공학박사 학위논문

Differential roles of two isoforms of Gcr1 transcription factor generated from spliced and un-spliced transcripts in *Saccharomyces cerevisiae*

*Saccharomyces cerevisiae*에서
스플라이싱의 유무에 따라 생성된 두 가지
Gcr1 전사 조절 인자의 역할 규명

2023 년 2 월

서울대학교 대학원
공과대학 화학생물공학부
차 승 우

Differential roles of two isoforms of Gcr1 transcription factor generated from spliced and un-spliced transcripts in *Saccharomyces cerevisiae*

by

Seungwoo Cha

Advisor : Professor Ji-Sook Hahn, Ph.D.

Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Seoul National University

February, 2023

School of Chemical and Biological Engineering

Graduate School

Seoul National University

Differential roles of two isoforms of Gcr1 transcription factor generated from spliced and un-spliced transcripts in *Saccharomyces cerevisiae*

지도교수 한 지 숙

이 논문을 공학박사 학위논문으로 제출함

2023년 2월

서울대학교 대학원

공과대학 화학생물공학부

차 승 우

차승우의 공학박사 학위논문을 인준함

2023 년 2 월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Abstract

Differential roles of two isoforms of Gcr1 transcription factor generated from spliced and un-spliced transcripts in *Saccharomyces cerevisiae*

Seungwoo Cha
School of Chemical and Biological Engineering
The Graduate School
Seoul National University

As a Crabtree-positive species, regulation of glycolysis in *Saccharomyces cerevisiae* is important for its adaptation to glucose, the most preferable carbon source. The most important transcription factor for glycolytic gene activation is Gcr1, which was recently revealed to have two isoforms, $Gcr1^{U}$ and $Gcr1^{S}$.

Gcr1^{S} is the major form produced from spliced mRNAs, whereas, Gcr1^{U} is produced from unspliced mRNA, using a translation start codon located in the middle of the intron. In this study, the differential roles of Gcr1^{U} and Gcr1^{S} were identified and applied to lactic acid production.

First, the CRISPR/Cas9-based genome editing strategy was used in order to generate the strains expressing only Gcr1^{U} or Gcr1^{S} isoform. Although these two strains did not show any growth defects and their binding target genes were almost the same, their DNA binding patterns were different. However, by investigating the effects of deleting *GCR2*, a coactivator of Gcr1, or functional domains of Gcr1 or Gcr2, it was revealed that Gcr1^{U} monomer forms an active complex with its coactivator Gcr2 homodimer, whereas Gcr1^{S} mainly acts as a homodimer without Gcr2.

Second, it was discovered that the USS (un-spliced form specific) domain, 55 residues at the N-terminus existing only in Gcr1^{U}, inhibits the homodimer formation of Gcr1^{U}. This domain even inhibits dimerzation of Gcr1^{S}, acting in trans. The Gcr1^{S} monomer inhibits the metabolic switch from fermentation to respiration by directly binding to the *ALD4* promoter, which

can be restored by overexpression of the *ALD4* gene, encoding a mitochondrial aldehyde dehydrogenase required for ethanol utilization. This respiratory defect also observed when $Gcr1^{S}$ mRNA is overexpressed, indicating that it is important to control the expression level of $GCR1^{S}$ for the respiratory metabolism.

Lastly, the respiratory defect caused by irregular expression of $Gcr1^{S}$ can be beneficial to produce NADH consuming pyruvate derivatives, by mimicking the microaerobic culture condition. The overexpression of $GCR1^{S}$ in the lactic acid producing background strain increased the titer even under aerobic condition. This result shows opposite trend that $GCR1^{WT}$ overexpression under aerobic condition rather decreased the lactic acid production, indicating that the overexpression of proper isoform of Gcr1 can improve the production of traget compound.

Keywords : *Saccharomyces cerevisiae*, Alternative splicing, CRISPR/Cas9, diauxic shift, Gcr1, Gcr2, glycolysis, glycolytic flux regulation, respiration

Student Number : 2016-21056

Contents

List of Tables

List of Figures

Chapter 1.

Research background and objective

Saccharomyces cerevisiae is the organism that is mostly used as a eukaryotic model system. Since it is a Crabtree-positive species, it consumes glucose and preferably produces ethanol as a fermentation product even under aerobic condition. When it consumes glucose and undergoes glycolysis, other cellular metabolisms such as respiration, TCA cycle, and other carbon source dissimilation pathways are downregulated, which is named 'glucose repression.' Therefore, sensing the external glucose and transmitting the signal to glycolytic pathway is important for *S. cerevisiae*.

The most important transcription factor for glycolytic gene activation is Gcr1, which was recently revealed to have two isoforms, $Gcr1^{U}$ and $Gcr1^{S}$. Actually, *GCR1* gene is known to produce at least seven spliced mRNA isoforms, among which only two generate in-frame functional proteins. One of the two major in-frame protein encoding mRNAs is named $Gcr1^{S}$. Even unspliced mRNA of *GCR1* can be translated using the start codon located in the middle of the intron, which is named $Gcr1^{U}$. In brief, two major isoforms of Gcr1 are produced: $Gcr1^{U}$ and $Gcr1^{S}$.

The first objective of this study was to investigate the role of each isoform of Gcr1. Although the previous study that revealed the two major isoforms of Gcr1 performed the same

study with plasmid expression, the CRISPR/Cas9 based genome editing was used in this study in order to minimize the perturbation due to the plasmid usage. The most native state expression of $Gcr1^U$ or $Gcr1^S$ isoforms did not show any growth defect in rich or minimal media and their target genes were same. However, by anlyzing the ChIP-seq and RNA-seq data during exponential phase and after diauxic shift, it was able to conclude that $Gcr1^U$ and $Gcr1^S$ mostly bind to the same targets, but with different binding strengths depending on growth phase. Additional protein domain or *GCR2* gene deletion revealed that $Gcr1^U$ mainly interacts with Gcr2, whereas $Gcr1^S$ forms a homodimer via leucine zipper domain.

The second objective of this study was to discover the role of $Gcr1^S$ isoform, especially in respiratory mechanism. Focusing on the USS domain, which only exists in the N-terminal of $Gcr1^U$ isoform, it was revealed to mask the leucine zipper domain of $Gcr1^U$ protein and inhibits homodimer formation. Even the USS domain could mask the leucine zipper domain of $Gcr1^S$ protein in *trans*-acting manner, resulting in the respiratory defect. Together with the domain deletion study, it was able to conclude that the impaired homodimerization of $Gcr1^S$ form causes a defect in respiratory metabolism. It was also confirmed that overexpression of $GCR1^S$ also makes the

cell not to undergo respiration, indicating that the delicate regulation of *GCR1*^{S} expression is important for the respiratory mechanism.

The last objective of this study was to apply the differential regulation of two isoforms of Gcr1 into the proudction of NADH consuming pyruvate derivative, lactic acid. The microaerobic culture had been used in order to decrease the respiratory metabolism so that unoxidized NADH can be used to convert pyruvate into lactic acid. Here, application of the respiratory defect caused by impaired Gcr1^{S} homodimer successfully increased the amount of lactic acid even without using microaerobic culture. Final 10.38 g/L of lactic acid was produced from 20 g/L glucose by overexpressing *GCR1*^{S}, whereas *GCR1*^{WT} overexpression rather decreased the production of lactic acid to 8.08 g/L under aerobic condition (170 rpm).

The objectives of this study are summarized as follows.

1. To investigate the differential role of Gcr1^{U} and Gcr1^{S} in most native state
2. To emphasize the delicate regulation of Gcr1^{S} and its effect to respiratory metabolism
3. To improve the production of lactic acid via applying the respiratory defect caused by *GCR1*^{S} overexpression

Chapter 2.

Literature review

2.1. Fermentation and respiration of *S. cerevisiae*

2.1.1. Aerobic fermentation

Definition of Crabtree effect

Crabtree effect is a metabolic phenomenon that cells repress respiratory pathway and prioritize fermentation at the high glucose concentration (1, 2). The most common cellular metabolism in aerobic condition is tricarboxylic acid (TCA) cycle, also known as the Krebs cycle or citric acid cycle, which produces energy by subsequent oxidative phosphorylation (Figure 1A) (3). However, in the case of Crabtree-positive organisms, cells accelerate glycolytic pathway even under aerobic conditions, resulting in enough amounts of ATPs from the substrate level phosphorylation. This decreases the necessity of oxidative phosphorylation, so the cells produce large amount of ethanol from pyruvate (4, 5). Many of the yeast species shows Crabtree-positive metabolism, but some species like *Kluyveromyces marxianus* and *Candida utilis* does not show Crabtree effect and called as Crabtree-negative or respiratory yeasts (6).

Production of ethanol

Saccharomyces cerevisiae is one of the most well-known Crabtree-positive microorganisms that has a strong tendency to metabolize fermentable carbon sources such as glucose into ethanol via glycolysis and fermentation even under aerobic conditions (2). After two pyruvate molecules are produced from glucose by glycolysis, they are used to produce ethanol and other cellular components or transported into mitochondria by mitochondrial pyruvate carrier complexes (MPC). Under fermentative conditions, MPC has low activity, so most of the pyruvate is converted into ethanol by pyruvate decarboxylase (Pdc1, 5, 6) and alcohol dehydrogenase (Adh1, 3, 4, 5, 6, 7, Sfa1) enzymes (7-9) (Figure 1A). Adh enzyme uses NADH as a cofactor during ethanol production, so that the NAD+ consumed during glycolysis can be regenerated (Figure 2) (8). Even though ethanol producing pathway is strong in *S. cerevisiae*, acetyl-CoA should be produced since they are key precursor of many essential cellular components. Therefore, acetaldehyde can also be converted into acetyl-CoA via acetate in cytosol, which is also called PDH bypass (Figure 1B) (10, 11).

Production of glycerol

Glycerol is also a main byproduct of fermentation of *S. cerevisiae*. During glycolysis, fructose-1,6-bisphosphate aldolase (Fba1) enzyme splits fructose-1,6-bisphosphate (FBP) into dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP). GAP undergoes the payoff phase of glycolysis, and DHAP is converted into glycerol 3-phosphate (G3P) by glycerol-3-phosphate dehydrogenase (Gpd1, 2) enzyme (12, 13). This Gpd enzyme also utilizes NADH as a cofactor like Adh enzymes, enables the balancing between NAD^+ and NADH (14). G3P is used as an important intermediate for lipid synthesis or switched to glycerol by glycerol 3-phosphate phosphatase (Gpp1, 2) enzyme (15, 16) (Figure 2).

Therefore, *S. cerevisiae* had been evolved as a Crabtree-positive organism that contains well organized metabolism by producing both ethanol and glycerol, even with the repression of TCA cycle.

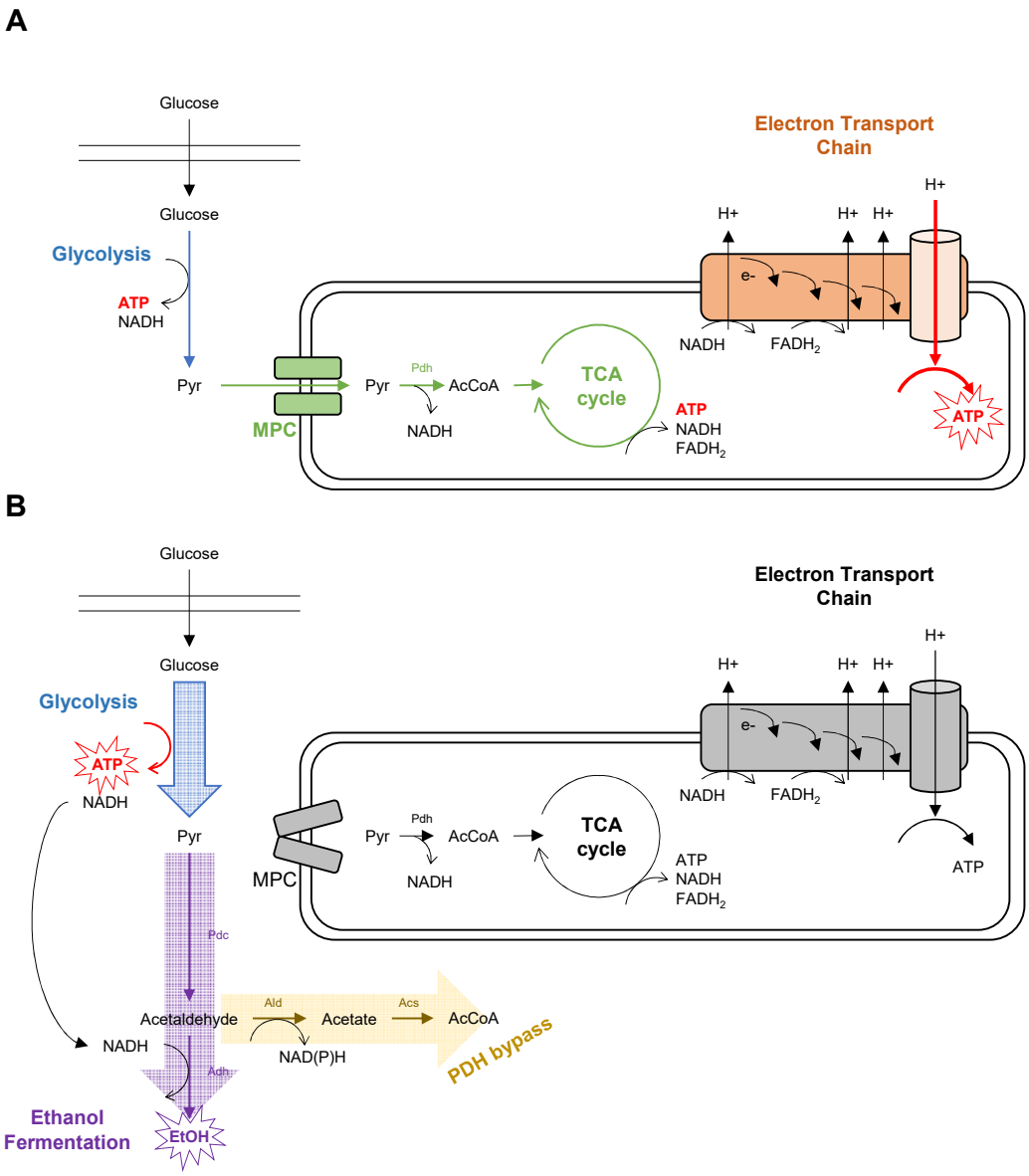

Figure 1. Schematic view of metabolic pathway of general respiration (A) and aerobic fermentation of *S. cerevisiae* (B)

(A) General cellular respiration under aerobic condition. After glycolysis, pyruvate goes into the mitochondria and undergoes TCA cycle, which produces NADH and FADH2. These electron donors provide proton motive force to electron transport chain, producing ATP synthesis.

(B) Aerobic ethanol fermentation pathway of S. cerevisiae. Crabtree effect blocks the respiration pathway, so pyruvate is converted into ethanol at cytosol. Acetyl CoA is produced by PDH bypass.

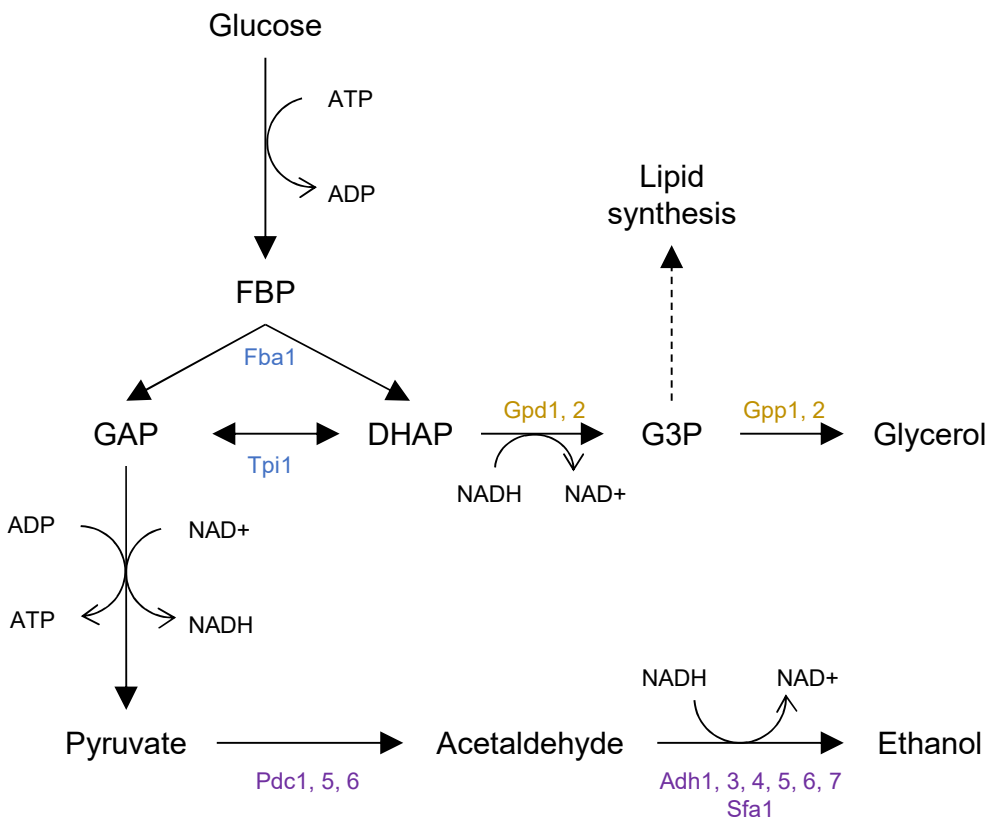

Figure 2. Simplified pathway of ethanol and glycerol production of *S. cerevisiae*

Metabolic pathway of aerobic fermentation that produces glycerol and ethanol from glucose. After glucose undergoes glycolysis, pyruvate molecule is converted into ethanol by Pdc and Adh enzymes. Additionally, glycerol is produced from DHAP through Gpd and Gpp enzymes.
FBP, fructose-1,6-bisphosphate; GAP, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate;, G3P, glycerol-3-phosphate

2.1.2. Aerobic respiration after diauxic shift

Unique features of respiration system in S*accharomyces cerevisiae* under aerobic condition

After consumption of fermentable carbon sources, *S. cerevisiae* switches its carbon metabolism to respiration to utilize its two major byproducts, ethanol, and glycerol, known as diauxic shift (17, 18). The most dramatic change is the expression level of *ADH2* gene that had been repressed during fermentation (19). After consuming all glucose, its expression derepressed and begins to utilize ethanol by converting it into acetaldehyde. Also, the Pdc enzymes are inhibited and MPC gets more activated, so that cytosolic pyruvate has more tendency to enter the mitochondria and converted into acetyl-CoA (7). These two fluxes allow the TCA cycle and produce NADH like a general oxidative phosphorylation. However, the absence of complex I in electron transport chain (ETC) results some several differences compared to the common respiratory pathway (20).

In common, NADH is preferentially used as a high energy containing electron donor which pumps out protons through the complex I. However, in *S. cerevisiae*, the complex I of electron transfer chain does not exists, so the NADH is not involved in the H+ gradient but only oxidized to NAD+ by two

kinds of alternative mitochondrial NADH dehydrogenase: located in external area (Nde1, 2) or in internal area (Ndi1) of mitochondrial membrane (21, 22). Cytosolic NADH can be oxidized by Nde1, 2, and mitochondrial NADH produced from TCA cycle is oxidized by Ndi1 (Figure 3). These unusual features cause lack of proton motive force (PMF) from mitochondrial NADH. Therefore, *S. cerevisiae* mainly uses two strategies mentioned below to compensate this lack of PMF with the ability to utilize cytosolic NADH.

Direct utilization of cytosolic NADH

In common, cytosolic NADH is impermeable to mitochondria. Therefore, shuttle pathways like malate-aspartate shuttle or G3P shuttle are used to couple cytosolic reducing cofactors into the mitochondria (23). In contrast, *S. cerevisiae*, two enzymes (Nde1, 2) located in the external area of mitochondrial membrane can directly oxidize cytosolic NADH and provide electrons to ETC even without the help of the shuttle system. When the ethanol is reutilized, ethanol is oxidized by Adh2 using NAD+ as a cofactor and produces cytosolic NADH (24, 25) (Figure 3). Also, the glycerol is oxidized to DHA by glycerol dehydrogenase (Gcy1) producing NAD(P)H (26). This accumulation of cytosolic NADH is

alleviated by Nde1 and Nde2 which oxidize NADH to regenerate cytosolic NAD+, and also by the G3P shuttle system explained below.

G3P shuttle coupled with FAD+ reduction

S. cerevisiae converts glycerol into DHAP via two pathways: oxidation of glycerol by Gcy1, as mentioned above, and phosphorylation of glycerol by glycerol kinase (Gut1) resulting in G3P (27). G3P is oxidized to DHAP by G3P dehydrogenase (Gut2) (28). Especially, Gut2 enzyme is located in the external area of mitochondrial membrane, linking the reduction of FAD+ with the oxidation of G3P. This non-standard mitochondrial G3P dehydrogenase (Gut2) enables bypassing the complex I pathway and efficiently transfer electrons (25, 29) (Figure 3). Gut2 forms a G3P shuttle system with Gpd1 and Gpd2 enzymes, contributing to the utilization of cytosolic NADH and regenerating the $FADH_2$ inside of the mitochondria (22, 29).

Single functional unit behavior of respiratory chain

In *S. cerevisiae*, this respiratory chain had been identified to behave as a single functional unit, rather than using quinone and cytochrome C redox pool. This kind of

supercomplex system which associates complex II, III, and IV together is not a general feature of ETC. However, concerning the Crabtree-positive feature of *S. cerevisiae*, this system is rather to be hypothesized as an ability to rapidly switch its carbon metabolism in response to the preferable carbon source (30-32).

In conclusion, *S. cerevisiae*, Crabtree-positive species which undergoes aerobic fermentation to produce ethanol and glycerol from glucose, re-utilizes its byproducts to obtain energy via respiration. Therefore, it is important to tightly regulate the glycolytic pathway and its transition to respiratory metabolism.

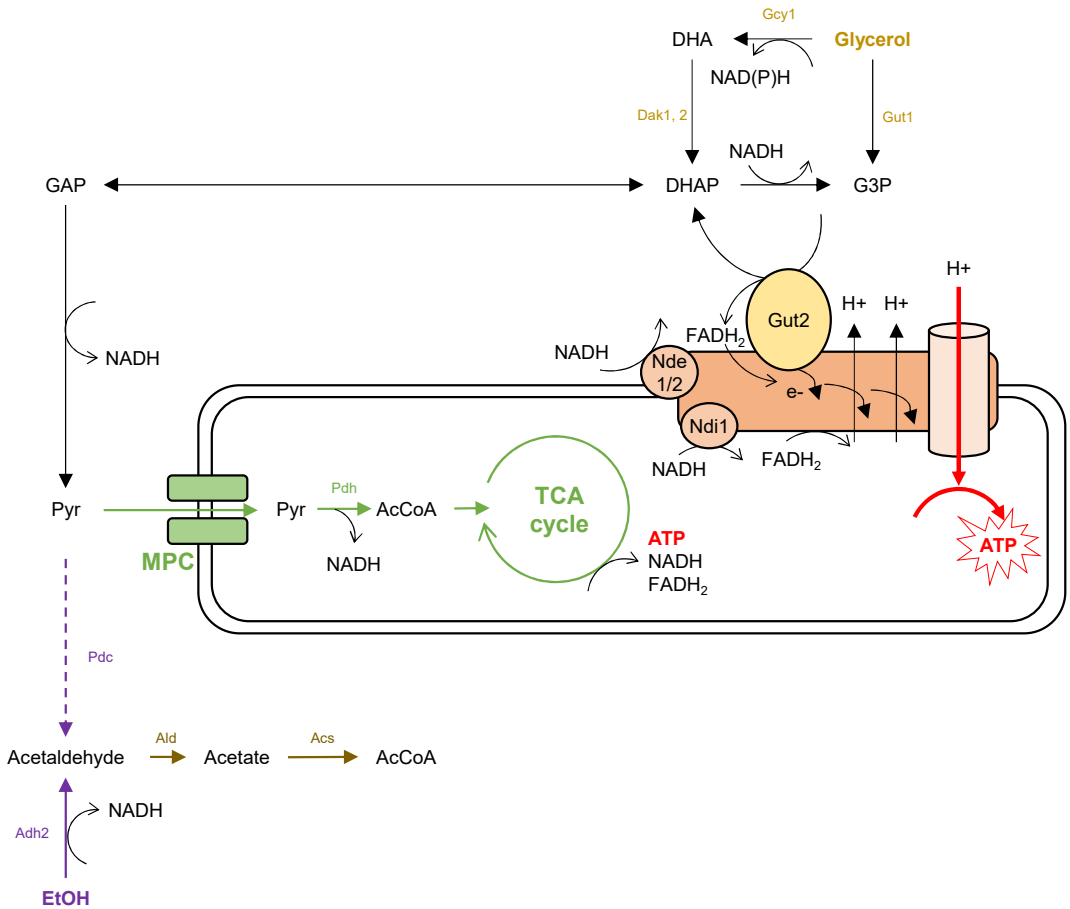

Figure 3. Aerobic respiration pathway after diauxic shift of *S. cerevisiae*

After *S. cerevisiae* consumes all glucose, it reutilizes ethanol and glycerol in order to obtain energy from aerobic respiration. Cytosolic NADH produced from ethanol oxidation is directly converted into NAD+ by Nde1/2 enzymes located in the external area of mitochondria. The mitochondrial NADH produced from TCA cycle is oxidized by Nde1 enzyme in the internal area of mitochondria. In the case glycerol utilization, Gut2 enzyme acts as a G3P shuttle that utilizes cytosolic NADH and regenerates $FADH_2$ in the mitochondria.

GAP, glyceraldehyde-3-phosphate; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate;, G3P, glycerol-3-phosphate; Pyr, pyruvate; AcCOA, acetyl-CoA

2.2. Cellular response to glucose in *S. cerevisiae*

2.2.1. Glucose sensing and glucose repression of *S. cerevisiae*

Since glucose is the most preferable carbon source of *S. cerevisiae*, cell contains several sensors and downstream signaling pathways in order to respond rapidly to the extracellular glucose. This transduction of signal activates the transcription of genes related to aerobic fermentation, also represses the metabolism for other carbon sources (33). Although *S. cerevisiae* can uptake and use several kinds of carbon sources such as sucrose, maltose, and galactose, the metabolisms for these carbon sources are inactivated during glucose fermentation (34-36). This phenomenon is named 'glucose repression', a key regulatory mechanism signal cascade of *S. cerevisiae* (37). Below, the three most studied pathways considered as a key regulatory mechanism in respond to glucose.

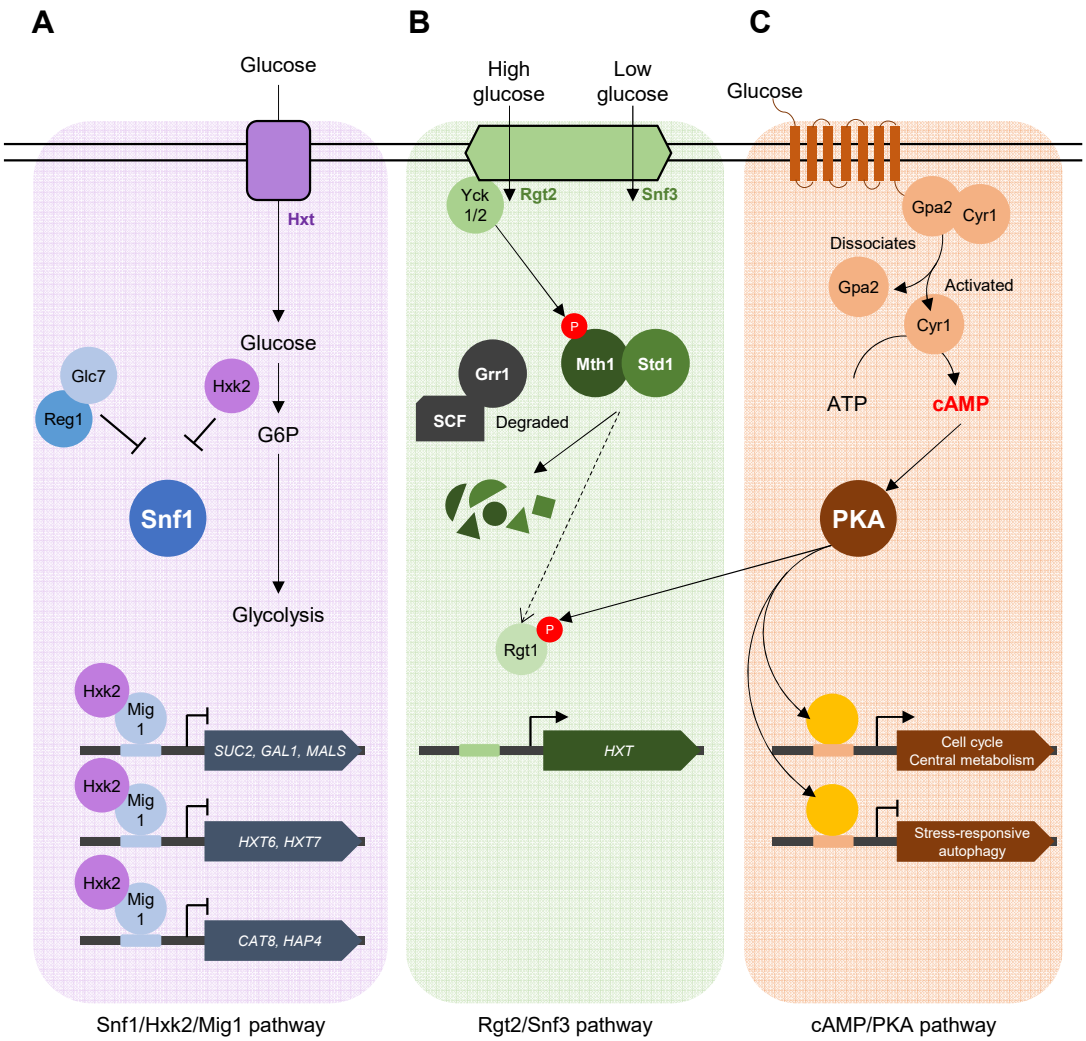

Figure 4. Three major glucose sensing pathways in *S. cerevisiae*

Three major glucose sensors and its signal transduction pathway involved in glucose repression. (A) Snf1/Hxk2/Mig1 pathway is activated by glucose transported by hexose transporter and enalbes the binding of transcriptional repressor, Mig1. (B) Rgt2/Snf3 recognizes high or low level of exteracellular glucose and disables the binding of transcriptional repressor Rgt1, resulting the derepression of hexose transporter genes. (C) cAMP/PKA pathway is related to GPCR that senses glucose, and activates the signal cascade that activates PKA. PKA is not only involved in Rgt1 phosphorylation, but also regulation of several metabolisms.

Snf1/Hxk2/Mig1 pathway (Figure 4A)

S. cerevisiae has been revealed to contain about 20 hexose transporter genes or their homology in its genome, and each of them has different affinity upon hexose and regulated in a different manner (38, 39). Several computational and experimental approaches identified that some of these hexose transporters also show ability to transport different kinds of hexose (glucose, mannitol, and sorbitol) (40). However, still the detailed functions of all HXT genes are not yet fully identified, and only Hxt1-7 had been well studied and defined as a major glucose transporters. When these seven genes are deleted, cells are known to fail to grow on glucose and have no glycolytic flux (41). Interestingly, these transporters show different affinity in accordance with glucose concentration: Hxt1 and Hxt3 are low affinity glucose transporter, Hxt6 and Hxt7 have high affinity to glucose, Hxt4 and Hxt5 show the moderate level of affinity, and Hxt2, expresses moderate to high affinity (42). The regulation of the genes encoding these hexose transporters is way more complicated. According to previous studies, *HXT1* is induced by high level of glucose, *HXT2* and *HXT4* are induced by low level of glucose, *HXT6* and *HXT7* are repressed by high level of glucose, and *HXT5* gene is expressed during the starvation condition. In the case of *HXT3* gene, its induction is

even independent to glucose concentration (43).

