heat shock transcription factor. Mol Cell Biol. 2004; 24: 5249-5256.
15. Eastmond DL, Nelson HC. Genome-wide analysis reveals new roles for the activation domains of the *Saccharomyces cerevisiae* heat shock transcription factor (Hsf1) during the transient heat shock response. J Biol Chem. 2006; 281: 32909-32921.
16. Giardina C, Lis JT. Dynamic protein-DNA architecture of a yeast heat shock promoter. Mol Cell Biol. 1995; 15: 2737-2744.
17. Chen J, Pederson D. A distal heat shock element promotes the rapid response to heat shock of the *HSP26* gene in the yeast *Saccharomyces cerevisiae*. J Biol Chem. 1993; 268: 7442-7448.
18. Hardy JA, Walsh ST, Nelson H. Role of an α-helical bulge in the yeast heat shock transcription factor. J Mol Biol. 2000; 295: 393-409.
19. Gibney PA, Lu C, Caudy AA, Hess DC, Botstein D. Yeast metabolic and signaling genes are required for heat-shock survival and have little overlap with the heat-induced genes. Proc Natl Acad Sci USA. 2013; 110: 4393-4402.
20. Martinez-Pastor M, Marchler G, Schüller C, Marchler-Bauer A, Ruis H, Estruch F. The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). EMBO J. 1996; 15: 2227-2235.
21. Boy-Marcotte E, Perrot M, Bussereau F, Boucherie H, Jacquet M. Msn2p and Msn4p control a large number of genes induced at the diauxic transition which are repressed by cyclic AMP in *Saccharomyces cerevisiae*. J Bacteriol. 1998; 180: 1044-1052.
22. Schmitt AP, McEntee K. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 1996; 93: 5777-5782.
23. Thevelein JM, De Winde JH. Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast *Saccharomyces cerevisiae*. Mol Microbiol. 1999; 33: 904-918.
24. Baroni MD, Monti P, Alberghina L. Repression of growth-regulated Gl cyclin expression by cyclic AMP in budding yeast. Nature. 1994; 371: 339-342.
25. Lin S, Ju, Defossez P, Antoine, Guarente L. Requirement of NAD and *SIR2* for life-span extension by calorie restriction in *Saccharomyces cerevisiae*. Science. 2000; 289: 2126-2128.
26. Boy-Marcotte E, Tadi D, Perrot M, Boucherie H, Jacquet M. High cAMP levels antagonize the reprogramming of gene expression that occurs at the diauxic shift in

Saccharomyces cerevisiae. Microbiology. 1996; 142: 459-467.

27. Görner W, Durchschlag E, Martinez-Pastor MT, Estruch F, Ammerer G, Hamilton B, et al. Nuclear localization of the C_2H_2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev. 1998; 12: 586-597.
28. Matsumoto K, Uno I, Oshima Y, Ishikawa T. Isolation and characterization of yeast mutants deficient in adenylate cyclase and cAMP-dependent protein kinase. Proc Natl Acad Sci USA. 1982; 79: 2355-2359.
29. Broek D, Samiy N, Fasano O, Fujiyama A, Tamanoi F, Northup J, et al. Differential activation of yeast adenylate cyclase by wild type and mutant *RAS* proteins. Cell. 1985; 41: 763-769.
30. Broek D, Toda T, Michaeli T, Levin L, Birchmeier C, Zoller M, et al. The *S. cerevisiae CDC25* gene product regulates the RAS/adenylate cyclase pathway. Cell. 1987; 48: 789-799.
31. Görner W, Durchschlag E, Wolf J, Brown EL, Ammerer G, Ruis H, et al. Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor. EMBO J. 2002; 21: 135-144.
32. Mitsuzawa H, Uno I, Oshima T, Ishikawa T. Isolation and characterization of temperature-sensitive mutations in the *RAS2* and *CYR1* genes of *Saccharomyces cerevisiae*. Genetics. 1989; 123: 739-748.
33. Chan SS-K, Kyba M. What is a master regulator? J Stem Cell Res Ther. 2013; 3.
34. Nishida N, Jing D, Kuroda K, Ueda M. Activation of signaling pathways related to cell wall integrity and multidrug resistance by organic solvent in *Saccharomyces cerevisiae*. Curr Genet. 2014; 60: 149-162.
35. Carvajal E, Van Den Hazel H, Cybularz-Kolaczkowska A, Balzi E, Goffeau A. Molecular and phenotypic characterization of yeast *PDR1* mutants that show hyperactive transcription of various ABC multidrug transporter genes. Mol Gen Genet. 1997; 256: 406-415.
36. Nourani A, Papajova D, Delahodde A, Jacq C, Subik J. Clustered amino acid substitutions in the yeast transcription regulator Pdr3p increase pleiotropic drug resistance and identify a new central regulatory domain. Mol Gen Genet. 1997; 256: 397-405.
37. Hashida-Okado T, Ogawa A, Endo M, Yasumoto R, Takesako K, Kato I. *AUR1*, a novel gene conferring aureobasidin resistance on *Saccharomyces cerevisiae*: a study of defective morphologies in Aur1p-depleted cells. Mol Gen Genet. 1996; 251: 236-244.
38. Kawahata M, Amari S, Nishizawa Y, Akada R. A positive selection for plasmid loss in *Saccharomyces cerevisiae* using galactose-inducible growth inhibitory sequences. Yeast.

1999; 15: 1-10.

39. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013; 31: 397-405.
40. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010; 79: 181-211.
41. Shibata A, Conrad S, Birraux J, Geuting V, Barton O, Ismail A, et al. Factors determining DNA double-strand break repair pathway choice in G2 phase. EMBO J. 2011; 30: 1079-1092.
42. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. Science. 2003; 300: 763.
43. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, et al. Rapid "Open-Source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell. 2008; 31: 294-301.
44. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009; 326: 1501.
45. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009; 326: 1509-1512.
46. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337: 816-821.
47. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide-sequence of the iap gene, responsible for alkaline-phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene-product. J Bacteriol. 1987; 169: 5429-5433.
48. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007; 315: 1709-1712.
49. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 2011; 471: 602-607.
50. Inui M, Miyado M, Igarashi M, Tamano M, Kubo A, Yamashita S, et al. Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. Sci Rep. 2014; 4.
51. Hwang WY, Fu YF, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013; 31: 227-229.
52. Friedland AE, Tzur YB, Esvelt KM, Colaiacovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. Nat Methods. 2013; 10:

741-743.

53. Jiang WY, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol. 2013; 31: 233-239.

54. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013; 41: 4336-4343.

Chapter I

Characterization of thermotolerant yeast isolated via stepwise adaption to heat stress

In order to identify mutations involved in thermotolerance in *S. cerevisiae*, I started from breeding of thermotolerant yeast. Since the genome sequences of yeast strains isolated during breeding and a parental strain would be very similar, comparative transcriptome and metabolome analyses between them were supposed to be informative to analyze the mechanisms associated with the acquisition of thermotolerance.

So far, several breeding approaches such as adaptation [1], genome shuffling [2], protoplast fusion [3], and mutagenesis techniques [4] have been used to isolate thermotolerant yeast. While these breeding methods have been applied to enhance the thermotolerance of yeasts, they are unable to simultaneously improve growth rate. In addition to thermotolerance, fast growth rate is desirable because elevated growth rates in *S. cerevisiae* are often associated with improved fermentation rate [5]. In this breeding, on the other hand, since cells were subjected to selective pressures of both heat and growth at all times, it simultaneously improved both thermotolerance and growth rate. To investigate the mechanisms of acquisition of thermotolerance, the thermotolerant strain was analyzed using DNA microarray and non-targeting metabolome analysis. These analyses showed that the strain acquired thermotolerance via the induction of genes encoding HSPs as well as by the accumulation of trehalose.

Materials and methods

Yeast strains and breeding method

S. cerevisiae MT8-1 (*MAT***a**, *ade*, *his3*, *leu2*, *trp1*, *ura3*) [6] was used as the parental strain for breeding. For breeding, MT8-1 was cultivated in YPD medium (1% yeast extract, 2% peptone, 2% glucose) at 32°C for 72 h, and then the optical density at 600 nm (OD_{600}) was measured. If the OD_{600} was lower than that of MT8-1 cultivated at 30°C for 72 h, the yeast cells were inoculated into fresh YPD medium for cultivation at 32°C for 72 h again. If the OD_{600} was

higher than that of MT8-1 cultivated at 30°C for 72 h, *S. cerevisiae* cells were regarded as adapted to 32°C. This step was repeated until the cells adapted to 32°C. Cells adapted to 32°C were cultivated in fresh YPD medium at 34°C for 72 h until the cells adapted to 34°C as the same way described above. Cells adapted to 34°C were then cultivated at 36°C and subsequently at 38°C until the cells adapted to the temperatures. The resultant strain after the breeding at 38°C was named YK60-1 (Fig. 1).

To determine cell growth, MT8-1 and YK60-1 were pre-cultivated in 5 mL of YPD medium at 30°C for 12 h and 8 h, respectively. Since MT8-1 cannot grow as fast as YK60-1, MT8-1 required a longer time (12 h) than YK60-1 (8 h) to reach a similar growth phase. Cells were harvested and inoculated into 10 mL of YPD to be adjusted to an OD_{600} of 0.1. MT8-1 and YK60-1 were cultivated at 30°C and 40°C at 170 rpm in a water bath shaker (Taitec, Tokyo, Japan), and cell growth was monitored by measuring the OD_{600} of the culture broth.

Spot assay

For spot assay, MT8-1 and YK60-1 were pre-cultivated in 5 mL of YPD medium at 30°C for 12 h and 8 h, respectively. Cells were inoculated into 10 mL of YPD to be adjusted to an OD_{600} of 0.1, and incubated at 30°C for 8 h. After cells were treated with or without mild heat shock (37°C, 45 min) in YPD medium, cells were subjected to sublethal heat shock (50°C, 30 min). Cells were harvested, and the OD_{600} was adjusted to 1.0. Cells were diluted 10-, 10^2-, 10^3-, and 10^4-fold with distilled water, and 5 μL of each sample was spotted on to a YPD plate. Photograph was taken after a 72-h incubation at 30°C.

RNA extraction and DNA microarray analysis

For RNA extraction, MT8-1 and YK60-1 were pre-cultivated at 30°C for 16 h. Cells were then inoculated into fresh YPD medium to be adjusted to an OD_{600} of 0.1 and cultivated at 30°C and 40°C for 8 h. Total RNA was extracted with Isogen-LS (Nippon Gene, Tokyo, Japan) according to the manufacturer's instructions. Labeling of samples for array analysis was performed using the GeneChip 3'IVT labeling assay (Affymetrix, Santa Clara, CA, USA) with 0.25 μg input RNA. Samples were hybridized to GeneChip Yeast Genome 2.0 microarrays according to Affymetrix protocols (n =1).

Data analysis

To detect differentially regulated genes between MT8-1 and YK60-1, I applied the limit fold change model [7]. Briefly, genes in the top 5% of the highest fold changes in each separated 200-gene group were selected to generate a heat map. For the heat map, fold changes were $\log_2$-transformated. The heat map was constructed using Gene Cluster 3.0 [8], which can perform hierarchical cluster analysis. Euclidean distance was used to measure similarities of transcriptional patterns among genes for the clustering. To visualize the clustering result from Gene Cluster 3.0, Java TreeView software [9] was used. To identify common biological functions among genes in each cluster, DAVID [10] (the Database for Annotation, Visualization and Integrated Discovery) was used. The *P* value threshold was set to 0.05. Transcriptional regulators were analyzed using YEASTRACT (http://www.yeastract.com), which can identify transcriptional regulators of given genes.

Sample preparation for GC-MS analysis

For metabolome analysis, MT8-1 and YK60-1 were pre-cultivated at 30°C for 16 h. Cells were inoculated into fresh YPD medium to be adjusted to an OD_{600} of 0.1, then cultivated at 30°C and 40°C for 8 h. Yeast cells were collected on a 0.45 μm membrane filter (Millipore, Bedford, MA, USA) and washed with ice-cold distilled water. The cells were freeze-dried for extraction of metabolites. Extraction and derivatization of metabolites were performed according to previous reports [11, 12].

Analysis of GC-MS data

In Ultra GC-Q/MS, the autosampler; AOC-20is series injector (Shimadzu, Kyoto, Japan), the gas chromatograph; GC-2010 Plus (Shimadzu), and the mass spectrometer; GCMS-QP2010 Ultra (Shimadzu) were used. The injection temperature was 230°C. The helium gas flow rate through the column was 1 mL/min. The column temperature was held at 80°C for 2 min isothermally and was then raised by 15°C/min to 330°C and maintained there for 6 min isothermally. The transfer line and ion source temperature were 250°C and 200°C, respectively. Ions were generated by a 70 kV electron impact (EI), and twenty scans per second were recorded over the mass range 85–500 m/z.

Identification and semi-quantification of metabolites

Peak detection and alignment were performed by MetAlign software (Wageningen UR, The Netherlands) [13]. Peak identification and prediction were performed with AIoutput2 (ver1.0.1) [14]. Principal component analysis (PCA) was carried out using the AIoutput2 software.

Results

Growth and thermotolerance of YK60-1

A thermotolerant yeast strain, YK-60-1, was bred from a *S. cerevisiae* strain, MT8-1, by stepwise adaptation to heat stress (Fig. 1). Successive cultures were performed at 32°C, 34°C, 36°C, and 38°C. Yeast cells had adapted to the each temperature after 9, 8, 21, 22 rounds of culture, respectively. YK60-1, which adapted to 38°C and grew faster than any other candidate strain was isolated. YK60-1 grew even at 40°C, at which temperature the parental strain MT8-1 could not grow at all (Fig. 2a). In addition, YK60-1 grew faster and better than MT8-1 at 30°C (Fig. 2b). YK60-1 also showed enhanced heat-shock tolerance compared to MT8-1 (Fig. 2c).

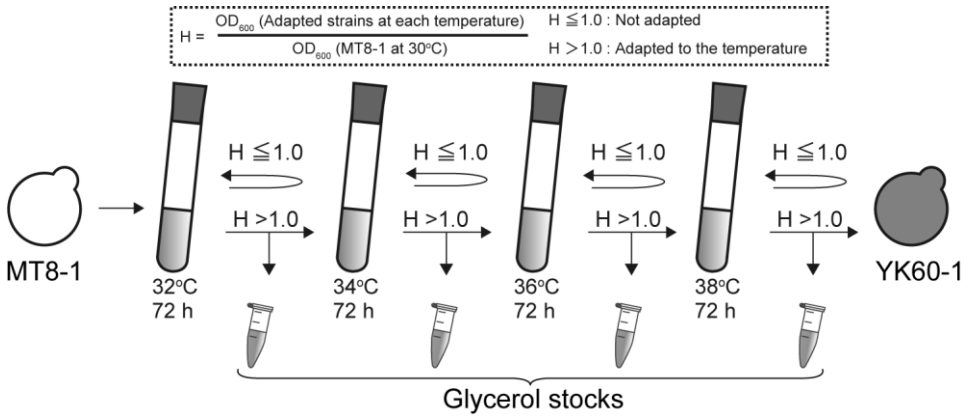

Fig. 1 Schematic of breeding of thermotolerant yeast MT8-1 was successively cultivated at 32°C for 72 h until the cells grew faster at 32°C than at 30°C, at which point they were regarded as adapted to 32°C. This stepwise adaptation method was performed at 34°C, 36°C, and 38°C in the same manner until the cells adapted to each temperature.