Right after the glucose is transported into the cytoplasm, it should be phosphorylated and converted into glucose-6-phosphate (G6P) in order not to lose the glucose from inside the cell. The most important kinase that phosphorylates glucose is hexokinase 2 (Hxk2), which is also known to downregulate the activity of Snf1 in coordination with Glc7 and Reg1 (44). Snf1 is the most well-known kinase which is mainly activated under low glucose condition (45). When the amount of glucose is enough, Hxk2 plays an important role to repress the activity of Snf1.

Hxk2 also stays inside the nucleus under high glucose condition and regulates the phosphorylation state of Mig1, which is the most important transcription factor for the glucose repression of S. cerevisiae (46, 47). When glucose presents, Mig1 mainly represses the glucose-repressed genes: genes involved in other carbon sources metabolism (*SUC2*, *GAL1*, and *MALS*), high affinity hexose transporters (*HXT6* and *HXT7)*, and other genes encoding transcription factors involved in the transcription of gluconeogenic genes (*CAT8*) or respiratory genes (*HAP4*) (48-52).

Rgt2/Snf3 pathway (Figure 4B)

Rgt2 and Snf3 are both major glucose sensors containing 12 transmembrane domains and involved in downstream signal transduction via C terminal domain. Although they are paralogues arose from the whole genome duplication, they sense different level of extracellular glucose (53, 54). Rgt2 detects the high glucose concentration and activates *HXT1* and *HXT6*, whereas Snf3 detects the low glucose concentration and induces the expression level of *HXT2* and *HXT4* (55). After these two sensors recognize the glucose, the membrane associated casein kinases, Yck1 and Yck2, phosphorylate Mth1 and Std1, and Grr1-SCF complex interacts with these phosphorylated substrates and ubiquitinates them, resulting the degradation of these proteins (56-58). Considering that Mth1/Std1 complex is a major co-repressor of a transcriptional repressor of *HXT* genes, Rgt1, the presence of glucose derepresses the *HXT* genes so that it can efficiently transport the glucose. When there is no glucose, the cascade does not work and Rgt1 represses *HXT* genes with Mth1/Std1 complex, so that unnecessary occupation of cell membrane does not occur (57, 59).

cAMP/PKA pathway (Figure 4C)

Gpr1 is one of the two G-protein coupled receptors (GPCRs) of S. cerevisiae, which contains seven transmembrane domains and coupled with Gα protein (Gpa2) and its negative regulator (Asc1 and Rgs2) (60, 61). While GPCR signal transduction does not limited to regulation of genes related with glucose metabolism, it largely mediates the phosphorylation of various proteins involved in the signal pathways mentioned above. After glucose binds with Gpr1, the G protein complex dissociates from Gpr1 and activates Cyr1 converts ATP into cAMP, also cause activation of protein kinase A (PKA) (62, 63). PKA is a well-known kinase that phosphorylates Rgt1, which is an important transcriptional repressor previously mentioned in Rgt2/Snf3 section (59, 64). Also, PKA is involved in the various regulatory pathway such as activation of cell cycle and central metabolism, or repression of transcription of stress-responsive and autophagy related genes (65, 66). Likewise, cAMP/PKA pathway is stimulated by glucose and regulates several downstream signaling pathways in order to adapt the cells to the nutrient-sufficient environment.

To conclude, *S. cerevisiae* contains a complex and elaborate signaling pathway that responds to glucose, preventing

the consumption of unnecessary energy so that the cell can fully focus on producing the energy according to the glucose. Still, other pathways including TOR (target of rapamycin) complex or 14-3-3 proteins (Bmh1 and Bmh2), are also investigated to participate in glucose signaling, the further detailed research is needed to fully understand the glucose signaling pathway (67-69).

2.2.2. Metabolic shift from fermentation to respiration

After *S. cerevisiae* consumes all glucose, the major byproduct during fermentation, ethanol, is used as a major carbon source. Therefore, the genes related with glucose fermentation should be repressed and respiratory pathway needs to be upregulated. Such a large whole cell reprogramming occurs when the metabolism shifts from fermentation to respiration. This re-wiring process must occur very rapid and precise manner, so that cells adopt to the carbon source change as soon as possible (17, 18). Snf1 and PKA are important mediator kinase proteins involved in regulation of several transcription factors or other mediators, resulting various transcriptional response.

Snf1 phosphorylates several zinc finger transcription factors related with diauxic shift (Mig1, Cat8, Sip4 and Adr1) and regulates their DNA binding ability (70-73). Phosphorylation of Mig1 inhibits the DNA binding of Mig1, resulting in the derepression of its target genes. This activates *CAT8* gene which encodes a positive regulator of *SIP4* and gluconeogenic genes. Also, Sip4 is an important transcription activator that also activates gluconeogenic genes with Cat8. Cat8 and Sip4 should be phosphorylated in order to activate their target genes (74, 75). In the case of Adr1, phosphorylation of DNA binding domain should be preceded for activation of genes required for ethanol and glycerol utilization (76, 77). Another derepressed target gene of Mig1, Hap4, is not a zinc finger transcription factor but it forms a complex with other Hap protein (Hap2, 3, and 5) to activate the genes encoding TCA-cycle and respiratory chain enzymes (Figure 5A) (78, 79).

Also, PKA regulates the phosphorylation state of Rgt1 transcription factor in Rgt2/Snf3 pathway. During ethanol fermentation, active PKA phosphorylates Rgt1 so that it cannot inhibit the transcription of *HXT* genes, but after diauxic shift, PKA gets inactivated and Rgt1 and its corepressors represses the expression of *HXT* genes (59, 64). In addition, Rim15 is also a kinase protein phosphorylated by PKA. Under starvation

status, dephosphorylated Rim15 enters the nucleus and stimulates the Msn2/4 or Gis1 transcription factors (80, 81). They induce transcription of stress-responsive element (STRE) or post-diauxic shift (PDS) genes, respectively, so that cell can obtain tolerance against caloric restriction (Figure 5B) (82, 83).

Even though such a dramatic shift has long been studied, the dynamics and physiological state of this change is not yet been fully understood. Cellular adaptation to respiration occurs in a complex manner, containing reconstruction of transcriptome, proteome, and metabolome. Several studies established a database for the change of cellular response upon carbon source starvation, but the comprehensive understanding of these large-scale omics database has not tried much yet (84-86). Since the development of technology now allows us to detect cellular changes more precisely, real-time approach for this dynamic reprogramming is expected to be achieved soon.

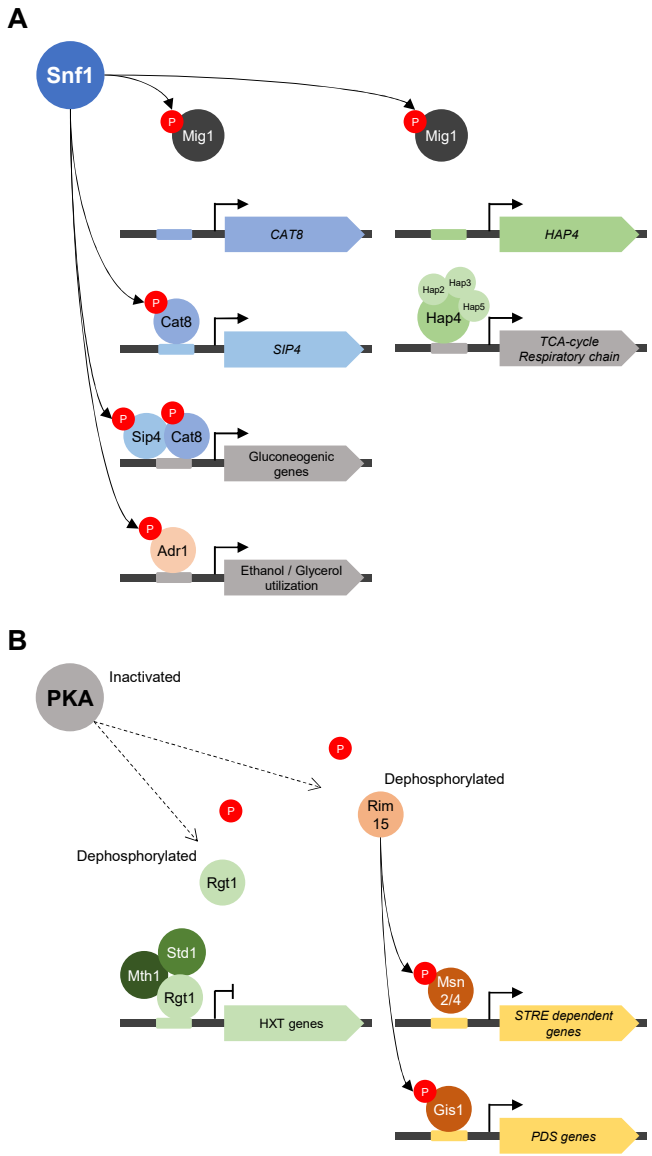

Figure 5. Regulatory pathway after diauxic shift.

Two major pathway that activates respiratroy pathway after diauxic shift. (A) Snf1 is activated and phosphorylates several zinc-finger transcirption factors, enables derepression or activation of target genes.
(B) Inactivated PKA results dephosphorylation of Rgt1 and Rim15, important transcirptional regulators involved in repression of HXT genes and activation of STRE and PDS genes, respectively.

2.3. Role of Gcr1 as a key regulator of glycolysis

2.3.1. General regulation of ethanol fermentation

As a Crabtree positive species, glycolysis is the most important pathway to utilize glucose and get energy in *S. cerevisiae*. After sensing extracellular glucose, the metabolism that digests glucose is activated accordingly. The genes encoding the twelve steps including glycolysis are the most strongly expressed genes during ethanol fermentation status. Currently, 21 glycolytic genes and 10 genes involved in ethanol production are annotated in *S. cerevisiae*, which indicates some of these pathways are regulated by multiple genes (Figure 6, Table 1). It had been once proposed that appearance of these paralogs from whole genome duplication (WGD) event plays a major role in strong tendency of aerobic ethanol fermentation in *S. cerevisiae* (87). However, many studies reported that deletion of minor paralogs did not severely affect the cell under several laboratory environments, indicating that multicopy glycolytic genes are not critical for Crabtree effect of *S. cerevisiae* (88).

Therefore, the regulation of glycolytic genes is important for glucose mediated fermentation in *S. cerevisiae.* The glycolytic and ethanol fermentation genes contain the binding sites of several key transcription factors, such as Reb1, Abf1, and Rap1 on its promoters (89, 90). Especially, Rap1 is

multifunctional regulator that has a wide effect in transcription of genes encoding ribosomal proteins, related to transcriptional machinery, and glycolytic genes (91, 92). However, there still is a missing link between glucose sensing signal pathway and regulation of glycolytic genes. The activation of PKA by external glucose is known to activate the transcription rate of Rap1 target genes, but its relationship with regulating glycolytic signal remains undiscovered (93).

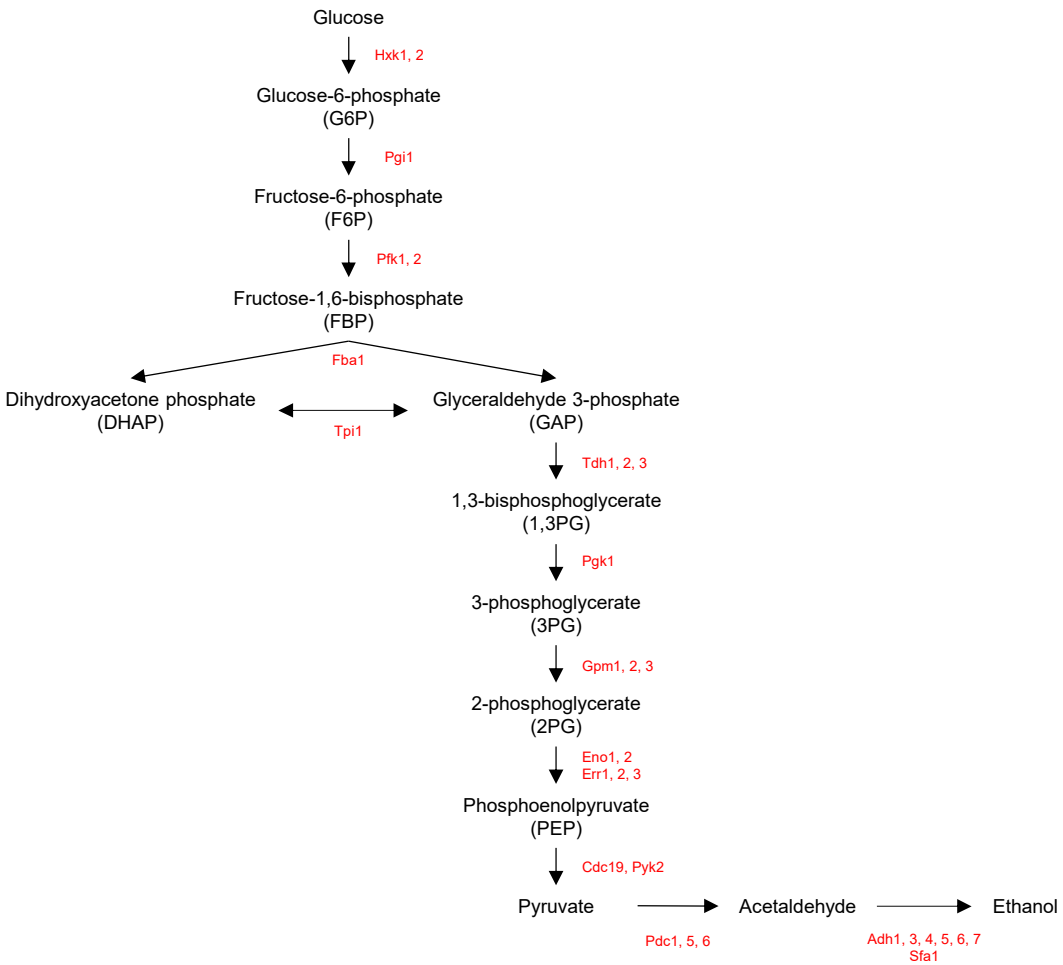

Figure 6. Pathway and enzymes involved in aerobic ethanol fermentation.

12 steps of metabolic pathway that utilize glucose to ethanol. 10 steps of glycolysis results pyruvate and additional two steps converts pyruvate into ethanol. The functions of each enzymes are listed in Table 1.

Table 1. List of enzymes involved in ethanol fermentation in *S. cerevisiae*

Function of enzyme	Name in *S. cerevisiae*
Hexokinase	Hxk1, 2
Glucose-6-phosphate isomerase	Pgi1
Phosphofructokinase	Pfk1, 2
Aldolase	Fba1
Triose phosphate isomerase	Tpi1
Glyceraldehyde phosphate dehydrogenase	Tdh1, 2, 3
Phosphoglycerate kinase	Pgk1
Phosphoglycerate mutase	Gpm1,2,3
Enolase	Eno1, 2, Err1, 2, 3
Pyruvate kinase	Cdc19, Pyk2
Pyruvate decarboxylase	Pdc1, 5, 6
Alcohol dehydrogenase	Adh1, 3, 4, 5, 6, 7, Sfa1

2.3.2. Regulation mechanism of Gcr1

With coordination with Rap1, Gcr1 is a well-known transcription factor which plays an important role to glycolytic gene expression, together with its co-activator Gcr2. Deletion of *GCR2* does not affect the growth of cell, but cell without *GCR1* gene shows severe growth defect. Gcr1 activates transcription of many glycolytic genes by binding to CT boxes (CWTCC) in the promoters (94, 95). In addition, it has been shown that Gcr1 lacking the C-terminal DNA binding domain can also activate transcription through interacting with Rap1 transcription factor bound to the promoters of glycolytic genes (96). Gcr2 works as a coactivator of Gcr1 without direct DNA binding (97) and both Gcr1 and Gcr2 can form homodimers through their leucine zipper (LZ) domains (98, 99). It has been suggested that a heterocomplex consisting of Gcr1 (monomer or dimer) and Gcr2 dimer is involved in transcriptional activation of glycolytic genes, whereas Gcr1 dimer is essential for the transcription of ribosomal protein (RP) genes (99, 100). However, the role for Gcr1 in RP gene expression has been controversial due to the growth defect accompanied by Gcr1 inactivation, which can indirectly affect RP gene expression (100). In fact, recent genome-wide ChIP-exo analysis showed that RP genes are not the direct targets of Gcr1 (86).

2.3.3. Unusual feature of *GCR1* gene and its splicing

The *GCR1* gene has an unusual feature of having a long intron, generating at least 7 different spliced isoforms of mRNA by alternative splicing (101, 102). Among the isoforms, only two were identified to generate in-frame proteins (102). The major spliced form is generated by splicing out of 739-nt intron using 5' splice site (SS) GUAUGG and 3'SS CAG, producing a Gcr1 protein of 789 amino acids, which I named $Gcr1^{S}$. The other form existing in much lower abundance uses 5'SS GUAUGA and 3'SS UAG located at 5-nt downstream and 17-nt downstream from the major 5'SS and 3'SS, respectively, generating a Gcr1 protein of 785 amino acids, which I named $Gcr1^{A}$. Accumulation of the spliced mRNAs increased later phase of cell growth around diauxic shift, whereas unspliced mRNA was observed as a major product at early growth phase (102). A start codon located in the middle of the intron of the unspliced mRNA is used as a translation start site, generating a Gcr1 protein of 844 amino acids, which I named $Gcr1^{U}$. Therefore, cells mainly produce two Gcr1 isoforms, $Gcr1^{U}$ and $Gcr1^{S}$ (Figure 7).

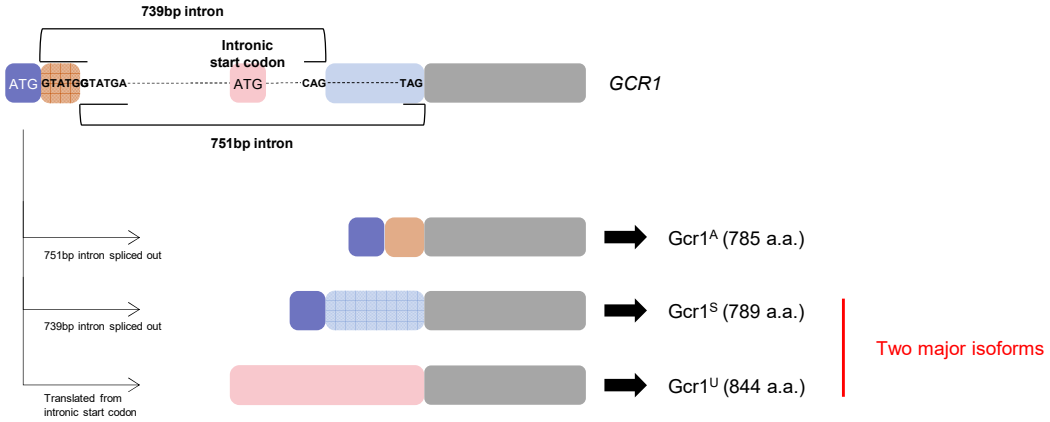

Figure 7. Unusual features of *GCR1* gene and formation of its major two isoforms

Among several alternative spliced transcripts, only two isoforms produce in-frame protein of Gcr1. Previously, GCR1 gene had been considered to have 751bp of intron in its gene, but recently more major transcript that splices out 739bp of intron by using another 5'SS and 3'SS was discovered. The first transcript with low abundance results Gcr1A protein, and major one produces Gcr1S protein. Also, the start codon located in the middle of intron is used as a translational start site, generating Gcr1U protein. Gcr1U and Gcr1S proteins are currently considered as a two major isoforms of Gcr1.

2.4. Regulation of metabolic flux for optimal production

2.4.1. Importance of regulating metabolic flux

Engineering of microbial cells to produce various kinds of chemicals has been receiving a great interest for past decades. This concept of making microbial cell factory has been widely researched aiming to maximize the cellular ability to make the compounds like bulk chemicals or biofuels (103-105). The common strategy is to strengthen the metabolic pathway to produce desired materials and to inhibit the competing pathways (106, 107). However, it is also important to optimize the central metabolism to keep the cell viable until it reaches the highest titer. Therefore, understanding not only the overall cellular metabolisms but also the balance among these metabolisms is a key factor for the microbial production.

Central carbon utilization pathway such as glycolysis had been commonly rerouted to product formation pathway to decrease the byproducts. However, the intermediates of glycolysis are linked with many other pathways: G6P is used in glycogen and trehalose metabolism, F6P and GAP is coupled with pentose phosphate pathway (PPP) and GAP is also related in producing glycerol (12, 108-110). Since those pathways are not significantly strong but essential for synthesis of cellular

components, simply reducing the side reaction is not an optimal condition for cell factory. To overcome limitation from rational engineering, computational method is gaining a lot of attention recently. Flux balance analysis (FBA) calculates the genome scale metabolic network and finds the optimal solution for maximized production (111-113). Although it needs some correction with the experimental data, it is promising that such approach can provide a novel perspective about balancing the flux distribution.

2.4.2. Previous studies of regulating metabolic flux

In order to prevent the loss of carbon, many studies have tried to redirect the glycolysis to other pathway for maximum production of desired products. By strengthening PPP from F6P, the titer of isoprenoid, aromatic chemicals, and shinorine has been increased (114-116). The production of N-acetylglucosamine, also the derivative starting from F6P, was also increased by reducing glycolytic flux (117).

In addition, pyruvate, the final product of glycolysis, is an important branchpoint to regulate the overall metabolism of *S. cerevisiae*. Since *S. cerevisiae* has a high tendency to convert pyruvate into ethanol, it is important to redirect the ethanol fermentation pathway to produce pyruvate-derived

chemicals. Several studies tried to eliminate the carbon leakage pathway to glycerol or ethanol, and successfully increased the production of D-lactate, acetoin, and 2,3-butanediol (118-120). In some study, substrate channeling has been used to redirect the pyruvate to produce the desired product even without eliminating the competitive pathway (121).

Recent mathematical approaches can be used for metabolic engineering, but the complexity of whole cell metabolism makes it difficult to apply to real cases. Several studies applied FBA to *S. cerevisiae*, but it is limited to some cases like maximizing the cell mass or analyzing the metabolism during different carbon sources (113, 122). The increment of bioethanol production had been tried with the computational analysis but producing chemicals which requires complex pathway including foreign genes has not been tried yet (113, 123).

Chapter 3.

Materials and methods

3.1. Yeast strains and culture conditions

S. cerevisiae strains used in this study are listed in Table 2. All strains were derived from *S. cerevisiae* BY4741 (MATa *his3*Δ1 *leu2*Δ0 *met15*Δ0 *ura3*Δ0). Cells were grown in YPD medium (1% yeast extract, 2% bacto-peptone, and 2% dextrose) or in synthetic complete (SC) medium (0.67% yeast nitrogen base without amino acids, 2% glucose, and 0.2% amino acids dropout mixture suitable for plasmid selection). Cell growth was detected by measuring the optical density (OD) at 600 nm with spectrophotometer (Varian Cary 50 UV-vis, Agilent Technologies, USA). OD600 of 0.5 pre-cultured cells were cultivated in 10 mL medium in a 50 mL Erlenmeyer flask at 30°C with shaking at 170 rpm.

For the production of desired product, SC medium containing 2% glucose was used for pre-culture and main culture. Cells were grown until they consume all glucose, and moved to 20 mL medium in a 100 mL Erlenmeyer flask at 30°C with shaking at 170 rpm with initial OD600 of 0.5. In the case of lactate production, 5 g/L of calcium carbonate were added in the media as a neutralizing reagent.

Table 2. Strains used in this study

Strain	Description	Genotype	Reference
BY4741	Wild type	*S. cerevisiae MATa his3Ä1 leu2Ä0 met15Ä0 ura3Ä0*	EUROSCARF
JHY9000	*gcr1Δ*	BY4741 *gcr1Δ*	This study
JHY9001	*GCR1*ΔAH	BY4741 *GCR1* (Δ1223-1297)	This study
JHY9002	*GCR1*ΔLZ1	BY4741 *GCR1* (Δ1550-1639)	This study
JHY9003	*GCR1*ΔSP	BY4741 *GCR1* (Δ1676-2527)	This study
JHY9004	*GCR* ΔDBD	BY4741 *GCR1* (Δ2642-3106)	This study
JHY9005	*GCR1 gcr2Δ*	BY4741 *gcr2Δ*	This study
JHY9006	*GCR1 GCR2*Δ2H	BY4741 *GCR2* (Δ748-1440)	This study
JHY9007	*GCR1 GCR2*ΔLZ2	BY4741 *GCR2* (Δ1489-1602)	This study
JHY9008	*GCR1 xrn1Δ*	BY4741 *xrn1Δ*	This study
JHY9009	*GCR1 upf1Δ*	BY4741 *upf1Δ*	This study
JHY9100	$GCR1^{U}$	BY4741 *GCR1* (Δ1-574)	This study
JHY9100OE	$GCR1^{U}$ *OE*	JHY9100 P_{TDH3}-$GCR1^{U}$	This study
JHY9101	$GCR1^{U}$ ΔAH	JHY9100 $GCR1^{U}$ (Δ649-723)	This study
JHY9102	$GCR1^{U}$ ΔLZ1	JHY9100 $GCR1^{U}$ (Δ976-1065)	This study
JHY9103	$GCR1^{U}$ ΔSP	JHY9100 $GCR1^{U}$ (Δ1102-1953)	This study
JHY9104	$GCR1^{U}$ ΔDBD	JHY9100 $GCR1^{U}$ (Δ2068-2532)	This study
JHY9105	$GCR1^{U}$ *gcr2Δ*	JHY9100 *gcr2Δ*	This study
JHY9106	$GCR1^{U}$ *GCR2*Δ2H	JHY9100 *GCR2* (Δ748-1440)	This study
JHY9107	$GCR1^{U}$ *GCR2*ΔLZ2	JHY9100 *GCR2* (Δ1489-1602)	This study
JHY9108	$GCR1^{U}$ *xrn1Δ*	JHY9100 *xrn1Δ*	This study
JHY9109	$GCR1^{U}$ *upf1Δ*	JHY9100 *upf1Δ*	This study
JHY9200	$GCR1^{S}$	BY4741 *GCR1* (Δ4-742)	This study

JHY9200OE	*GCR1*^{S} *OE*	JHY9200 P_{TDH3}-*GCR1*^{S}	This study
JHY9201	*GCR1*^{S} △AH	JHY9200 *GCR1*^{S} (△484-558)	This study
JHY9202	*GCR1*^{S} △LZ1	JHY9200 *GCR1*^{S} (△811-900)	This study
JHY9202A	*GCR1*^{S} △LZ1+ ALD4	JHY9200 *GCR1*^{S} (△811-900) *ura3Δ0*::P_{TEF1}-*ALD4*-T_{CYC1}	This study
JHY9203	*GCR1*^{S} △SP	JHY9200 *GCR1*^{S} (△937-1788)	This study
JHY9204	*GCR1*^{S} △DBD	JHY9200 *GCR1*^{S} (△1903-2367)	This study
JHY9205	*GCR1*^{S} *gcr2Δ*	JHY9200 *gcr2Δ*	This study
JHY9206	*GCR1*^{S} *GCR2*△2H	JHY9200 *GCR2* (△748-1440)	This study
JHY9207	*GCR1*^{U} *GCR2*△LZ2	JHY9200 *GCR2* (△1489-1602)	This study
JHY9208	*GCR1*^{S} *xrn1Δ*	JHY9200 *xrn1Δ*	This study
JHY9209	*GCR1*^{S} *upf1Δ*	JHY9200 *upf1Δ*	This study
JHY9210	*GCR1*^{S}+USS	JHY9200 *ura3Δ0*::P_{GCR1}-*GCR1*^{U} (1-168)-T_{GCR1}	This study
JHY9211	*GCR1*^{S}+USS^{F12L}	JHY9200 *ura3Δ0*::P_{GCR1}-*GCR1*^{U} $(1\text{-}168)^{T34C}$-T_{GCR1}	This study
JHY9212	*GCR1*^{S}+USS^{L50P}	JHY9200 *ura3Δ0*::P_{GCR1}-*GCR1*^{U} $(1\text{-}168)^{T149C}$-T_{GCR1}	This study
JHY9210F	*GCR1*^{S}+USS-5 Flag	JHY9200 *ura3Δ0*::*GCR1*^{U} (1-168)-*5Flag* ::*hphMX6*	This study
JHY9211F	*GCR1*^{S}+USS^{F12L} -5Flag	JHY9200 *ura3Δ0*::*GCR1*^{U} $(1\text{-}168)^{T34C}$-*5Flag* ::*hphMX6*	This study
JHY9212F	*GCR1*^{S}+USS^{L50P} -5Flag	JHY9200 *ura3Δ0*::*GCR1*^{U} $(1\text{-}168)^{T149C}$-*5Flag* ::*hphMX6*	This study
	GCR1-TAP	BY4741 *GCR1-TAP*::*his3MX6*	(124)
JHY9302	*GCR1-5Flag*	BY4741 *GCR1*^{WT}-*5Flag*::*hphMX6*	This study
JHY9310	*GCR1*^{U}-*TAP*	BY4741 *GCR1*^{U}-*TAP*::*his3MX6*	This study
JHY9311	*GCR1*^{U}-*TAP* *GCR2-5Flag*	JHY9310 *GCR2-5Flag*::*hphMX6*	This study
JHY9312	*GCR1*^{U}-*5Flag*	BY4741 *GCR1*^{U}-*5Flag*::*hphMX6*	This study
JHY9320	*GCR1*^{S}-*TAP*	BY4741 *GCR1*^{S}-*TAP*::*his3MX6*	This study
JHY9321	*GCR1*^{S}-*TAP* *GCR2-5Flag*	JHY9320 *GCR2-5Flag*::*hphMX6*	This study

JHY9322	*GCR1*^{S}*-5Flag*	BY4741 *GCR1*^{S}*-5Flag::hphMX6*	This study
JHY9322U1	*GCR1*^{S}*-5Flag* + USS	JHY9322 *ura3Δ0::*P_{GCR1}*-GCR1*^{U} (1-168)-T_{GCR1}	This study
SP1130	LA production strain	*S. cerevisiae* CEN.PK2-1C *pdc1*Δ::P_{CCW12}*-BtLDH* *cyb2*Δ::P_{CCW12}*-BtLDH* *gpd1*Δ::P_{CCW12}*-PsjLDH* *adh1*Δ::P_{CCW12}*-PsjLDH* *ald6*Δ::*loxP* *leu2*Δ::P_{TDH3}*-mhpF* *pdc6*Δ::P_{TDH3}*-eutE*	(126)
SP1130-U	SP1130 Gcr1^{U}	SP1130 *GCR1* (△1-574)	This study
SP1130-S	SP1130 Gcr1^{S}	SP1130 *GCR1* (△4-742)	This study
SP1141	LA production strain with Gcr1 overexpression	SP1130 *his3*Δ::P_{TDH3}*-GCR1*	(126)
SP1141-U	SP1141 Gcr1^{U}	SP1130-U *his3*Δ::P_{TDH3}*-GCR1*^{U}	This study
SP1141-S	SP1141 Gcr1^{S}	SP1130-S *his3*Δ::P_{TDH3}*-GCR1*^{S}	This study

3.2. Construction of plasmids and yeast strains

Genome editing using CRISPR/Cas9 system was performed as previously described (127). Briefly, Coex413-Cas9 vector containing proper gRNA and donor DNA were introduced into *S. cerevisiae* and selected on SC-His medium. The genome-edited strains were confirmed by both PCR and sequencing. The gRNA and primer sequences for donor DNA are listed in Table 3 and 5.

Plasmids used in this study are listed in Table 4. To produce JHY9100 and JHY9200 strains, expressing only $Gcr1^{U}$ and $Gcr1^{S}$, respectively, plasmids containing each *GCR1* form were generated first. DNA fragment containing the *GCR1* ORF flanked by 500-bp upstream (P_{GCR1}) and 500-bp downstream (T_{GCR1}) regions was PCR-amplified and cloned between *Spe*I and *Xho*I sites of pRS413, generating p413Gcr1WT. Next, $p413Gcr1^{U}$ and $p413Gcr1^{S}$ harboring *GCR1*(Δ1-574) and *GCR1*(Δ4-742) were generated by site directed mutagenesis of p413Gcr1WT and used as PCR templates to produce donor DNAs to generate JHY9100 and JHY9200 strains using the CRISPR/Cas9 system. This strategy also applicated when producing SP1130-U and SP1130-S strains from their background strains.

JHY9210, 9211, and 9212 strains were generated by

integrating expression cassettes for *GCR1*^{U} (1-168), *GCR1*^{U} (1-168)^{T34C}, and *GCR1*^{U} (1-168)^{T149C}, respectively, at *ura3Δ0* site in the genome of JHY9200 using Coex416-Cas9-gURA3Δ0 plasmid. To generate the expression cassettes, P_{GCR1} and T_{GCR1} were cloned into pRS416 vector using *Sac*I/*Xba*I and *Xho*I/*Kpn*I sites, respectively, resulting in p416GCR1PT. Next, p416GCR1-USS was generated by cloning *GCR1*^{U}(1-168) between the *Xba*I and *Xho*I sites of p416GCR1PT. p416GCR1-USS^{F12L} and p416GCR1-USS^{L50P} containing the expression cassettes for *GCR1*^{U} (1-168)^{T34C}, and *GCR1*^{U} (1-168)^{T149C}, respectively, were generated by site directed mutagenesis of p416GCR1-USS. Donor DNA containing the expression cassette flanked by 35-bp homology regions targeting the *ura3Δ0* site was amplified by PCR and transformed into JHY9200. Donor DNAs for other strains were produced by PCR without template DNA using primers containing 35-bp homology regions targeting the integration site and 20-bp overlapped base pairs between two primers. To overexpress *ALD4* gene, PCR-amplified *ALD4* ORF was cloned into p416TEF between *Spe*I and *Sal*I sites, generating p416TEF-ALD4. PCR-amplified expression cassette containing P_{TEF1}-*ALD4*-T_{CYC1} was transformed with Coex416-Cas9-gURA3 Δ0 resulting in JHY9202A strain.

JHY9310 and JHY9320 containing *GCR1^{U}-TAP* and *GCR1^{S}-TAP*, respectively, were generated by homologous recombination by introducing DNA fragment PCR amplified from *S. cerevisiae GCR1-TAP* strain (124) into JHY9100 and JHY9200, respectively. Strains with 5Flag-tagged *GCR1* (JHY9322), *GCR2* (JHY9311 and 9321) and USS domain (JHY9210F, 9211F and 9212F) were generated by using DNA fragments amplified from pFA6a-5Flag-hphMX6 vector as a template. JHY9322U1, 9322U2 and 9322U3 strain was generated by integrating the USS expression cassette into JHY9322 strain, as described for the production of JHY9210~9212 strains.

When generating SP1141-U and SP1141-S strains that overexpress *GCR1^{U}* or *GCR1^{S}* in the at *his3Δ0* site in the genome, marker based integration method was used like the previous study (126). The ORFs of GCR1^{U} and GCR1^{S} were amplified from p413Gcr1^{U} and p413Gcr1^{S} respectively and inserted between the *Xba*I/*Xho*I sites of pRS416GPD vector, generating pRS416GPD-GCR1^{U} and pRS416GPD-GCR1^{S}. P_{TDH3}-GCR1$^{U/S}$-T_{CYC1} fragment was amplified from these two vectors using *c_GCR1* F and *c_GCR1* R primers, and cloned into *Kpn*I site of pUC57-URA3 (126). The *GCR1* isoform overexpressing cassettes containing *URA3* selection marker was amplified from pUC57-URA3-*GCR1^{U}* and pUC57-URA3-*GCR1^{S}*

vectors, with *i_GCR1* F and *i_GCR1* R primers (126). The amplified cassettes were transformed to SP1130-U and SP1130-S strains and selected on SC-URA plate, generating SP1141-U and SP1141-S.

Table 3. 20-bp gRNA sequences used in this study

Name	gRNA sequence [a] (5' to 3')	Description
gGCR1	AATTAACTACAGAAAATATCagg	Used for deletion of *GCR1* and the leucine zipper domain of *GCR1*
gGCR1$^{U/S}$	TGCGTCTGTCTGCGTACAAGagg	Used for producing JHY9100, JHY9200 strains
gGCR1AH	TGCTTCAAATAAGACAGCCAtgg	Used for deletion of the alpha helix domain *GCR1*
gGCR1SP	TTGGGCTTGTCCGTTGGGCTtgg	Used for deletion of the serine-proline rich domain of *GCR1*
gGCR1DBD	TCCTACTAGAAGAATTATTAtgg	Used for deletion of the DNA binding domain of *GCR1*
gURA3Δ0	CGAGATTCCCGGGAGCTTTAtgg	Used for insertion at *ura3Δ0* site
gGCR1USS	CTTTTCAAGGGACAAATAACagg	Used for random mutagenesis of *GCR1*^{U} (1-168) in JHY9105
gGCR2	TGGCTAAAAATGCTAAAAATggg	Used for deletion of *GCR2* and the 2H domain of *GCR2*
gGCR2LZ	TCCTTAACCACTGCGTCTCTtgg	Used for deletion of the leucine zipper domain of *GCR2*
gXRN1	AAATGGTAAAGACAATAACGtgg	Used for deletion of *XRN1*
gUPF1	GAACTCAGCAGTAAAACCAGtgg	Used for deletion of *UPF1*
gGCR1prom	AGATTTAAACTCCGTACACCcgg	Used for substituting the promoter of *GCR1*

a. PAM sequences (NGG) are denoted as lower case letters.

Table 4. Plasmids used in this study

Plasmid	Relevant characteristics	Reference
Plasmids for *S. cerevisiae*		
pFA6a-5Flag-hphMX6	Vector for genomic 5xFlag epitope tagging, hygromycin antibiotic marker	(128)
pRS413	CEN/ARS plasmid, *HIS3* marker	(129)
p413Gcr1^{WT}	P_{GCR1}-*GCR1*-T_{GCR1} cloned between *Spe*I and *Xho*I sites of pRS413	This study
p413Gcr1^{U}	pRS413 containing P_{GCR1}-*GCR1* (Δ1-574)-T_{GCR1} generated by site directed mutagenesis of p413Gcr1^{WT}	This study
p413Gcr1^{S}	pRS413 containing P_{GCR1}-*GCR1* (Δ4-742)-T_{GCR1} generated by site directed mutagenesis of p413Gcr1^{WT}	This study
pRS416	CEN/ARS plasmid, *URA3* marker	(129)
p416TEF	P_{TEF1} and T_{CYC1} cloned at *Sac*I/*Xba*I and *Xho*I/*Kpn*I sites, respectively, of pRS416	(130)
p416GPD	P_{TDH3} and T_{CYC1} cloned at *Sac*I/*Xba*I and *Xho*I/*Kpn*I sites, respectively, of pRS416	(130)
p416TEF-ALD4	*ALD4* cloned at *Spe*I/*Sal*I sites of p416TEF	This study
p416GCR1PT	P_{GCR1} and T_{GCR1} cloned at *Sac*I/*Xba*I and *Xho*I/*Kpn*I sites, respectively, of pRS416	This study
p416GCR1-USS	pRS416 containing P_{GCR1}-*GCR1*^{U}(1-168)-T_{GCR1} generated by cloning *GCR1*^{U} (1-168) between *Xba*I and *Xho*I sites of p416GCR1PT	This study
p416GCR1-USS^{F12L}	pRS416 containing P_{GCR1}-*GCR1*^{U}(1-168)^{T34C}-T_{GCR1} generated by directed mutagenesis of p416GCR1-USS	This study
p416GCR1-USS^{L50P}	pRS416 containing P_{GCR1}-*GCR1*^{U}(1-168)^{T149C}-T_{GCR1} generated by directed mutagenesis of p416GCR1-USS	This study
pRS416GPD-GCR1	*GCR1* ORF cloned between the *Xba*I and *Xho*I sites of p416GPD	(126)
pRS416GPD-GCR1^{U}	*GCR1*^{U} ORF cloned between the *Xba*I and *Xho*I sites of pRS416GPD	This study
pRS416GPD-GCR1^{S}	*GCR1*^{S} ORF cloned between the *Xba*I and *Xho*I sites of pRS416GPD	This study
pUC57-*GCR1*	P_{TDH3}-*GCR1*-T_{CYC1} cloned into the *Kpn*I site of pUC57-URA3	(126)
pUC57-*GCR1*^{U}	P_{TDH3}-*GCR1*^{U}-T_{CYC1} cloned into the *Kpn*I site of pUC57-URA3	This study
pUC57-*GCR1*^{S}	P_{TDH3}-*GCR1*^{S}-T_{CYC1} cloned into the *Kpn*I site of pUC57-URA3	This study
Plasmids for CRISPR/Cas9-mediated genome editing		
Coex413-Cas9	Coex413 containing P_{TDH3}-*Cas9*-T_{TPI1}	(127)
Coex413-Cas9-gGCR1	Coex413-Cas9 with gGCR1 gRNA	This study
Coex413-Cas9-gGCR1$^{U/S}$	Coex413-Cas9 with gGCR1$^{U/S}$ gRNA	This study
Coex413-Cas9-gGCR1AH	Coex413-Cas9 with gGCR1AH gRNA	This study
Coex413-Cas9-gGCR1SP	Coex413-Cas9 with gGCR1SP gRNA	This study

Coex413-Cas9-gGCR1DBD	Coex413-Cas9 with gGCR1DBD gRNA	This study
Coex413-Cas9-gGCR1USS	Coex413-Cas9 with gGCR1USS gRNA	This study
Coex413-Cas9-gGCR2	Coex413-Cas9 with gGCR2 gRNA	This study
Coex413-Cas9-gGCR2LZ	Coex413-Cas9 with gGCR2LZ gRNA	This study
Coex413-Cas9-gXRN1	Coex413-Cas9 with gXRN1 gRNA	This study
Coex413-Cas9-gUPF1	Coex413-Cas9 with gUPF1 gRNA	This study
Coex413-Cas9-gGCR1prom	Coex413-Cas9 with gGCR1prom gRNA	This study
Coex416-Cas9	Coex416 containing P_{TEF1}-*Cas9*-T_{TPI1}	This study
Coex416-Cas9-gURA3△0	Coex416-Cas9 with gURA3△0 gRNA	This study

Table 5. Primers used in this study [a, b]

Forward primer (5' to 3')	Reverse primer (5' to 3')	Usage
Primer sequence for plasmid (for *S. cerevisiae*) generating		
gcgACTAGTCCCGGATGAGAAACCTTTAAAATAAGCG	gcgCTCGAGCTGCTAGGTTTTTATATTGATTTTTGACAACAGAG	Cloning of Gcr1^{WT} using *Spe*I/*Xho*I
GCCTTTGATATATATTGAAAATGAATTTTCTGACTCAGGC	GCCTGAGTCAGAAAATTCATTTTCAATATATATCAAAGGC	Generation of p413Gcr1^{U} by site directed mutagenesis
TTTGATATATATTGAAAATGCAAACAAGTGTTGATAGTAC	GTACTATCAACACTTGTTTGCATTTTCAATATATATCAAA	Generation of p413Gcr1^{S} by site directed mutagenesis
gcgGAGCTCCCCGGATGAGAAACCTTTAAAATAAGCG	gcgTCTAGATTTCAATATATATCAAAGGCCAATTAAATATAACCGTCGTTTG	Cloning of P_{GCR1} using *Sac*I/*Xba*I
gcgCTCGAGGTTTATTGAGGTTGTCCGCGACAATAG	gcgGGTACCCTGCTAGGTTTTTATATTGATTTTTGACAACAGAG	Cloning of T_{GCR1} using *Xho*I/*Kpn*I
gcgTCTAGAATGAATTTTCTGACTCAGGCTATGTCAGAAAC	gcgCTCGAGTTACTGTTCCAATTGAGAAAGTAGGGCATTTAATTGG	Cloning of *GCR1*^{U} (1-168) using *Xba*I/*Xho*I
GAATTTTCTGACTCAGGCTATGTCAGAAACTCTTCAAGGGACAAATAACAGGATAAAACG	CGTTTTATCCTGTTATTTGTCCCTTGAAGAGTTTCTGACATAGCCTGAGTCAGAAAATTC	Generation of p416GCR1-USS^{F12L} by site directed mutagenesis
CTAACCAATTAAATGCCCCACTTTCTCAATTGG	CCAATTGAGAAAGTGGGGCATTTAATTGGTTAG	Generation of p416GCR1-USS^{L50P} by site directed mutagenesis
gcgACTAGTATGTTCAGTAGATCTACGCTCTGC	gcgGTCGACTTACTCGTCCAATTTGGCACGGAC	Cloning of *ALD4* into p416TEF using *Spe*I/*Sal*I
gcgGGTACCTCATTATCAATACTCGCCATTTCA	cccGGTACCCAAATTAAAGCCTTCGAGCG	*c_GCR1* F and *c_GCR1* R primers to generate pUC57-*GCR1*, *GCR1*^{U} and *GCR1*^{S} (126)
Primer sequence for generating donor DNA by overlapping PCR		
CAAGTGACAAACGACGGTTATATTTAATTGGCCTTTGATATATATTGAAAGTTTATTGAG	GCTTCGTTATTTTGTTGAAGGAACTATTGTCGCGGACAACCTCAATAAACTTTCAATATA	*gcr1* deletion
TTGAACCTTCTAAAGAGTTGATCGATTTGGTATTTCCATGGCTGTCTTATGGCAGGGTTG	AATAAGGGATGTTGACTCAGATATCGGAGATCCTGAATGGCAACCCTGCCATAAGACAGC	Deletion of the alpha helix domain of *GCR1*
AGGAAGTGAGTCAAAAAGTTGATTCTTACTTTATGGAATTATCAAAAAAACTGTTGCAGA	GACTTTGATCCGATCGCCTGATTTCCTGAAAGCAATTGTCTCTGCAACAGTTTTTTTGAT	Deletion of the leucine zipper domain of *GCR1*
GTAACCAAATTTTGCTGTTGCAGAGACAATTGCTTTCAGGAAATCAGGCGATTCATAATT	GTCTTTGTCACTGTACCACTTGAATTTGCAGCCTCAGTAGAATTATGAATCGCCTGATTT	Deletion of the serine-proline rich domain of *GCR1*
CGCAGAGTTCTTCTAAGTTTGAAATTATAAATAAAAAGGATACGAAGGCGTAAGTTTATT	TCGTTATTTTGTTGAAGGAACTATTGTCGCGGACAACCTCAATAAACTTACGCCTTCGTA	Deletion of the DNA binding domain of *GCR1*

AAGGAACCTGAGAACACAAAGAGTATTTGACGAAA AGTTACACTCACATACACGATAATA	CACCAGAAAATTAAAAGAGAAAGCAATATATGTTA AACATTATTATCGTGTATGTGAGTG	*gcr2* deletion
GACGCAGGAAGGGAAACTCCTTAAATACATCAACT AAAGGCTCCCCATCACCATTGAAAG	ATTCTTTGGCCTTTTTCAGAAATGATTGCGTCATA AGCTTCTTTCAATGGTGATGGGGAG	Deletion of the 2H domain of *GCR2*
GCCCATTGAAAGAAGCTTATGACGCAATCATTTCT GAAAAAGGCCAAAGATGACACGATA	CAGAAAATTAAAAGAGAAAGCAATATATGTTAAA CATTATTATCGTGTCAAGGCCAAAGA	Deletion of the leucine zipper domain of *GCR2*
AAAAATCAACACTTGTAACAACAGCAGCAACAAAT ATATATCAGTACGGTAACATACGAC	GATATACTATTAAAGTAACCTCGAATATACTTCGT TTTTAGTCGTATGTTACCGTACTGA	*xrn1* deletion
AGCAAGACCGAATATACTTTTTATATTACATCAAT CATTGTCATTATCAATTCGGTGAAC	TTTGTATCACAAGCCAAGTTTAACATTTTATTTTA ACAGGGTTCACCGAATTGATAATGA	*upf1* deletion
Primer sequence for generating donor or integration cassette DNA by PCR from cloned plasmids		
CAAGTGACAAACGACGGTTATATTTAATTG	TTATTTGTCCCTTGAAAAGTTTCTGACATAG	Generation of JHY9100 strain using p413Gcr1^{U} as a template
CAAGTGACAAACGACGGTTATATTTAATTG	ATGGAATAAAAGTTTGAGCTCGTGCTG	Generation of JHY9200 strain using p413Gcr1^{S} as a template
TTAATGTGGCTGTGGTTTCAGGGTCCATAAAGCTT CCCGGATGAGAAACCTTTAAAATAA	TTTAGTATACATGCATTTACTTATAATACAGTTTT CTGCTAGGTTTTTATATTGATTTTT	*ura3Δ0* site insertion with P_{GCR1}, T_{GCR1}
TTAATGTGGCTGTGGTTTCAGGGTCCATAAAGCTT ATAGCTTCAAAATGTTTCTACTCC	TTTAGTATACATGCATTTACTTATAATACAGTTTT GCAAATTAAAGCCTTCGAGCGTCCC	*ura3Δ0* site insertion with P_{TEF1}, T_{CYC1}
TTCTTCTTTTTGGCACTTGGTTATGTGATAATATC TCATTATCAATACTCGCCATTTC	TTTTCAGAATGCGTGTTATGATCATACCATACCAT TCGAAACTAAGTTCTGGTG	Promoter change of *GCR1* into P_{TDH3}
GGTTTCCCGACTGGAAAGC	CAGTCACGACGTTGTAAAA	*i_GCR1* F and *i_GCR1* R primers to integrate *GCR1*^{U} or *GCR1*^{S} overexpression cassette (126)
Primer sequence for epitope tag for for *S. cerevisiae*		
GAAGAGAAATTAAAGTATTGCAAAAGGCGACATAA TACACCATCTCGGATCCCCGGGTTAATTAA	GTTTAAACGAGCTCGAATTCGTTTATTGAGGTTGT CCGCGACAATAGTTCCTTCAACAAAATAAC	Generation of C-terminal epitope tag at *GCR1*
GTTGTGTTAGAAGTATGTTAAGGGATTTACAAAGA CGGATCCCCGGGTTAATTAA	AGAGAAAGCAATATATGTTAAACATTATTATCGTG GAATTCGAGCTCGTTTAAAC	Generation of C-terminal epitope tag at *GCR2*
CTAACCAATTAAATGCCCTACTTTCTCAATTGGAA CAGCGGATCCCCGGGTTAATTAA	CTATTGTCGCGGACAACCTCAATAAACCTCGAGTT AGAATTCGAGCTCGTTTAAAC	Generating JHY9200U1, JHY9200U2 from JHY9200
CTAACCAATTAAATGCCCCACTTTCTCAATTGGAA CAGCGGATCCCCGGGTTAATTAA	CTATTGTCGCGGACAACCTCAATAAACCTCGAGTT AGAATTCGAGCTCGTTTAAAC	Generating JHY9200U3 from JHY9200

[a] Restriction enzyme sites are underlined.

[b] Mutated nucleotides are shown in red.

3.3. Screening suppressor mutants of JHY9105

To select *GCR1-USS* mutants, which can suppress the growth defect of JHY9105 (*GCR1*^{U}*gcr2*Δ), USS mutant library was generated by error-prone PCR of *GCR1*^{U} (1-168) DNA fragment. Error-prone PCR was conducted with 5 mM of $MgCl_2$ or each 1 mM of dCTP/dTTP, using Taq polymerase (BioFACTTM, Biofact, Korea). The mutant library was introduced into JHY9105 as donor DNA with Coex413-Cas9-gGCR1USS expressing a gRNA targeting the *GCR1-USS* locus. The transformants were spread on SC-His medium and suppressor mutants were selected based on the bigger colony size, and the mutated sequences were analyzed by DNA sequencing of the USS domain.

3.4. mRNA extraction and cDNA synthesis

Total RNA was extracted from yeast cells using the hot phenol method (131). For cDNA synthesis, 1 μg of heat-denatured total RNA was mixed with total 30 μL reaction mixture (containing 4 μL oligo dT, 2 μL M-MLV reverse transcriptase, and 4 μL each of 10 mM dNTPs) and incubated at 42°C for 60 min, and then reverse transcription reaction was terminated by heating at 75°C for 15 min.

3.5. Quantitative reverse transcription PCR (qRT-PCR) analysis

The relative amount of target mRNA was determined by qRT-PCR of the synthesized cDNA. 5 μL of cDNA was amplified by SYBR Green I master mix (Roche Life Science, Germany) and gene-specific primers with 45 cycles of 95°C for 20 s, 60°C for 20 s, and 72°C for 20 s using a Lightcycler 480 II system (Roche Life Science, Germany). The Crossing point (Cp) values were processed using Light Cycler 480 software version 1.5. Expression levels of target genes were normalized by selected reference gene, *TFC1*. Primers used in qRT-PCR was listed in Table 6.

3.6. RNA-seq and analysis

RNA-seq libraries were prepared using a TruSeq Standard mRNA Sample Prep Kit (Illumina) according to the manufacturer's protocol. Paired-end sequencing with 100 cycles was performed using a HiSeq2500 (Illumina) instrument according to manufacturer's protocol. The quality of raw reads was assessed with FastQC (version 0.11.9); the quality scores were >Q30, which indicated high quality. Clean reads with high quality scores were processed using the Tuxedo protocol (133) with TopHat2 (version 2.1.1) (135) and Cufflinks (135). Reads

for each sample were aligned to the yeast reference genome (sacCer3 assembly) using TopHat2. Gene expression quantification was performed using Cufflinks, and fragments per kilobase of transcript per million reads mapped (FPKM) was calculated as the expression value. Differential expression analysis between exponential phase and diauxic shift of $Gcr1^{U}$ and $Gcr1^{S}$ with two replicates were performed using Cuffdiff (135), with the cut-off set at $p<0.01$ and ≥ 1.5-fold change. Their expression pattern of targeted genes of $Gcr1^{U}$ and $Gcr1^{S}$ was visualized as heatmap by using MeV (http://mev.tm4.org). Expressions of genes were shown as Z-score for FPKM.

3.7. *In vivo* TAP pull-down assay and immunoblotting

Cells were grown until half of the glucose was consumed and lysed in lysis buffer [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2mM $MgCl_2$, 0.1% NP40] supplemented with 0.1 % protease inhibitor cocktail (Calbiochem, USA) and 1 mM PMSF using acid-washed glass bead. After repeating 30 sec on / 90 sec off cycle for 10 times, cell debris were centrifuged down for 20 min and total protein concentration of the supernatant was analyzed by Bradford assay. 800 μg of proteins were used for TAP pull down with 20 μL of IgG Sepharose 6

Fast Flow resin (GE healthcare, USA) for 1 h, washed 3 times with lysis buffer, and eluted by boiling with 5x sample buffer. Samples were resolved by size on 6% SDS-PAGE gel and analyzed by western blotting with anti-DDDDK antibody (MBL life science, USA) for flag tag and anti-mouse IgG antibody (Sigma-Aldrich, USA) for TAP tag. Blotted membrane was treated with proper HRP-conjugated secondary antibody and visualized by G::box Chemi-XL (Syngene, USA).

3.8. Chromatin immunoprecipitation (ChIP) qPCR and ChIP-seq

Cells were cross-linked with final 1% of formaldehyde for 25 min followed by 5 min quenching with 250 mM glycine. Harvested cells were washed with ice-cold TBS [50 mM Tris-HCl (pH 7.4), 150 mM NaCl] three times and ChIP-lysis buffer [50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.2% SDS] once. Cells were lysed in 200 μL lysis buffer supplemented with 0.1% protease inhibitor cocktail (Calbiochem, USA) and 1 mM PMSF using acid-washed glass beads and periodical vortexing. 800 μL of lysis buffer was added to the lysates and sonicated for 20 s using sonicator (Vibra-cell, Sonics & materials inc., USA) with amplitude 22%, 12 times for

6 h samples and 14 times for 12 h samples. Crude lysates were centrifuged for 20 min to eliminate debris. 200 μL of lysis buffer was supplemented to equal volume of sonicated lysates and incubated at 4°C for overnight with anti-DDDDK antibody followed by 1 h incubation with the Protein G Plus agarose bead (Santa Cruz Biotechnology, USA). For TAP-tagged protein, the lysates were incubated with IgG-sepharose beads (GE healthcare, Sweden) at 4°C for 1 h. Beads were washed with lysis buffer without SDS, twice in high salt lysis buffer [50 mM HEPES-KOH (pH 7.5), 500 mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate], once with LiCl wash buffer [Tris-HCl (pH 8.0), 250 mM LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% Sodium deoxycholate], twice with TE buffer [Tris-HCl (pH 8.0), 1 mM EDTA]. DNA was eluted from beads by incubating with elution buffer [1% SDS, 250 mM NaCl] for 30 min, 65°C. Eluent was transferred into fresh tube and treated with RNase for 1 h, and Proteinase K for 2 h. After reversal of crosslinking by overnight incubation at 65°C with 100 mM NaCl, DNA was purified with DNA purification kit (Qiagen). Input samples were prepared with the same procedure except for the beads-binding and elution steps.

For ChIP-qPCR, Fold enrichment of DNA binding Gcr1-TAP or Gcr2-5Flag was determined by qPCR using

Roche LightCycle 480 II. Concentrations of each target promoter DNA fragment in immuoprecipitated samples were divided by input samples first and normalized by *ACT1* promoter. Primers used in ChIP-qPCR was listed in table 6.

For ChIP-seq sample preparation, triplicate cultured cells were harvested and sonicated each and mixed up to 3 mL of total volume. For each protein tag, two sets of 600 μL samples from the mixture were incubated with proper antibody and bead followed by reverse crosslinking. Reverse crosslinked mixtures were treated with Phenol-choloform-isoamylalcohol (25:24:1) and precipitated by ethanol and glycogen at -80°C. Two sets of dried pellets were dissolved in water and combined for high DNA concentration and used for analysis.

The sequencing libraries were prepared from ChIP DNA fragments (1 to 5 ng) of $Gcr1^{U}$ and $Gcr1^{S}$ with two replicates using ThruPLEX DNA-Seq Kit (TaKaRa) according to the manufacturer's protocol. In brief, DNA fragments were subjected to steps of end-repair, 3′A-tailing, and adapter ligation. Then DNA was PCR amplified (15 cycles) and purified. Single-end sequencing with 50 cycles was performed using a HiSeq2500 (Illumina) instrument according to manufacturer's protocol.

Read quality was assessed using FastQC (v0.10.1) (136),

showing about 90% bases above Q30 across all samples. Reads were aligned to the yeast reference genome (sacCer3 assembly) using BWA (v0.7.15) (137) with the allowance of two mismatches, and redundant reads with identical coordinates were filtered out using Picard (v2.92) (https://broadinstitute.github.io/picard/index.html). ChIP peaks was called using HOMER (138) 'findPeaks' with -style 'factor'. Called peaks were filtered with the following conditions; (1) peak score ≥ 100, (2) poisson p-value threshold relative to local tag count < 1E-10, (3) fold enrichment over local tag count ≥ 2, and (4) relation to genes. Peaks overlapped between exponential phase and diauxic shift of $Gcr1^{U}$ and $Gcr1^{S}$ were further selected to compare their binding pattern in the two mutants, and ChIP reads fallen in peaks were then collected using BEDTools 'intersectbed' (139). Moreover, peaks that are differentially enriched between the two experiments were examined by using HOMER 'getDifferentialPeaks.' The resulting peak calls were annotated using HOMER annotatePeaks with a pre-configured genome annotation provided from HOMER. The enrichments of $Gcr1^{U}$ and $Gcr1^{S}$-bound regions in the yeast genome were plotted using seqMiner (140) with ChIP reads fallen in peaks.

3.9. Data availability

The ChIP-seq and RNA-seq data have been deposited in the NCBI Sequence Read Archive (SRA) under the accession numbers of BioProject ID PRJNA639179 (SRA accession numbers: SRR12006023-SRR12006028 and SRR12006035-SRR12006036 for ChIP seq, SRR12006021-SRR12006022 and SRR12006029-SRR12006036 for RNA-seq).

3.10. HPLC analysis

To determine the concentrations of metabolites, 600 μL of culture supernatants were filtered through a 0.22 μm syringe filter and analyzed by high performance liquid chromatography (HPLC) with BioRad Aminex HPX-87H column. 5 mM H_2SO_4 was used as a mobile phase at a flow rate 0.6 mL/min and column and refractive index (RI) detector temperature were maintained at 60°C and 35°C, respectively.

Table 6. Primers for qRT-PCR and ChIP

Forward primer (5' to 3')	Reverse primer (5' to 3')	Target gene or locus
Primer sequence for qRT-PCR		
GCTGCTGGTAACACCGTCATCATTGG	CCACCACCAGTAGAGACATGGGAG	*PGK1*
CTGCTCAAGACTCTTTCGCTGCCAAC	CCGGCGTAGACAGCCTTGTCACCC	*ENO2*
CGAAAAGGAACCTGTCTCTGACTGGAC	CGGTAGAGACTTGCAAAGTGTTGGAG TGACC	*PYK1* (*CDC19*)
GCCAGCTGGTGCCAAGTGTTGT	TCCATCTTTTCGTAAATTTCTGGCAAG GTA	*ADH1*
GCCAGCCGGTGCAAAGTGCTCC	CTCCATCTTTTCGTAAATTTCTGGTA AACTG	*ADH2*
GAGCACAGGTTTGAAGGTGGCCAAG	GACTTCTTCACCCATTTCTCTACCGTA AC	*ALD6*
GGGTTGGCTGCTGGTATTCACACCT	CAGACATTTCCCTGCCCAAACCA	*ALD4*
GGATGTGAACGAGCGCCCATTATG	CCACGGCAACTTCCCAATACTTC	*GUT1*
Primer sequence for ChIP		
TCGGATCCTCAAAACCCTTAAAAAC	CGTGGCGGGTAAAGAAGAAAATGG	Promoter of *ACT1* (negative control)
GCGGGAAAGGGTTTAGTACCACATG	TGTCACACGATTCGGACAATTCTG	Promoter of *PGK1*
GGTACGGCTGTTATCCAGCGATGC	GTCAATTGTCACCGACAAACCCCCC	Promoter of *ENO2*
CGGATATCCTTTTGTTGTTTCCGG	GGGAGACCAACGAAGGTATTATAG	Promoter of *ADH*1
CAGGAATGTTCCACGTGAAGC	CAATGAGCTCTGAAGACGAATTG	Promoter of *ADH*2
CAAGACTTTTAGAACGGATAAGGTG	CCGGCCACAACTCAAACCAC	Promoter of *GUT1*
GGTACGGCTGTTATCCAGCGATGC	GTCAATTGTCACCGACAAACCCCCC	Promoter of *ALD4*

Chapter 4.