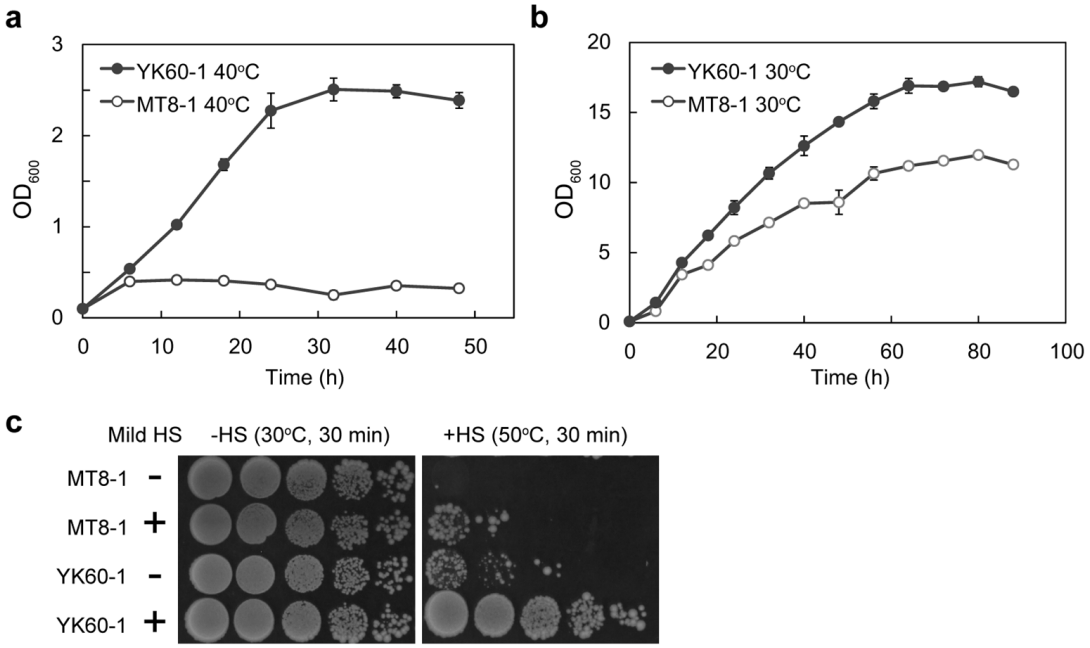

Fig. 2 Growth of YK60-1 Growth curves of MT8-1 and YK60-1 in the YPD medium at 40°C (a) and 30°C (b). (c) Growth of MT8-1 and YK60-1 on YPD plate at 30°C after heat shock. MT8-1 and YK60-1 were treated with or without mild heat shock (45 min incubation at 37°C) and grown on a YPD plate for 72 h at 30°C after a short sublethal heat shock (30 min incubation at 50°C). "+" indicates that cells were subjected to a mild heat treatment and "–" indicates that cells were incubated at 30°C instead of heat treatment. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements.

DNA microarray analysis

In order to investigate the mechanisms by which MT8-1 acquired thermotolerance, DNA microarray analyses of MT8-1 and YK60-1 incubated at 30°C and 40°C were performed. An arbitrary fold change cutoff was not applied to identify differentially regulated genes, since this method often overestimates fold-changes of genes with low transcript levels and underestimates regulation changes of genes with high transcript levels [7]. Instead, I applied the limit fold change model [7]. As displayed in the Venn diagram (Fig. 3a), compared with MT8-1 incubated at 30°C, the following differentially regulated genes were identified: 126 (98+15+5+8) upregulated genes and 151 (90+38+2+21) downregulated genes in MT8-1 incubated at 40°C; 162 (84+5+65+8) upregulated genes and 116 (88+2+5+21) downregulated genes in YK60-1

incubated at 30°C; and 141 (53+15+65+8) upregulated genes and 135 (71+38+5+21) downregulated genes in YK60-1 incubated at 40°C. I performed hierarchical clustering to classify the 643 differentially regulated genes based on similarities in their pattern of regulation across the different samples (Fig. 3b). Of the 643 genes, genes responding to environmental stimulus and heat (cluster A) were highly upregulated in YK60-1 incubated at 30°C and marginally upregulated in YK60-1 incubated at 40°C compared with MT8-1 incubated at 30°C and 40°C. These highly induced genes included *HSP42*, *HSP78*, and *SSA4*, which were reported to be highly induced in response to heat stress [15].

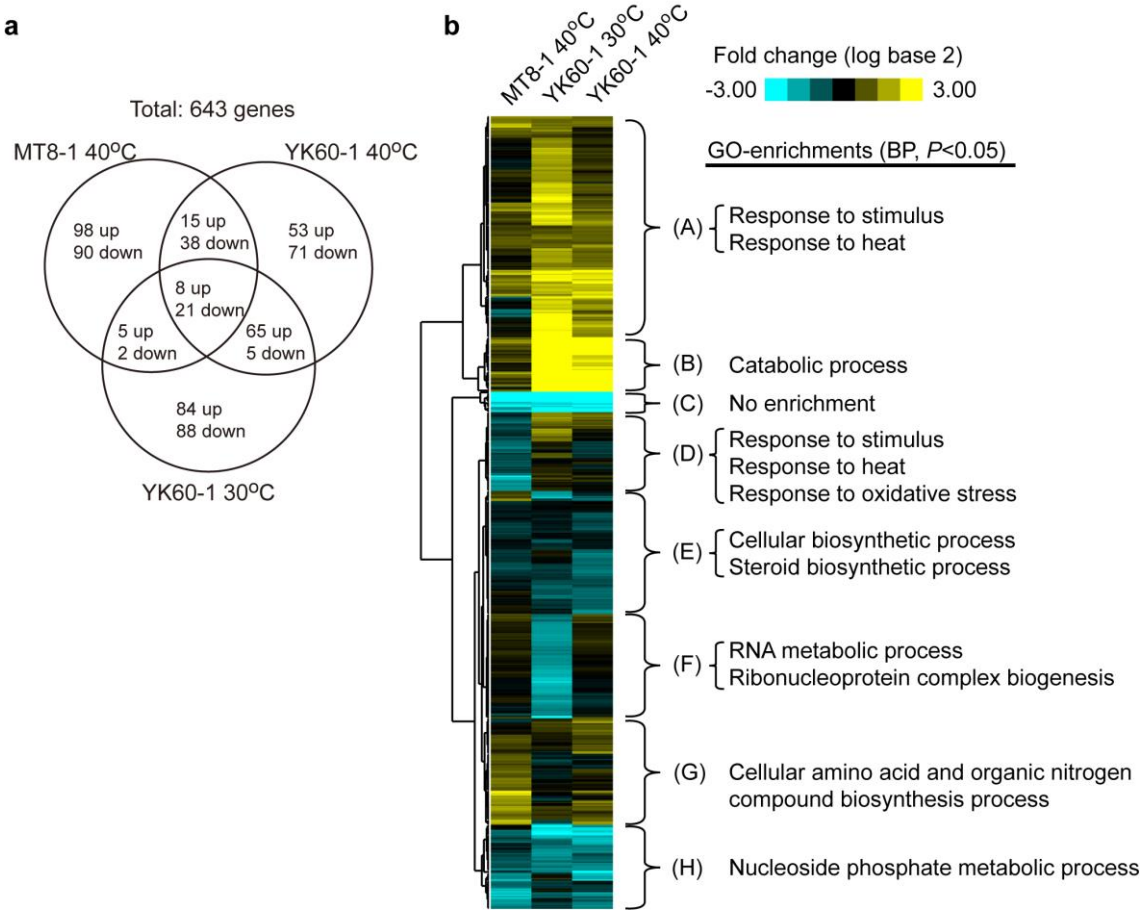

Fig. 3 Comparative analysis of mRNA profiles from DNA microarrays of MT8-1 and YK60-1 (a) Venn diagram showing the number of genes upregulated or downregulated relative to MT8-1 incubated at 30°C. (b) Hierarchical clustering of 643 genes. Upregulated and downregulated genes compared with MT8-1 incubated at 30°C were clustered. Significantly enriched ($P < 0.05$) biological processes (BP) based on GO (gene ontology) are indicated for all the 8 gene clusters obtained.

Cluster D represents genes downregulated in MT8-1 at 40°C and marginally upregulated in YK60-1 at 30°C and 40°C and it contains genes induced by stress, such as *PRX* [16, 17]. The transcriptome analyses showed that YK60-1 highly expressed stress-responsive genes even in a basal state at 30°C, which would be the thermotolerant mechanism of YK60-1.

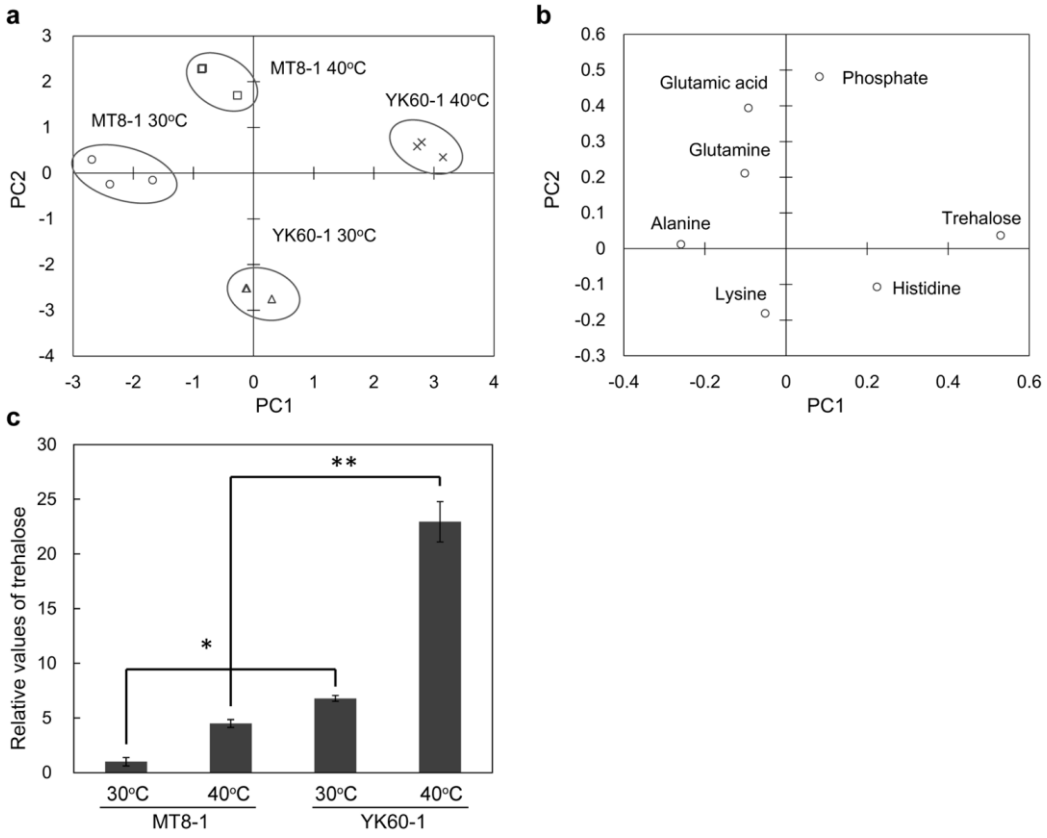

Figure 4 Metabolome analyses of MT8-1 and YK60-1 (a) Principal component analysis (PCA) score plot of MT8-1 incubated at 30°C and at 40°C; and YK60-1 incubated at 30°C and at 40°C. (b) Corresponding loading plots of principal component 1 (PC1) and PC2. Metabolites highly contributing to the group separation are shown. The cumulative contribution rate for PC1 to PC2 is 64%, indicating the loading plot presented by PC1 and PC2 express 64% of the information from original data. (c) Semi-quantitative analysis of trehalose. Relative trehalose value of MT8-1 incubated at 30°C and at 40°C, and YK60-1 incubated at 30°C and at 40°C are shown relative to the value of MT8-1 30°C as 1. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements. * $P < 0.05$, ** $P < 0.01$ were determined by one-way analysis of variance (ANOVA) followed by Tukey's test, which was used for multiple comparison.

Metabolome analysis

In order to further investigate the mechanisms of acquisition of thermotolerance in YK60-1, metabolome analysis was performed via non-targeted approach, in which all peak data from GC-MS chromatograms were used for multivariate analysis after alignment and normalization. Principal component analysis (PCA) was carried out to evaluate the reproducibility of the metabolome analysis. Clear separation among each strain and condition can be observed in the principal component analysis (PCA) score plot, which validated the quality and reproducibility of the analysis (Fig. 4A). The corresponding loading plot represented the characteristic metabolites in YK60-1 (Fig. 4B). The score and direction of a given metabolite in the loading plot indicate the strength and direction of its contribution to the group separation in the PCA score plot. Loading plot suggested that trehalose, a metabolite responsive for thermotolerance, was one of the specific metabolites in YK60-1 incubated at 40°C [18, 19]. Relative values of trehalose concentration were 4.5-fold, 6.8-fold, and 23-fold in MT8-1 at 40°C, YK60-1 at 30°C, and YK60-1 at 40°C, respectively, compared to the value of MT8-1 at 30°C as 1 (Fig. 4C). Indeed, the trehalose concentration in YK60-1 incubated at 40°C was 5.1-fold higher than that in MT8-1 incubated at 40°C (Fig. 4C). Furthermore, the basal concentration of trehalose in YK60-1 incubated at 30°C was 6.8-fold higher than in MT8-1 incubated at 30°C (Fig. 4C).

Discussion

A thermotolerant *S. cerevisiae* strain, YK60-1, was successfully bred via stepwise adaptation of a parental strain MT8-1 (Fig. 1). YK60-1 grew at a higher rate at 30°C and exhibited greater thermotolerance at 40°C than MT8-1 (Fig. 2a, b). DNA microarray analysis, showed that YK60-1 induced more stress-responsive genes than MT8-1 even in the absence of stresses (Fig. 3a, b). Many of them have been report to be regulated by Msn2p and Msn4p [20-22]. Msn2/4p are stress responsive transcriptional activators regulating expressions of various genes involved in stress tolerance [23]. The basal expression of these stress responsive genes in YK60-1 indicates that the changes in signaling pathways constitutively induce Msn2/4p, which would result in thermotolerance. Supporting this, many trehalose synthetic genes, which are under the controls of Msn2/4p, were upregulated according to Microarray analysis (Fig. 5). Constitutive activation of Msn2/4p would lead to the elevated level of cytoprotective trehalose in YK60-1 (Fig. 4c).

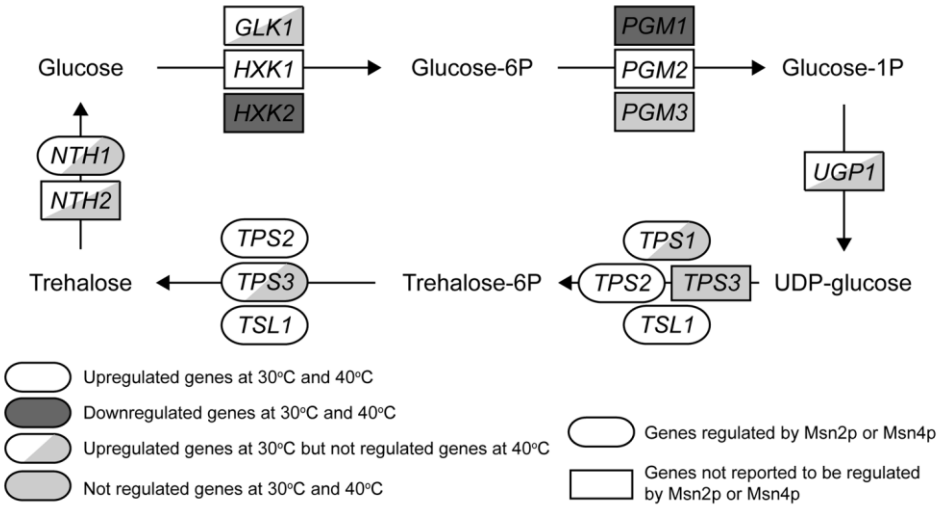

Fig. 5 Schematic model for regulation of the trehalose biosynthetic pathway in YK60-1
Genes upregulated or downregulated over 2-fold compared with those of MT8-1 incubated at 30°C were regarded as upregulated or downregulated genes, respectively.

Experimental evolution is one of the effective methods to connect genotypes to phenotypes [26]. The outcomes obtained from an experimental evolution potentially facilitate the rational engineering of productive strains for bioindustry. However, identification of mutations associated with the phenotype sometimes requires laborious works, especially when evolved strains have acquired variety of mutations. Although the development of deep-sequence technologies has enabled comprehensive whole-genome analyses of microorganisms [27], it has remained difficult to select candidate mutations involved in the phenotypes from total mutations. In contrast, I preserved intermediate populations that had adapted to the each temperature from 32°C to 38°C (Fig. 1). These intermediate strains would allow us to track the dominant strains at each temperature and to analyze the genomic mutations which the adapted strains have acquired.

The stepwise adaptation method successfully improved the thermotolerance and growth rate of parental strain MT8-1 (Fig. 2) and these features of YK60-1 will be favorable for bioproduction applications. This breeding system can be easily applied to improve a variety of tolerances such as acid tolerance, osmotic tolerance, and ethanol tolerance. In conclusion, stepwise adaptation will expand the range of yeast bioproduction applications under stress conditions.