Determination of differential activation mechanisms of two isoforms of Gcr1

4.1. Introduction

Gcr1 is one of the most important transcription factors supporting *S. cerevisiae* cell growth through high-level expression of glycolytic genes. However, the presence of two major isoforms of Gcr1, $Gcr1^U$ (844 amino acids) and $Gcr1^S$ (789 amino acids), generated from the unspliced and spliced forms of *GCR1* mRNA, respectively, has not been recognized until recently (102).

When the *GCR1* gene was first cloned, a *GCR1* ORF corresponding to that generating $Gcr1^U$ was predicted without realizing the presence of an intron (141). Later, *GCR1* cDNA, generating $Gcr1^A$ (785 amino acids), was cloned from a cDNA library (142). Therefore, previous studies on Gcr1 were conducted with either $Gcr1^U$ or $Gcr1^A$ without considering the presence of two isoforms. Although the spliced mRNA generating $Gcr1^A$ was turned out to be a minor isoform (101, 102), it was initially annotated as a spliced form (142) and has been used in most of previous studies on Gcr1 ever since. It turned out that the spliced mRNA species generating $Gcr1^A$ represents only very minor portion of the *GCR1* mRNA isoforms generated by alternative splicing compared with that generating $Gcr1^S$ (101, 102). Amino acid sequences of $Gcr1^A$ and $Gcr1^S$ are almost identical except for a few N-terminal amino

acids; VCT from position 2 to 4 of $Gcr1^{A}$ is replace by QTSVDST in Gcr1S (101). Considering that these N-terminal amino acids are not critical for the known Gcr1 function, $Gcr1^{A}$ and $Gcr1^{S}$ might have similar characteristics.

In a previous study based on expression of these two isoforms from episomal plasmids in *GCR1* deletion mutant, cells showed normal growth only when both Gcr1 isoforms were expressed, suggesting that each isoform might have a complementary role to support cell growth (102). However, it has not yet been elucidated how these two isoforms play different roles.

In this chapter, to understand the differential functions of the $Gcr1^{U}$ and $Gcr1^{S}$ isoforms, I used CRISPR/Cas9-mediated genome editing, which allows a minimum of genomic perturbation and native-level expression of the genes. By generating *S. cerevisiae* strains expressing either $Gcr1^{U}$ or $Gcr1^{S}$ in most native genomic level, I was able to figure out more detailed regulatory mechanisms between two isoforms.

4.2. No phenotypic difference was observed between strains producing only $Gcr1^{U}$ or $Gcr1^{S}$

In order to generate strains producing only $Gcr1^{U}$ or $Gcr1^{S}$ in a native state, CRISPR/Cas9-mediated genome editing

was conducted in *S. cerevisiae* BY4741 strain. $Gcr1^U$ strain, producing only the unspliced form of Gcr1 protein, was generated by deleting exon1 (a start codon) and 574-bp intron region upstream of the intronic start codon of the *GCR1* gene (Figure 8A). $Gcr1^S$ strain, producing only the spliced form of Gcr1 protein, was generated by deleting 739-bp intron region between the major 5′ and 3′ splicing sites (Figure 8A). Compared with the $Gcr1^S$ protein, the $Gcr1^U$ protein has additional 55 amino acids at the N-terminus, which was named USS (Un-spliced form specific) domain (Figure 8A).

In a previous study using plasmid-based gene expression, cells expressing $Gcr1^U$ or $Gcr1^S$ alone showed growth defects compared with cells expressing both isoforms (102). Therefore, I first examined growth rates of our strains in rich YPD medium containing 2% glucose. However, unlike the previous study, both strains showed the same growth rates as wild-type BY4741 strain (Figure 8B). I also tested different culture conditions including SC minimal medium, different concentrations of glucose (0, 0.1, and 30%), different carbon sources (3% glycerol, 2% ethanol, and 2% glycerol + 2% lactate), various environmental stress conditions (heat shock at 37°C, cold shock at 25 °C, 500 mM NaCl, 1 M sorbitol, and 1 mM/10 mM 3-amino-1,2,4-triazole), but $Gcr1^U$ and $Gcr1^S$ strains

showed no significant difference in growth compared with wild type (Figure 9).

To confirm the expression of the specific Gcr1 isoform in each strain, I also generated strains expressing $Gcr1^{U}$ or $Gcr1^{S}$ tagged with TAP at the C-terminus. $Gcr1^{U}$-TAP and $Gcr1^{S}$-TAP strains also showed no growth defect as compared with wild-type Gcr1-TAP strain (Figure 10A). In Western blotting analysis, wild-type Gcr1-TAP and Gcr1-5Flag strains showed two Gcr1 protein bands corresponding to $GGcr1^{U}$ and $Gcr1^{S}$ (Figure 10B, C). However, the $Gcr1^{U}$-TAP and $Gcr1^{S}$-TAP (or $Gcr1^{U}$-5Flag and $Gcr1^{S}$-5Flag) strains showed only its respective isoform, confirming that the normal growth phenotypes of the $Gcr1^{U}$ and $Gcr1^{S}$ strains are not due to the concurrent production of two Gcr1 isoforms.

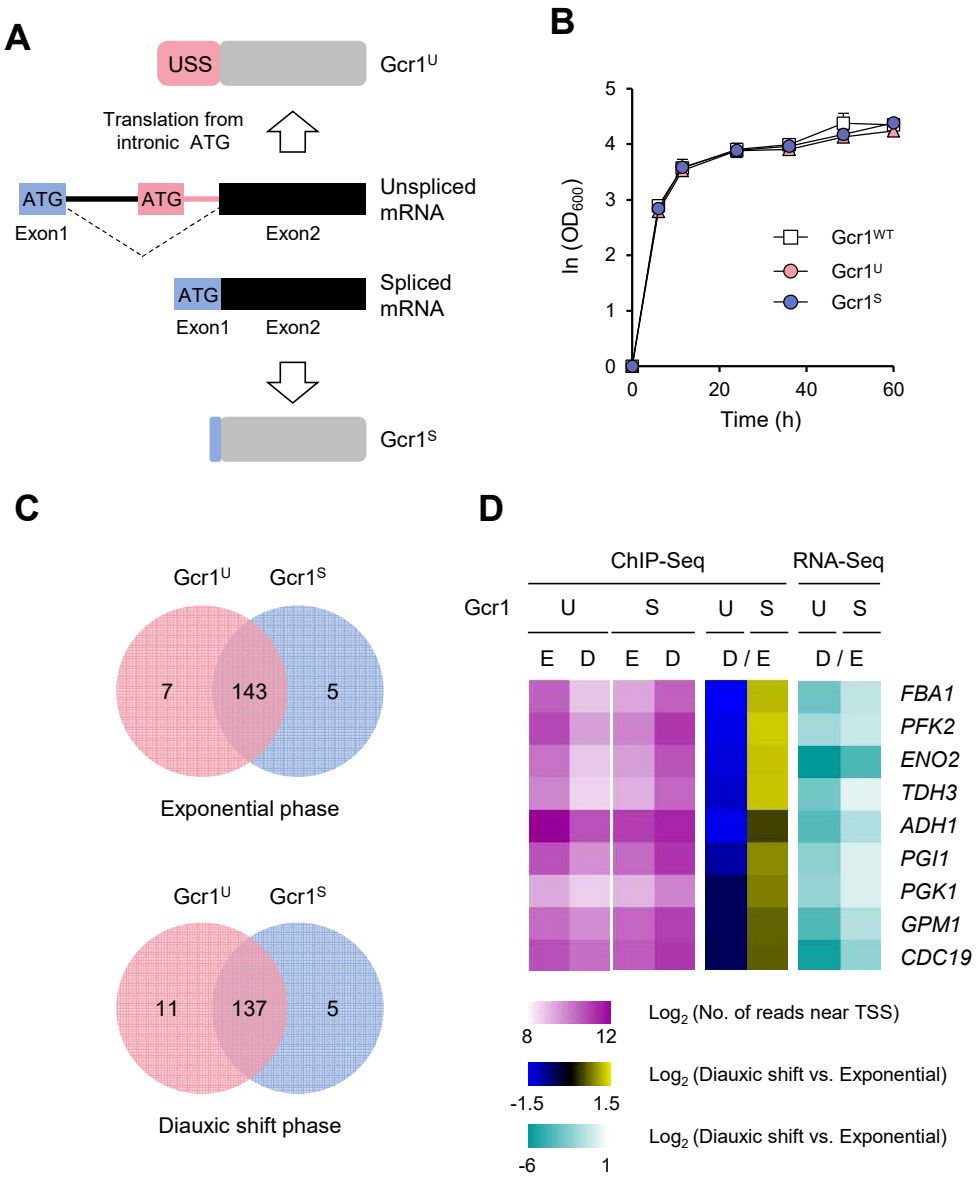

Figure 8. Gcr1^{U} and Gcr1^{S} has no phenotypic difference and mostly bind to the same targets, but with different affinities depending on growth phase

Figure 8. $Gcr1^U$ and $Gcr1^S$ has no phenotypic difference and mostly bind to the same targets, but with different affinities depending on growth phase

(A) Schematic view of production of $Gcr1^U$ and $Gcr1^S$ from un-spliced and spliced forms of mRNA, respectively. Without splicing, translation starts from the intronic start codon (pink box) and generates $Gcr1^U$. When splicing occurs using the major 5′ and 3′ splicing sites, the start codon in the exon 1 (blue box) is used to generate $Gcr1^S$. USS indicates un-spliced form-specific domain.

(B) Growth curves of $Gcr1^{WT}$, $Gcr1^U$, and $Gcr1^S$ strains. Cells were grown in YPD media containing 2% glucose. Error bars indicate the standard deviations of the three independent experiments.

(C) Venn diagrams show the number of overlapping target genes of $Gcr1^U$ and $Gcr1^S$ identified by ChIP-Seq analysis at exponential and diauxic shift phases. The binding pattern of target genes are visualized by heatmap in Figure 11 and listed in Table 7. Genes used in the Venn diagram are listed in Table 8.

(D) Relative abundance of ChIP-seq 'reads' related to TSS of nine targeted genes at exponential (E) and diauxic shift (D) phases were visualized for $Gcr1^U$ (U) and $Gcr1^S$ (S) strains. The log2 fold-changes (D vs. E) of 'reads' abundance were visualized for each Gcr1 isoform. Target gene expression levels were examined by RNA-seq and log2 fold-changes (D vs. E) of FPKM were visualized as heat maps for each Gcr1 isoform. The color scales and the names of the nine target genes are indicated.

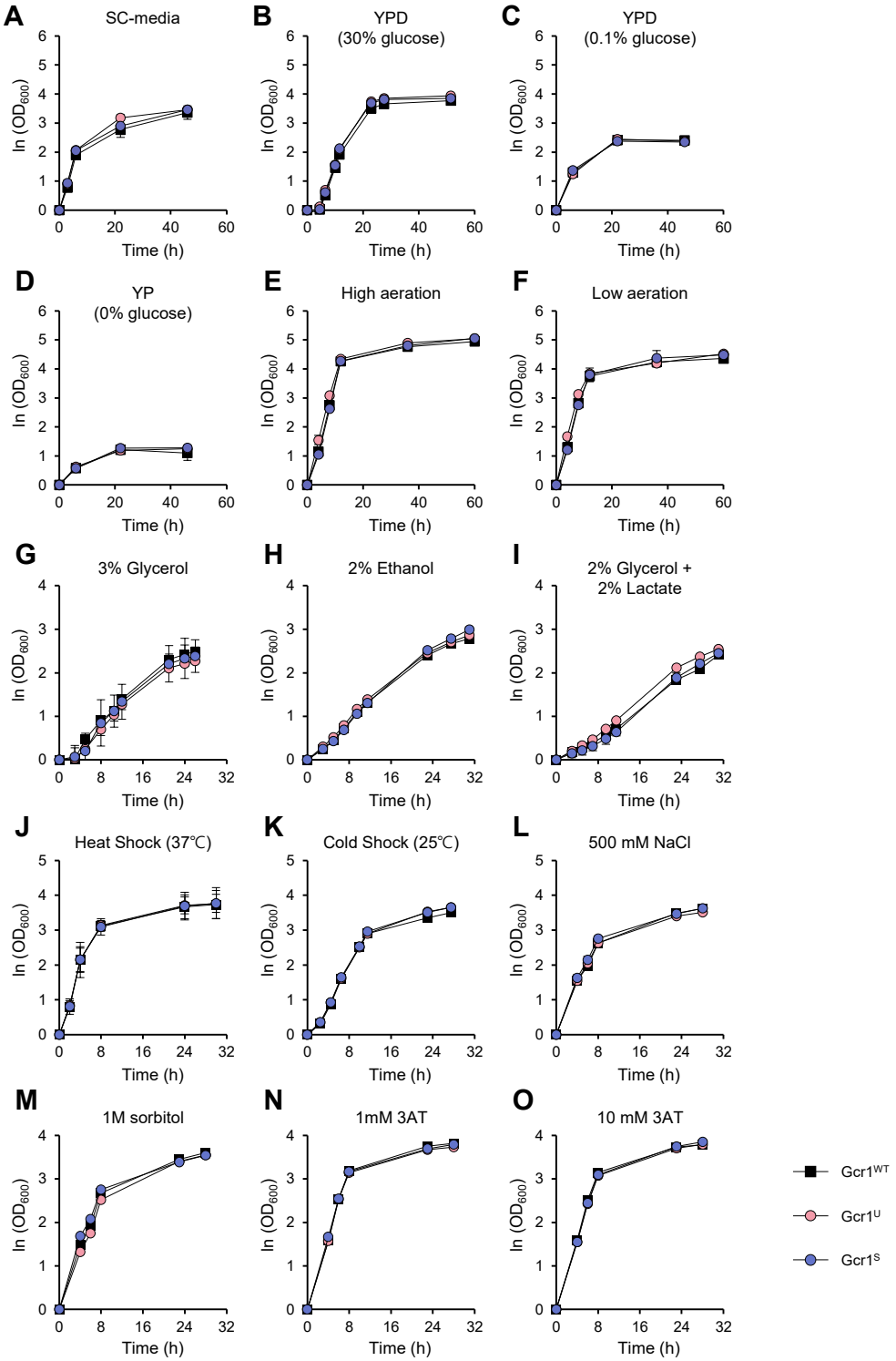

Figure 9. Growth curves of $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$ strains under various stress conditions

Figure 9. Growth curves of $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$ strains under various stress conditions

Three strains were grown in SC media containing 2% glucose (A), YPD media containing high (B) or low (C) concentration of glucose, or no glucose (D), and high (E, 5 mL culture in 50 mL flask) or low (F, 50 mL culture in 50 mL flask) aeration. Different carbon sources (G, 3% glycerol; H, 2% ethanol; I, 2% glycerol + 2% lactate), temperature change (J, heat shock 37℃; K, cold shock 25℃), osmotic stresses (L, 500 mM NaCl; M, 1M sorbitol), and chemical stresses (N, 3-amino triazole 1 mM; O, 10 mM) were also tested. The error bars indicate the standard deviations of two independent experiments.

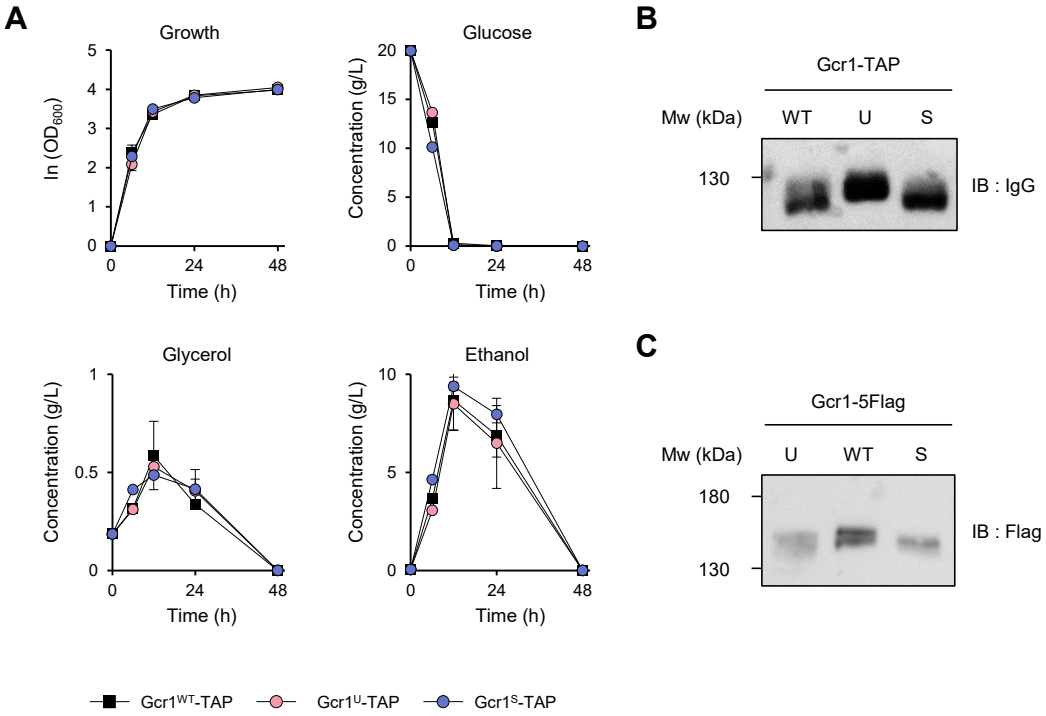

Figure 10. Properties of TAP-tagged Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S} strains

(A) Growth curves and metabolite profiles of strains expressing TAP-tagged Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S}. Error bars indicate standard deviations of three independent experiments.

(B, C) Immunoblotting analysis of TAP (B) or Flag (C) - tagged Gcr1 proteins. Strains expressing TAP or Flag-tagged Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S} were grown in YPD media until the exponential phase, and Gcr1-TAP or Gcr1-Flag proteins were detected by immunoblotting using anti-IgG or anti-DDDDK antibody.

4.3. $Gcr1^U$ and $Gcr1^S$ mostly bind to the same targets, but with different intensities depending on growth phase

Normal growth of $Gcr1^U$ and $Gcr1^S$ strains suggest that the functions of $Gcr1^U$ and $Gcr1^S$ are largely indistinguishable from each other, at least under our laboratory culture conditions. Therefore, I next investigated whether $Gcr1^U$ and $Gcr1^S$ regulate any other sets of target genes. Genome-wide binding targets of $Gcr1^U$ and $Gcr1^S$ were examined at exponential and post-diauxic shift phases by ChIP-seq analysis. I identified 155 genes showing 'reads' enrichment around the transcription start sites (TSS) (Figure 11, Table 7). Most of the identified genes overlapped between $Gcr1^U$ and $Gcr1^S$ during both growth phases (Figure 8C, Table 8). Among the identified binding sites, I further selected biologically meaningful targets of Gcr1 by excluding 50 unnamed genes, 26 tRNA genes, and seven genes near telomeres. In agreement with previous studies showing Gcr1-dependent regulation of transposable elements (Ty) (143-145), 33 peaks were located near transposable elements. After filtering out those sites, the 39 remaining binding targets of $Gcr1^U$ and $Gcr1^S$ mainly represented glycolytic genes consistent with the known role of Gcr1 (Table 9).

Although I could not identify any meaningful target genes specific for either $Gcr1^U$ or $Gcr1^S$ the DNA binding affinities for each Gcr1 isoform showed a different tendency depending on growth phase. ChIP-seq analysis of $Gcr1^U$ showed a decrease in the 'reads' abundance of the targets from the exponential to the diauxic shift phases, whereas ChIP-seq analysis of $Gcr1^S$ showed an opposite trend (Figure 8D). I also examined expression levels of the target genes in $Gcr1^U$ and $Gcr1^S$ strains by using RNA-seq analysis. Expression of the glycolytic genes decreased after a diauxic shift in both $Gcr1^U$ and $Gcr1^S$ strains, reflecting the contribution of glycolysis in the presence of glucose. However, the fold-changes in target gene expression (diauxic shift/exponential) were greater in the $Gcr1^U$ strain than in the $Gcr1^S$ strain (Figure 8D). Therefore, $Gcr1^U$ may work mainly during the exponential phase, but stronger $Gcr1^S$ binding to the target gene promoters after diauxic shift might support residual expression of glycolytic genes even after glucose depletion. Taken together, $Gcr1^U$ and $Gcr1^S$ might play differential roles not by regulating different sets of target genes, but by differential binding to the same genes depending on growth conditions.

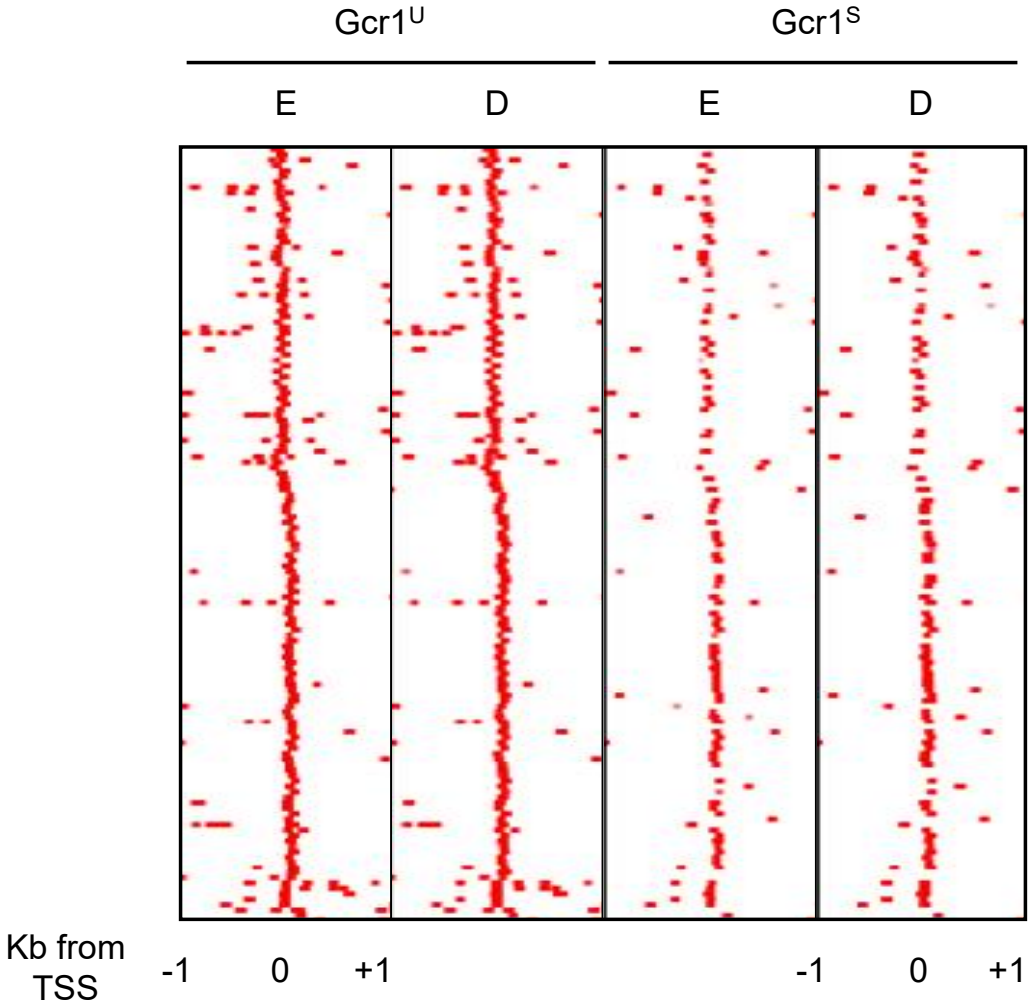

Figure 11. Heatmap of Gcr1^{U} and Gcr1^{S} occupancies from two independent ChIP-seqs at exponential phase and diauxic shift phases

Regions are sorted by the pattern of distribution of ChIP-seq reads mapped to peaks of Gcr1^{U} and Gcr1^{S}. The x axis indicates Gcr1 occupied regions within 2 kb of gene TSS and the y axis indicates each gene. E, exponential phase; D, diauxic shift Phase.

Table 7. The list of gene IDs, gene names and clusters used for heatmap of Gcr1 ChIP-seq analysis

Order	Chromosome	TSS start (bp)	TSS end (bp)	Gene ID	Gene Name	Strand (Gene)	Cluster
1	chrII	259567	260566	YBR012C	N/A*	-	C1
2	chrII	326060	327059	YBR044C	TCM62	-	C1
3	chrII	36053	37052	YBL099W	ATP1	+	C1
4	chrII	226636	227635	YBL004W	UTP20	+	C1
5	chrXIV	547100	548099	YNL042W	BOP3	+	C1
6	chrXIV	599231	600230	YNL019C	N/A	-	C1
7	chrXIV	631062	632061	YNR001C	CIT1	-	C1
8	chrXIV	663270	664269	YNR018W	RCF2	+	C1
9	chrXIV	750009	751008	YNR064C	N/A	-	C1
10	chrXIV	569477	570476	YNL034W	N/A	+	C1
11	chrXIV	601775	602774	YNL018C	N/A	-	C1
12	chrXIV	783542	784541	YNR077C	N/A	-	C1
13	chrIII	147635	148634	YCR017C	CWH43	-	C1
14	chrIII	123004	124003	YCR006C	N/A	-	C1
15	chrI	141174	142173	YAL003W	EFB1	+	C1
16	chrI	11952	12951	YAL065C	N/A	-	C1
17	chrI	113360	114359	YAL021C	CCR4	-	C1
18	chrI	139760	140759	YAL004W	N/A	+	C1
19	chrI	207367	208366	YAR053W	N/A	+	C1
20	chrI	70786	71785	YAL038W	CDC19	+	C1
21	chrXVI	571379	572378	YPR007C	REC8	-	C1
22	chrXVI	942880	943879	YPR203W	N/A	+	C1
23	chrV	135279	136278	YEL010W	N/A	+	C1
24	chrV	526082	527081	YER171W	RAD3	+	C1
25	chrIV	83270	84269	YDL210W	UGA4	+	C1
26	chrIV	802223	803222	YDR170C	SEC7	-	C1
27	chrVII	845655	846654	YGR174W-A	N/A	+	C1
28	chrVII	356377	357376	YGL081W	N/A	+	C1
29	chrXII	822593	823592	YLR346C	CIS1	-	C1
30	chrVII	706502	707501	YGR109C	CLB6	-	C1
31	chrVII	774193	775192	YGR143W	SKN1	+	C1
32	chrXII	5486	6485	YLL066W-A	N/A	+	C1
33	chrVII	779399	780398	YGR144W	THI4	+	C1
34	chrXII	687203	688202	YLR272C	YCS4	-	C1
35	chrXII	818312	819311	YLR344W	RPL26A	+	C1
36	chrVII	1083864	1084863	YGR296W	YRF1-3	+	C1
37	chrXII	1070870	1071869	YLR466C-A	N/A	-	C1
38	chrVII	661358	662357	YGR090W	UTP22	+	C1
39	chrVII	735759	736758	YGR122C-A	N/A	-	C1
40	chrVIII	5401	6400	YHL048W	COS8	+	C1
41	chrVIII	105055	106054	YHR001W	OSH7	+	C1
42	chrVIII	561682	562681	YHR219C-A	N/A	-	C1
43	chrXIII	807549	808548	YMR271C	URA10	-	C1
44	chrXIII	919080	920079	YMR322C	SNO4	-	C1
45	chrXIII	258417	259416	YML006C	GIS4	-	C1
46	chrXIII	674767	675766	YMR205C	PFK2	-	C1
47	chrXI	73866	74865	YKL197C	PEX1	-	C1

48	chrXIII	923803	924802	YMR326C	N/A	-	C1
49	chrXIII	420030	421029	YMR076C	PDS5	-	C1
50	chrXV	123001	124000	YOL103W	ITR2	+	C1
51	chrXV	160595	161594	YOL086C	ADH1	-	C1
52	chrXV	1090498	1091497	YOR396C-A	N/A	-	C1
53	chrXV	117455	118454	YOL104C	NDJ1	-	C1
54	chrXV	226076	227075	YOL055C	THI20	-	C1
55	chrVI	95011	96010	YFL022C	FRS2	-	C1
56	chrVI	225959	226958	YFR034C	PHO4	-	C1
57	chrV	139764	140763	YEL009C	GCN4	-	C1
58	chrV	311200	312199	YER075C	PTP3	-	C1
59	chrV	430450	431449	YER132C	PMD1	-	C1
60	chrV	99769	100768	YEL027W	VMA3	+	C1
61	chrV	117381	118380	YEL020C-B	N/A	-	C1
62	chrII	347302	348301	YBR055C	PRP6	-	C2
63	chrII	478338	479337	YBR119W	MUD1	+	C2
64	chrII	613901	614900	YBR196C	PGI1	-	C2
65	chrXIV	62944	63943	YNL302C	RPS19B	-	C2
66	chrXIV	374693	375692	YNL133C	FYV6	-	C2
67	chrIII	136746	137745	YCR012W	PGK1	+	C2
68	chrIII	168000	168999	YCR027C	RHB1	-	C2
69	chrIII	227318	228317	YCR063W	BUD31	+	C2
70	chrIII	316189	317188	YCR108C	N/A	-	C2
71	chrXVI	338621	339620	YPL112C	PEX25	-	C2
72	chrXVI	411254	412253	YPL075W	GCR1	+	C2
73	chrXVI	731749	732748	YPR102C	RPL11A	-	C2
74	chrXVI	743689	744688	YPR109W	N/A	+	C2
75	chrV	491959	492958	YER159C	BUR6	-	C2
76	chrV	86216	87215	YEL034C-A	N/A	-	C2
77	chrV	305323	306322	YER074W	RPS24A	+	C2
78	chrV	354140	355139	YER097W	N/A	+	C2
79	chrV	441820	442819	YER137C	N/A	-	C2
80	chrIV	1524934	1525933	YDR543C	N/A	-	C2
81	chrIV	410825	411824	YDL022W	GPD1	+	C2
82	chrIV	488663	489662	YDR022C	ATG31	-	C2
83	chrIV	575474	576473	YDR062W	LCB2	+	C2
84	chrIV	619646	620645	YDR088C	SLU7	-	C2
85	chrIV	644931	645930	YDR098C	GRX3	-	C2
86	chrIV	945807	946806	YDR242W	AMD2	+	C2
87	chrIV	1075172	1076171	YDR306C	N/A	-	C2
88	chrIV	1256848	1257847	YDR390C	UBA2	-	C2
89	chrIV	1461555	1462554	YDR506C	GMC1	-	C2
90	chrVII	253860	254859	YGL136C	MRM2	-	C2
91	chrXII	10726	11725	YLL065W	N/A	+	C2
92	chrXII	214457	215456	YLR035C	MLH2	-	C2
93	chrVII	883811	884810	YGR192C	TDH3	-	C2
94	chrXII	167803	168802	YLR010C	TEN1	-	C2
95	chrVII	401288	402287	YGL054C	ERV14	-	C2
96	chrXII	731542	732541	YLR303W	MET17	+	C2
97	chrVII	541203	542202	YGR028W	MSP1	+	C2
98	chrVII	730822	731821	YGR120C	COG2	-	C2

99	chrXII	1065572	1066571	YLR464W	N/A	+	C2
100	chrVII	1003963	1004962	YGR255C	COQ6	-	C2
101	chrIX	253926	254925	YIL055C	N/A	-	C2
102	chrIX	183128	184127	YIL096C	BMT5	-	C2
103	chrIX	241776	242775	YIL065C	FIS1	-	C2
104	chrVIII	450327	451326	YHR174W	ENO2	+	C2
105	chrVIII	62564	63563	YHL023C	NPR3	-	C2
106	chrXI	83704	84703	YKL189W	HYM1	+	C2
107	chrVIII	85061	86060	YHL009C	YAP3	-	C2
108	chrXIII	768040	769039	YMR247C	RKR1	-	C2
109	chrVIII	382752	383751	YHR141C	RPL42B	-	C2
110	chrXIII	195756	196755	YML041C	VPS71	-	C2
111	chrVIII	388727	389726	YHR144C	DCD1	-	C2
112	chrVIII	475341	476340	YHR185C	PFS1	-	C2
113	chrXIII	290134	291133	YMR012W	CLU1	+	C2
114	chrXI	202541	203540	YKL127W	PGM1	+	C2
115	chrXV	24295	25294	YOL157C	IMA2	-	C2
116	chrXI	327488	328487	YKL060C	FBA1	-	C2
117	chrXIII	321018	322017	YMR023C	MSS1	-	C2
118	chrXI	140692	141691	YKL165C	MCD4	-	C2
119	chrXV	216138	217137	YOL060C	MAM3	-	C2
120	chrXV	227614	228613	YOL054W	PSH1	+	C2
121	chrXV	704225	705224	YOR192C-C	N/A	-	C2
122	chrX	424110	425109	YJL006C	CTK2	-	C2
123	chrXI	164386	165385	YKL152C	GPM1	-	C2
124	chrX	537772	538771	YJR055W	HIT1	+	C2
125	chrXV	1075784	1076783	YOR390W	FEX1	+	C2
126	chrX	745262	746261	YJR162C	N/A	-	C2
127	chrXI	517198	518197	YKR041W	N/A	+	C2
128	chrVI	161488	162487	YFR009W	GCN20	+	C2
129	chrVI	221419	222418	YFR031C-A	RPL2A	-	C2
130	chrII	5010	6009	YBL111C	N/A	-	C3
131	chrII	808057	809056	YBR301W	PAU24	+	C3
132	chrII	7734	8733	YBL108C-A	PAU9	-	C3
133	chrXIV	546113	547112	YNL042W-B	N/A	+	C3
134	chrXIV	602477	603476	YNL017C	N/A	-	C3
135	chrI	20566	21565	YAL064W	N/A	+	C3
136	chrI	141432	142431	YAL005C	SSA1	-	C3
137	chrXVI	743175	744174	YPR108W-A	N/A	+	C3
138	chrV	569908	570907	YER188C-A	N/A	-	C3
139	chrVII	402687	403686	YGL051W	MST27	+	C3
140	chrVII	402437	403436	YGL052W	N/A	+	C3
141	chrIV	538468	539467	YDR040C	ENA1	-	C3
142	chrIV	1017319	1018318	YDR278C	N/A	-	C3
143	chrIV	1080200	1081199	YDR309C	GIC2	-	C3
144	chrIV	1411373	1412372	YDR477W	SNF1	+	C3
145	chrXII	1061919	1062918	YLR461W	PAU4	+	C3
146	chrXII	97486	98485	YLL024C	SSA2	-	C3
147	chrVIII	214250	215249	YHR054C	N/A	-	C3
148	chrVIII	451154	452153	YHR173C	N/A	-	C3
149	chrVIII	465932	466931	YHR180W-A	N/A	+	C3

150	chrVIII	5798	6797	YHL048C-A	N/A	-	C3
151	chrXV	300047	301046	YOL013W-A	N/A	+	C3
152	chrX	423013	424012	YJL007C	N/A	-	C3
153	chrXV	463469	464468	YOR072W-B	N/A	+	C3
154	chrX	638942	639941	YJR115W	N/A	+	C3
155	chrXI	171130	172129	YKL148C	SDH1	-	C3
* Not available							

Table 8. The list of ChIP-seq target genes of $Gcr1^{U}$ (U) and $Gcr1^{S}$ (S) at exponential (E) and diauxic shift (D) phases

Growth phase	E			D		
Gcr1 isoform	U	S	overlap	U	S	overlap
	YAL003W	YNL042W	YBR012C	YNR064C	YNL042W	YBR012C
	YPR102C	YAL005C	YBR044C	YGR143W	YAL005C	YBR044C
	YPR109W	YPR108W-A	YBL099W	YFR034C	YJL007C	YBL099W
	YER137C	YOL013W-A	YBL004W	YBR055C	YOR072W-B	YBL004W
	YFR031C-A	YJR115W	YNL019C	YNL133C	YJR115W	YNL019C
	YNL042W-B		YNR001C	YPL112C		YNR001C
	YGL052W		YNR018W	YEL034C-A		YNR018W
			YNR064C	YER137C		YNL034W
			YNL034W	YDR306C		YNL018C
			YNL018C	YMR012W		YNR077C
			YNR077C	YNL042W-B		YCR017C
			YCR017C			YCR006C
			YCR006C			YAL003W
			YAL065C			YAL065C
			YAL021C			YAL021C
			YAL004W			YAL004W
			YAR053W			YAR053W
			YAL038W			YAL038W
			YPR007C			YPR007C
			YPR203W			YPR203W
			YEL010W			YEL010W
			YER171W			YER171W
			YDL210W			YDL210W
			YDR170C			YDR170C
			YGR174W-A			YGR174W-A
			YGL081W			YGL081W
			YLR346C			YLR346C
			YGR109C			YGR109C
			YGR143W			YLL066W-A
			YLL066W-A			YGR144W
			YGR144W			YLR272C
			YLR272C			YLR344W
			YLR344W			YGR296W
			YGR296W			YLR466C-A
			YLR466C-A			YGR090W
			YGR090W			YGR122C-A
			YGR122C-A			YHL048W
			YHL048W			YHR001W
			YHR001W			YHR219C-A
			YHR219C-A			YMR271C
			YMR271C			YMR322C
			YMR322C			YML006C
			YML006C			YMR205C
			YMR205C			YKL197C
			YKL197C			YMR326C
			YMR326C			YMR076C
			YMR076C			YOL103W
			YOL103W			YOL086C
			YOL086C			YOR396C-A

			YOR396C-A			YOL104C
			YOL104C			YOL055C
			YOL055C			YFL022C
			YFL022C			YEL009C
			YFR034C			YER075C
			YEL009C			YER132C
			YER075C			YEL027W
			YER132C			YEL020C-B
			YEL027W			YBR119W
			YEL020C-B			YBR196C
			YBR055C			YNL302C
			YBR119W			YCR012W
			YBR196C			YCR027C
			YNL302C			YCR063W
			YNL133C			YCR108C
			YCR012W			YPL075W
			YCR027C			YPR102C
			YCR063W			YPR109W
			YCR108C			YER159C
			YPL112C			YER074W
			YPL075W			YER097W
			YER159C			YDR543C
			YEL034C-A			YDL022W
			YER074W			YDR022C
			YER097W			YDR062W
			YDR543C			YDR088C
			YDL022W			YDR098C
			YDR022C			YDR242W
			YDR062W			YDR390C
			YDR088C			YDR506C
			YDR098C			YGL136C
			YDR242W			YLL065W
			YDR306C			YLR035C
			YDR390C			YGR192C
			YDR506C			YLR010C
			YGL136C			YGL054C
			YLL065W			YLR303W
			YLR035C			YGR028W
			YGR192C			YGR120C
			YLR010C			YLR464W
			YGL054C			YGR255C
			YLR303W			YIL055C
			YGR028W			YIL096C
			YGR120C			YIL065C
			YLR464W			YHR174W
			YGR255C			YHL023C
			YIL055C			YKL189W
			YIL096C			YHL009C
			YIL065C			YMR247C
			YHR174W			YHR141C
			YHL023C			YML041C
			YKL189W			YHR144C
			YHL009C			YHR185C
			YMR247C			YKL127W
			YHR141C			YOL157C
			YML041C			YKL060C

			YHR144C			YMR023C
			YHR185C			YKL165C
			YMR012W			YOL060C
			YKL127W			YOL054W
			YOL157C			YOR192C-C
			YKL060C			YJL006C
			YMR023C			YKL152C
			YKL165C			YJR055W
			YOL060C			YOR390W
			YOL054W			YJR162C
			YOR192C-C			YKR041W
			YJL006C			YFR009W
			YKL152C			YFR031C-A
			YJR055W			YBL111C
			YOR390W			YBR301W
			YJR162C			YBL108C-A
			YKR041W			YNL017C
			YFR009W			YAL064W
			YBL111C			YER188C-A
			YBR301W			YGL051W
			YBL108C-A			YGL052W
			YNL017C			YDR040C
			YAL064W			YDR278C
			YER188C-A			YDR309C
			YGL051W			YDR477W
			YDR040C			YLR461W
			YDR278C			YLL024C
			YDR309C			YHR054C
			YDR477W			YHR173C
			YLR461W			YHR180W-A
			YLL024C			YHL048C-A
			YHR054C			YKL148C
			YHR173C			
			YHR180W-A			
			YHL048C-A			
			YJL007C			
			YOR072W-B			
			YKL148C			

Table 9. The list of filtered target genes of $Gcr1^{U}$ (U) and $Gcr1^{S}$ (S) at exponential (E) and diauxic shift (D) phases

Gene ID	Gene Name	Property	Cluster	E		D	
				U	S	U	S
YAL038W	CDC19	glycolytic enzyme	C1	o	o	o	o
YMR205C	PFK2	glycolytic enzyme	C1	o	o	o	o
YBR196C	PGI1	glycolytic enzyme	C2	o	o	o	o
YCR012W	PGK1	glycolytic enzyme	C2	o	o	o	o
YGR192C	TDH3	glycolytic enzyme	C2	o	o	o	o
YHR174W	ENO2	glycolytic enzyme	C2	o	o	o	o
YKL060C	FBA1	glycolytic enzyme	C2	o	o	o	o
YKL152C	GPM1	glycolytic enzyme	C2	o	o	o	o
YOL086C	ADH1	glycolytic enzyme	C1	o	o	o	o
YOL157C	IMA2		C2	o	o	o	o
YDL022W	GPD1		C2	o	o	o	o
YOL060C	MAM3		C2	o	o	o	o
YKL127W	PGM1		C2	o	o	o	o
YPL075W	GCR1		C2	o	o	o	o
YDR477W	SNF1		C3	o	o	o	o
YKL148C	SDH1		C3	o	o	o	o
YNR018W	RCF2		C1	o	o	o	o
YCR017C	CWH43		C1	o	o	o	o
YAL021C	CCR4		C1	o	o	o	o
YER171W	RAD3		C1	o	o	o	o
YLR346C	CIS1		C1	o	o	o	o
YHR001W	OSH7		C1	o	o	o	o
YFL022C	FRS2		C1	o	o	o	o
YEL009C	GCN4		C1	o	o	o	o
YER132C	PMD1		C1	o	o	o	o
YBR119W	MUD1		C2	o	o	o	o
YNL302C	RPS19B		C2	o	o	o	o
YER074W	RPS24A		C2	o	o	o	o
YHR141C	RPL42B		C2	o	o	o	o
YDR062W	LCB2		C2	o	o	o	o
YGL136C	MRM2		C2	o	o	o	o
YIL065C	FIS1		C2	o	o	o	o
YDR040C	ENA1		C3	o	o	o	o
YDR309C	GIC2		C3	o	o	o	o
YLL024C	SSA2		C3	o	o	o	o
YAL003W	EFB1		C1	o		o	o
YPR102C	RPL11A		C2	o		o	o
YFR031C-A	RPL2A		C2	o		o	o
YAL005C	SSA1		C3		o		o
YBR044C	TCM62	Ty element	C1	o	o	o	o
YBL099W	ATP1	Ty element	C1	o	o	o	o
YBL004W	UTP20	Ty element	C1	o	o	o	o
YNL042W	BOP3	Ty element	C1		o		o
YNR001C	CIT1	Ty element	C1	o	o	o	o
YPR007C	REC8	Ty element	C1	o	o	o	o
YDR170C	SEC7	Ty element	C1	o	o	o	o
YGR109C	CLB6	Ty element	C1	o	o	o	o
YGR143W	SKN1	Ty element	C1	o	o	o	
YGR144W	THI4	Ty element	C1	o	o	o	o
YLR272C	YCS4	Ty element	C1	o	o	o	o

YLR344W	RPL26A	Ty element	C1	o	o	o	o
YKL197C	PEX1	Ty element	C1	o	o	o	o
YOL103W	ITR2	Ty element	C1	o	o	o	o
YOL104C	NDJ1	Ty element	C1	o	o	o	o
YOL055C	THI20	Ty element	C1	o	o	o	o
YER075C	PTP3	Ty element	C1	o	o	o	o
YER159C	BUR6	Ty element	C2	o	o	o	o
YDR098C	GRX3	Ty element	C2	o	o	o	o
YDR242W	AMD2	Ty element	C2	o	o	o	o
YLR035C	MLH2	Ty element	C2	o	o	o	o
YLR010C	TEN1	Ty element	C2	o	o	o	o
YGL054C	ERV14	Ty element	C2	o	o	o	o
YLR303W	MET17	Ty element	C2	o	o	o	o
YGR028W	MSP1	Ty element	C2	o	o	o	o
YHL009C	YAP3	Ty element	C2	o	o	o	o
YMR247C	RKR1	Ty element	C2	o	o	o	o
YML041C	VPS71	Ty element	C2	o	o	o	o
YHR144C	DCD1	Ty element	C2	o	o	o	o
YOL054W	PSH1	Ty element	C2	o	o	o	o
YJR055W	HIT1	Ty element	C2	o	o	o	o
YFR009W	GCN20	Ty element	C2	o	o	o	o
YGL051W	MST27	Ty element	C3	o	o	o	o
YDL210W	UGA4	tRNA	C1	o	o	o	o
YGR090W	UTP22	tRNA	C1	o	o	o	o
YMR271C	URA10	tRNA	C1	o	o	o	o
YML006C	GIS4	tRNA	C1	o	o	o	o
YMR076C	PDS5	tRNA	C1	o	o	o	o
YFR034C	PHO4	tRNA	C1	o	o	o	
YEL027W	VMA3	tRNA	C1	o	o	o	o
YBR055C	PRP6	tRNA	C2	o	o	o	
YNL133C	FYV6	tRNA	C2	o	o	o	
YCR027C	RHB1	tRNA	C2	o	o	o	o
YCR063W	BUD31	tRNA	C2	o	o	o	o
YPL112C	PEX25	tRNA	C2	o	o	o	
YDR022C	ATG31	tRNA	C2	o	o	o	o
YDR088C	SLU7	tRNA	C2	o	o	o	o
YDR390C	UBA2	tRNA	C2	o	o	o	o
YDR506C	GMC1	tRNA	C2	o	o	o	o
YGR120C	COG2	tRNA	C2	o	o	o	o
YGR255C	COQ6	tRNA	C2	o	o	o	o
YIL096C	BMT5	tRNA	C2	o	o	o	o
YHL023C	NPR3	tRNA	C2	o	o	o	o
YKL189W	HYM1	tRNA	C2	o	o	o	o
YHR185C	PFS1	tRNA	C2	o	o	o	o
YMR012W	CLU1	tRNA	C2	o	o	o	
YMR023C	MSS1	tRNA	C2	o	o	o	o
YKL165C	MCD4	tRNA	C2	o	o	o	o
YJL006C	CTK2	tRNA	C2	o	o	o	o
YGR296W	YRF1-3	Near telomere	C1	o	o	o	o
YHL048W	COS8	Near telomere	C1	o	o	o	o
YMR322C	SNO4	Near telomere	C1	o	o	o	o
YOR390W	FEX1	Near telomere	C2	o	o	o	o
YBR301W	PAU24	Near telomere	C3	o	o	o	o
YBL108C-A	PAU9	Near telomere	C3	o	o	o	o
YLR461W	PAU4	Near telomere	C3	o	o	o	o

4.4. Deletion of specific Gcr1 domains or *GCR2* revealed the phenotypic differences between $Gcr1^U$ and $Gcr1^S$ strains

To test the additional hypothesis that the regulatory mechanisms of $Gcr1^U$ and $Gcr1^S$ were different, I deleted previously identified domains of Gcr1 including the alpha helix (AH), leucine zipper (LZ1), serine-proline rich (SP), and DNA binding domains (DBD) (98,99) using CRISPR/Cas9-mediated genome editing, which allowed seamless editing of each domain into the genome (Figure 12A). In addition, *GCR2*, which encodes a coactivator of Gcr1, or its 2H or LZ domain (LZ2) was deleted in cells expressing $Gcr1^{WT}$, $Gcr1^U$, or $Gcr1^S$, respectively (Figure 12A).

As previously reported (100), the *GCR1* deletion strain showed a severe growth defect in YPD media (Figure 12C). However, deletion of the SP domain did not affect growth rates against all three Gcr1 backgrounds (Figure 12D). On the other hand, deletion of the AH domain and DBD led to reduced growth rates in all three Gcr1 backgrounds, suggesting that these domains affect Gcr1 activity irrespective of the isoforms (Figure 12E and 12F). In agreement with previous data showing that Gcr1 which lacks a DBD can support activation of glycolytic genes through interaction with Rap1 (96) (Figure 12

B), the Gcr1ΔDBD strains showed slightly higher specific growth rates than the *gcr1*Δ strain. However, deletion of the LZ1 domain, which plays an important role in Gcr1 homodimerization (99), showed differential effects depending on the Gcr1 isoform. LZ1 deletion led to a significant reduction in cell growth rate in the $Gcr1^{S}$ strain, but not in the $Gcr1^{WT}$ and $Gcr1^{U}$ strains (Figure 12G). Moreover, the $Gcr1^{S}$ΔLZ1 strain showed a lower final cell density than the other strains (Figure 12G). Therefore, homodimerization may be essential for the activity of $Gcr1^{S}$, but not for $Gcr1^{U}$.

On the contrary, *GCR2* deletion mainly affected $Gcr1^{U}$, but not $Gcr1^{S}$ and $Gcr1^{WT}$ (Figure 12H). The $Gcr1^{U}$*gcr2*Δ strain showed a lower growth rate than the other two strains, but final cell density was not affected by *GCR2* deletion (Figure 12 H). Deletion of the LZ domain of Gcr2 (LZ2), which is involved in Gcr2 homodimerization, exhibited the same effects as *GCR2* deletion, resulting in a severe growth defect only in the $Gcr1^{U}$ background strain (Figure 12I). On the other hand, deletion of the 2H domain of Gcr2 led to mild growth defects in all strains (Figure 12J). These results indicate that $Gcr1^{U}$ acts as a transcription factor with the help of Gcr2 homodimer, but $Gcr1^{S}$ can be functional without Gcr2.

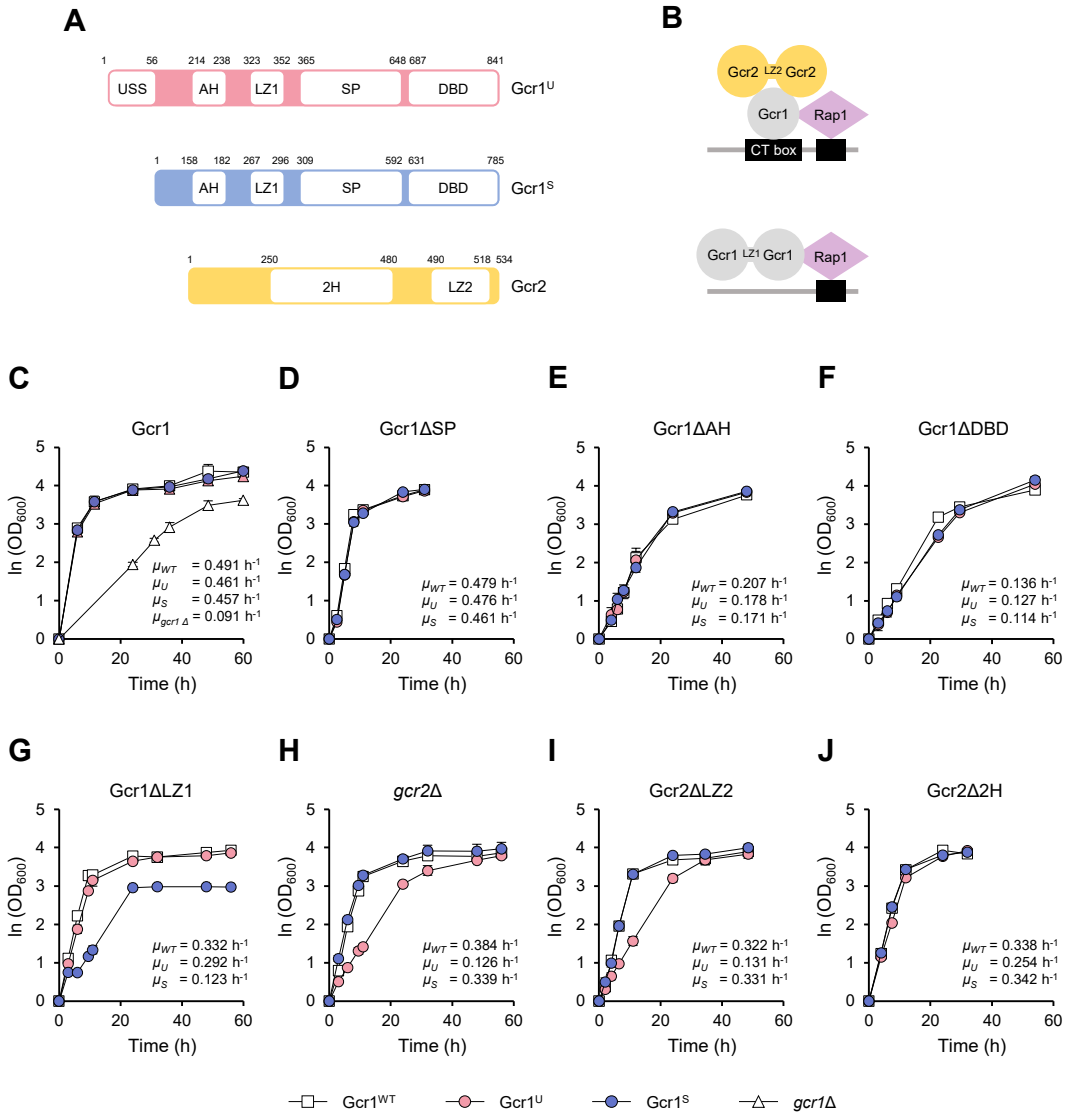

Figure 12. Deletion of specific domains of Gcr1 and Gcr2 revealed the phenotypic differences between $Gcr1^{U}$ and $Gcr1^{S}$ strains

Figure 12. Deletion of specific domains of Gcr1 and Gcr2 revealed the phenotypic differences between $Gcr1^{U}$ and $Gcr1^{S}$strains

(A) Functional domains of Gcr1 and Gcr2. AH, alpha helix; LZ1, leucine zipper of Gcr1; SP, serine-proline rich; DBD, DNA binding domain; 2H, Gcr2 region homologous to Gcr1; LZ2, leucine zipper of Gcr2.

(B) Previously suggested working models of Gcr1 and Gcr2. It is hypothesized that a Gcr1 (monomer or dimer) associates with a Gcr2 dimer to activate glycolytic genes, whereas the Gcr1 homodimer activates RP genes through interaction with Rap1.

(C-J) Growth curves and specific growth rates of $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$ strains, and strains with the indicated mutations on three different Gcr1 backgrounds. The growth curve of the gcr1Δ strain was plotted with those of the $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$ strains shown as in Figure 1B, on a log scale. Cells were grown in YPD media containing 2% glucose. The results of three (C, G, H, I) or two (D, E, F, J) independent experiments were averaged and plotted with standard deviations.

Taken together, I hypothesized a working model that $Gcr1^U$ and $Gcr1^S$ might be activated through different regulatory mechanisms (Figure 13A). $Gcr1^S$ activity was reduced by deleting the LZ1 domain, suggesting that $Gcr1^S$ mainly works as a homodimer connected through its LZ1 domain. In contrast, $Gcr1^U$ activity was reduced by deletion of the GCR2 gene or the LZ2 domain of Gcr2, suggesting that $Gcr1^U$ works as a monomer that forms a heterocomplex with the Gcr2 homodimer. To verify this working model, I investigated whether $Gcr1^U$ has higher Gcr2-binding affinity than $Gcr1^S$. I generated strains expressing $Gcr1^S$-TAP or $Gcr1^U$-TAP together with Gcr2-5Flag. In agreement with our working model, the TAP-pull down experiment showed stronger binding of Gcr2 to $Gcr1^U$ than to $Gcr1^S$ (Figure 13B). I also investigated the binding affinity between Gcr1 and Gcr2 by using a ChIP experiment. Because Gcr2 can bind to DNA only through interacting with Gcr1, the DNA binding affinity of Gcr2 reflects its binding affinity to Gcr1. Although $Gcr1^U$ and $Gcr1^S$ showed similar binding affinities to target promoters of the glycolytic genes, Gcr2 showed a higher DNA binding affinity in cells expressing $Gcr1^U$ as compared with cells expressing $Gcr1^S$ (Figure 13C). These experiments support the idea that $Gcr1^U$ is the major binding partner of Gcr2.

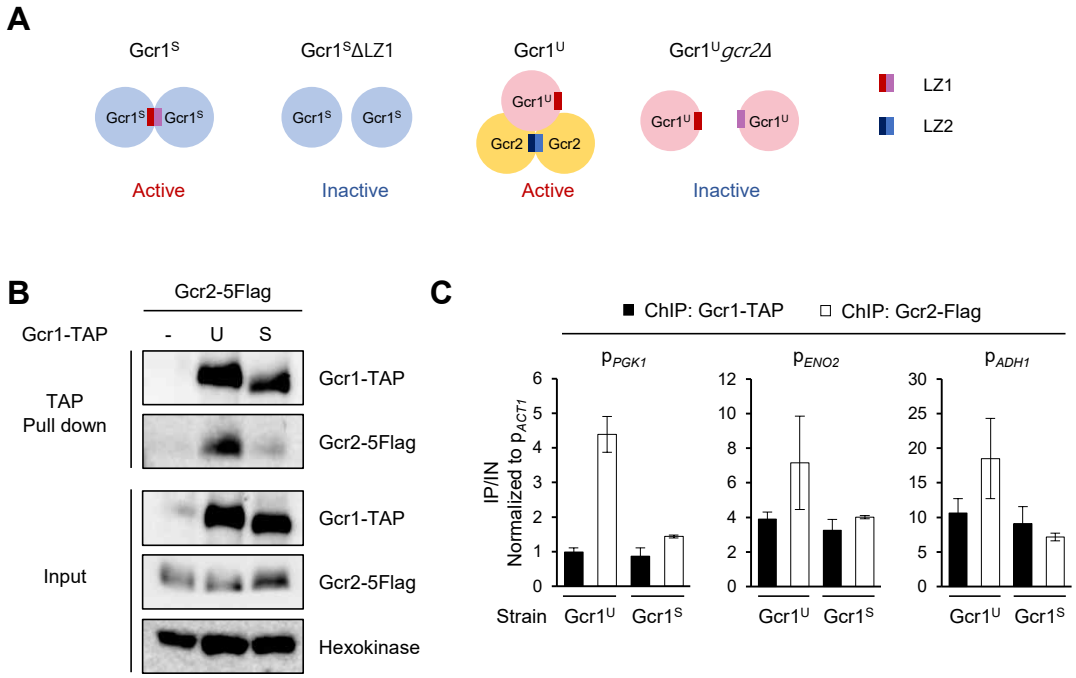

Figure 13. Working models of Gcr1^{U} and Gcr1^{S}

(A) Working models of Gcr1^{U} and Gcr1^{S}. Gcr1^{S} mainly works as a homodimer linked through the LZ1 domain and is thereby inactivated by deletion of the LZ1 domain (Gcr1^{S}ΔLZ1). In contrast, Gcr1^{U} mainly acts as a monomer forming a heterocomplex with Gcr2 dimer, and is inactivated in the absence of Gcr2 (Gcr1^{U}*gcr2*Δ).

(B) Stronger Gcr2 binding to Gcr1^{U} than to Gcr1^{S}. The interaction between Gcr1 and Gcr2 was detected by in vivo TAP pull-down assay using strains co-expressing Gcr1^{U} or Gcr1^{S}-TAP and Gcr2-5Flag. Gcr1-TAP and Gcr2-5Flag were detected by immunoblotting with IgG and anti-Flag antibody, respectively. Hexokinase was used as a loading control.

(C) Higher Gcr2 binding to the promoters of glycolytic genes in the Gcr1^{U} strain than in the Gcr1^{S} strain. Binding of Gcr1^{U} or Gcr1^{S}-TAP and Gcr2-5Flag to the indicated target promoters were detected by ChIP experiments using strains co-expressing Gcr1^{U}-TAP and Gcr^{2}-5Flag (Gcr1^{U}) or Gcr1^{S}-TAP and Gcr2-5Flag (Gcr1^{S}) and are indicated as fold enrichments normalized to the *ACT1* promoter. Each value represents the average ± standard deviations from two independent experiments.

4.5. Conclusions

In this chapter, I used CRSIPR/Cas9 system in order to generate two isoforms of Gcr1 ($Gcr1^U$ and $Gcr1^S$) under the most native regulation. Inconsistent with previous study, strains that only produce $Gcr1^U$ or $Gcr1^S$ single isoform did not show any growth defect in rich or minimal culture media. Other common stresses such as glucose concentration, aeration, different carbon sources, temperature, osmotic stress and chemical stress also did not cause any growth defect between two isoforms of Gcr1.

Interestingly, although it was demonstrated that $Gcr1^U$ and $Gcr1^S$ binds to the same target genes, their binding pattern is quite different. The affinity of $Gcr1^U$ decreases after diauxic shift, whereas the intensity of $Gcr1^S$ binding increases. Also the expression level of target genes showed different patterns: in $Gcr1^U$ strain, the expression of target genes were decreased more dramatically than $Gcr1^S$ strain after diauxic shift. To summarize, $Gcr1^U$ and $Gcr1^S$ might play differential roles not by regulating different sets of target genes, but by differential binding to the same genes depending on growth conditions.

Next, I performed domain deletion in each isoform generating strains, and successfully identified their main binding partners. $Gcr1^U$ mainly interacts with its co-activator, Gcr2, and

$Gcr1^{S}$ mainly forms homodimer via LZ1 domain. Additional biochemical assays like TAP-pull down and ChIP-qPCR in $Gcr1^{U}$ or $Gcr1^{S}$ strains support the working model of two different isoforms, described in Figure 13A.

Chapter 5.

Role of Gcr1^{S} in respiratory metabolism after diauxic shift

5.1. Introduction

Considering the domain and *GCR2* deletion study of previous chapter, it is likely that proper formation of Gcr1 homodimer or Gcr1-Gcr2 complex is important (Figure 12C-J). $Gcr1^U$ is more likely to interact with Gcr2 rather than forming a homodimer, whereas $Gcr1^S$ seems to act mainly as a homodimer without Gcr2 (Figure 13A). Afterwards, I focused on the only difference between the $Gcr1^U$ and $Gcr1^S$ proteins: the additional 55 N-terminal residues of $Gcr1^U$, named USS domain.