References

1. Balakumar S, Arasaratnam V, Balasubramaniam K. Isolation and improvement of a thermotolerant *Saccharomyces cerevisiae* strain. World J Microbiol Biotechnol. 2001; 17: 739-746.
2. Shi D, Wang C, Wang K. Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae*. J Ind Microbiol Biotechnol. 2009; 36: 139-147.
3. Kida K, Kume K, Morimura S, Sonoda Y. Repeated-batch fermentation process using a thermotolerant flocculating yeast constructed by protoplast fusion. J Ferment Bioeng. 1992; 74: 169-173.
4. Sridhar M, Kiran Sree N, Venkateswar Rao L. Effect of UV radiation on thermotolerance, ethanol tolerance and osmotolerance of *Saccharomyces cerevisiae* VS $_1$ and VS_3 strains. Bioresour Technol. 2002; 83: 199-202.
5. Van Hoek P, Van Dijken JP, Pronk JT. Regulation of fermentative capacity and levels of glycolytic enzymes in chemostat cultures of *Saccharomyces cerevisiae*. Enzyme Microb Technol. 2000; 26: 724-736.
6. Tajima M, Nogi Y, Fukasawa T. Primary structure of the *Saccharomyces cerevisiae GAL7* gene. Yeast. 1985; 1: 67-77.
7. Mutch DM, Berger A, Mansourian R, Rytz A, Roberts MA. The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data. BMC Bioinformatics. 2002; 3: 17.
8. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci. 1998; 95: 14863-14868.
9. Saldanha AJ. Java Treeview-extensible visualization of microarray data. Bioinformatics. 2004; 20: 3246-3248.
10. Alvord G, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007; 8: R183.
11. Pongsuwan W, Fukusaki E, Bamba T, Yonetani T, Yamahara T, Kobayashi A. Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting. J Agric Food Chem. 2007; 55: 231-236.

12. Tsugawa H, Bamba T, Shinohara M, Nishiumi S, Yoshida M, Fukusaki E. Practical non-targeted gas chromatography/mass spectrometry-based metabolomics platform for metabolic phenotype analysis. J Biosci Bioeng. 2011; 112: 292-298.
13. Lommen A. MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Anal Chem. 2009; 81: 3079-3086.
14. Tsugawa H, Tsujimoto Y, Arita M, Bamba T, Fukusaki E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). BMC Bioinformatics. 2011; 12: 131.
15. Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. Mol Cell. 2010; 40: 253-266.
16. Monteiro G, Pereira GAG, Netto LES. Regulation of mitochondrial thioredoxin peroxidase I expression by two different pathways: one dependent on cAMP and the other on heme. Free Radical Biol Med. 2002; 32: 278-288.
17. Westerbeek-Marres CAM, Moore MM, Autor AP. Regulation of manganese superoxide dismutase in *Saccharomyces cerevisiae*. Eur J Biochem. 2005; 174: 611-620.
18. Singer MA, Lindquist S. Multiple effects of trehalose on protein folding in vitro and in vivo. Mol Cell. 1998; 1: 639-648.
19. Attfield PV. Trehalose accumulates in *Saccharomyces cerevisiae* during exposure to agents that induce heat shock response. FEBS Lett. 1987; 225: 259-263.
20. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, et al. Transcriptional regulatory code of a eukaryotic genome. Nature. 2004; 431: 99-104.
21. Ni L, Bruce C, Hart C, Leigh-Bell J, Gelperin D, Umansky L, et al. Dynamic and complex transcription factor binding during an inducible response in yeast. Genes Dev. 2009; 23: 1351-1363.
22. Teixeira MC, Monteiro P, Jain P, Tenreiro S, Fernandes AR, Mira NP, et al. The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*. Nucleic Acids Res. 2006; 34: D446-D451.
23. Schmitt AP, McEntee K. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 1996; 93: 5777-5782.
24. Zähringer H, Thevelein JM, Nwaka S. Induction of neutral trehalase Nth1 by heat and

osmotic stress is controlled by STRE elements and Msn2/Msn4 transcription factors: variations of PKA effect during stress and growth. Mol Microbiol. 2002; 35: 397-406.

25. Zhao XQ, Bai F. Zinc and yeast stress tolerance: micronutrient plays a big role. J Biotechnol. 2012; 158: 176-183.

26. Caspeta L, Chen Y, Ghiaci P, Feizi A, Buskov S, Hallstrom BM, et al. Altered sterol composition renders yeast thermotolerant. Science. 2014; 346: 75-78.

27. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature. 2012; 489: 513-518.

Chapter II

Molecular analysis of adaptational strategies to heat stress

In Chapter I, I performed stepwise breeding under heat stress and obtained a thermotolerant strain YK60-1 as well as intermediate strains. Here, I performed whole-genome sequencing of 18 intermediate strains. It is sometimes very difficult to identify key mutations involved in thermotolerance from all candidate mutations, especially when many mutations are found between parental and bred strains. On the other hand, the genome analyses of intermediate strains in each temperature would help us identify crucial mutations for thermotolerance.

Even though the signaling pathways responsible for stress tolerance have been studied, fine regulations of the pathways have not been achieved. Therefore, I examined how the mutations affected the signaling pathways to induce thermotolerance for the better understanding of mechanisms in thermotolerant yeasts.

Materials and methods

Yeast strains

S. cerevisiae MT8-1 (*MAT*a, *ade*, *his3*, *leu2*, *trp1*, *ura3*) [1] was used as the parental strain for breeding. One-point mutations were introduced into the *CDC25* gene of MT8-1 using pAUR135 vector (Takara Bio, Shiga, Japan) following the manufacturer's instructions. The inserts were introduced into the pAUR135 vector by the In-Fusion HD Cloning Kit (Clontech, CA, USA). Primers used to construct pAUR135 vectors are listed in Table 1.

Whole-genome sequencing and RNA sequencing

Single colonies were isolated from the revived glycerol stocks. Randomly picked single colonies were separately cultivated in YPD medium for 24 h and genomic DNA was extracted by the Dr. GenTLE (from yeast) high-recovery kit (Takara Bio) and purified with AMpure XP beads (Beckman Coulter, CA, USA). Extracted genomic DNA was subjected to library construction using the Nextera DNA Sample Prep Kit (Illumina Inc., San Diego, CA, USA). For

Table 1 Primers used in this study

Name	Sequence (5' - 3')
F *CDC25* latter	GGCTAGAGGATCCCCCTGAACTACGCAGCAAGAATGATGAAG
R *CDC25* latter	ACGACGGCCAGTGAACCTTCAGAGTTTCACTAGAGGCGC
F *CDC25* anterior	GGCTAGAGGATCCCCGGCTCACAAGATGTTCCTTAAAGAGAATC
R *CDC25* anterior	ACGACGGCCAGTGAACATATGTGTATGGGTCAATATCCAAGAGC
F *ACT1*	TCGTTCCAATTTACGCTGGTT
R *ACT1*	ACCGGCCAAATCGATTCTC
F *HSP12*	TTTGGCAGACCAAGCTAGAGATT
R *HSP12*	CGGCATCGTTCAACTTGGA
F *HSP104*	TTGAGGCCATCAAGCAACAA
R *HSP104*	AGCGCCACGAGAGTCAATTC
F *TPS1*	TACAGGTTGCAGTGCCAAGTCG
R *TPS1*	ATTGTGCGGCACCTGTGAACTC
F *TPS2*	GCAGTCCTACTGCCAACAGAAA
R *TPS2*	CTCGTTGACTTGTTGTTCCAATCT

mRNA extraction, MT8-1 and *CDC25* mutants were pre-cultivated in YPD medium for 24 h. Cells were then inoculated into fresh YPD medium to be adjusted to an OD_{600} of 0.1 and cultivated at 30^{o}C for 12 h. Total RNA was extracted with Isogen-LS (Nippon Gene, Tokyo, Japan) according to the manufacturer's instructions. Library construction was performed using KAPA Hyper Prep Kit (Kapa Biosystems, Woburn, MA, USA). Library qualities were confirmed by an Agilent 2100 Bioanalyzer and High Sensitivity DNA Chips (Agilent Technologies, Palo Alto, CA, USA). The libraries were sequenced on Illumina MiSeq (75 bp nucleotide paired-end sequence). The average depth of each genome analysis was 30×.

Data analysis

Paired-end reads were quality-checked with FastQC and mapped to the *S. cerevisiae* genome (sacCer3) using Burrows-Wheeler Aligner [2] for whole-genome analysis, and TopHat2 [3] for RNA-seq. To identify genomic mutations, Genome Analysis Toolkit [4] was used. The fastq files were available in the DDBJ data base (Accession number: DRA004175). Some mutations including the *CDC25* mutations were also confirmed by DNA sequencing using ABI PRISM 310 (Thermo Fisher Scientific, Waltham, MA, USA). The phylogenetic tree was

constructed based on irreversible Camin-Sokal model [5] for all mutational events (base substitutions, insertions, and deletions). Gene expressions were estimated as fragments per kilobase of exon model per million mapped fragments (FPKM) using Cufflinks [6]. The data was utilized as the average of two independent experiments. Hierarchical clustering was performed using Cluster 3.0 [7] and visualized using Java TreeView software [8]. To identify common biological functions among genes in each cluster, DAVID [9] was used. The *P* value threshold was set to 0.05. YEASTRACT (http://www.yeastract.com) was used to identify stress responsible transcription factors and transcriptional activators involved in gene inductions with more than 1.5-fold changes in transcriptional level compared to parental MT8-1.

cAMP assay

Yeast cells were cultivated in the same way as for the extraction of mRNA. Cells were collected by centrifugation at 3000×*g* for 5 min at 4°C and washed with ice-cold water. The cell pellet was resuspended in 6% ice-cold trichloroacetic acid and 200 mg of 0.5-mm glass beads (TOMY SEIKO, Tokyo, Japan) were added. After a freeze thaw with liquid nitrogen, cells were smashed 3 times using BeadSmash 12 (WAKENYAKU, Kyoto, Japan) at 4°C with 4000 agitations per min for 1 min. The solution was centrifuged at 2000×*g* for 15 min at 4°C and the supernatant was collected. HCl was added to the supernatant to a final concentration of 10 mM. The cAMP solution was extracted four times with diethyl ether and dried by lyophilization. The extracted cAMP was measured by cAMP-Screen System (Thermo Fisher Scientific). To evaluate glucose-responsiveness, cells were incubated in 0.1 mM EDTA (pH 6.0 buffered with 10 mM MES) for 2 h at 30°C. Following the addition of glucose to a final concentration of 2%, each sample was collected at 0 min, 0.5 min, 1 min, 2 min, and 3 min, and subjected to cAMP extraction.

Real-time PCR

Total RNAs were extracted in the same way as described above. Syntheses of cDNAs were performed using a High Capacity cDNA Transcription kit (Thermo Fisher Scientific) with 1 μg of the total RNA as a template according to the manufacture's protocol. Primers used in this study are listed in Table 1. For the quantitative PCR, the *ACT1* gene was used as an endogenous control to normalize the expression data for each gene. Amplification was carried

out using Power SYBR green PCR Master Mix (Thermo Fisher Scientific) in the 7500 Real-Time PCR System (Thermo Fisher Scientific).

Ethanol fermentation

For ethanol fermentation, each strain was pre-cultivated in YPD medium for 24 h. Cells were inoculated into fresh YPD and YPG (1% yeast extract, 2% peptone, 2% galactose) medium to be adjusted to an OD_{600} of 0.5 and cultivated aerobically at 30^{o}C and 39^{o}C. Sampling was performed at every 12 h interval until 48 h. Measurement of ethanol was performed as previously reported [10].

Results

Identification of a key gene for thermotolerance of yeast

In order to identify mutations that contribute to thermotolerance in *S. cerevisiae*, I sequenced the whole-genome of the parental strain and 18 intermediate strains isolated in the stepwise breeding (DDBJ accession number DRA004175). Forty-nine genomic mutations including SNPs, insertions, and deletions were identified in the intermediate strains by comparing them with the genomic sequence of MT8-1. Inversions and other rearrangements were not observed in this analysis of the 49 obtained mutations. The genome analyses revealed that the yeast cells had acquired four different point mutations in the *CDC25* gene. I constructed a phylogenetic tree according to each of the mutational events in the intermediate strains to analyze how the *CDC25* mutants had developed in the adaptational steps (Fig. 1). The phylogenetic tree suggests that there were at least five events (T943P, G1459C, N1393T, and twice in W1416C) in which the bred strains had obtained mutations in the *CDC25* gene. The phylogenetic tree implied that the mutants harboring the *CDC25* (W1416C) mutation appeared independently at 34°C and 36°C. It is highly rare that the several strains have acquired the mutations in the same gene locus, since genomic mutations basically occur randomly in 1.2×10^{7} base-pair genome of *S. cerevisiae*. Therefore, I considered that the frequently appeared *CDC25* mutations in the stepwise adaptation would play a crucial role in thermotolerance.

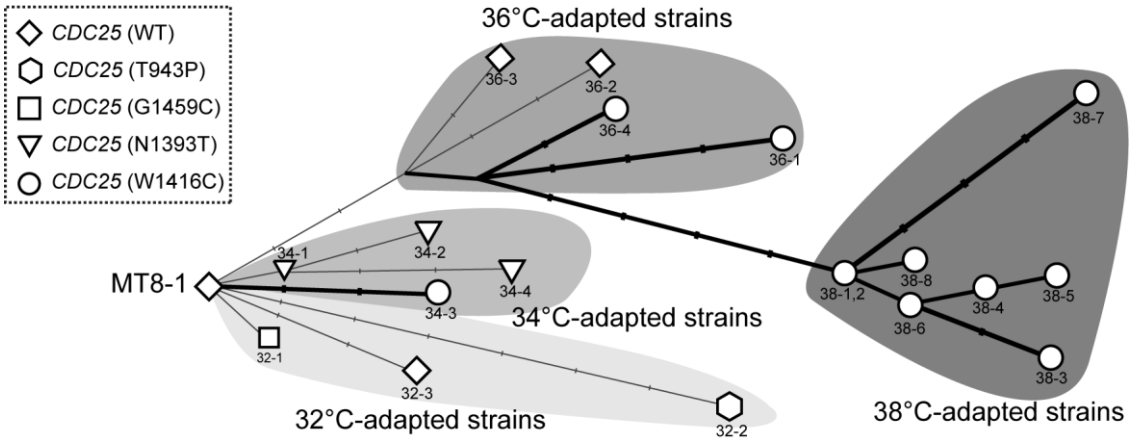

Fig. 1 Comparative genome analysis of the strains adapted to the heat stress The small scales indicate each mutational event (base substitutions, insertions, and deletions). The bold lines indicate the cell lines that had acquired a *CDC25* (W1416C) mutation. *CDC25* (W1416C) appeared independently in two populations.

To examine this hypothesis, I introduced each *CDC25* mutation into the parental strain and reconstructed *CDC25* mutants based on MT8-1 (hereafter referred to as *CDC25* ^{T943P}, *CDC25* ^{G1459C}, *CDC25* ^{N1393T}, and *CDC25* ^{W1416C}). I used pAUR135 vector system for the introduction of mutations. All reconstructed *CDC25* mutants grew better at 38°C and 39°C in contrast to MT8-1 (Fig. 2a). Cdc25p, known as guanine nucleotide exchange factor (GEF), indirectly regulates intracellular cAMP levels and thus, the cAMP/PKA signaling pathway. The cAMP/PKA pathway is responsible for inactivation of the Msn2p/Msn4p transcriptional activators that control general stress responses in *S. cerevisiae*. Since lower levels of intracellular cAMP activate Msn2p/Msn4p, I considered that the mutations in Cdc25p decreased the GEF activity and intracellular cAMP levels. In fact, reconstructed *CDC25* mutants showed lower intracellular cAMP levels than the wild-type parental strain (Fig. 2b). The decreased activity of *CDC25* mutants was also confirmed by measuring the glucose-responsiveness (Fig. 2c). Addition of glucose to reconstructed *CDC25* mutants did not trigger a rapid increase in the cAMP levels unlike parental MT8-1, indicating that the mutated Cdc25p is involved in lowering of the intracellular cAMP levels. To further validate the activation of Msn2p and Msn4p in reconstructed *CDC25* mutants, the transcriptional levels of *HSP12*, *HSP104*, *TPS1*, and *TPS2* were measured (Fig. 2d). These genes are induced by Msn2p/Msn4p through the upstream

stress-responsive elements (STREs) [11]. Hsp12p is a membrane protein involved in membrane organization under heat stress [12]. Hsp104p acts as molecular chaperone responsible for refolding of denatured and aggregated proteins [13]. Tps1p and Tps2p are the subunits of trehalose-6-P synthase complex [14]. Reconstructed *CDC25* mutants showed higher transcriptional levels of these genes, suggesting that the enhanced Msn2p and Msn4p activities resulted in thermotolerance of the *CDC25* mutants.