I hypothesised that the USS domain could play a role in inhibiting Gcr1 dimerization while facilitating the interaction with Gcr2. Considering the fact that cell growth is little affected by deletion of *GCR2* in the $Gcr1^S$ strain, the growth defect of $Gcr1^U$*gcr2*Δ strain might be due to inhibition of $Gcr1^U$ dimerization by the USS domain, possibly through intramolecular interaction masking the LZ1 domain (Figure 14). If this is the case, the USS domain might contain some residues essential for inhibiting Gcr1 dimerization. Mutation of residues in USS domain could rescue the growth defect of the $Gcr1^U$*gcr2*Δ strain by allowing homodimerization of $Gcr1^U$, thus activating $Gcr1^U$ without Gcr2 (Figure 14). With this strategy, I was able to examine the effect of defective $Gcr1^S$ homodimer in cellular metabolism.

5.2. The USS domain inhibits homodimerization of $Gcr1^U$ protein

First, I performed random mutagenesis of the USS domain in the $Gcr1^U$*gcr2*Δ strain using the CRISPR/Cas9 system, and isolated two suppressor mutants (F12L and L50P) with improved growth. Moreover, the USS domain with these point mutations is expected to lose its ability to inhibit Gcr1 dimerization (Figure 14). If the USS domain acts through intramolecular interaction, the USS domain alone could play the same inhibitory role in a trans-acting manner. To test this possibility, I examined whether the USS domain could inhibit dimerization of $Gcr1^S$, which would lead to a growth defect in the $Gcr1^S$ strain, mimicking the $Gcr1^S$ΔLZ1 strain (Figure 14A). In agreement with this hypothesis, when a DNA fragment encoding the USS domain was inserted into the *ura3*Δ0 locus in the $Gcr1^S$ strain and expressed under the control of the native *GCR1* promoter and terminator, the cell growth pattern was very similar to that of the $Gcr1^S$ΔLZ1 strain, exhibiting a reduced growth rate and decreased final cell density (Figure 15 A). However, the growth rate of the $Gcr1^S$ strain was not affected when the USS^{F12L} or USS^{L50P} suppressor mutant was expressed (Figure 15B), further confirming that these mutant USS domains cannot inhibit Gcr1 dimerization (Figure 14A).

The protein expression levels of USS^{F12L} and USS^{L50P} were even higher than that of the wild type USS (Figure 15B), indicating that the lack of $Gcr1^{S}$ inhibition by these mutants was not due to defects in their expression.

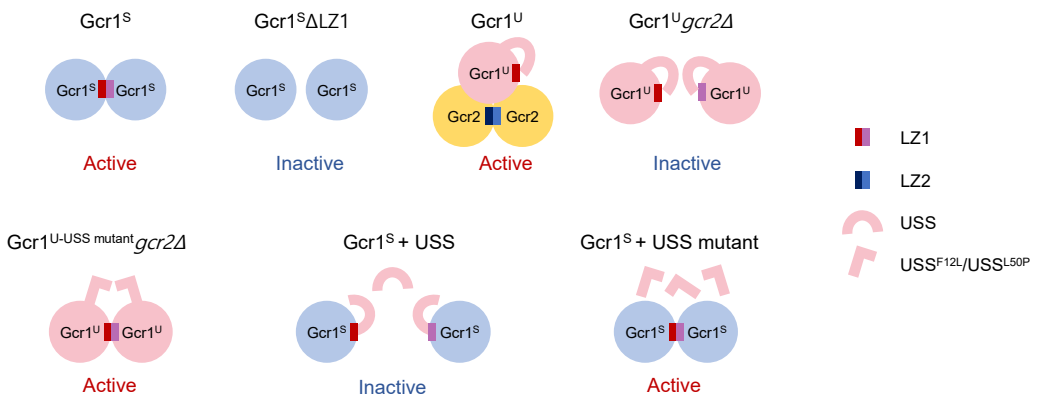

Figure 14. Revised working model of Gcr1^{U} and Gcr1^{S}

The working model of Gcr1^{U} and Gcr1^{S}, supplemented to the model in Figure 13. The USS domain inhibits dimerization of Gcr1^{U}. The mutants USS^{F12L} and USS^{L50P} may suppress Gcr1^{U}*gcr2*Δ by allowing dimierization of Gcr1^{U}. The USS domain, but not the USS mutants, inhibit dimerization Gcr1^{S} in a trans-acting manner, thus inactivating Gcr1^{S}.

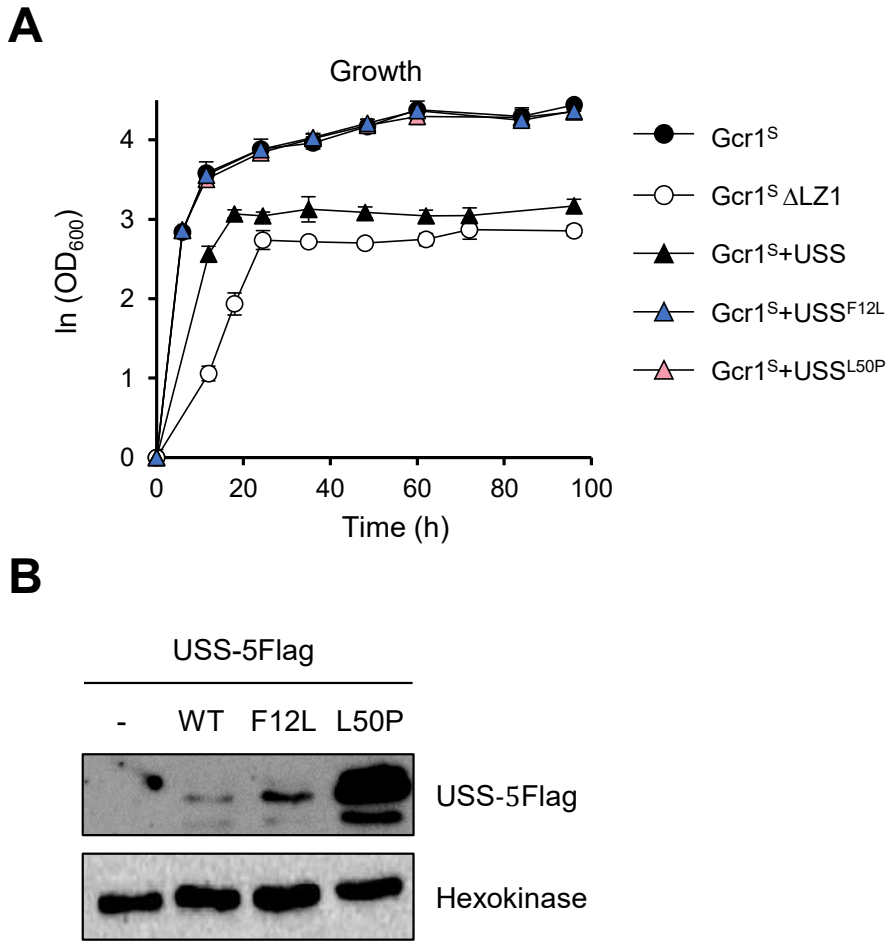

Figure 15. Effect of expressing USS domain and its suppressor mutants

(A) Growth curves of $Gcr1^{S}$ strains expressing the USS domain or its suppressor mutant F12L or L50P in comparison with $Gcr1^{S}$ and $Gcr1^{S\Delta}$LZ1 strains (n=3, average ± standard deviations).

(B) Expression levels of the USS domains. Cell lysates of the $Gcr1^{S}$ strain expressing wild-type or mutant USS-5Flag domain were subjected to immunoblotting with anti-Flag antibody to detect expression levels of the USS domains. Hexokinase was used as a loading control and the $Gcr1^{S}$ strain without the USS domain was used as a negative control.

5.3. Cells expressing Gcr1^{S}ΔLZ1 showed a defect in respiration

Both Gcr1^{U}*gcr2*Δ and Gcr1^{S}ΔLZ1 strains showed reduced specific growth rates in the exponential growth phase, reflecting reduced expression of glycolytic genes (Figure 12G, H). However, the Gcr1^{S}ΔLZ1 strain, but not the Gcr1^{U}*gcr2*Δ strain, showed diminished final cell density, suggesting dissimilar carbon metabolic pathways of the two strains. Therefore, I examined metabolite profiles in strains expressing different Gcr1 forms. In agreement with their similar growth rates, Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S} strains showed similar patterns of glucose uptake, production of ethanol and glycerol, and utilization of ethanol and glycerol via respiration after glucose depletion (Figure 16A). Alternatively, the Gcr1^{U}gcr2Δ and *gcr1*Δ strains with reduced growth and glucose uptake rates, showed lower ethanol production and higher glycerol production levels as compared with the wild type strain, but metabolized ethanol and glycerol normally after glucose depletion (Figure 16A). The Gcr1^{S}ΔLZ1 strain also accumulated higher concentrations of glycerol than wild type (Figure 16A). However, the Gcr1^{S}ΔLZ1 strain showed a defect in respiratory consumption of glycerol and ethanol after glucose depletion (Figure 16A).

I also confirmed the respiratory defect in the Gcr1^{S}ΔLZ1

strain by analyzing metabolite profiles of the $Gcr1^{S}$ strain expressing the USS domain, which mimics the growth phenotype of the $Gcr1^{S}\Delta LZ1$ strain. Similar to the $Gcr1^{S}\Delta LZ1$ strain, the $Gcr1^{S}$ strain expressing the USS domain showed a defect in metabolizing glycerol and ethanol after the diauxic shift (Figure 16B). However, expression of the mutant USS^{F12L} or USS^{L50P} domain did not affect the respiration of the $Gcr1^{S}$ strain, confirming that inhibition of $Gcr1^{S}$ dimerization leads to a respiratory defect. Because such a respiratory defect was not observed in the *gcr1*Δ strain (Figure 16A), the monomer form of $Gcr1^{S}$, mainly produced in the $Gcr1^{S}\Delta LZ1$ strain or in the $Gcr1^{S}$ strain expressing the USS domain, might exert a dominant negative effect on the expression of genes involved in respiration after diauxic shift.

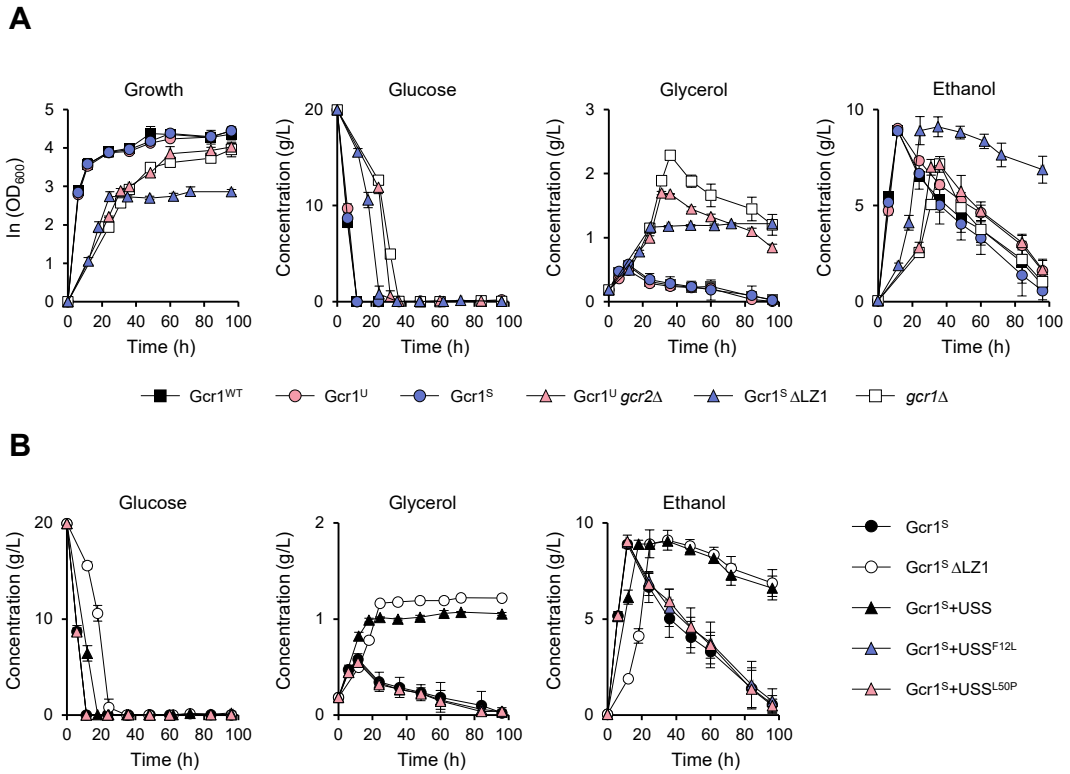

Figure 16. Growth curves and metabolite profiles of strains expressing various Gcr1 isoforms and mutants

(A) Cell growth curves and concentrations of metabolites (glucose, glycerol, and ethanol) in the medium during the cell growth in YPD medium (average ± standard deviations, n=3).

(B) Metabolite profiles of $Gcr1^{S}$ strains expressing the USS domain or its suppressor mutant F12L or L50P in comparison with $Gcr1^{S}$ and $Gcr1^{S}$ΔLZ1 strains (average ± standard deviations, n=3).

5.4. Gcr1^{S}ΔLZ1 showed a defect in inducing the respiratory genes after diauxic shift

To confirm the effects of various Gcr1 isoforms and mutants on expression of genes involved in carbon metabolism, I investigated transcription of genes involved in glycolysis (*PGK1*, *ENO2*, *PYK1*), ethanol production (*ADH1*), ethanol utilization (*ADH2* and *ALD4*), and glycerol utilization (*GUT1*) in strains expressing different Gcr1 isoforms or mutants (Figure 17A). Gene expression levels were analyzed during the exponential and diauxic shift phases. Because of the discrete growth rates of each strain, I determined sampling time points based on the glucose concentrations remaining in the medium. Exponential growth phase samples were taken when the remaining glucose concentration was 10 g/L, and the diauxic shift phase samples were obtained when the cells consumed the entire glucose supply.

In all strains evaluated (wild type, Gcr1^{U}, Gcr1^{S}, *gcr1*Δ, Gcr1^{U}*gcr2*Δ, and Gcr1^{S}ΔLZ1), expression of *PGK1*, *ENO2*, *PYK1*, and *ADH1* genes, which are target genes of Gcr1, decreased upon diauxic shift. Among the wild type, Gcr1^{U}, and Gcr1^{S} strains, Gcr1^{S} showed the highest expression levels of these genes throughout the growth phase (Figure 17B). Instead, the Gcr1^{U} strain exhibited the lowest expression levels of the

glycolytic genes after diauxic shift (Figure 17B), suggesting that Gcr1^{S} is mainly responsible for glycolytic gene expression after glucose depletion. These results are consistent with our RNA-seq results showing greater growth-dependent fold-changes in target gene expression levels in the Gcr1^{U} strain than in the Gcr1^{S} strain (Figure 8D). However, based on the similar growth rates of wild-type, Gcr1^{U}, and Gcr1^{S} strains, such differences in glycolytic gene expression levels might not be critical for cell growth, at least under our culture conditions. As expected from their slow growth rates, the *gcr1*Δ, Gcr1^{U}*gcr2*Δ, and Gcr1^{S}ΔLZ1 strains displayed lower expression levels of the glycolytic genes and *ADH1* throughout the growth phase.

In contrast, *ADH2*, *ALD4*, and *GUT1* genes involved in respiratory consumption of ethanol and glycerol were induced upon diauxic shift, exhibiting similar expression patterns in wild-type, Gcr1^{U}, and Gcr1^{S} strains (Figure 17B). The Gcr1^{U}gcr2Δ strain also showed similar induction patterns of these genes upon diauxic shift, but expression levels of *ADH2* and *ALD4* were higher than those of wild type. Such induction was not observed in the Gcr1^{S}ΔLZ1 strain (Figure 17B). Therefore, the respiratory defect in the Gcr1^{S}ΔLZ1 strain might be due to the failure to induce respiratory genes upon diauxic

shift. The $Gcr1^{S}$ strain expressing the USS domain, which mimics the $Gcr1^{S}\Delta LZ1$ strain in terms of cell growth and metabolite profiles, showed gene expression patterns similar to those of the $Gcr1^{S}\Delta LZ1$ strain, exhibiting defects in induction of respiratory genes as well as in expressing the glycolytic genes (Figure 17C). In agreement with the inactivity of USS^{F12L} and USS^{L50P} in preventing Gcr1 dimerization, expression of these mutant USS domains did not affect induction of respiratory genes upon diauxic shift (Figure 17C). These results further support the dominant negative role for inactive $Gcr1^{S}$ monomer in induction of respiratory genes.

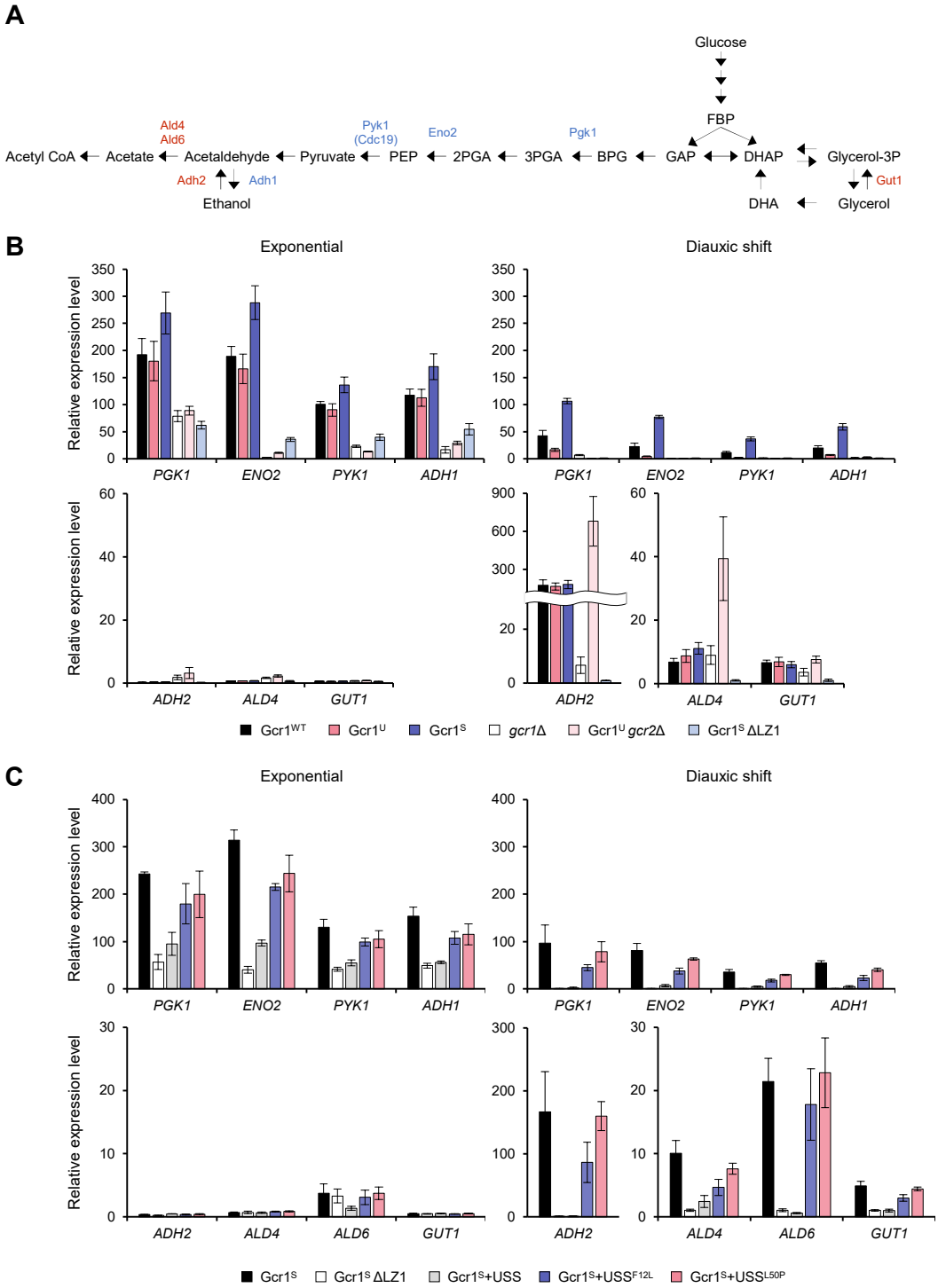

Figure 17. Expression of genes involved in glycolysis, ethanol production, and consumption of ethanol and glycerol in strains expressing various Gcr1 isoforms and mutants

Figure 17. Expression of genes involved in glycolysis, ethanol production, and consumption of ethanol and glycerol in strains expressing various Gcr1 isoforms and mutants

(A) Metabolic pathways of glycolysis, ethanol fermentation, and utilization of ethanol and glycerol. Enzymes encoded by Gcr1 target genes involved in glycolysis (Pgk1, Eno2, and Pyk1) and ethanol production (Adh1) are shown in blue. Enzymes involved in the utilization of ethanol (Adh2, Ald4, and Ald6) and glycerol (Gut1) are shown in red.

(B) Gene expression levels in strains expressing various Gcr1 isoforms and mutants. Cells were grown in YPD media and gene expression levels at the exponential and diauxic shift phases were detected by qRT-PCR normalized to the mRNA levels of *TFC1*. Each value represents the average ± standard deviations (n=3) of the relative fold-change in expression, normalized to the expression level of the $Gcr1^{S}\Delta LZ1$ strain at the diauxic shift phase.

(C) Gene expression levels in the $Gcr1^{S}$ strain expressing various USS domains. Cells were grown in YPD media and gene expression levels at the exponential and diauxic shift phases were detected by qRT-PCR normalized to the mRNA levels of *TFC1*. Each value represents the average ± standard deviations (n=3) of the relative fold-change in expression, normalized to the expression level of the $Gcr1^{S}\Delta LZ1$ strain at the diauxic shift phase.

5.5. The respiratory defect of Gcr1^{S}ΔLZ1 strain could be restored by overexpressing *ALD4*

Considering the normal induction of the respiratory genes in *gcr1*Δ upon diauxic shift (Figure 17B), Gcr1 seems unnecessary for the induction of these genes. Therefore, an inactive Gcr1^{S} monomer might affect expression of respiratory genes either directly or indirectly. Respiratory genes, *ADH2*, *ALD4*, *ALD6*, and *GUT1*, were not detected as Gcr1-binding targets in our ChIP-seq analysis, but other recent ChIP-exo analysis identified *ALD4* as a target where Gcr1, but not Gcr2, binds upon glucose limitation (86). Therefore, I examined whether Gcr1^{S} dimers and Gcr1^{S} monomers (Gcr1^{S}△LZ1) could bind to respiratory gene promoters. To find any differences in DNA binding activities between Gcr1^{S} and Gcr1^{S}△LZ1 by using ChIP experiments, I created strains expressing Gcr1^{S}-5Flag or Gcr1^{S}△LZ1-5Flag, but the Gcr1^{S}△LZ1-5Flag strain showed a different growth phenotype compared with Gcr1^{S}△LZ1 strain, which might be due to perturbation of protein function by the tag itself. Therefore, I instead expressed the USS domain in a Gcr1^{S}-5Flag strain to mimic the phenotype of Gcr1^{S}△LZ1 strain. In agreement with our ChIP-Seq experiment, no consequential binding of Gcr1^{S} was detected to the *ADH2*, *GUT1*, and *ALD4* promoters (Figure 18A). Gcr1^{S} co-expressed

with the USS domain also did not bind to the *ADH2* and *GUT1* promoters, suggesting that the Gcr1^{S} monomer might indirectly affect the expression of these genes. However, when the USS domain was co-expressed with Gcr1^{S}, binding to the *ALD4* promoter was detected at the diauxic shift phase, but not at the exponential phase (Figure 18A). Therefore, in accord with the previous study (86), Gcr1^{S} might bind to the *ALD4* promoter upon glucose limitation. Stronger binding of the Gcr1^{S} monomer than the Gcr1^{S} dimer seems to negatively affect the transcriptional induction of the *ALD4* gene upon diauxic shift, possibly by inhibiting the binding of other transcription factors. Furthermore, Gcr1^{S} co-expressed with the USS domain also showed enhanced binding to the *PGK1* promoter, suggesting that the Gcr1^{S} monomer has a higher DNA binding affinity than the Gcr1^{S} dimer in general. Consistent with the results of our ChIP-seq experiments (Figure 8D), Gcr1^{S} binding to the *PGK1* promoter increased from the exponential to the diauxic shift phase (Figure 18A), although the *PGK1* transcription diminished after diauxic shift (Figure 17B and 17C). Considering the reduced expression of glycolytic genes after diauxic shift even in the *gcr1*Δ strain (Figure 17B), other transcription regulators might also affect glycolytic gene expression after glucose depletion.

The ChIP experiment suggests that *ALD4* is a direct target affected by $Gcr1^S$ monomer. Because *ALD4* encodes a mitochondrial aldehyde dehydrogenase required for ethanol utilization, I examined whether *ALD4* overexpression could rescue the respiratory defect of the $Gcr1^S$△LZ1 strain. When *ALD4* was overexpressed using the *TEF1* promoter in the $Gcr1^S$△LZ1 strain, ethanol utilization was recovered, resulting in a higher final cell density (Figure 18B). Although Ald4 is only involved in ethanol degradation, glycerol utilization also recovered, suggesting that overexpression of *ALD4* alone can trigger cellular metabolic reprogramming to respiratory growth.

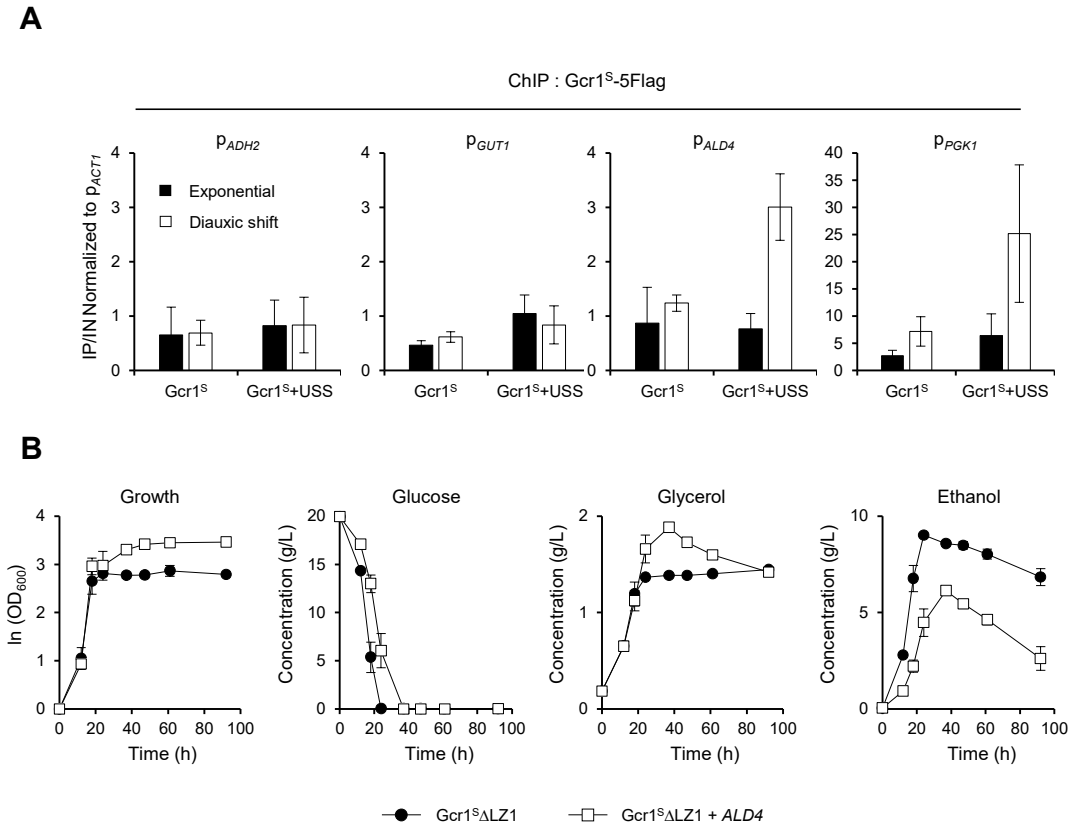

Figure 18. Restoration of the respiratory growth defect of Gcr1^{S}△LZ1 strain by overexpressing *ALD4*

(A) Binding of the Gcr1^{S} monomer to the *ALD4* promoter after diauxic shift. Gcr1^{S} binding to the indicated promoters at the exponential and diauxic shift phases were monitored by ChIP experiments using Gcr1^{S}-5Flag strains with or without USS overexpression, and indicated as fold enrichments normalized to the *ACT1* promoter. Each value represents the average ± standard deviations from three independent experiments.

(B) Growth curves and metabolite profiles of the Gcr1^{S}△LZ1 strain with or without overexpression of *ALD4* (n=3, average ± standard deviations).

5.6. Accumulation of *GCR1*^{S} mRNA also showed respiratory defect

It has been already studied that *GCR1* gene is a target of nonsense-mediated mRNA decay (NMD) (101, 102, 146). NMD is an important pathway that gets rid of the mRNA transcripts containing premature termination codons (PTC). The wrong transcripts that contain PTC is recognized by Upf1-3 complex proteins and degraded by 5′→3′ exoribonuclease (Xrn1) (147). Previous result showed that *xrn1* deletion is more responsible for the accumulation of *GCR1* transcripts than *upf1* deletion (102). In order to figure out the effect of accumulation of each isoform of *GCR1* gene, *XRN1* and *UPF1* was deleted in cells expressing Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S} respectively.

Consistent with previous study, *XRN1* deletion causes growth defect in Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S} strains compared with native states (Figure 8B, 19A) (148, 149). Although *xrn1* deletion in the Gcr1^{U} generating strain showed a decreased growth, its metabolism after diauxic shift still occurred. However, in Gcr1^{WT} and Gcr1^{S} generating strains, *xrn1* deletion not only slows down the fermentative growth but also defects the respiratory growth. Considering the previous result that *xrn1* deletion causes the accumulation of mRNA generating Gcr1^{U} and Gcr1^{S} proteins, the respiratory defect of Gcr1^{WT}

strain is expected mainly due to the accumulation of $Gcr1^S$ mRNA (102). On the contrary, *upf1* deletion did not show growth defect in all background strains (Figure 19B). Taken together that amount of *GCR1* spliced transcripts did not significantly affected by *upf1* deletion, it seems that PTC is not a major regulatory mechanism of regulating the splicing of *GCR1* mRNA (102). Therefore, it can be concluded that the inadequate elimination of spliced *GCR1* causes both fermentative and respiratory defect to cells, and especially the accumulated $GCR1^S$ mRNA is mainly responsible for the respiration defect.

Next, I confirmed that whether the simple overexpression of native level of $Gcr1^S$ also affects the respiratory growth. By changing their own promoter into the strong promoter, P_{TDH3}, by CRISPR/Cas9-mediated genome editing, I successfully produced the strains that natively overexpressing each of two isoforms. As expected, only $GCR1^S$ overexpressing strain showed the defective respiratory growth and metabolism (Figure 20). $GCR1^U$ overexpressing strain did not show any different phenotype compared to the strains expressing $Gcr1^U$ and $Gcr1^S$. This is consistent with the result that unnecessary accumulation of $GCR1^S$ causes the respiratory defect (Figure 20 A).

To be concluded, deletion of exonuclease, *XRN1*, causes

the increased amount of both *GCR1*^{U} and *GCR1*^{S} mRNA, but only accumulation of *GCR1*^{S} seems to be responsible for the respiratory defect. This suggests that proper regulation of *GCR1*^{S} level is crucial for the cellular respiratory metabolism.

5.7. Conclusions

In this chapter, more detailed study to understand the role of two isoforms of Gcr1. First, focusing on the Gcr1^{U} form specific protein domain, USS domain, the inhibitory effect of this domain in homodimerization of Gcr1^{U} had been demonstrated. Also the suppressor mutants that alleviates the defective growth of Gcr1^{U}*gcr2*Δ were discovered. Especially, expression of USS domain, not mutatnt USS domain, can even mask the leucine zipper domain and cause impaired Gcr1^{S} homodimerization in *trans*-acting manner.