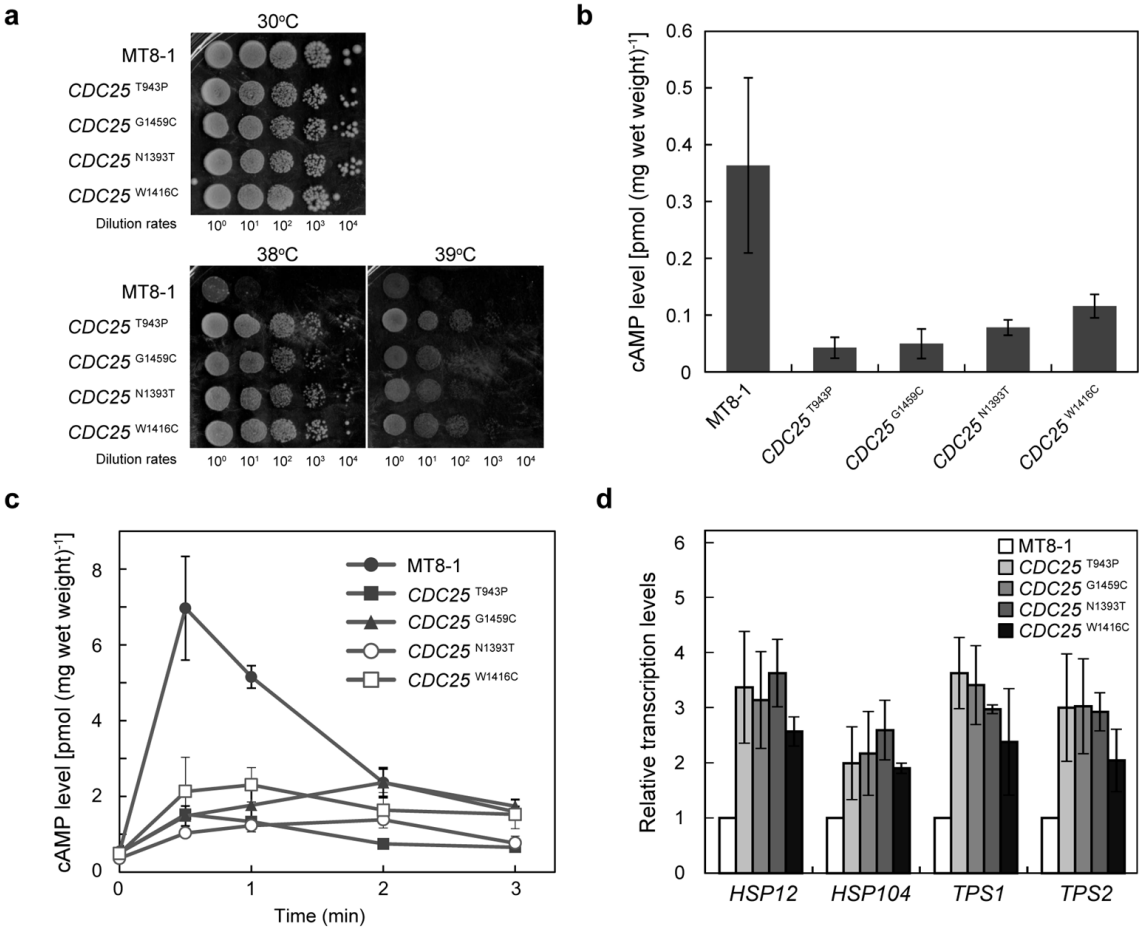

Fig. 2 Characteristics of reconstructed CDC25 mutants based on MT8-1 (a) Spot assay. Cells (OD600 = 1.0) were diluted 10-, 10^2-, 10^3-, and 10^4-fold and spotted onto a YPD plate. (b) Basal cAMP levels. (c) Induced cAMP levels by glucose. Following the addition of glucose, the cAMP levels were measured at each time point. (d) Relative transcription levels of *HSP12*, *HSP104*, *TPS1*, and *TPS2*. The error bars show standard error of the mean (SEM) based on 3 independent measurements.

Transcriptome analysis

Transcriptome analyses were performed to further investigate the transcriptional dynamics in the reconstructed *CDC25* mutants. I classified the expression levels of the *CDC25* mutants compared to that of MT8-1 by hierarchical clustering based on similarities in their expression patterns (Fig. 3a). Of the 5920 genes, the genes involved in response to stress and heat were highly upregulated in all the reconstructed *CDC25* mutants compared to the gene expression in parental MT8-1 (Cluster A). Genes in Cluster A include *HSP12*, *HSP104*, *TPS1*, and *TPS2*. Cluster B shows genes involved in membrane compositions. These upregulated genes in Cluster A and B indicate that the *CDC25* mutants have acquired thermotolerance through upregulating genes involved in stress response and organelle membrane to overcome heat stress.

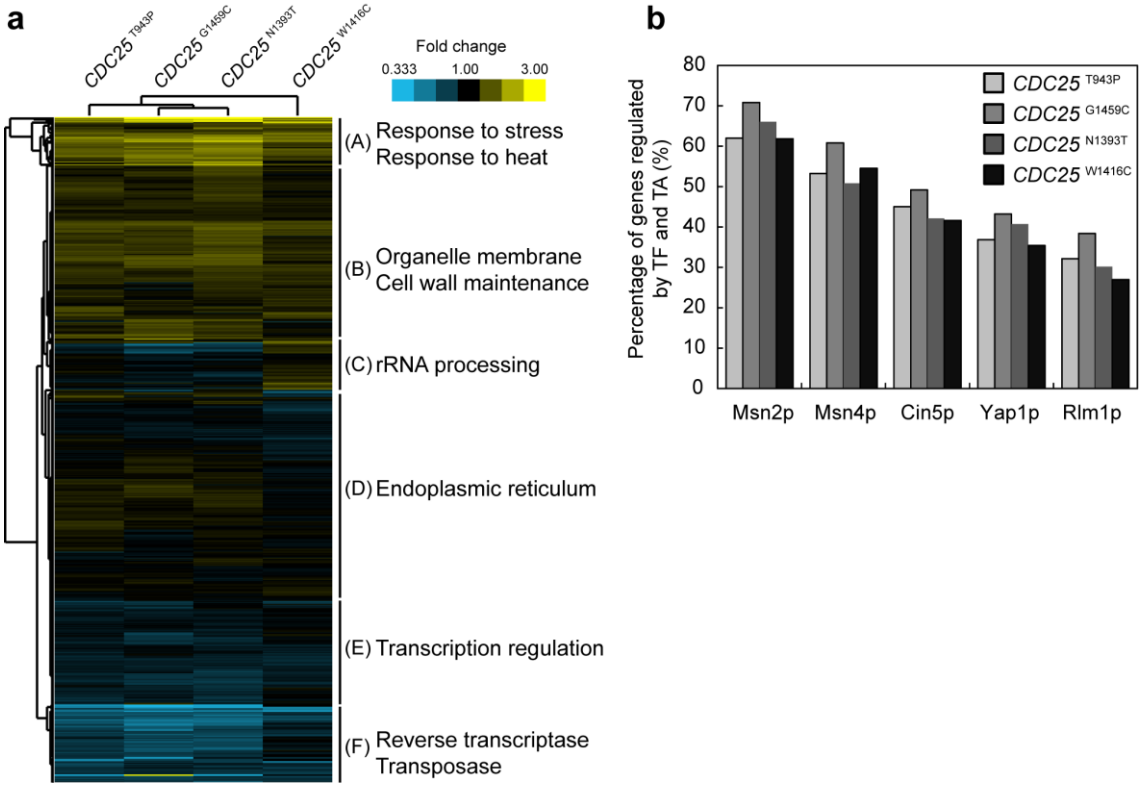

Fig. 3 mRNA profiles of reconstructed *CDC25* mutants in comparison with that of MT8-1

(a) Hierarchical clustering analysis. The gene expressions of the reconstructed *CDC25* mutants compared to that of MT8-1 were classified by hierarchical clustering. (b) Transcription factor (TF) and Transcriptional activator (TA) analyses involved in the expressions of more than 1.5-fold induced genes compared to that in MT8-1. The y-axis indicates the percentage of the genes controlled by each of the transcriptional regulator.

To identify activated transcriptional regulators in the *CDC25* mutants, I analyzed transcription factors and transcriptional activators regulating the expression of induced genes in the reconstructed *CDC25* mutants (Fig. 3b). Approximately 50–60% of the genes that were induced by more than 1.5 fold compared to that in parental MT8-1 were regulated by Msn2p and Msn4p. This result further supports the idea that the depressed activity of Cdc25p reduces intracellular cAMP levels that activates Msn2p and Msn4p through the cAMP/PKA pathway.

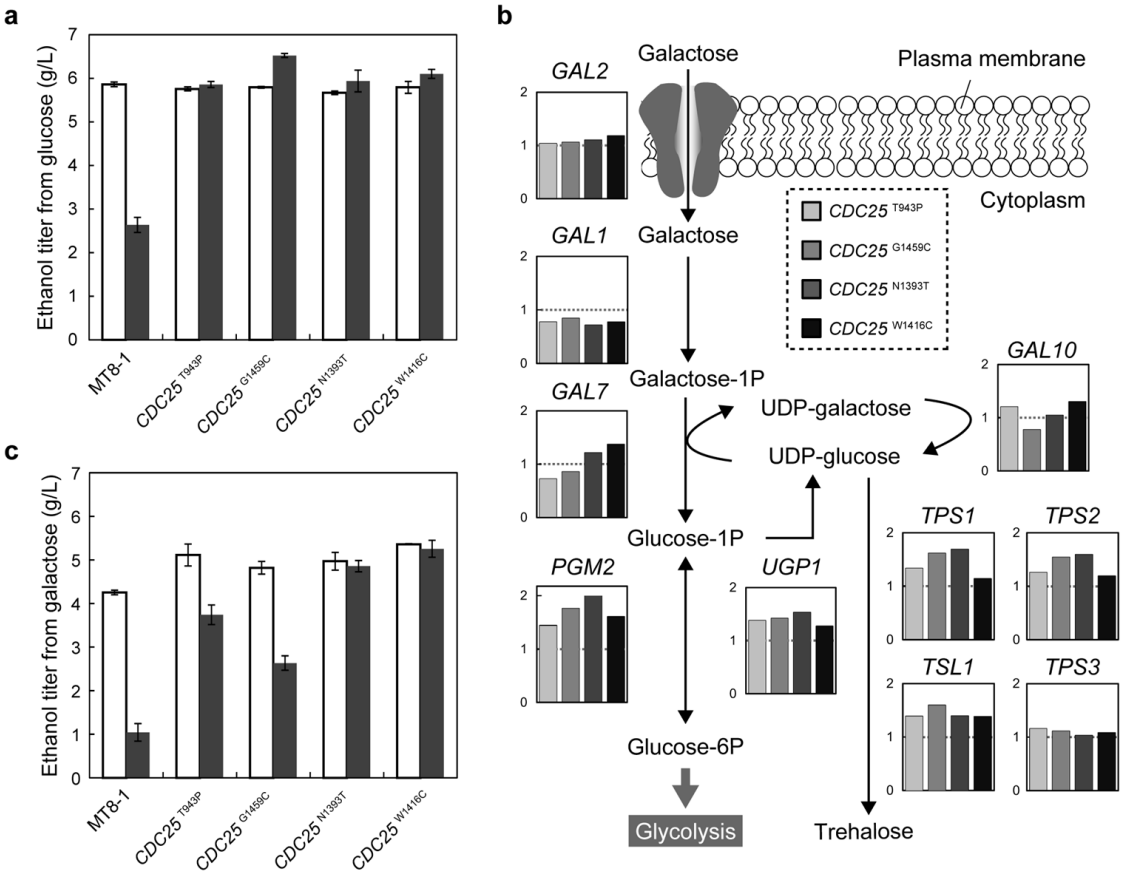

Fig. 4 Ethanol fermentation from glucose and galactose (a) Ethanol titers in 24 h from glucose at 30°C (open bars) and 39°C (filled bars). (b) Galactose fermentation pathway that partially overlaps with the trehalose synthesis pathway. The graphs represent the relative transcriptional levels of the reconstructed *CDC25* mutants compared to that of MT8-1 at 30°C. Dashed lines indicate the transcriptional levels of MT8-1. (c) Ethanol titers in 24 h from galactose at 30°C (open bars) and 39°C (filled bars). The error bars show standard error of the mean (SEM) based on 3 independent measurements.

Ethanol fermentation from glucose

Next, I examined the ethanol-fermentation abilities of the *CDC25* mutants at high temperature (39°C). Parental MT8-1 and the reconstructed *CDC25* mutants showed similar growth rates and ethanol productivities at 30°C in glucose medium (Fig. 4a). At 39°C, however, the ethanol productivity of parental MT8-1 was greatly reduced due to heat stress. On the other hand, the reconstructed *CDC25* mutants retained ethanol productivity at the temperature. The ethanol titer of reconstructed strain *CDC25* ^{G1459C} in 24 h was 2.5 fold higher than that of parental MT8-1 at 39°C. The reconstructed *CDC25* mutants had developed not only thermotolerance but also efficient ethanol fermentation ability under heat stress.

Ethanol fermentation from galactose

In *S. cerevisiae*, galactose is metabolized through the Leloir pathway that partially overlaps with the trehalose synthesis pathway (Fig. 4b). Trehalose is a metabolite that plays an important role in thermotolerance [15]. During diauxic shift, *PGM2* and *UGP1* are induced by Msn2p and Msn4p. *PGM2* is a key enzyme for both the trehalose synthesis pathway and Leloir pathway. Therefore, I considered that the activation of Msn2p and Msn4p in the reconstructed *CDC25* mutants induced *PGM2* that would allow efficient galactose fermentation. According to transcriptome analysis performed at 30°C, *PGM2* and *UGP1* levels were upregulated in the reconstructed *CDC25* mutants compared to that in parental MT8-1 (Fig. 4b). In addition, as expected, ethanol titers in galactose medium were slightly higher in the reconstructed *CDC25* mutants at 30°C (Fig. 4c). I measured ethanol titers of the reconstructed *CDC25* mutants at 39°C, because thermotolerance of the reconstructed *CDC25* mutants was considered to contribute to the ethanol productions under the heat stress. The reconstructed *CDC25* mutants, especially *CDC25* ^{N1393T} and *CDC25* ^{W1416C} produced more ethanol from galactose at 39°C compared to parental MT8-1 (Fig. 4c). The ethanol titers of *CDC25* ^{W1416C} in 24 h was 5.1 fold higher compared to that of MT8-1 at 39°C. The whole-genome sequencing of intermediate strains revealed four kinds of one-point mutations in the *CDC25* gene which contributed to thermotolerance (Fig. 2a), faster growth in galactose medium, and efficient ethanol fermentation from glucose and galactose under heat stress (Fig. 4a, c).

Discussion

In the stepwise adaptation, I preserved and used the intermediate strains during breeding steps for the genome analysis (Fig. 1a). According to the phylogenic tree (Fig. 1), the 36°C- and 38°C-adapted populations were not derived from 32°C- and 34°C-adapted strains which I analyzed (Fig. 1). The 36°C- and 38°C-adapted populations would have developed from a relatively minor 34°C-adapted strain. The whole-genome analyses of intermediate strains allowed us to track individual cells and provided information about the advantageous mutations for the survival in the environment. Because reconstructed *CDC25* ^{N1393T} and *CDC25* ^{W1416C} grew better than *CDC25* ^{G1459C} and *CDC25* ^{T943P} at 39°C, only the intermediate strains with *CDC25* (N1393T) and *CDC25* (W1416C) mutations were isolated after 34°C-adapted populations. However, it is still unclear why only the bred mutants with the *CDC25* (W1416C) mutation were observed in the 38°C-adapted population (Fig. 1), even though reconstructed *CDC25* ^{N1393T} grew better than *CDC25* ^{W1416C} at 39°C. One possible explanation for this could be that there are additional and beneficial mutations that only a strain with the *CDC25* (W1416C) mutation had acquired in the stepwise breeding. The identification of such mutations will also facilitate the construction of thermotolerant strains.

Previous reports suggested that the cAMP-depressing mutations such as *cdc25-21* and *cdc25-22* contributed to thermotolerance in *S. cerevisiae* [16]. These *CDC25* mutations reside in the C-terminal domain that functions as GEF [17]. The N1393T, W1416C, and G1459C mutations discovered in this study are also located in the same domain. These mutations would affect the GEF activity and decrease the intracellular cAMP levels. The *cdc25-21* and *cdc25-22* mutants, however, were not able to grow as fast as the wild-type strain, because the cAMP/PKA pathway controls cell-cycle progression [18]. On the other hand, the *CDC25* mutants reconstructed in this study grew as fast as parental MT8-1 in glucose medium, and faster than parental MT8-1 in galactose medium. The fine balance between thermotolerance and growth rates would be brought from the stepwise adaptation, because the yeast cells isolated during this adaptation process were subjected to selective pressures of both heat and growth at all times. Evolutionary engineering is one of the most effective methods of fine-tuning signaling pathways through unpredicted mutations. *CDC25* mutants, which showed thermotolerance and similar growth rates to the parental strain, would be advantageous for fermentation, since elevated growth rates are often associated with better fermentation rates [19].

As galactose is abundant sugars in nonfood crops as well as glucose, efficient utilization of glucose and galactose is crucial [20]. However, *S. cerevisiae* grows in galactose media at approximately half the rate than that in glucose media [21]. In *S. cerevisiae*, metabolism of galactose takes place through the Leloir pathway [22]. Metabolome analysis in a previous report indicated that the accumulation of intermediate products such as galactose-1-phosphate and glucose-1-phosphate inhibited galactose fermentation [23]. Galactose uptake and fermentation are improved by the overexpression of *PGM2*, which encodes phosphoglucomutase that mediates the conversion between glucose-1-phosphate and glucose-6-phosphate [24]. In all the reconstructed *CDC25* mutants, *PGM2* as well as other trehalose synthetic genes were induced at higher levels than those in parental MT8-1 (Fig. 4b). These gene expressions are regulated by Msn2p and Msn4p [25]. The downregulated cAMP/PKA pathway in the reconstructed *CDC25* mutants led to thermotolerance and efficient galactose utilization at 30°C, and even at 39°C.

In all the four reconstructed *CDC25* mutants, the upregulated genes with more than 1.5-fold change compared to parental MT8-1, were mainly regulated by Msn2p, Msn4p, Cin5p (also called Yap4p), Rlm1p and Yap1p (Fig. 3b). This indicates that the depressed activity of Cdc25p and thus, the lower level of cAMP, activated these transcriptional regulators. Increasing intracellular level of cAMP is known to antagonize the activities of Msn2p and Msn4p through the inhibition of Bcy1p and the activation of cAMP-dependent protein kinase (PKA) composed of Tpk1p, Tpk2p, and Tpk3p (Fig. 5) [26]. The kinase is responsible for the inactivation of various transcriptional regulators including Msn2p and Msn4p [27].

Cin5p is a transcription factor that regulates pleiotropic drug resistance and salt tolerance [28, 29]. The phosphorylation and stability of Cin5p is positively regulated by glycogen synthase kinase 3 (GSK3) such as Rim11p, Mck1p, and Yol128c (Fig. 5) [30]. GSK3 activities are inhibited by PKA, indicating that lower intracellular cAMP levels led to the activation of Cin5p.

Rlm1p is a transcription factor controlling the cell wall integrity pathway [31]. The GSK3 mutants such as mrk1, mck1, or yol128c showed sensitivity to SDS, zymolyase, and calcofluor white, which cause cell wall stress [32]. The mutants that lacked cAMP signaling showed thick cell walls, indicating the activation of cell wall integrity pathway. The activated GSK3 might directly or indirectly activate Rlm1p (Fig. 5).

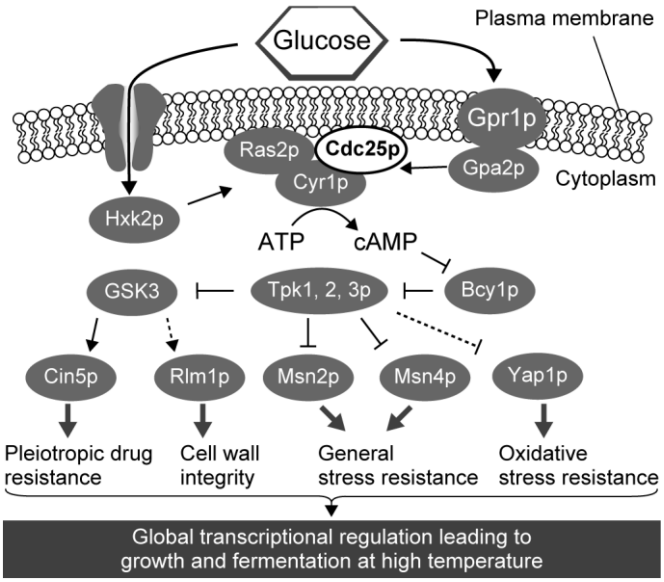

Fig. 5 Signaling pathway in the reconstructed *CDC25* mutants Arrows and bars represent positive and negative regulations, respectively. GSK indicates Rim11p, Mck1p, and Yol128c. Dashed arrows and bars represent putative or indirect regulations.

Yap1p is a transcription factor required for oxidative stress tolerance [33]. The cAMP/PKA pathway inhibits Yap1p [34]. In addition, in the absence of Bcy1p, the inhibitory subunit of PKA negatively affects the function of Yap1, suggesting that the cAMP/PKA pathway at least indirectly regulates Yap1p activity [35].

The point mutations in *CDC25* were likely to enhance multiple genes through these transcriptional regulators. In fact, *CDC25* N1393T, the reconstructed mutant that showed the fastest growth at 39°C, induced 376 gene expressions with more than 1.5-fold compared to parental MT8-1. The multiple transcriptional regulators responsible for several stress tolerance mechanisms would contribute to the growth at the high temperatures, because high temperature causes not only heat stress, but also various other kinds of stresses including cell wall stress and oxidative stress (Fig. 5) [32, 36].

In conclusion, I found four *CDC25* mutations involved in thermotolerance and efficient glucose and galactose utilization at high temperature. Introduction of point mutations into *CDC25* conferred thermotolerance. Since Cdc25p acts upstream of the cAMP/PKA pathway, the *CDC25* mutations could comprehensively alter the transcriptional profiles through various

transcriptional regulators, which would result in thermotolerance (Fig. 5). By altering the activities of transcriptional regulators, we will be able to comprehensively alter transcriptional profiles. This strategy does not need laborious genetic manipulation to organize several gene expressions; it only requires the introduction of a point mutation into a gene located upstream of the signaling pathways for global transcriptional changes. The identified *CDC25* mutations will be beneficial for rational design of thermotolerant *S. cerevisiae*.

References

1. Tajima M, Nogi Y, Fukasawa T. Primary structure of the *Saccharomyces cerevisiae GAL7* gene. Yeast. 1985; 1: 67-77.
2. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754-1760.
3. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14: R36.
4. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297-1303.
5. Camin JH, Sokal RR. A method for deducing branching sequences in phylogeny. Evolution. 1965; 19: 311-326.
6. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols. 2012; 7: 562-578.
7. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci. 1998; 95: 14863-14868.
8. Saldanha AJ. Java Treeview-extensible visualization of microarray data. Bioinformatics. 2004; 20: 3246-3248.
9. Alvord G, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007; 8: R183.
10. Nakanishi A, Bae JG, Fukai K, Tokumoto N, Kuroda K, Ogawa J, et al. Effect of pretreatment of hydrothermally processed rice straw with laccase-displaying yeast on ethanol fermentation. Appl Microbiol Biotechnol. 2012; 94: 939-948.

11. Boy-Marcotte E, Perrot M, Bussereau F, Boucherie H, Jacquet M. Msn2p and Msn4p control a large number of genes induced at the diauxic transition which are repressed by cyclic AMP in *Saccharomyces cerevisiae*. J Bacteriol. 1998; 180: 1044-1052.
12. Welker S, Rudolph B, Frenzel E, Hagn F, Liebisch G, Schmitz G, et al. Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. Mol Cell. 2010; 39: 507-520.
13. Sanchez Y, Lindquist SL. *HSP104* required for induced thermotolerance. Science. 1990; 248: 1112-1115.
14. Virgilio C, Burckert N, Bell W, Jeno P, Boller T, Wiemken A. Disruption of *TPS2*, the gene encoding the 100-kDa subunit of the trehalose-6-phosphate synthase/phosphatase complex in *Saccharomyces cerevisiae*, causes accumulation of trehalose-6-phosphate and loss of trehalose-6-phosphate phosphatase activity. Eur J Biochem. 2005; 212: 315-323.
15. Singer MA, Lindquist S. Multiple effects of trehalose on protein folding in vitro and in vivo. Mol Cell. 1998; 1: 639-648.
16. Folch-Mallol JL, Martinez LM, Casas SJ, Yang R, Martinez-Anaya C, Lopez L, et al. New roles for *CDC25* in growth control, galactose regulation and cellular differentiation in *Saccharomyces cerevisiae*. Microbiology. 2004; 150: 2865-2879.
17. Lai CC, Boguski M, Broek D, Powers S. Influence of guanine-nucleotides on complex-formation between Ras and Cdc25 proteins. Mol Cell Biol. 1993; 13: 1345-1352.
18. Baroni MD, Monti P, Alberghina L. Repression of growth-regulated Gl cyclin expression by cyclic AMP in budding yeast. Nature. 1994; 371: 339-342.
19. Van Hoek P, Van Dijken JP, Pronk JT. Regulation of fermentative capacity and levels of glycolytic enzymes in chemostat cultures of *Saccharomyces cerevisiae*. Enzyme Microb Technol. 2000; 26: 724-736.
20. Mussatto SI, Mancilha IM. Non-digestible oligosaccharides: a review. Carbohydr Polym. 2007; 68: 587-597.
21. Velagapudi VR, Wittmann C, Schneider K, Heinzle E. Metabolic flux screening of *Saccharomyces cerevisiae* single knockout strains on glucose and galactose supports elucidation of gene function. J Biotechnol. 2007; 132: 395-404.
22. Frey PA. The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. The FASEB Journal. 1996; 10: 461-470.
23. Hong KK, Vongsangnak W, Vemuri GN, Nielsen J. Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis.

Proc Natl Acad Sci. 2011; 108: 12179-12184.

24. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J. Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. Appl Environ Microbiol. 2005; 71: 6465-6472.
25. Vianna CR, Silva CLC, Neves MJ, Rosa CA. *Saccharomyces cerevisiae* strains from traditional fermentations of Brazilian cachaça: Trehalose metabolism, heat and ethanol resistance. Antonie Van Leeuwenhoek. 2008; 93: 205-217.
26. Görner W, Durchschlag E, Martinez-Pastor MT, Estruch F, Ammerer G, Hamilton B, et al. Nuclear localization of the C_2H_2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. Genes Dev. 1998; 12: 586-597.
27. Görner W, Durchschlag E, Wolf J, Brown EL, Ammerer G, Ruis H, et al. Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor. EMBO J. 2002; 21: 135-144.
28. Delling U, Raymond M, Schurr E. Identification of *Saccharomyces cerevisiae* genes conferring resistance to quinoline ring-containing antimalarial drugs. Antimicrob Agents Chemother. 1998; 42: 1034-1041.
29. Mendizabal I, Rios G, Mulet JM, Serrano R, de Larrinoa IF. Yeast putative transcription factors involved in salt tolerance. FEBS Lett. 1998; 425: 323-328.
30. Pereira J, Pimentel C, Amaral C, Menezes RA, Rodrigues-Pousada C. Yap4 PKA-and GSK3-dependent phosphorylation affects its stability but not its nuclear localization. Yeast. 2009; 26: 641-653.
31. Jung US, Levin DE. Genome-wide analysis of gene expression regulated by the yeast cell wall integrity signalling pathway. Mol Microbiol. 1999; 34: 1049-1057.
32. Griffioen G, Swinnen S, Thevelein JM. Feedback inhibition on cell wall integrity signaling by Zds1 involves Gsk3 phosphorylation of a cAMP-dependent protein kinase regulatory subunit. J Biol Chem. 2003; 278: 23460-23471.
33. Schnell N, Krems B, Entian K, Dieter. The *PAR1* (*YAP1*/*SNQ3*) gene of *Saccharomyces cerevisiae*, ac-jun homologue, is involved in oxygen metabolism. Curr Genet. 1992; 21: 269-273.
34. Charizanis C, Juhnke H, Krems B, Entian K, D. The oxidative stress response mediated via Pos9/Skn7 is negatively regulated by the Ras/PKA pathway in *Saccharomyces cerevisiae*. Mol Gen Genet. 1999; 261: 740-752.
35. Hasan R, Leroy C, Isnard AD, Labarre J, Boy‐Marcotte E, Toledano MB. The control of the yeast H_2O_2 response by the Msn2/4 transcription factors. Mol Microbiol. 2002; 45: 233-241.

36. Davidson JF, Whyte B, Bissinger PH, Schiestl RH. Oxidative stress is involved in heat-induced cell death in *Saccharomyces cerevisiae*. Proc Natl Acad Sci. 1996; 93: 5116-5121.

Chapter III

Development of novel genome editing technology for the reconstruction of thermotolerant yeast

I have successfully identified thermotolerance-associated *CDC25* mutantions in the previous chapter. The introduction of *CDC25* mutants, however, was time-consuming and inefficient [1, 2]. This genome-integration method relies on a selection step based on marker gene and a marker-removing step by counter-selection. It is time-consuming in which plasmid construction, integration, and maker-removing steps take about 2 weeks. For rapid construction of yeast mutants, some marker genes such as auxotrophic makers have been used for targeted genomic mutagenesis [3]. However, these techniques can not rule out the possibility of the influences of the marker genes on the phenotypes.

For efficient, marker-less, targeted genomic-mutagenesis technology, clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system has been developed [4]. In the current CRISPR/Cas9 system, Cas9 nuclease targets to a specific genomic site by the function of guide RNA (gRNA) [4]. Cas9/gRNA complex recognizes a targeting site bearing a complementary 20-nt sequence of gRNA which is immediately followed by a protospacer adjacent motif (PAM), typically “NGG”. This system leads to double strand breaks (DSBs) at 3 base pairs (bp) upstream of the PAM by the function of RuvC and HNH nuclease domains. DSBs induce cellular death as well as two independent repairing pathways; homology-directed repair (HDR) and non-homologous end joining (NHEJ) [5, 6]. NHEJ is an error-prone repairing pathway, which leaves an insertion or deletion at the cleavage site. HDR is a high-fidelity repairing pathway depending on donor DNA. Therefore, target-specific cleavage by this system can stimulate gene knockout through NHEJ and flexible gene designs through HDR.

Despite the high efficiency and robustness of the CRISPR/Cas9 system, there will be two major problems in this system. Firstly, editable bases through HDR would be limited to only in the inside of the 20 bp gRNA-targeting site and PAM sequence. When mutations are introduced into the outside areas of PAM and gRNA-targeting sequences by HDR, these recognition sequences remain intact even after HDR, which potentially causes re-cleavage by the CRISPR/Cas9 system again. The re-cleavage will continuously induce HDR until the PAM or gRNA-targeting sequence is disrupted by error-prone NHEJ. Therefore, the CRISPR/Cas9

system has been thought to be unable to precisely edit the outside areas of PAM and gRNA-targeting sequences by HDR, while it can precisely edit the inside of these sequences with nearly 100% efficiencies in yeast [7]. Secondary, the specificity of this system is imperfect [4, 8, 9]. Unintended mutations sometimes have been observed in off-target sites that share sequence homology to the on-target site [10]. For stable and precise genome editing, a new genome editing technology which is capable of genome-wide and target-specific editing would have to be developed.

For robust and genome-wide base editing technology, I developed a novel CRISPR Nickase system to overcome the limitations in the CRISPR/Cas9 system. The constructed CRISPR Nickase system was able to edit bases regardless of the inside or outside of PAM and gRNA-targeting sequences, while the CRISPR/Cas9 system could edit only the inside of these sequences. In addition, this new system was highly specific to the on-target site in contrast to the CRISPR/Cas9 system. Coupling with yeast gap repair cloning (GRC), it was able to construct yeast mutants including thermotolerant *CDC25* mutants only within 5 days. CRISPR Nickase system serves a versatile and powerful genomics tool to construct valuable strains for bioproductions.