The irregular growth defect of Gcr1^{S}ΔLZ1 was identified as a respiratory defect. This strain could not consume ethanol or glycerol and several genes related to utilize ethanol and glycerol were not activated after diauxic shift. This respiratory defect also occurred when *XRN1* was deleted in Gcr1^{S} strain, which results the accumulation of *GCR1*^{S} mRNA. Also the overexpression of *GCR1*^{S}, not *GCR1*^{U}, caused the same defect, suggesting that delicate and tight regulation of *GCR1*^{S}

transcription is especially important for cellular respiratory metabolism.

Additional analysis revealed that *ALD4*, which encodes mitochondrial aldehyde dehydrogenase, is repressed unexpectedly by the monomer form of $Gcr1^{S}$. The impaired homodimerization, or overexpression may generate the excessive amount of $Gcr1^{S}$ monomer, which results unnecessary binding to the promoters of respiratory genes and defects the respiratory metabolism. Interestingly, overexpression of *ALD4* in $Gcr1^{S}\Delta LZ1$ can restore the respiratory defect, indicating that *ALD4* can solely trigger the reprogramming of cellular metabolism to respiration.

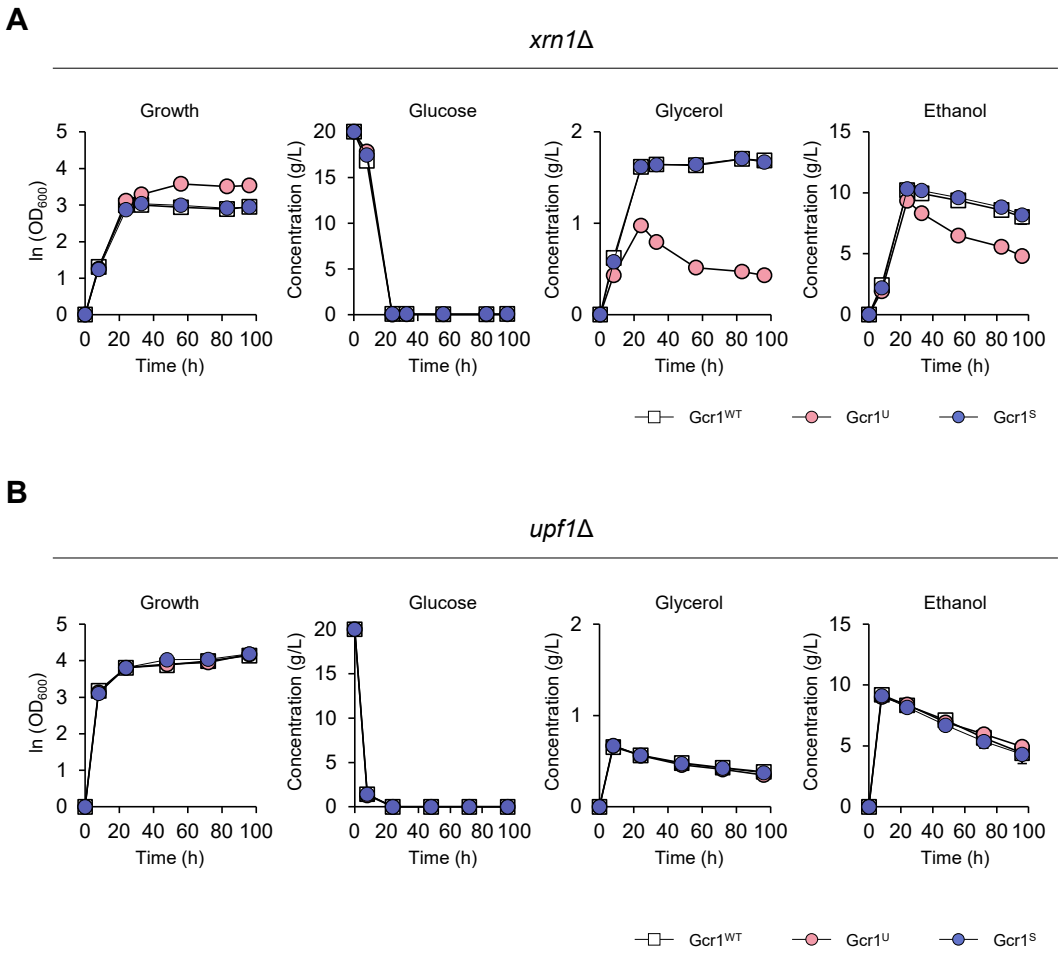

Figure 19. Growth curves and metabolite profiles of $Gcr1^{WT}$, $Gcr1^{U}$ and $Gcr1^{S}$ strains that lacks *XRN1* or *UPF1*

(A) Growth curves and metabolite profiles of the *xrn1*△ strains (n=2, average ± standard deviations).

(B) Growth curves and metabolite profiles of the *upf1*△ strains (n=2, average ± standard deviations).

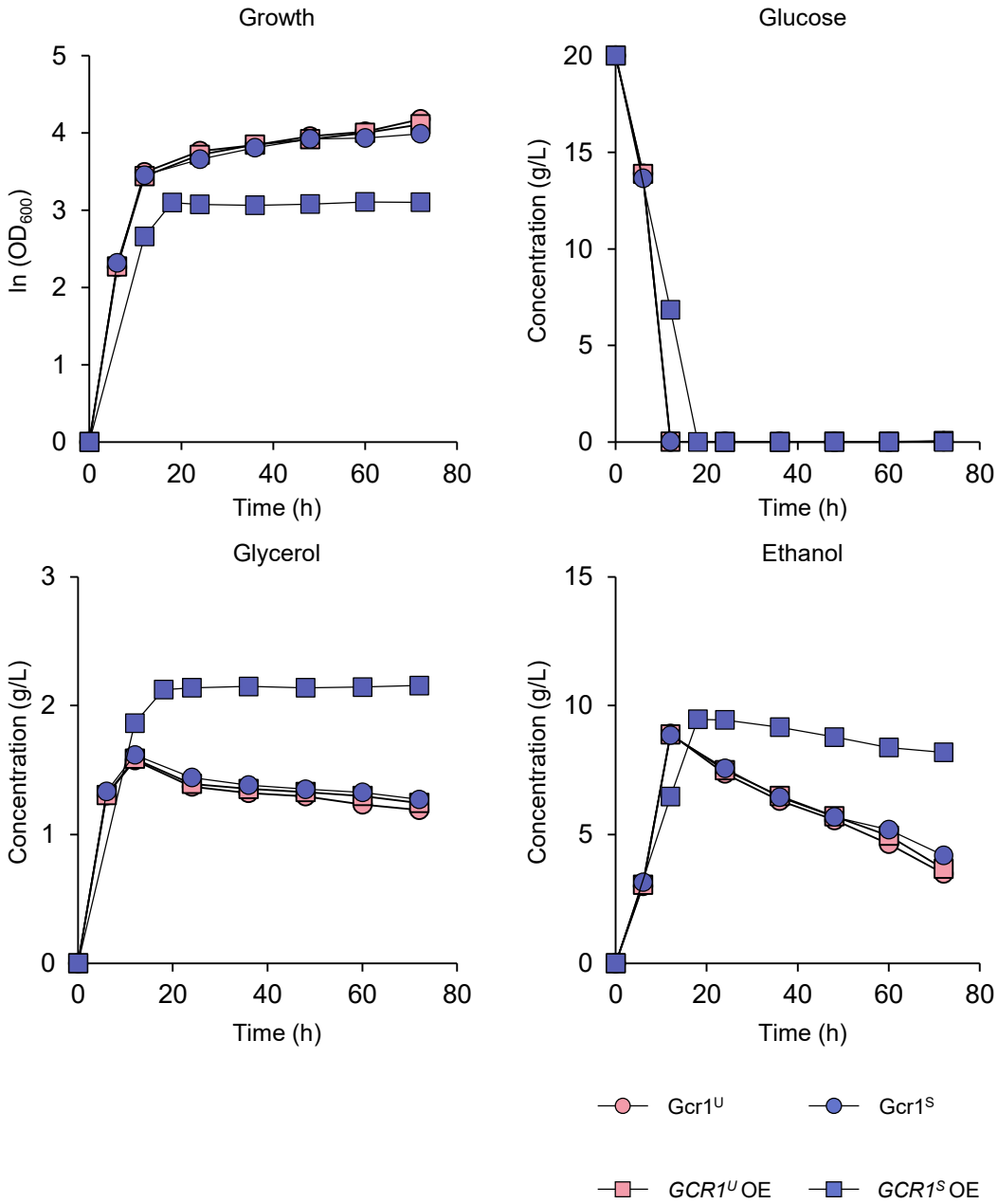

Figure 20. Growth curves and metabolite profiles of strains that Gcr1^{U}, Gcr1^{S} strains and *GCR1^{U}*, *GCR1^{S}* overexpressing strains.

Growth curves and concentrations of glucose, glycerol and ethanol for four strains : Gcr1^{U}, Gcr1^{S}, *GCR1^{U}* overexpressing, and *GCR1^{S}* overexpressing strains (n=2, average ± standard deviations).

Chapter 6.

Effective production of lactic acid by overexpressing Gcr1 isoforms

6.1. Introduction

Since *S. cerevisiae* converts most of the cytosolic pyruvate into ethanol, eliminating ethanol fermentation pathway and focusing on the production of pyruvate-derived chemical is a widely used strategy for microbial cell factory system. In previous studies, the *adh1-5* and *gpd1-2* genes were deleted in order to block the carbon leakage to ethanol and glycerol (120). Starting with this strain, 2, 3-butandiol, lactic acid, and acetoin was successfully produced with increased titer (118-120, 128).

Lactic acid (LA) is one of the common pyruvate derivatives used as a monomer for biodegradable polymer, poly lactic acid (PLA). Previous study had successfully increased the amount of LA by eliminating competing pathway and overexpressing *GCR1*, under microaerobic condition (126). In this study, SP1130 strain was used as a background strain: deleting *PDC1*, *PDC6*, *GPD1*, *ADH1* to eliminate the competing pathway and increase the pyruvate pool and introducing the bacterial acetaldehyde dehydrogenase enzymes so that increased supply of acetyl-CoA can compensate the cellular growth (Figure 21). Although this study confirmed the enhanced production of LA by *GCR1* overexpression, they did not focus on which isoform of Gcr1 is responsible for the improvement.

Since pyruvate is rapidly utilized in *S. cerevisiae*, it is

hard to distinguish the glycolytic flux or pyruvate pool in the native state. Therefore, deleting the natural ethanol and glycerol pathway expected to help determining the difference of glycolytic flux by decelerating the cellular metabolism. In this chapter, I edited previously used background strain (SP1130) to generate solely $Gcr1^{U}$ or $Gcr1^{S}$ in a native state. The final product of glycolysis, pyruvate, was converted to the desired product by expressing *ldh* enzymes. As a result, I was able to indirectly observe the role of two isoforms of Gcr1. Also I aimed to observe the role of overexpressed Gcr1 in producing LA under aerobic condition.

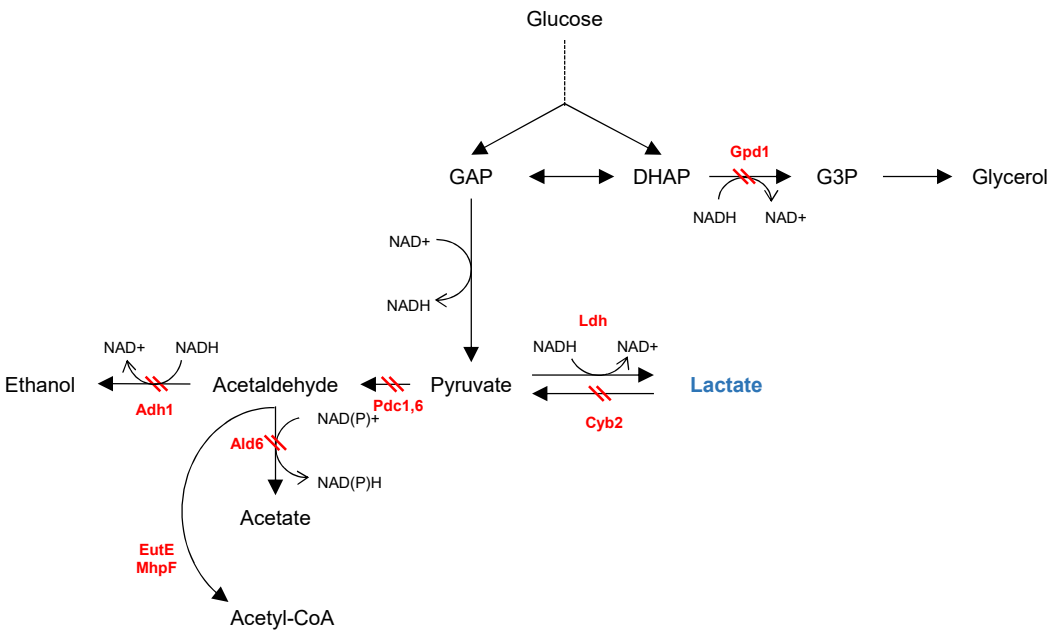

Figure 21. Engineered pathway of SP1130 strains.

Metabolic pathway of two lactic acid producing strain, SP1130. Genes related to produce byproducts and utilize lactate were eliminated. In addition, acetaldehyde utilizing enzymes from *E. coli* were expressed.
GAP, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate;, G3P, glycerol-3-phosphate

6.2. Differential regulation between two isoforms of Gcr1 in LA producing strain

To mimick the strategy of SP1130 strain, the ethanol and glycerol production pathway was eliminated by *adh1* and *gpd1* deletion, respectively. To increase the pool of pyruvate, pathway from pyruvate to acetate was blocked by *PDC1*, *PDC6*, and *ALD6* deletion. Next, gene coding LA utilizing enzyme (*CYB2*) was eliminated and acetaldehyde was redirected to acetyl-CoA by heterologous enzymes, EutE and MhpF (126, Figure 21). After generating the strains that only generates $Gcr1^{U}$ or $Gcr1^{S}$ in the SP1130 background, I used minimal media (SC) as a culture media not to make the metabolism excessively accelerated so that I can observe the difference between two isoforms. Under aerobic culture condition (170 rpm), $Gcr1^{U}$ and $Gcr1^{S}$ clearly showed the different phenotype (Figure 22A). The final cell density of $Gcr1^{U}$ strain is higher than $Gcr1^{S}$ and even than $Gcr1^{WT}$, whereas glucose consumption rates were similar among three strains. Also, the lactic acid amount was increased in SP1130-S strain (9.62 g/L) and decreased in SP1130-U strain (8.79 g/L) compared with SP1130 strain (9.06 g/L) (Figure 22C).

Considering the faster consumption of ethanol in SP1130-U strain than other two strains, it can be possibly

concluded that $Gcr1^{U}$ and $Gcr1^{S}$ have different roles in utilizing glucose into cellular metabolism under minimal media culture condition (Figure 22A). $Gcr1^{U}$ has more tendency to utilize glucose into cell mass, whereas $Gcr1^{S}$ shows trend to redirect glucose into byproducts. Also, when I focus on the amount of product, $Gcr1^{S}$ is slightly more beneficial for producing the target compound.

6.3. Improvement of lactic acid production by overexpressing Gcr1 isoforms

In the previous study using *GCR1* overexpressed SP1130 strain, the microaerobic condition (90 rpm) used to increase the glucose consumption and LA production. Under microaerobic culture condition, it is known that intracellular cofactor imbalance cannot be alleviated by respiration so that accumulated NADH from glycolysis can force the flux to produce lactic acid, regenerating NAD^{+} (150). However, when the SP1130 and SP1141 cells were grown in the aerobic condition (170 rpm), the overexpression of $GCR1^{WT}$ had a negative effect in producing LA and showed increased final cell density (Figure 22B, 23). This indicates that under aerobic condition, which leads to NADH consumption by aerobic respiration, overexpression of both isoforms are not beneficial to

produce LA. Therefore, I hypothesized that the respiratory defect caused by defective $Gcr1^{S}$ dimer can be applied as a similar strategy even under aerobic condition.

When the *GCR1*^{S} was overexpressed in SP1130-S by genomic integration, resulting SP1141-S, the overall cellular metabolism including glucose consumption and ethanol production gets severely decreased. However, as I expected, the production of LA was increased even more than wild type $GCR1^{WT}$ overexpressed strain (SP1141), from 9.62 g/L to 10.38 g/L (Figure 22C, 23). Consistent with the results from previous chapter, overexpression of *GCR1*^{S} causes defect in both fermentative and respiratory growth (Figure 20). Considering the unnecessary inhibitory function of $Gcr1^{S}$ monomer, the respiratory defect from *GCR1*^{S} overexpression may accumulate the NADH in the cytosol, so that NADH cannot be oxidized by respiration but by converting pyruvate into LA, which consumes NADH.

In the case of *GCR1*^{U} overexpressing strain, SP1141-U, also showed the increase in LA production (8.79 g/L to 9.48 g/L) (Figure 22B, 22C). However, concerning the similar ethanol production and slightly increased glucose consumption rate, the overexpression of *GCR1*^{U} does not seem to cause the specific metabolic defect like the case of *GCR1*^{S}, but simply increasing

the glycolytic flux (Figure 22B, 23). This increased flux of glycolytic flux may bring out the faster glucose consumption and increased LA production. This result is also consistent with previous chapter, that overexpressed *GCR1*^{U} did not cause any defect in growth or metabolism under native pathway (Figure 20). Therefore, the increased glycolytic flux in the strain with weakened ethanol or glycerol producing strain may utilize remaining carbon source to LA in SP1141-U strain. The decreased final cell density can be also explained by increased LA amount, which is toxic for cell growth (Figure 23).

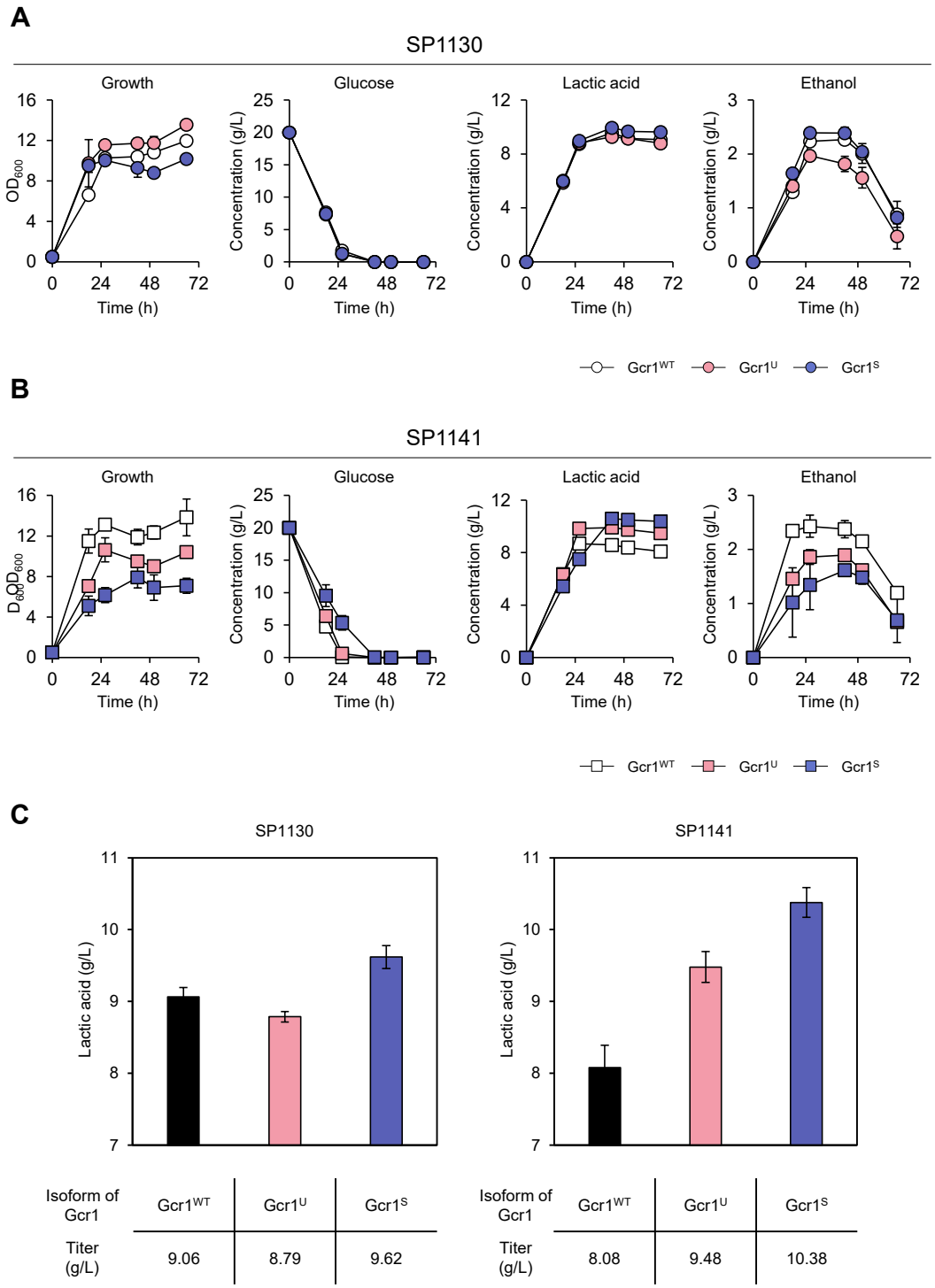

Figure 22. Growth, glucose consumption and amount of lactic acid produced in SP1130 and SP1141 strains expressing only $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$.

Figure 22. Growth, glucose consumption and amount of lactic acid andn ethanol produced in SP1130 and SP1141 strains expressing only Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S}.

(A,B) Growth curves, concentrations of glucose and lactate for three strains that (A) expresses the native level of *GCR1* : SP1130 (producing Gcr1^{WT}), SP1130-U (producing Gcr1^{U}) and SP1130-S (producing Gcr1^{S}), or (B) the overexpressed level of *GCR1* : SP1140 (producing Gcr1^{WT}), SP1141-U (producing Gcr1^{U}) and SP1141-S (producing Gcr1^{S}). Each value represents the average ± standard deviations of two independent experiments.

(C) The titer of lactic acid produced after 66 hours of culture, taken from Figure 22A lactate section. The error bar indicates the standard deviations of two i ndependent experiments.

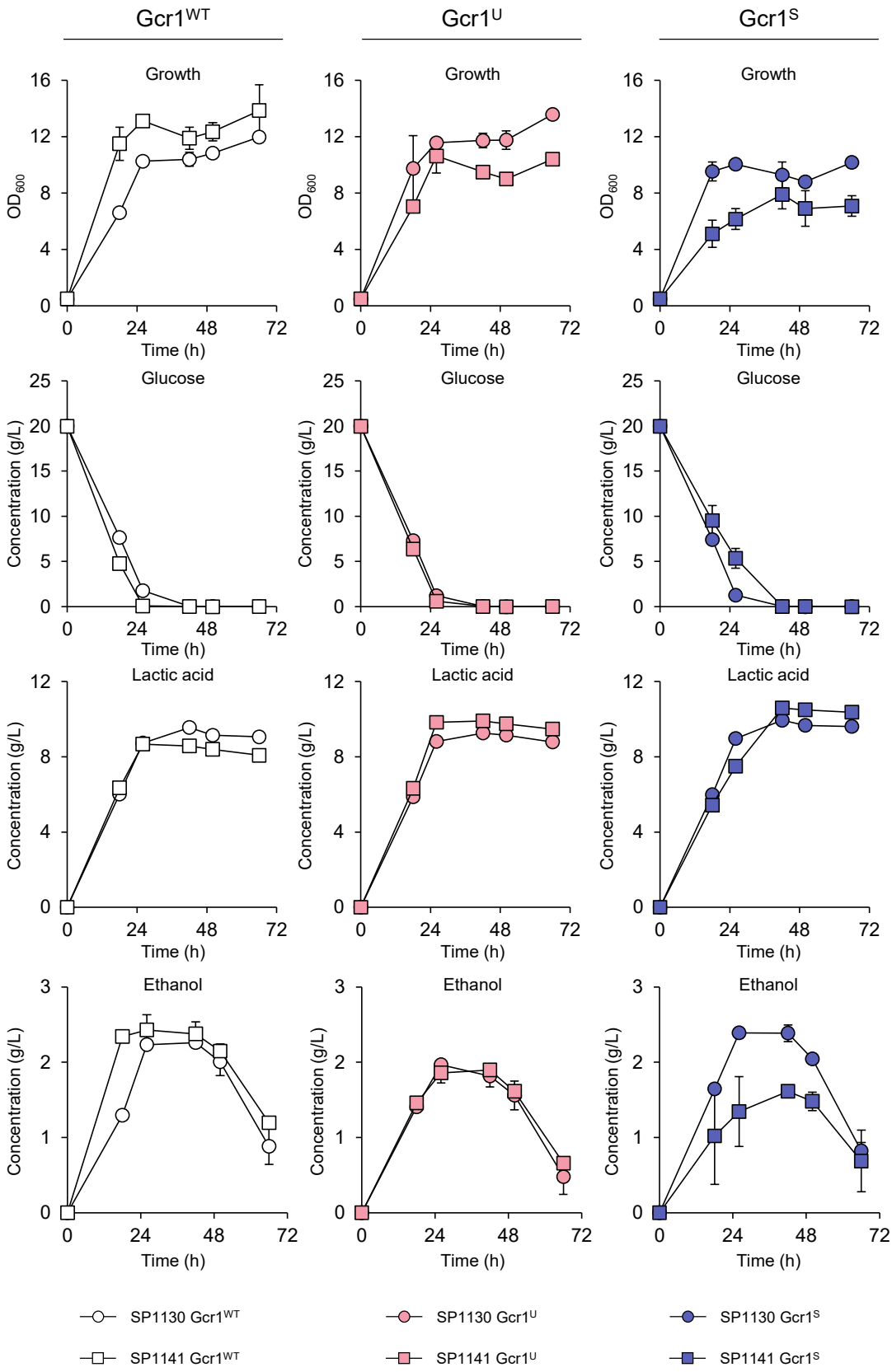

Figure 23. Growth, glucose consumption and amount of lactic acid produced in SP1130 and SP1141 strains expressing only Gcr1^{WT}, Gcr1^{U}, and Gcr1^{S}. Data from Figure 22 were sorted according to the isoforms of Gcr1.

Figure 23. Growth, glucose consumption and amount of lactic acid and ethanol produced in SP1130 and SP1141 strains expressing only $Gcr1^{WT}$, $Gcr1^{U}$, and $Gcr1^{S}$. Data from Figure 23 were sorted according to the isoforms of Gcr1.

Growth curve and concentration of glucose, lactic acid, and ethanol of six strains (SP1130, SP1130-U, SP1130-S, SP1141, SP1141-U and SP1141-S), sorted by isoforms of Gcr1.

6.4. Conclusions

In this chapter, I applied the concept of previous study that change the culture condition into microaerobic in order to make the respiration flux less major. In the microaerobic condition, overexpressing *GCR1*^{WT} clearly beneficial for consuming glucose and producing LA. However, although general aerobic condition (170 rpm) accelerated the growth and glucose consumption, the LA production level was rather decreased, may be due to the too fast glycolytic metabolism.

Considering that both microaerobic culture and Gcr1^{S} monomer expression are disadvantageous in terms of respiration, I decided to overexpress the Gcr1^{S} in the lactic acid producing strain and culture it under aerobic condition. The cells successfully showed the respiratory defect and the production of lactic acid also increased by overexpressing *GCR1*^{S}. Overexpression of Gcr1^{U} also increased the amount of lactic acid produced, but considering the ethanol metabolism is almost same in SP1130-U and SP1141-U strains, it is just due to the slightly accelerated glycolytic flux, not by the defective respiratory metabolism.

Also, SP1130-U and SP1130-S strains showed different metabolic profile and growth pattern, suggesting that these two isoforms of Gcr1 clearly have differential roles in regulating

metabolism. $Gcr1^{U}$ expressing strain show more tendency to rewire the glucose into the cellular growth, whereas $Gcr1^{S}$ expressing strain focuses on producing target compound or ethanol, with lower final cell density. Although this phenotype did not occur in the wild type strains, eliminating some byproduct pathway and make the enough pyruvate pool can help slow down the fermentation metabolism and distinguish the differential role between $Gcr1^{U}$ and $Gcr1^{S}$.

Chapter 7.

Overall discussion and recommendations

7.1. CRISPR/Cas9-mediated genome editing minimized the effects of non-native expression levels of Gcr1 isoforms on cell growth

In this study, I investigated differential roles of $Gcr1^U$ and $Gcr1^S$ by generating strains producing only one isoform of Gcr1. $Gcr1^U$ and $Gcr1^S$ strains and other strains producing Gcr1 or Gcr2 mutants were generated by CRISPR/Cas9-mediated genome editing with minimum genomic perturbations. In a previous study, when each isoform alone was expressed from a CEN-based low-copy-number plasmid in *GCR1* deletion strain, cells showed growth defects compared with cells expressing both isoforms (102). Unexpectedly however, our strain producing only $Gcr1^U$ or $Gcr1^S$ did not show any noticeable growth defects under normal and various stress conditions I tested. Such different results might be mainly due to the difference in *GCR1* expression levels depending on the experimental designs. Compared with the plasmid-based expression, CRISPR/Cas9-mediated seamless genome editing allows almost native-level expression of each *GCR1* isoform from its natural chromosomal location, which might help more precise investigation of the Gcr1 isoforms. Although plasmid-based gene expression has been widely used because of its convenience, gene expression levels could be affected by

plasmid copy numbers and stability. Such non-natural gene expression levels could possibly lead to a biased interpretation of the gene function, especially when fine-tuned expression level is crucial for its function. Since Gcr1 is one of the key transcription factors controlling cell growth, slight perturbation of its expression level might affect cell growth. In fact, I observed some changes in cell growth rates when $Gcr1^U$ or $Gcr1^S$ was expressed from a plasmid vector (data not shown). Our study demonstrates the importance of proper control of gene expression levels when studying the gene function.

7.2. $Gcr1^U$ and $Gcr1^S$ work as different protein complexes

Although strains producing only $Gcr1^U$ or $Gcr1^S$ did not show any growth defects, I could identify different working models for $Gcr1^U$ and $Gcr1^S$ through investigating the deletion effects of various Gcr1 functional domains and Gcr2. Our genetic and biochemical evidences suggest that $Gcr1^U$ mainly works as a monomer forming a heterocomplex with Gcr2 dimer, whereas $Gcr1^S$ works as a homodimer without Gcr2 binding. The N-terminal 55-amino acids USS domain, existing only in $Gcr1^U$, was shown to inhibit Gcr1 homodimerization, playing a key role in determining the formation of different Gcr1

complexes. The USS domain could even inhibit and $Gcr1^S$ dimerization in a *trans*-acting manner, suggesting that intramolecular interaction of the USS domain might prevent the LZ1-dependent dimerization of $Gcr1^U$. Cells expressing $Gcr1^U\Delta$ LZ1 showed just slightly reduced growth rate, indicating that $Gcr1^U$ monomer can interact with Gcr2 to form an active complex. However, cells expressing $Gcr1^S\Delta$LZ1 showed a severe growth defect, suggesting that $Gcr1^S$ monomer cannot form an active heterocomplex with Gcr2. Although pull-down and ChIP experiments revealed that $Gcr1^S$ has weaker Gcr2 binding activity than $Gcr1^U$, significant binding of Gcr2 to the target promoters was still observed in $Gcr1^S$ strain, suggesting that the USS domain is not absolutely necessary for Gcr2 binding. Therefore, the USS domain might contribute to Gcr2-dependent activation of $Gcr1^U$. In this case, Gcr2 binding to $Gcr1^S$ monomer might not be enough to induce proper conformational changes leading to $Gcr1^S$ activation.