Materials and methods

Strains and DNA manipulation

S. cerevisiae MT8-1 (*MAT***a**, *ade*, *his3*, *leu2*, *trp1*, *ura3*) [11] was used as a host strain for genome editing. The *rad*- knockout strains were obtained from the yeast gene knockout collection [12]. *Escherichia coli* strain DH5α [F^-, *end*A1, *hsd*R17(r_k^-/m_K^+), *sup*E44, *thi*-1, λ^-, *deo*R, *rec*A1, *gyr*A96, *pho*A, ϕ80d*lac*ZΔM15, Δ(*lac*ZYA-*arg*F)U169] (Toyobo, Osaka, Japan) was used as a host for the recombinant DNA manipulation. Donor fragments and 20-nt gRNA expressing DNA were inserted into backbone vectors by In-Fusion HD Cloning Kit (Clontech, CA, USA) and Ligation high (Toyobo), respectively. *E. coli* transformants were grown in Luria–Bertani medium [1% (w/v) tryptone, 0.5% (w/v) yeast extract, and 1% (w/v) sodium chloride] containing 50 μg/mL ampicillin. Introductions of one-point mutations by the conventional two-step integration/excision method were conducted using pAUR135 vector (Takara Bio, Shiga, Japan) following to manufacturer's instructions.

Genome editing by conventional CRISPR Cas9 system

Genome editing was performed as previously reported [7]. Double strand donor DNA was used by annealing 80 nt-single strand oligonucleotides. Yeast cells were transformed with 250 ng of plasmids and 500 pmol of donor DNA by using the Frozen-EZ yeast transformation II kit (Zymo Research, CA, USA). Transformants on selective plate medium were replica-plated onto SD plate medium supplemented 60 mg/L canavanine and appropriate amino acids without arginine.

Base editing by CRISPR Nickase system

Transformation was performed as described above. Transformants were selected and grown on SDC+AW [0.67% (w/v) yeast nitrogen base w/o amino acids, 2% glucose, 2% casamino acid (BD BioSciences, CA, USA), 2.0×10^{-3}% adenine, and 2.0×10^{-3}% L-tryptophan] plate medium. The colonies on the selective medium were suspended in distilled water. Suspension was adjusted to optical density at 600 nm (OD_{600}) of 10^{-5} ($\fallingdotseq1.1\times10^{3}$ cells) in 5 mL SDC+AW liquid medium and grown for 48 h at 30°C. Cultivated cells were spread on YPD [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose] plate medium. For *CAN1* mutagenesis analyses, colonies on YPD plate medium were replica plated to SD plate medium supplemented 60 mg/L canavanine and appropriate amino acids without arginine to evaluate genome editing efficiencies. For the sequence analyses, colonies were randomly picked and DNA fragments containing a mutation site were directly amplified by PCR, which were analyzed by Sanger sequencing (Eurofins Genomics, Tokyo, Japan).

Calculation of editable bases

Yeast genome used for the calculation of editable bases was obtained from *Saccharomyces genome* database (http://www.yeastgenome.org/). All bases (1.21×10^{7} bp) were classified based on whether editable or not. “GG” and “CC” in PAM sequences and following 20 bp- or 12 bp- gRNA-targetable sequences in the genome were defined as editable. Genome coverage (%) was calculated by dividing the number of editable bases by the total bases in the genome. In a similar manner, editable bases by CRISPR Nickase system were determined based on the distances from a nicking site.

Gap repair cloning

DNA fragments were amplified to add more than 30 bp homology arms on the each edge. For amplification of a Cas9 D10A and gRNA containing fragment, 2 ng templates were used. Donor DNA fragments containing mutation was amplified directly from yeast cells. Amplified DNA fragments were treated with DpnI (Toyobo) for 2 h at 37°C to degrade template plasmids completely. Purification was performed with Qiaquick PCR purification kit (QIAGEN, Hilden, Germany) and 0.2 pmol of DNA fragments were used for transformation. Transformants were cultivated as described above. Sequencing of *CDC25* mutants was performed by randomly picking colonies on YPD plate medium.

Fluorescent analysis

At least more than 20 colonies on selective plate medium were suspended in distilled water at the same time. Suspension was adjusted to an OD_{600} of 10^{-5} in 5 mL SDC+AW liquid medium and grown for 48 h at 30°C. Cells were then inoculated into fresh YPD medium to be adjusted to an OD_{600} of 0.1 and cultivated at 30^{o}C for 12 h. The Flow cytometry analysis was performed using a JSAN desktop cell sorter (Bay bioscience, Kobe, Japan). The fluorescence of cells was analyzed using AppSan software (Bay bioscience), and were measured with excitation at 488 nm and emission from 512 to 558 nm. A total of 5.0×10^{4} cells was recorded per sample. In order to quantify the proportion of cells producing non-functional EGFP, I asked whether the fluorescence intensities were below the threshold. The threshold was determined based on each negative control distribution, and was defined as within the mediumn +2 standard deviations of fluorescence intensity. Cells that did not exceed to the threshold were regarded as non-fluorescent.

Results

Limitation of CRISPR/Cas9 system

In order to investigate whether CRISPR/Cas9 system can precisely edit the outside areas of recognition sequences, I used a previously reported CRISPR/Cas9 system in *S. cerevisiae* [7]. In this system, a centromeric single-copy plasmid constitutively expresses Cas9 under the strong *TEF1* promoter and a multi-copy 2 μ plasmid constitutively expresses gRNA under the *snR52* promoter. gRNA was constructed to target a *CAN1* gene, which encodes a arginine transporter

[13]. Non-sense mutation in the *CAN1* gene confers canavanine resistance with yeast cells, while the functional *CAN1* gene causes canavanine sensitivity. Using canavanine sensitivity and resistance, I evaluated the genome editing efficiencies of CRISPR/Cas9 system by introducing stop codons in the PAM sequence, gRNA-targeting sequence, and an outside area of the PAM and gRNA-targeting sequences. Double strand donor DNA was used to mediate the editing of the *CAN1* gene. When CRISPR/Cas9 system introduced stop codons at PAM and gRNA-targeting sequences, almost all yeast cells acquired canavanine resistance (Fig. 1c), which were similar efficiencies to the previous report [7]. A catalytically inactive Cas9 (Dead Cas9) did not induce genome editing. All sequenced colonies obtained intended stop codons in the *CAN1* gene (Fig. 1d). On the other hand, when a stop codon was introduced at the outside area (9 bp downstream from the cleavage site) by using Cas9, only a few cells acquired canavanine resistance (Fig. 1c). In addition, many cells obtained unnecessary mutations in the gRNA-targeting sequences as well as the intended premature stop codon (Fig. 1d). Only one colony underwent the intended genome editing.

Colony forming units (CFUs) varied greatly depending on the mutation sites in the donor DNA (Fig. 1e). Production of Cas9 in the absence of donor DNA severely decreased CFU, because DSB induced cellular death as well as NHEJ. On the other hand, donor DNA containing the mutation in the PAM sequence protected cells producing Cas9 from cellular death caused by DSB. However, donor DNA containing the mutation at the outside area led to cellular death, suggesting that continuous cleavages by Cas9 occurred as expected.

Development of CRISPR Nickase system

Since CRISPR/Cas9 system had the limitation of the editable bases, I attempted to develop a new genome editing technology that is free from the limitation. To achieve this, I used Cas9 nickases, Cas9 D10A and Cas9 H840A [14], instead of Cas9 nuclease (Fig. 1b). Cas9 D10A cleaves the gRNA-targeting strand, while Cas9 H840A cleaves the non-targeted strand [15]. Nickase introduces single strand breaks (nicks) that can be precisely repaired in the cells [16]. Nick can induce homologous recombination, but not NHEJ [14]. Therefore, continuous re-recognition and re-cleavage by Cas9 nickase/gRNA after HDR would not stimulate NHEJ, resulting in the precise genome editing. However, simple replacement of Cas9 nuclease with Cas9 nickases (Cas9 D10A and Cas9 H840A) did not induce genome editing at all (Fig. 1c).

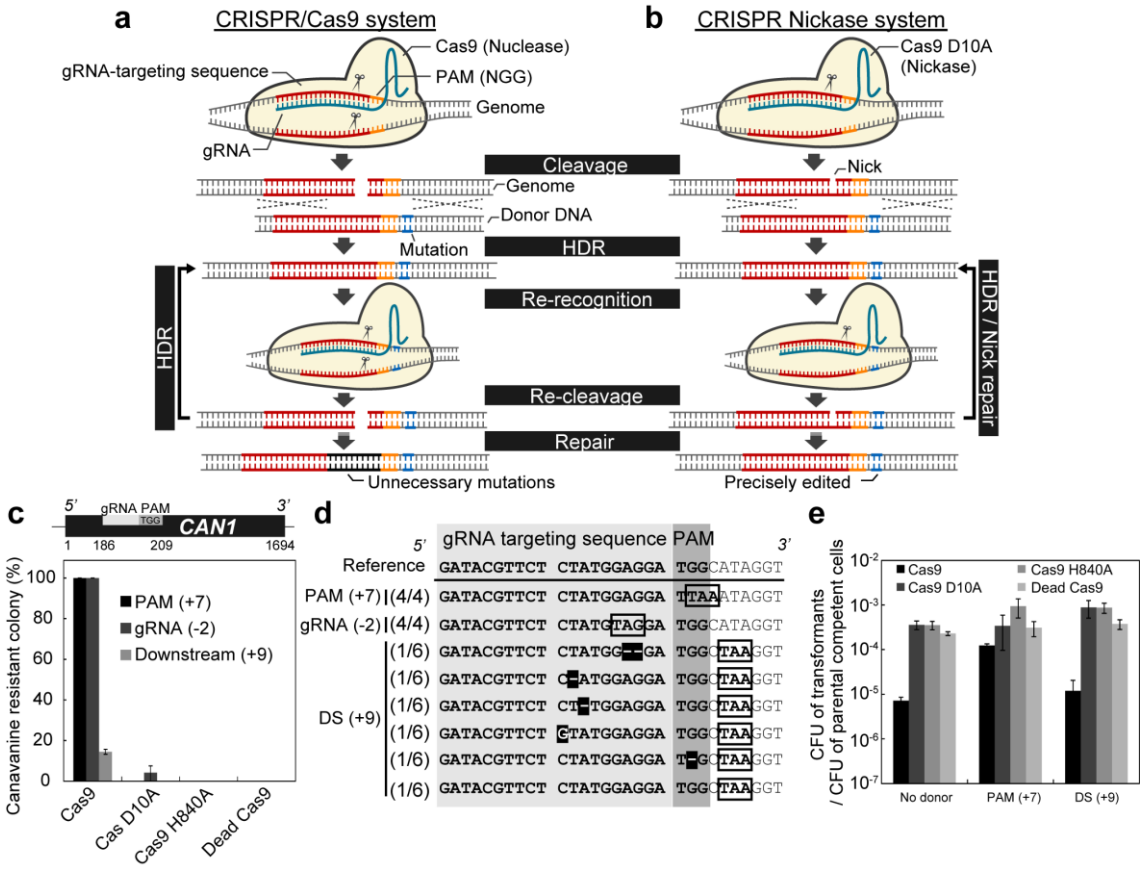

Fig. 1 Limitation of CRISPR/Cas9 system (a) Scheme of genome editing at the outside areas of PAM and gRNA-targeting sequences by CRIPSR/Cas9 system and (b) CRISPR Nickase system. (c) Genome editing efficiencies at the *CAN1* gene by conventional method [7]. The targeted position is described above the graph. (d) Sequencing of edited *CAN1* gene. Cas9 precisely edited PAM and gRNA-targeting sequences, but not a downstream (DS) site from the cleavage site. Introduced stop codons (squared) and unintended mutations (inverted) are represented. The numbers of observed sequences against the numbers of total sequenced strains (e.g. 4/4) are shown. (e) Transformation efficiencies. The numbers of transformants on selective medium divided by the number of competent cells counted on non-selective medium are represented to evaluate the toxicity of CRISPR systems. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements.

I thought that transient donor introduction was not enough for the nick-mediating HDR. Nick is precisely filled in cells and does not induce HDR as efficiently as DSB [17]. Therefore, I

integrated donor DNA fragment into the plasmid expressing Cas9 nickase and gRNA to prevent losses of donor by degradation and cell divisions (Fig. 2a). About 1000 bp of donor DNA was used in this study, because more than 700 bp donor length has been reported to serve enough HDR efficiencies in *S. cerevisiae* [18]. In addition, transformants that harbored the plasmids were further cultivated in liquid medium for 48 h. Since HDR is induced in late S and G2 phases when sister chromatids are available as HDR template [19], promoting cell cycles by cell divisions was expected to confer more chances with cells to repair nicks by HDR.

At first, I tested this system by producing Cas9 nuclease as a positive control (Fig. 2b). In this vector design, Cas9-production induced efficient and precise genome editing at the PAM sequence, suggesting that this system worked as expected (Fig. 2b, 3). When the outside areas were edited, Cas9 introduced intended stop codons as well as unintended mutations around the cleavage sites (Fig. 3). Production of Cas9 without donor DNA or with donor DNA containing mutations at the outside area caused severe decreases in CFUs, while donor DNA containing the mutation in the PAM sequence protected cells (Fig. 2c). These results emphasize the fact that prevention of re-recognition of Cas9/gRNA complex after HDR is crucial for accurate genome editing. When Cas9 nickases were produced to edit the PAM sequence, approximately 50% cells obtained canavanine resistance and the intended stop codon in the *CAN1* gene (Fig. 2 b, 3). Notably, Cas9 nickases were able to introduce mutations at the outside areas with about 50% efficiencies, without any unintended mutations (Fig. 3). Furthermore, it successfully edited bases up to 53 bp upstream and 50 bp downstream from the nicking site (Fig. 2b). There were not any major biases in efficiencies between two nickases, Cas9 D10A and Cas9 H840A, indicating that the nicking strand did not affected the HDR efficiencies. CFUs of cells producing Cas9 nickases were not varied by mutation sites in donor DNA (Fig. 2c). This system could precisely edit bases at the different site of the *CAN1* gene (Fig. 2d). I termed this base editing technology using Cas9 D10A or Cas9 H840A as “CRISPR Nickase system”.

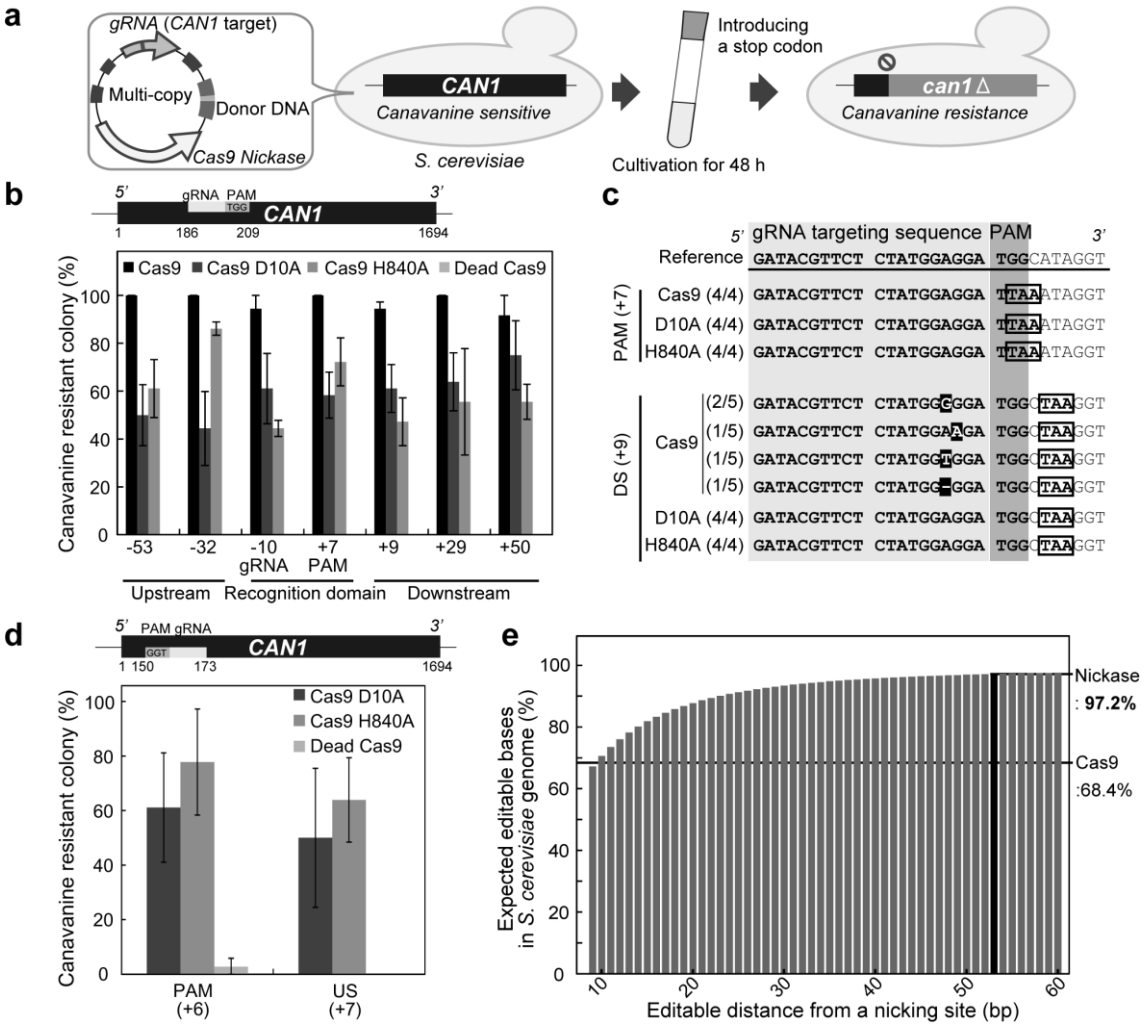

Fig. 2 Construction of CRISPR Nickase system (a) Scheme of the experiment. Cas9 nickase- and gRNA-expressing cassettes were integrated into a single plasmid that had a donor DNA. Transformants were cultivated in selective medium for 48 h from an OD_{600} of 10^{-5} ($\fallingdotseq 1.1\times10^{3}$ cells). (b) Genome editing efficiencies at the *CAN1* gene. Efficiencies were calculated based on canavanine assays. The numbers indicate the distances (bp) of edited sites from the cleavage site. (c) Transformation efficiencies. The numbers of transformants on selective medium divided by the number of competent cells counted on non-selective medium are represented to evaluate the toxicity of CRISPR systems. (d) Editing efficiencies at the different targeting sequence on the opposite strand of the *CAN1* gene. (e) Theoretical editable bases by CRISPR Nickase system. X-axis indicates the hypothetical editable distances from cleavage sites. Y-axis shows theoretical editable bases (%) in the *S. cerevisiae* genome. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements.

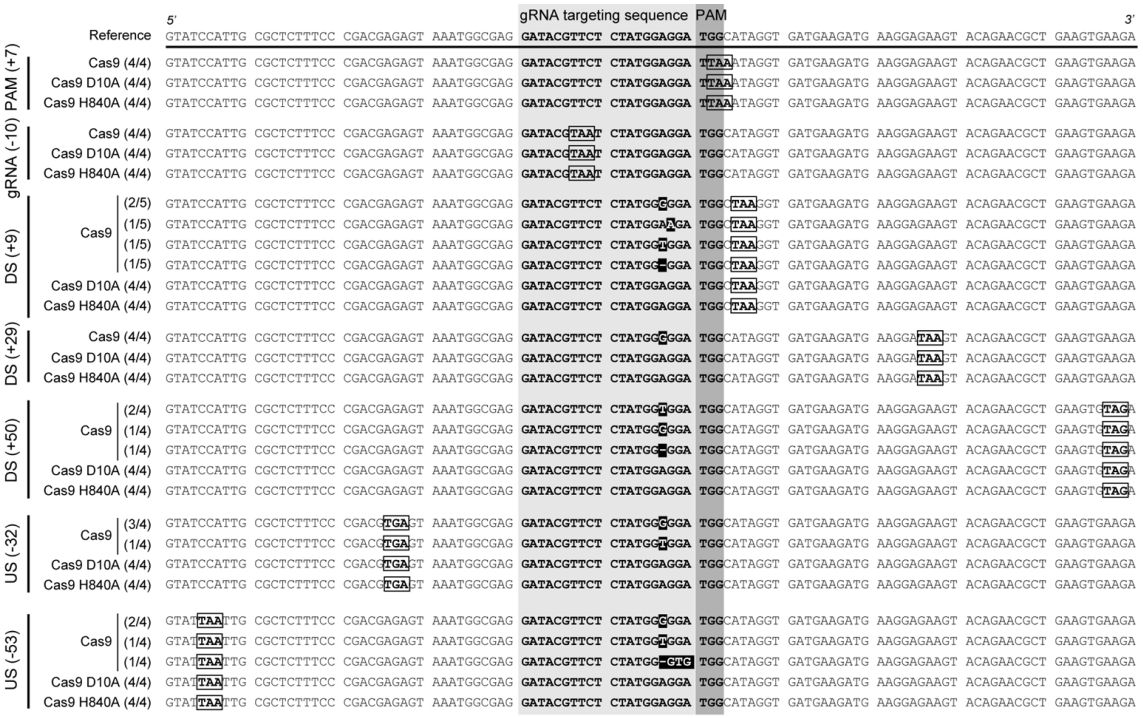

Fig. 3 Sequence analysis Introduced stop codons (squared) and unintended mutations (inverted) are represented. The numbers of observed sequences against the numbers of total sequenced strains are shown. DS and US represent downstream and upstream from the cleavage site, respectively.

These results suggest that CRISPR/Cas9 system can edit only the inside of gRNA-targeting sequence (20 bp) and PAM sequence (NGG; 2 bp). According to *in silico* calculation, editable bases by CRISPR/Cas9 system accounts for 68.4% in whole *S. cerevisiae* genome (Fig. 2e). The rest of 31.6% areas were the areas that were not covered by any PAM or potential gRNA-targeting sequences. In contrast, the CRISPR Nickase system, which can precisely edit much broader areas can theoretically edit 97.2% bases in the *S.cerevisiae* genome, according to a similar *in silico* calculation (Fig. 2e). This system overcame the major limitation of CRISPR/Cas9 system.

Specificity of CRISPR Nickase system

CRISPR/Cas9 system has been reported to induce unintended mutations at off-target sites that have sequence similarities with the on-target site [8, 10, 20]. On the other hand,

CRISPR Nickase system was supposed not to induce off-target editing, because induction of nicks at the off-target site unlikely causes NHEJ [21]. In order to investigate the specificity of CRISPR/Cas9 and CRISPR Nickase systems, I integrated an artificial off-target site at the upstream of an *EGFP* gene. The off-target site had the same sequence with the on-target site in the indigenous *CAN1* gene (Fig. 4a). Since on-target site in the *CAN1* gene was repaired through HDR using donor DNA containing the mutation in the PAM sequence, Cas9 and Cas9-nickase procutions did not cause cellular death in the absence of the off-target site (Fig. 4b).

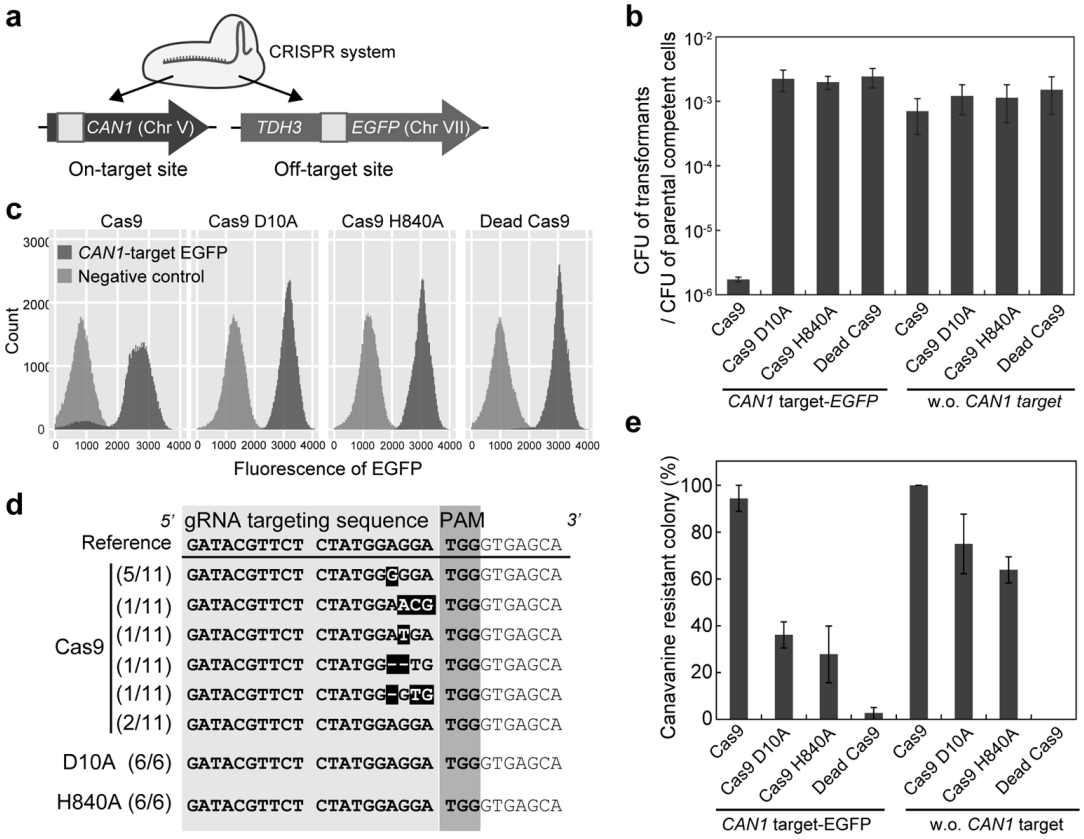

Fig. 4 Specificity of CRISPR Nickase system (a) Scheme of specificity assay. (b) Transformation efficiencies. The numbers of transformants on selective medium divided by the number of competent cells counted on non-selective medium are represented to evaluate the toxicity of CRISPR systems. (c) Fluorescent analysis of EGFP. Histograms are represented based on the fluorescent intensities of the cells. (d) Sequencing of the off-target site. Unintended mutations are inverted. The numbers of observed sequences against the numbers of total sequenced strains are shown. (e) Genome editing efficiencies at the on-target *CAN1* site. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements.

In the presence of off-target site, however, CFU of Cas9 producing cells was severely decreased due to DSB at the off-target site, while those of nickase- producing cells were not decreased. This fact indicates that CRISPR Nickase system did not induce cellular toxicity and DSBs at the off-target site. Fluorescence analysis revealed that some cells producing Cas9 lost fluorescence (Fig. 4c). Unnecessary mutations were found even in the rest of fluorescent cells (Fig. 4d). In contrast, CRISPR Nickase system induced neither fluorescent losses nor unintended mutations at the nicking site (Fig. 4c, d). In addition, this system edited the on-target site without any errors (Fig. 4e). These results suggest that CRISPR Nickase system can distinguish on-target and off-target sites even if the both sequences are exactly the same.

Rapid constructions of mutants

CRISPR Nickase system was revealed to have the advantages for genome-wide base editing and specificity over CRISPR/Cas9 system. This system, however, still had the difficulty for the rapid constructions of mutants. In this system, custom-made plasmids bearing donor DNA and gRNA expression cassette have to be constructed based on the target site, which usually takes about one-week efforts (Fig. 5a). In order to achieve the rapid construction of mutants, I combined yeast gap repair cloning (GRC) with CRISPR Nickase system (Fig. 5b). GRC is a genetic technique that depends on the ability of yeast to assemble DNA fragments to construct plasmids in the cells [22]. It only requires short overlapping sequences at the both edges of the fragments that can be easily added by PCR. Parts of plasmid fragments were amplified to attach a target-specific gRNA sequence and mutated sequence as well as overlapping sequences to each fragment and directly introduced into yeast cells. I assembled the Cas9, Cas9 D10A, Cas9 H840A, or Dead Cas9 expressing cassette as well as gRNA expressing cassette and donor DNA to introduce stop codons at the PAM sequence of the *CAN1* gene. CRISPR/Cas9 system as well as CRISPR/Nickase system coupling with GRC achieved precise genome editing at the *CAN1* gene in only 5 days (Fig. 5c). In addition to the *can1Δ* mutant, I constructed thermotolerant *CDC25* mutants. Importantly, CRISPR Nickase system coupling with GRC precisely introduced mutations regardless of the inside (W1416C) or the outside (T943P, G1459C, and N1393T) of PAM and gRNA-targeting sequences. The *CDC25* mutants exhibited thermotolerance (Fig. 5d). The phenotypes did not differ from those of *CDC25* mutants constructed by the two-step insertion/excision technique (Fig. 5d).

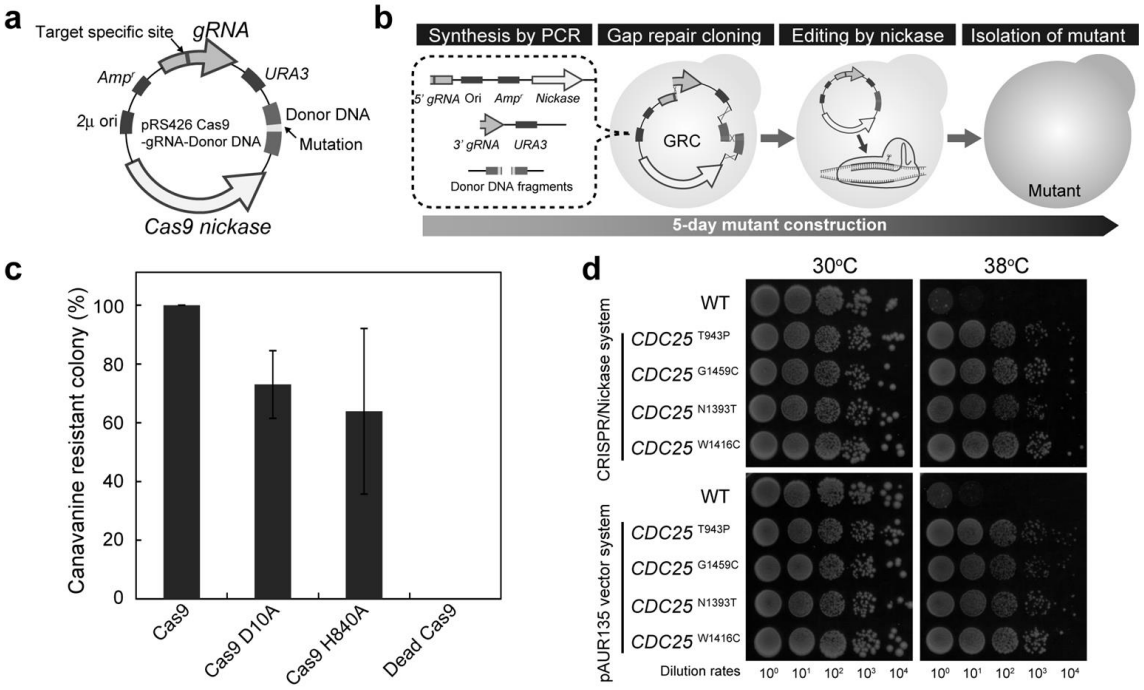

Fig. 5 CRISPR Nickase system coupling with GRC (a) Plasmid design for CRISPR Nickase system. (b) Scheme of CRISPR Nickase system coupling with GRC to construct mutants in 5 days. (c) Efficiencies of genome editing at the *CAN1* gene coupling with GRC. (d) Thermotolerance of *CDC25* mutants constructed through CRISPR Nickase system coupling with GRC. Cells (OD_{600} = 1.0) were diluted 10-, 10^2-, 10^3-, and 10^4-fold and spotted onto YPD plates. Since the major differences in efficiencies between Cas9 D10 and Cas9 H840A had not been observed so far, Cas9 D10A was used in this experiment. The *CDC25* mutants constructed through CRISPR Nickase system and a conventional two-step integration/excision method are compared. The error bars are shown as ± standard error mean (SEM) based on 3 independent measurements.

Investigation of key proteins for nick-repair via HDR

The key proteins for the repair of nicks via HDR have not been elucidated in *S. cerevisiae*. Rad51p and Rad52p play important roles in repairs of DSBs via HDR [23]. Consistent with this fact, Cas9 production in each *rad51* and *rad52* knockout mutant strain greatly inhibited colony formations even in the presence donor DNA (Fig. 6a). In *rad52Δ* strain, CRISPR Nickase system could not induce genome editing at all, suggesting that Rad52p plays critical roles in

repairs of nicks via HDR (Fig. 6b). This fact indicates that promotion of cell cycles in the liquid medium enhanced genome editing efficiencies by proving the opportunities with cells to undergo S/G2 phases at which *RAD52* is expressed [24]. On the other hand, CRISPR Nickase system was able to induce genome editing in the absence of Rad51p. Supporting this, this system could induce genome editing without Rad54p and Rad56p, which function with Rad51p as "recombinosome complex" (Fig. 6b) [23]. These results strongly suggest that repairs of single strand breaks via HDR are dependent on Rad52p but are independent of Rad51p.

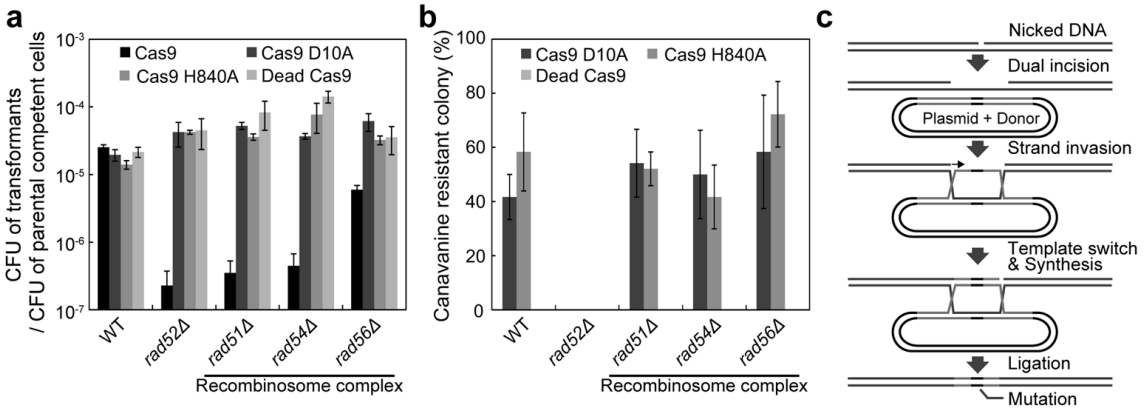

Fig. 6 Essential proteins for the repair of nicks via HDR (a) Transformation efficiencies in *rad*-deleted strains. The targeting site by Cas9, Cas9 D10A, Cas9 H840A, and dead Cas9 was the same as in Fig. 2b. The donor DNA in the plasmids contained a stop codon in the PAM sequence. The numbers of transformants on selective medium divided by the number of competent cells counted on non-selective medium are represented to evaluate the necessities of the Rad-proteins for the repairs of DSB and nick. (b) Genome editing efficiencies at the *CAN1* gene by CRISPR Nickase system in *rad*-deleted strains. (c) Working model of HDR at single strand breaks. The error bars are shown as ± standard error mean (SEM) based on more than 3 independent measurements.

Discussion

In this study, I successfully developed the genome-wide base editing technology, CRISPR Nickase system. This system was able to edit bases where the conventional CRISPR/Cas9 system was unable to edit. Coupling with GRC, I was able to construct the thermotolerant

CDC25 mutants within only 5 days. To my best knowledge, this is the fastest method to construct yeast mutants without using any maker genes.

CRISPR/Cas9 system was unable to edit the outside areas of PAM and gRNA-targeting sequences due to the continuous cleavages (Fig. 1d, 3). One possible way to avoid continuous cleaves may be to induce Cas9 expression only transiently to prevent prolonged Cas9 cleavages. However, mutation frequencies achieved by the temporal expression of Cas9 were less than 0.1% [7], probably because transient induction would not serve enough Cas9 protein for efficient genome editing. Previous report suggested that Cas9 abundance in the cells was the one of the most determinant factors to decide editing efficiency [25]. For efficient genome-wide base editing, transcription activator-like effector nuclease (TALEN) would be suited [26, 27]. Since TALEN technology does not require any motif sequences like the PAM sequence, TALEN can be constructed to target all areas in genome in contrast to CRISPR/Cas9 system. However, unlike CRISPR/Cas9 system it needs a correct alignment of binding modules which requires time-consuming DNA manipulations. Although use of other types of CRISPR-related systems with different PAM requirement will expand the editable bases, the utilization of different systems based on the targeting-sites is laborious and makes it low through-put. As a similar approach to CRISPR Nickase system, double nicking strategy using the Cas9 nickases has been developed [28]. This strategy uses a paired Cas9 nickases which introduce nicks at each DNA strand, causing site-specific DSBs. This technique is highly specific because DSBs occur only when two nickases simultaneously work. However, at least either recognition sequence should be altered after HDR not to induce unwanted mutations at the cleavage site, which limits editable bases. On the other hand, CRISPR Nickase system easily and rapidly edited the PAM, gRNA-targeting sequences, and the outside areas of these sequences in a precise manner. This system is robust and versatile tool for genome-wide base editing.

I revealed that nick repairing via HDR was independent of Rad51p, although HDR at DSB was profoundly dependent on Rad51p (Fig. 6a) [23]. Little has been known about the mechanisms of HDR at nicks in human cells and even in *S. cerevisiae*. A previous study showed that transcription-coupled repair (TCR) played a part of the roles in HDR at nicks in human cells [29]. TCR preferentially detects and repairs damages on the transcribed DNA strands [30]. The study also showed that the inhibition of Rad51p enhanced the TCR efficiency. However, any biases in the efficiencies between Cas9 D10A and Cas9 H840A were not observed in this

study. In addition, this study revealed that Rad51p was not involved in HDR at nicks at all. There will be differences in the nick repair mechanisms via HDR between human and *S. cerevisiae* cells. In the absence of Rad51p, break-induce replication (BIR) has been reported to be able to repair DSBs using homologous donor DNA [31]. BIR is a recombinant-dependent mechanism which is initiated from DNA synthesis [32]. In BIR, however, site-specific local recombination has not been observed. Instead, the repair process continues to the end of the template DNA, typically up to telomere domains when donor is chromosomal DNA [31]. Nevertheless, in this study, integrations of plasmid backbone were not found. This fact suggests that BIR was not involved in HDR at nicks in *S. cerevisiae*. I consider that post-replication repair (PRR) [33] or a similar mechanism would play a role in HDR at nicks (Fig. 6C). In this model, the first driving force of the recombination is a dual incision, which is observed in nucleotide excision repair [34]. Subsequently, the intact single strand DNA invades the undamaged donor DNA, which was harbored in the plasmid in this study. The nicked strand switches the template and extends the strand according to the donor DNA, as is observed in PRR [33]. Consequently, the nick induces the genome editing. Supporting this hypothesis, PRR has been reported to be conducted in the absence of Rad51p, but inhibited in the absence of Rad52p [35].

In this study, I successfully developed a novel genome-wide base editing system. This is the first approach to edit the outside areas of PAM and gRNA-targeting sequences. When to edit the genomic areas where are not covered by PAM or gRNA-targeting sequences, this system can be the first choice in the genome editing. In addition, this is one of the fastest genetic techniques to induce genome editing in yeast. Considering the editable areas, high-fidelity, and rapidness for the mutant constructions, CRISPR Nickase system will be the robust tool that overcomes the limitations in CRISPR/Cas9 system.

References

1. Rothstein R. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. 1991; 194: 281-301.
2. Kawahata M, Amari S, Nishizawa Y, Akada R. A positive selection for plasmid loss in *Saccharomyces cerevisiae* using galactose-inducible growth inhibitory sequences. Yeast. 1999; 15: 1-10.

3. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science. 1999; 285: 901-906.
4. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337: 816-821.
5. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010; 79: 181-211.
6. Shibata A, Conrad S, Birraux J, Geuting V, Barton O, Ismail A, et al. Factors determining DNA double-strand break repair pathway choice in G2 phase. EMBO J. 2011; 30: 1079-1092.
7. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013; 41: 4336-4343.
8. Fu YF, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol. 2014; 32: 279-284.
9. Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target effects in CRISPR/Cas9-mediated genome engineering. Mol Ther. 2015; 4: e264.
10. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, Thapar V, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol. 2015; 33: 187-197.
11. Tajima M, Nogi Y, Fukasawa T. Primary structure of the *Saccharomyces cerevisiae GAL7* gene. Yeast. 1985; 1: 67-77.
12. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science. 1999; 285: 901-906.
13. Ahmad M, Bussey H. Yeast arginine permease: nucleotide sequence of the *CAN1* gene. Curr Genet. 1986; 10: 587-592.
14. Cong L, Ran FA, Cox D, Lin SL, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339: 819-823.
15. Standage-Beier K, Zhang Q, Wang X. Targeted large-scale deletion of bacterial genomes using CRISPR-nickases. Acs Synth Biol. 2015; 4: 1217-1225.
16. Caldecott KW. Single-strand break repair and genetic disease. Nat Rev Genet. 2008; 9: 619-631.
17. Smith AM, Takeuchi R, Pellenz S, Davis L, Maizels N, Monnat RJ, et al. Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI

LAGLIDADG homing endonuclease. Proc Natl Acad Sci USA. 2009; 106: 5099-5104.

18. Manivasakam P, Weber SC, McElver J, Schiestl RH. Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. Nucleic Acids Res. 1995; 23: 2799-2800.
19. Heyer W-D, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. Annu Rev Genet. 2010; 44: 113-139.
20. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature. 2016; 529: 490-495.
21. Certo MT, Ryu BY, Annis JE, Garibov M, Jarjour J, Rawlings DJ, et al. Tracking genome engineering outcome at individual DNA breakpoints. Nat Methods. 2011; 8: 671-676.
22. Eckert-Boulet N, Pedersen ML, Krogh BO, Lisby M. Optimization of ordered plasmid assembly by gap repair in *Saccharomyces cerevisiae*. Yeast. 2012; 29: 323-334.
23. Paques F, Haber JE. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev. 1999; 63: 349-404.
24. Chen F, Nastasi A, Shen Z, Brenneman M, Crissman H, Chen DJ. Cell cycle-dependent protein expression of mammalian homologs of yeast DNA double-strand break repair genes Rad51 and Rad52. Mutat Res. 1997; 384: 205-211.
25. Steyer B, Carlson-Stevermer J, Angenent-Mari N, Khalil A, Harkness T, Saha K. High content analysis platform for optimization of lipid mediated CRISPR-Cas9 delivery strategies in human cells. Acta Biomater. 2015; 34: 143-158.
26. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009; 326: 1501.
27. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009; 326: 1509-1512.
28. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013; 154: 1380-1389.
29. Davis L, Maizels N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. Proc Natl Acad Sci USA. 2014; 111: 924-932.
30. Li S. Transcription coupled nucleotide excision repair in the yeast *Saccharomyces cerevisiae*: The ambiguous role of Rad26. DNA Repair. 2015; 36: 43-48.
31. Malkova A, Ivanov EL, Haber JE. Double-strand break repair in the absence of *RAD51* in yeast: A possible role for break-induced DNA replication. Proc Natl Acad Sci USA. 1996; 93: 7131-7136.

32. Bosco G, Haber JE. Chromosome break-induced DNA replication leads to nonreciprocal translocations and telomere capture. Genetics. 1998; 150: 1037-1047.
33. Prakash L. Characterization of postreplication repair in *Saccharomyces cerevisiae* and effects of rad6, rad18, rev3 and rad52 mutations. Mol Gen Genet. 1981; 184: 471-478.
34. Swanson RL, Morey NJ, Doetsch PW, Jinks-Robertson S. Overlapping specificities of base excision repair, nucleotide excision repair, recombination, and translesion synthesis pathways for DNA base damage in *Saccharomyces cerevisiae*. Mol Cell Biol. 1999; 19: 2929-2935.
35. Mott C, Symington LS. *RAD51*-independent inverted-repeat recombination by a strand-annealing mechanism. DNA Repair. 2011; 10: 408-415.

Conclusions

The goal of this study was to identify mutations that contribute to the thermotolerance of *S. cerevisiae* and to reconstruct thermotolerance. There have been two main challenges in the current research of thermotolerance in yeast. First, few mutations related to thermotolerance have been identified, even though thermotolerant yeasts play essential roles in bioproduction. Second, the current method of introducing mutations into the yeast genome is inefficient and time-consuming. In this study, I addressed both of these problems by analyzing the genome sequences of strains isolated during stepwise adaptation, and developing a new genome-wide base editing technology.

In Chapter I, I performed a stepwise adaptation from a non-thermotolerant strain and then isolated thermotolerant strains. A DNA microarray analysis of the one of the most thermotolerant strains, YK60-1, revealed the constitutive induction of stress-responsive genes, such as those encoding heat shock proteins and trehalose synthetic genes. Additionally, the non-targeting metabolome analysis showed that YK60-1 accumulated much more trehalose than MT8-1.

In Chapter II, I presented whole genome analyses of intermediate strains that had adapted from 32°C to 38°C in the adaptation steps. The analyses discovered *CDC25* mutations that contributed to thermotolerance. The pathway analyses of the reconstructed *CDC25* mutants showed that mutated Cdc25p caused the downregulation of the intracellular level of cAMP, which led to thermotolerance.

In Chapter III, I developed a new genome editing technology that was capable of reconstruction of thermotolerant strains in much shorter period of time and with higher efficiencies than the current technique. The novel CRISPR/Nickase system was able to precisely and specifically edit the outside areas of PAM and guide RNA targeting sequences, in contrast with the CRISPR/Cas9 system.

In this study, I successfully identified mutations involved in the thermotolerance of yeast, which potentially reduces the costs for bioproduction. I also developed a new genome-wide base editing technology for the fast and efficient phenotypic designs of productive yeast strains.

Acknowledgements

This thesis was submitted by the author to Kyoto University for the Doctor degree of Agriculture. The studies presented here have been carried out under the direction of Professor Mitsuyoshi Ueda in the Laboratory of Biomacromolecular Chemistry, Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, during 2011-2017.

I would like to express greatest appreciation to Mitsuyoshi Ueda who offered insightful comments, meaningful discussion, and continuous support throughout the course of my study.

I would like to thank associate professor Kouichi Kuroda who always gave me detailed and insightful comments to my study. I am grateful to assistant professor Wataru Aoki and late Hironobu Morisaka for their critical comments, which greatly improved my study. I also appreciate Yoshiaki Katsuyama. I could not do anything with his huge efforts to breed thermotolerant strains. I express my appreciation to the secretary Fukuko Suzuki for her support of my laboratory life.

I am also grateful to all my colleagues in the laboratory for wide-ranging discussion, encouragements, and supports. I would like to thank financial supports from the research fellowship of the Japan Society for the Promotion of Science for Young Scientists.

Lastly, I would like to express my appreciation to my parents and my wife Miki who always warmly supported me and gave a chance to climb to the shoulders of giants.

Atsushi SATOMURA

Laboratory of Biomacromolecular Chemistry
Division of Applied Life Sciences
Graduate School of Agriculture
Kyoto University

Publications

Chapter I

Atsushi Satomura, Yoshiaki Katsuyama, Natsuko Miura, Kouichi Kuroda, Ayako Tomio, Takeshi Bamba, Eiichiro Fukusaki, Mitsuyoshi Ueda, Acquisition of thermotolerant yeast *Saccharomyces cerevisiae* by breeding via stepwise adaptation. *Biotechnology Progress*, 29 (5), 1116–1123, 2013.

Chapter II

Atsushi Satomura, Natsuko Miura, Kouichi Kuroda, Mitsuyoshi Ueda, Reconstruction of thermotolerant yeast by one-point mutation identified through whole-genome analyses of adaptively-evolved strains. *Scientific Reports*, 6, 23157, 2016.

Chapter III

Atsushi Satomura, Ryosuke Nishioka, Hitoshi Mori, Kosuke Sato, Kouichi Kuroda, Mitsuyoshi Ueda, Precise genome-wide base editing by CRISPR Nickase system in yeast. *Submitted*.

Other publication

Atsushi Satomura, Kouichi Kuroda, Mitsuyoshi Ueda, Environmental stress tolerance engineering by modification of cell surface and transcription factor in *Saccharomyces cerevisiae*. *Current Environmental Engineering*, 1 (3), 149-156, 2014.

Atsushi Satomura, Kouichi Kuroda, Mitsuyoshi Ueda, Generation of a functionally distinct *Rhizopus oryzae* lipase through protein folding memory. *PLOS ONE*, 10 (5) e0124545, 2015.

Ryosuke Nishioka, Atsushi Satomura, Junki Yamada, Kouichi Kuroda, Mitsuyoshi Ueda, Rapid preparation of mutated influenza hemagglutinins for influenza virus pandemic prevention. *AMB Express*, 6 (10), s13568, 2015.