The presence of these two forms of Gcr1 complex has been suggested in a previous study carried out with $Gcr1^A$ (97, 99). Previously, it has been suggested that Gcr1 homodimer and Gcr1-Gcr2 heterocomplex are involved in the regulation of the ribosomal protein (RP) genes and glycolytic genes, respectively (99). Although it was a controversial issue whether Gcr1

regulates the RP genes, now it seems to be clear that the effects of Gcr1 on RP gene expression are an indirect effects correlated with the cell growth. Recent ChIP-exo (86) as well as our ChIP-seq analyses revealed that the RP genes are not direct targets of Gcr1. Assuming that $Gcr1^{A}$ mainly works as a homodimer like $Gcr1^{S}$, cells expressing the dimerization-defective $Gcr1^{A}$ would have a growth defect concomitant with reduced expression of the RP genes, which in turn might have led to the misinterpretation that Gcr1 homodimer is required for RP gene transcription.

Although I revealed that $Gcr1^{U}$ and $Gcr1^{S}$ form different types of complex, it is not clear yet why cells should have both forms of Gcr1 complex. In line with the normal growth phenotypes of $Gcr1^{U}$ and $Gcr1^{S}$ strains, ChIP-seq experiments revealed that the binding targets of $Gcr1^{U}$ and $Gcr1^{S}$ are almost identical. However, each Gcr1 complex might have some differences in DNA binding activity, transcriptional activation activity, or interaction with other transcription factors. Since the level of spliced *GCR1* mRNA producing $Gcr1^{S}$ was shown to increase at later growth phase, $Gcr1^{S}$ and $Gcr1^{U}$ might have different roles depending on growth phase (102). In fact, I observed some growth-dependent differences between $Gcr1^{S}$ and $Gcr1^{U}$ in terms of target gene expression and DNA binding

affinity. Although expression of glycolytic genes decreased after diauxic shift, $Gcr1^S$ DNA binding affinity increased upon diauxic shift, exhibiting the opposite trend compared with $Gcr1^U$. In addition, $Gcr1^S$ strain showed higher expression levels of glycolytic genes than $Gcr1^U$ strain, especially after diauxic shift. Considering the similar growth rates of $Gcr1^S$ and $Gcr1^U$ strains, the observed differences between $Gcr1^S$ and $Gcr1^U$ might not be critical for cell growth at least under our culture conditions. These results are consistent with the fact that most of introns in *S. cerevisiae* can be deleted without any growth defects under normal conditions, but several intron deletions cause minor phenotypes under specific growth conditions (151). Therefore, in wild type cells producing both isoforms, changes in the ratio of $Gcr1^U$ and $Gcr1^S$ might contribute to sophisticated regulation of cell growth depending on environmental conditions. It needs further studies to understand how the splicing is regulated and what the specific roles of the two types of Gcr1 complex are.

$Gcr1^S\Delta LZ1$ strain producing inactive $Gcr1^S$ monomer showed a unique phenotype of respiration defect involving the failure of inducing respiratory genes after diauxic shift. Although most of the respiratory genes are not direct targets of Gcr1, I detected direct binding of $Gcr1^S$ monomer ($Gcr1^S$

coexpressed with the USS domain) to the *ALD4* promoter at diauxic shift phase. Furthermore, the respiratory defect of $Gcr1^S$ ΔLZ1 strain was restored by overexpressing *ALD4* from a heterologous promoter, suggesting that inactivation of the *ALD4* induction might be one of the major reasons for the respiratory defect of $Gcr1^S$ΔLZ1 strain. It is not clear yet how the binding of $Gcr1^S$ monomer to the *ALD4* promoter inhibits transcription, but it could be through inhibiting other essential transcription factors. Considering the fact that *GCR1* deletion does not affect the respiratory growth, Gcr1 might not be essential for the activation of the respiratory genes after diauxic shift. However, the dominant negative effect of $Gcr1^S$ monomer on the respiratory growth suggests a potential regulatory role for Gcr1 isoforms in transition from the fermentative to respiratory growth. Based on the fact that overexpression of a single *ALD4* gene was enough to restore the respiratory defect of $Gcr1^S$ΔLZ1 strain, global metabolic regulation could be achieved by fine-tuning of a few essential target genes.

7.3. Regulation of *GCR1* by alternative splicing

GCR1 gene has unusually long intron of 739 nt and produces multiple spliced isoforms by alternative splicing (101, 102, 142). The intron-containing genes in *S. cerevisiae* comprise

only about 5% of the genome (152). Although alternative splicing is extensively used in metazoans to increase proteome diversity from a single gene, there are just a few known examples of alternative splicing in *S. cerevisiae* producing functional proteins with different roles. Alterative splicing of *PTC7* (153) and *SRC1* (154) generate proteins with different cellular localizations (nuclear envelope or mitochondria) and different folding patterns in the membrane, respectively. Also mitochondrial genes have very complex splicing patterns among multiple introns, encoding various essential proteins related to the respiratory chain complex (155). *GCR1* is the first example of producing two functionally different transcription factors by splicing and intron retention.

It has been shown that many intron-containing genes in *S. cerevisiae* generate non-productive mRNA species containing premature termination codon (PTC) by alternative splicing, which are degraded by nonsense-mediated mRNA decay (NMD) system (156). Therefore, alternative splicing in *S. cerevisiae* might be mainly used as a regulatory mechanism to down-regulate gene expression during stress. In fact, the 5 alternatively spliced *GCR1* transcripts with the PTC were shown to be degraded by NMD, and heat shock affected the selection of the splicing sites (101). In addition, unspliced *GCR1*

mRNA is also a target of NMD as a quality control system as observed for other intron-containing genes (101). Therefore, splicing might regulate *GCR1* at multiple levels including the formation of two different functional proteins and condition-specific degradation of mRNAs to suppress gene expression. Such complicated regulatory mechanisms of Gcr1 transcription factor might reflect the importance of sophisticated regulation of glycolysis for the survival of *S. cerevisiae* in ever-changing natural environment.

7.4. Application of lactic acid producing strain to distinguish the differential role of Gcr1 isoforms

Although I could not clearly determine the separate roles of two isoforms of Gcr1, the SP1130 strain enables the indirect observation of different phenotype between $Gcr1^{U}$ and $Gcr1^{S}$ strains (Figure 22A). Since the glycolytic flux is so strong in *S. cerevisiae*, so-called 'overflow metabolism', cells can utilize the glucose more effectively and grow fast (157). However, this highly efficient glucose consuming system makes the study difficult because distinguishable changes occur in a very short time. Also, intermediates of glycolysis are known to act as a flux sensor, suggesting that the glycolysis itself is a very

important metabolism, but also related to the various regulatory mechanisms (158, 159). Therefore, elimination of ethanol and glycerol production pathway in SP1130 strain is expected to be advantageous to observe the difference by decelerating the overall glycolytic flux. Though, since those pathways are so critical in the metabolism, there also exists a possibility that this difference between $Gcr1^{U}$ and $Gcr1^{S}$ is an indirect expression of the phenotype. Therefore, understanding the substantial role of two isoforms is far more challenging without application of computational approaches. Flux balance analysis (FBA) has recently been used to understand the dynamics at the starting point of glucose induction (160), providing a wider aspect of this kind of study. However, combining the computational method with experimental studies may take some time.

Microaerobic culture had been widely used in production of target compounds in several bacteria in order to reduce the aerobic respiration and rewire the cellular metabolism into production (150, 161). Although this change of culture condition has significant advantage, the difficulties of controlling the oxygen and the accumulation of glycerol that causes osmotic stress is a limitation of microaerobic condition (162). As studied above, the increased aeration rather decreased the amount of

lactic acid, indicating that the oxygen transfer rate must be carefully controlled (Figure 23). Thus, application of same principle that repressing the respiration even under aerobic condition by overexpressing *GCR1*^{S} seems clearly beneficial for producing NADH consuming chemical compounds.

Appendix.

Regulation of acetate tolerance by small ORF-encoded polypeptides modulating efflux pump specificity in *Methylomonas* sp. DH-1

A.1. Introduction

Methane is an abundant and low-cost carbon source available from natural gas and biogas. In addition, methane is a greenhouse gas with a greater effect on global warming than carbon dioxide. Therefore, there is growing interest in utilizing methane as a next-generation feedstock (1, 2). Methanotrophic bacteria, which utilize methane as a sole carbon and energy source, are promising hosts for the biological conversion of methane into value-added chemicals. Recently, various chemicals, including lactic acid, succinic acid, indole 3-acetic acid, and cadaverine, were successfully produced through the metabolic engineering of methanotrophs (3-7). Even though the metabolic pathways of various methanotrophs have been predicted based on genomic sequencing, transcriptome and metabolome analyses, and metabolic modeling, little is known regarding transcriptional regulatory networks. So far, only a few gene-specific transcription factors have been characterized in methanotrophs, including MmoD, involved in the regulation of methane monooxygenase (MMO) genes, and EctR1, involved in ectoine biosynthesis in *Methylomicrobium alcaliphilum* 20Z (8, 9). However, genome-wide studies to identify target genes and functions of transcriptional regulators have not yet been reported.

Organic acids, such as lactic, succinic, 3-hydroxy propionic, itaconic, and citric acids, are widely used in the food, cosmetics, and pharmaceutical industries, also serving as building blocks for polymer production (10). In addition, weak monocarboxylic acids, such as acetic, propionic, sorbic, and benzoic acids, are widely used as food and beverage preservatives, which inhibit microbial cell growth. Therefore, understanding the tolerance mechanisms against weak organic acids is of great relevance for the microbial production of organic acids through increasing acid tolerance and for efficient microbial control. Undissociated forms of organic acids in acidic medium can diffuse into cells and dissociate into protons and anions in the neutral cytosol (11, 12). Both protons and anions perturb normal cellular functions, thus inducing cellular defense mechanisms, which vary depending on the chemical structures of anions (13-15).

To improve lactate production from methane, previous study developed a lactate-tolerant strain JHM80 through adaptive laboratory evolution of *Methylomonas* sp. DH-1 (4). In that study, they determined that overexpression of the *watR* gene (AYM39_21120/AYM39_RS21130), which encodes a LysR-type transcriptional regulator (LTTR), is partly responsible for the lactate tolerance of JHM80. LTTR is one of

the largest family of bacterial regulators with diverse functions (16, 17). LTTR has a conserved N-terminal helix-turn-helix (HTH) motif responsible for DNA binding and a C-terminal effector binding domain, which recognizes various signaling molecules regulating LTTR activity. The *watR* operon, consisting of the *watR* and two downstream genes, is overexpressed in JHM80 due to a 2-bp (TT) deletion in the promoter region (4). Overexpression of the two downstream genes did not affect lactate tolerance, suggesting that overexpressed WatR may enhance lactate tolerance via the activation or repression of its target genes.

In this appendix, I investigated the role of WatR in regulating stress responses against weak organic acids. By investigating genome-wide binding targets of WatR and WatR-dependent transcriptional regulation, I propose a novel role for previously unannotated small open reading frames (smORFs) in acetate tolerance.

A.2. Materials and Methods

Strains and culture conditions

All strains used in this appendix study are listed in Table A.1. Strains derived from *Methylomonas* sp. DH-1 (KCTC13004BP) were cultured in 3-mL nitrate mineral salts (NMS) medium (0.49 g/L $MgSO_4$, 1.0 g/L KNO_3, 0.23 g/L $CaCl_2 \cdot 2H_2O$, 3.8 mg/L Fe-EDTA, 0.5 mg/L Na_2MoO_4, 10 μM $CuSO_4 \cdot 5H_2O$, with the addition of trace element solution, vitamin stock and phosphate stock solution: recipes of these solutions are in Table A.2.) with 20% (v/v) methane in a 30-mL serum bottle capped with a butyl rubber stopper at 30 °C with shaking at 170 rpm (4). For chromatin immunoprecipitation (ChIP) and ChIP-seq experiments, strains were cultured in 50 mL NMS medium with 20% (v/v) methane in a 500 mL baffled flask sealed with rubber type screw cap.

Table A.1. Strains used in appendix

Strain	Description	Genotype	Reference
Methylomonas sp. DH-1	Wild type strain		(50)
JHM15	*watR* operon deletion in DH-1	DH-1 *Δ(watR-smtM-rstM)::kan*^{R}	This study
JHM16	*fliE* deletion in DH-1	DH-1 *ΔfliE::kan*^{R}	This study
JHM161	*watS1* overexpression in DH-1	DH-1 *ΔfliE::*P_{mxaF}*-watS1-*T_{rrnB}*-kan*^{R}	This study
JHM162	*watS2* overexpression in DH-1	DH-1 *ΔfliE::*P_{mxaF}*- watS2-*T_{rrnB}*-kan*^{R}	This study
JHM163	*watS3* overexpression in DH-1	DH-1 *ΔfliE::*P_{mxaF}*- watS3-*T_{rrnB}*-kan*^{R}	This study
JHM164	*watS4* overexpression in DH-1	DH-1 *ΔfliE::*P_{mxaF}*- watS4-*T_{rrnB}*-kan*^{R}	This study
JHM165	*watS5* overexpression in DH-1	DH-1 *ΔfliE::*P_{mxaF}*- watS5-*T_{rrnB}*-kan*^{R}	This study
JHM17	*watS1* deletion in DH-1	DH-1 *ΔwatS1::kan*^{R}	This study
JHM18	WatABO pump deletion in DH-1	DH-1 *Δ(watPAB-watO)::amp*^{R}	This study
JHM181	*fliE* deletion in JHM18	JHM18 *ΔfliE::kan*^{R}	This study
JHM182	*watS1* deletion in JHM18	JHM18 *ΔwatS1::kan*^{R}	This study
JHM183	*watS1* overexpression in JHM18	JHM18 *ΔfliE*::P_{mxaF}*- watS1-*T_{rrnB}*-kan*^{R}	This study
JHM80	Evolved strain from DH-1		(4)
JHM82	*watR* operon deletion in JHM80	JHM80 *Δ(watR-smtM-rstM)::kan*^{R}	(4)
JHM87	WatABO pump deletion in JHM80	JHM80 *Δ(watABP-watO)::kan*^{R}	This study
JHM16WF	*watR-Flag* overexpression in DH-1	DH-1 *ΔfliE*::P_{EFTu}*- watR*-Flag-T_{rrnB}*-kan*^{R}	This study
JHM80WF	*watR-Flag* tagging in JHM80	JHM80 *watR*-Flag-T_{rrnB}*-Kan*^{R}	This study
JHM161T	*watS1*-T7 overexpression in DH-1	DH-1 *ΔfliE*::P_{mxaF}*- watS1*-T7-T_{rrnB}*-kan*^{R}	This study

Table A.2. Recipes of the stock solutions of NMS

Ingredient	Amount
1000X trace element solution	
$FeSO_4 \cdot 7H_2O$	500 mg/L
$ZnSO_4 \cdot 7H_2O$	400 mg/L
$MnCl_2 \cdot 7H_2O$	20 mg/L
$CoCl_2 \cdot 6H_2O$	50 mg/L
$NiCl_2 \cdot 6H_2O$	10 mg/L
H_3BO_3	15 mg/L
EDTA	250 mg/L
100X vitamin stock	
Biotin	2.0 mg/L
Folic acid	2.0 mg/L
Thiamine HCl	5.0 mg/L
Ca pantothenate	5.0 mg/L
Vitamin B12	0.1 mg/L
Riboflavin	5.0 mg/L
Nicotinamide	5.0 mg/L
100X phosphate stock solution	
KH_2PO_4	26 g/L
Na_2HPO_4	32.8 g/L

Plasmid construction

Plasmids and primers used in this appendix study are listed in Table A.3. and Table A.4. Plasmids for deletion were generated based on the pDel2-fliE plasmid by replacing the chromosome targeting sequences for fliE with 1-kb upstream and downstream sequences of the target genes. To generate pDel-watABPO(A) plasmid, ampicillin resistance gene (*AmpR*) was PCR amplified from pCM184 (18) and cloned between *Apa*I and *Pac*I site, replacing kanamycin resistance gene (*KanR*) of Del2-watABPO(K). For DNA integration via substituting *fliE*, plasmid pFliE-mxaF containing [U_{fliE}-T_{rrnB}-P_{mxaF}-T_{rrnB}-*KanR*-D_{fliE}] cassette was generated by inserting P_{mxaF} promoter using *Mau*BI and *Bam*HI sites, and T_{rrnB} terminator using *Asc*I and *Mau*BI sites between the U_{fliE}-T_{rrnB} cassette of pDel2-fliE plasmid. T_{rrnB} terminator was inserted right after the U_{fliE} cassette to prevent transcription from *fliE* promoter after genome integration. The gene of interest were cloned between the promoter and terminator using *Bam*HI and *Spe*I sites for overexpression. To make pFliE-EFTu-watR-Flag plasmid, P_{mxaF} of pFliE-mxaF was subtituted to P_{EFTu} by *Mau*BI and *Bam*HI, and Flag tag containing *watR* ORF was cloned with *Bam*HI and *Spe*I. pWatR-G4S-Flag was designed to insert the Flag tag sequence

with G4S linker before the stop codon of the *watR* ORF (19). The upstream homology region was amplified with reverse primer containing G4S linker, Flag tag sequence, and stop codon, and then cloned into pDel2 using *Not*I and *Spe*I.

Genetic manipulation of *Methylomonas* sp. DH-1

Gene deletion or insertion in *Methylomonas* sp. DH-1 and JHM80 strains were performed as previously described via homologous recombination into the chromosome, using electroporation (4).

Quantitative reverse transcription PCR (qRT-PCR) and RNA-seq

Total RNA of *Methylomonas* sp. DH-1, JHM15, JHM80, JHM82 and JHM16WF were extracted as previously described with minor modifications (4). For qRT-PCR analysis, 5 μL of cDNA (diluted 1:200) was amplified by SYBR Green I master mix (Roche-Applied Science, USA) and analyzed with gene-specific primers. The Crossing point (Cp) values were processed using Light Cycler 480 software version 1.5 and expression levels of each target genes were normalized by that of *mxaF* (AYM39_RS15615).

Table A.3. Plasmids used in appendix

Plasmid	Description	Reference
Plasmids for gene deletion in *Methylomonas* sp. DH-1		
pDel2-WSR	pDel2-U_{watR}-[T_{rrnB}-Kan^R]-D_{rstM}	(7)
pDel2-fliE	pDel2-U_{fliE}-[T_{rrnB}-Kan^R]-D_{fliE}	(7)
pDel2-watABPO(A)	pDel2-D_{watO}-[T_{rrnB}-Amp^R]-D_{watA}	This study
pDel2-watABPO(K)	pDel2-D_{watO}-[T_{rrnB}-Kan^R]-D_{watA}	This study
pDel2-watS1	pDel2-U_{watS1}-[T_{rrnB}-Kan^R]-D_{watS1}	This study
Plasmids for gene expression in *Methylomonas* sp. DH-1		
pFliE-mxaF	pDel2-U_{fliE}-T_{rrnB}-[P_{mxaF}-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS1	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS1*-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS2	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS2*-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS3	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS3*-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS4	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS4*-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS5	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS5*-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-watS1-T7	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{mxaF}-*watS1*-T7-T_{rrnB}-Kan^R]-D_{fliE}	This study
pFliE-EFTu-watR-Flag	pFliE-mxaF-U_{fliE}-T_{rrnB} -[P_{EFTu}-*watR*-G4S-Flag-T_{rrnB}-Kan^R]-D_{fliE}	This study
pWatR-G4S-Flag	pDel2-*watR*-G4S-Flag-[T_{rrnB}-Kan^R]-D_{watR}	This study
Plasmids for gene expression in *E. coli*		
pGEX-4T-1-WatR	pGEX-4T-1-P_{tac}-*GST*-*watR*	This study

Table A.4. Primers used for plasmid and strain construction in appendix

Forward primer (5' to 3')	Reverse primer (5' to 3')	Usage
Primers used for gene deletion in *Methylomonas* sp. DH-1		
actGCGGCCGCTCACCAGCCTGTTAGGTG	actACTAGTATATTGTCGGCGAAATTCGT	Cloning U_{watO} of pDel2-watABPO(K) with *Not*I/*Spe*I
actGGGCCCTCGCTTCGGCTAACAGTG	actGAGCTCCCCAACAAAGCGATAGCG	Cloning D_{watA} of pDel2-watABPO(K) with *Apa*I/*Sac*I
gcgATTTAAATCGAATTGGACAGTCCCGCCAC	gcgACTAGTGGTCTACTCCAAAAGTTGAACACATGGC	Cloning U_{watS1} of pDel2-watS1 with *Swa*I/*Spe*I
actGGGCCCCCGCTTGCGCCCGGCGCGGG	gcgGAGCTCGCCGTCAAGCGGGCTGCACTTTATC	Cloning D_{watS1} of pDel2-watS1 with *Apa*I/*Sac*I
actGGGCCCGCGGAACCCCTATTTGT	gcgTTAATTAATCAAGAAGATCCTTTGATC	Cloning Amp^R of pDel2-watABPO(A) with *Apa*I/*Pac*I
Primers used for gene expression in *Methylomonas* sp. DH-1		
gactGCGGCCGCTAGCGAAAACCAACGTGACG	actCGCGCGCGGATATTGACTCGTGTGTATTGCCTGCC	Cloning U_{fliE} of pFliE-mxaF with *Not*I/*Mau*BI
tataGGGCCCGTAGCGGAAGCCGGCCA	tataGGCGCGCCGGCACGCCTTGAACTT	Cloning D_{fliE} of pFliE-mxaF with *Apa*I/*Asc*I
actGGCGCGCCCAAATAAAACGAAAGGCTCAGTCGG	actCGCGCGCGATTTGTCCTACTCAGGAGAGCGTTCACCG	Cloning T_{rrnB} of pFliE-mxaF with *Asc*I/*Mau*BI
gggCGCGCGCGCTGATTTTGTTTGCCACAGGC	gcgGGATCCGAATCCTCCTAAGTTGTTTATTAGAGTGC	Cloning P_{mxaF} of pFliE-mxaF with *Mau*BI/*Bam*HI
gcgGGATCCATGACTAACGTACAAAAAGACATTCTGAACG	gcgACTAGTTCAGTCTTTTTGATTTTTCCGGTTTAGGTTG	Cloning *watS1* of pFliE-watS1 with *Bam*HI/*spe*I
actGGATCCATGAGCAAATCTAACAATCTTTATCGAG	gcgACTAGTTTAAGCGCGGAACGAGCC	Cloning *watS2* of pFliE-watS2with *Bam*HI/*spe*I
gcgGGATCCATGAACCAGTCAAAATTTAAAGACATCG	gcgACTAGTTTAAATCCGGACCGGTTTGGCAAC	Cloning *watS3* of pFliE-watS3 with *Bam*HI/*spe*I
gcgGGATCCATGAATACCCCGACGTTCTATCG	gcgACTAGTTCAGATGCGTACCGGTTTGG	Cloning *watS4* of pFliE-watS4 with *Bam*HI/*spe*I
gcgGGATCCATGGAACTTGTTATCCAACCGGTTC	gcgACTAGTTCAGGCAGCGATAGAGGCTGTAGAG	Cloning *watS5* of pFliE-watS5 with *Bam*HI/*spe*I
gcgGCGGCCGCGCAACTTCCAACGCCACCGG	gcgACTAGTTCACTTGTCATCGTCATCCTTGTAATCagaacc accccgccTGGGTTCATGCCGATCCG	Cloning *watR*-G4S-Flag of pWatR-G4S-Flag with *Not*I/*Spe*I
gcgGGGCCCGCGAACTGCAACGCAAGTGGG	gcgGAGCTCGGCCGGCGCGGAAGTGGGCTGG	Cloning *watR* downstream 1 kb to pWatR-G4S-Flag with *Apa*I/*Sac*I
Primers used for gene expression in *E. coli*		
gcgGAATTCATGGACAAACTAACCAGCATGAACG	gcgGCGGCCGCTCATGGGTTCATGCCGATCC	Cloning *watR* to pGEX-4T-1 vector with *Eco*RI/*Not*I

* Underlined upper case sequences indicate restriction enzyme sites

* Underlined lower case sequences indicate G4S linker sequence

For RNA-seq, two sets of total RNA from *Methylomonas* sp. DH-1 and JHM80 were isolated. 1 μg of total RNA was proceeded to rRNA depletion using NEBNext rRNA depletion kit (Bacteria) (#7850, NEB). Resulted mRNA was used for sequencing library construction by TruSeq Stranded mRNA Library Prep kit (#20020594, Illumina). All experiments were performed following manufacturer's instructions. The prepared sequencing library was sequenced using NovaSeq 6000 (Illumina). The sequencing adapter removal and quality-based trimming on raw data was performed by Trimmomatic v. 0.36 with default parameter (20). Cleaned reads were mapped to reference genome (*Methylomonas* sp. DH-1, GCF_001644685.1) using hisat2 v. 2.2.1 with '--no-spliced-alignment' option (21). For counting reads which mapped to each CDS, featureCounts in Subread package was used (22). Finally, normalization of retrieved counts and fold change calculation between groups were performed by DESeq2 package (23).

ChIP and ChIP-seq analyses

ChIP assay was conducted as previously described with minor modifications using JHM16WF and JHM80WF strains harboring *watR*-Flag (24). Final 2.7% of formaldehyde was

added to 50 mL cell culture using 5 mL syringe and crosslinked for 25 min, and quenched with 250 mM glycine for 5 min. After washing once with ice-cold NMS, twice with TBS, and once with lysozyme buffer [10 mM Tris-HCl (pH 8.0), 20% sucrose, 50 mM NaCl, 10 mM EDTA], cells were resuspended to 500 μL of lysozyme buffer containing 10 mg/mL of lysozyme (Thermo Scientific, USA), 1 mM PMSF and 0.1% protease inhibitor cocktail. After cell lysis with shaking at 37 ℃ for 30 min, 500 μL of 2x ChIP lysis buffer [100 mM HEPES-KOH (pH 7.5), 300 mM NaCl, 2 mM EDTA, 2% Triton X-100, 0.2% sodium deoxycholate, 0.4% SDS] was added and sonicated 12 times for 20 s (Vitra-cell, Sonics & materials inc, USA) with amplitude 22%. Crude lysates were centrifuged for 20 min and 100 μL of supernatant was used as input. 2 μL of anti-DDDDK antibody was added to proper amounts of lysate (400 μL for JHM80WF strain and 900 μL for JHM16WF strain) and immunoprecipitated overnight at 4oC, followed by 2 h incubation with 20 μL Protein A Plus agarose bead (Santa Cruz Biotechnology, USA). After washing the beads, DNA was eluted from beads and treated with RNase and proteinase K. Crosslink was reversed by overnight incubation at 65 ℃ with 100 mM NaCl, and DNA was purified using Qiagen DNA purification kit. The occupancy of WatR on the target promoter

was calculated by dividing the amount of PCR product from the IP sample into input sample, compared to the negative control (*glgA* ORF).

For ChIP-seq analysis, the eluted DNA after ChIP was extracted with phenol-chloroform-isoamylalcohol (25:24:1) and precipitated with ethanol and glycogen at -80 ℃. 1 ng of prepared DNA was proceeded to sequencing library construction by using NEBNext Ultra II DNA Library Prep Kit for Illumina (#E7645, NEB) following manufacturer's instructions. The sequencing adapter removal and quality-based trimming on raw data was performed by Trimmomatic v. 0.36 with default parameter (20). Cleaned reads were mapped to reference genome using bowtie2 v. 2.4.2 with default parameter. Peak calling was performed by findpeaks command in homer v. 4.10.3 using "-style factor" parameter (25). Resulted peaks were annotated by annotatePeaks.pl in homer package. Peaks was transformed to bed file using pos2bed.pl in homer package for detailed analysis. The conserved motifs from peaks were found by MEME-ChIP v4.9.0 (26).

Data availability

The ChIP-seq and RNA-seq data have been deposited in the GEO repository under the accession number of GSE206217

(GSE206215 for RNA-seq and GSE206216 for ChIP-seq).

Electrophoretic mobility shift assay (EMSA)

EMSA was performed with biotin-labeled DNA probe, which was prepared via PCR amplification using 5′-biotin modified primer and gel extraction. Total 20 μL of mixture, containing 20 fmole probe and 0.4 to 1.2 μg of GST-WatR protein purified from E. coli in binding buffer [100 mM Tris-HCl (pH 7.5), 500 mM KCl, 10 mM DTT, 60% glycerol, 5 mM EDTA, 500 μg/mL BSA, 500 μg/mL salmon sperm DNA] was incubated at room temperature for 20 min. After adding 5x sample buffer without SDS [60 mM Tris-HCl (pH 6.8), 0.1% bromophenol blue, 5% 2-mercaptoethanol, 25% glycerol], the mixture was resolved via electrophoresis in 6% native polyacrylamide gel with 0.5X TBE buffer [40 mM Tris base, 45 mM Boric acid, and 1 mM EDTA]. DNA was transferred to hybond-N+ membrane (GE healthcare Amersham, USA), and crosslinked for 10 min using UV lamp (Korea Ace Sci., Korea) device under 254 nm wavelength. The membrane was incubated with HRP-conjugated streptavidin (Thermofisher scientific, USA) and the signals were visualized by G::box Chemi-XL (Syngene, USA).

A.3. Differential roles of WatR in tolernace to weak organic acids

In the previous study, it wasy demonstrated that overexpression of *watR* due to a mutation in its promoter region is partly responsible for lactate tolerance in the JHM80 strain (4). I further examined whether WatR is involved in the regulation of tolerance against other weak organic acids, including formate, acetate, and propionate, by growing cells in the presence these acids. The mutant JHM80 strain exhibited higher tolerance against formate and propionate than its wild-type counterpart (Figure A.1). However, the tolerance phenotypes were abolished by deleting the *watR* operon (JHM82), suggesting that *watR* overexpression can enhance tolerance against propionate and formate, as well as lactate. In addition, deletion of the *watR* operon in the wild-type strain (JHM15) led to a decrease in lactate, propionate, and formate tolerance, further confirming the role for WatR in the stress response induced by these acids (Figure A.1). It had been previously confirmed that *watR* deletion in JHM80 reduced tolerance in the presence of 8 g/L lactate, while deletion of *watR* barely affected lactate tolerance at a lower concentration of 0.3 g/L (Figure A.1). Even without the *watR* gene, the JHM80 Δ*watR* strain exhibited higher lactate tolerance than the

wild-type strain (Figure A.1). In contrast, the JHM80 Δ*watR* strain was more sensitive to propionate and formate than the watR-deleted wild-type strain (Figure A.1). These phenotypes of JHM80 Δ*watR* might be due to additional mutations in the JHM80 strain conferring a selective advantage for lactate tolerance, which has not yet been characterized (4).

Intriguingly, the effect of WatR on acetate tolerance showed the opposite tendency when compared with the other acids, i.e., JHM80 exhibited lower tolerance to acetate than the wild type but the tolerance increased when *watR* was deleted (Figure A.1). Deletion of *watR* in the wild type also increased acetate tolerance, suggesting that WatR may negatively affect acetate tolerance. Taken together, WatR has differential effects on tolerance depending on the type of weak organic acids: it increases tolerance against lactate, propionate, and formate while decreasing tolerance against acetate.

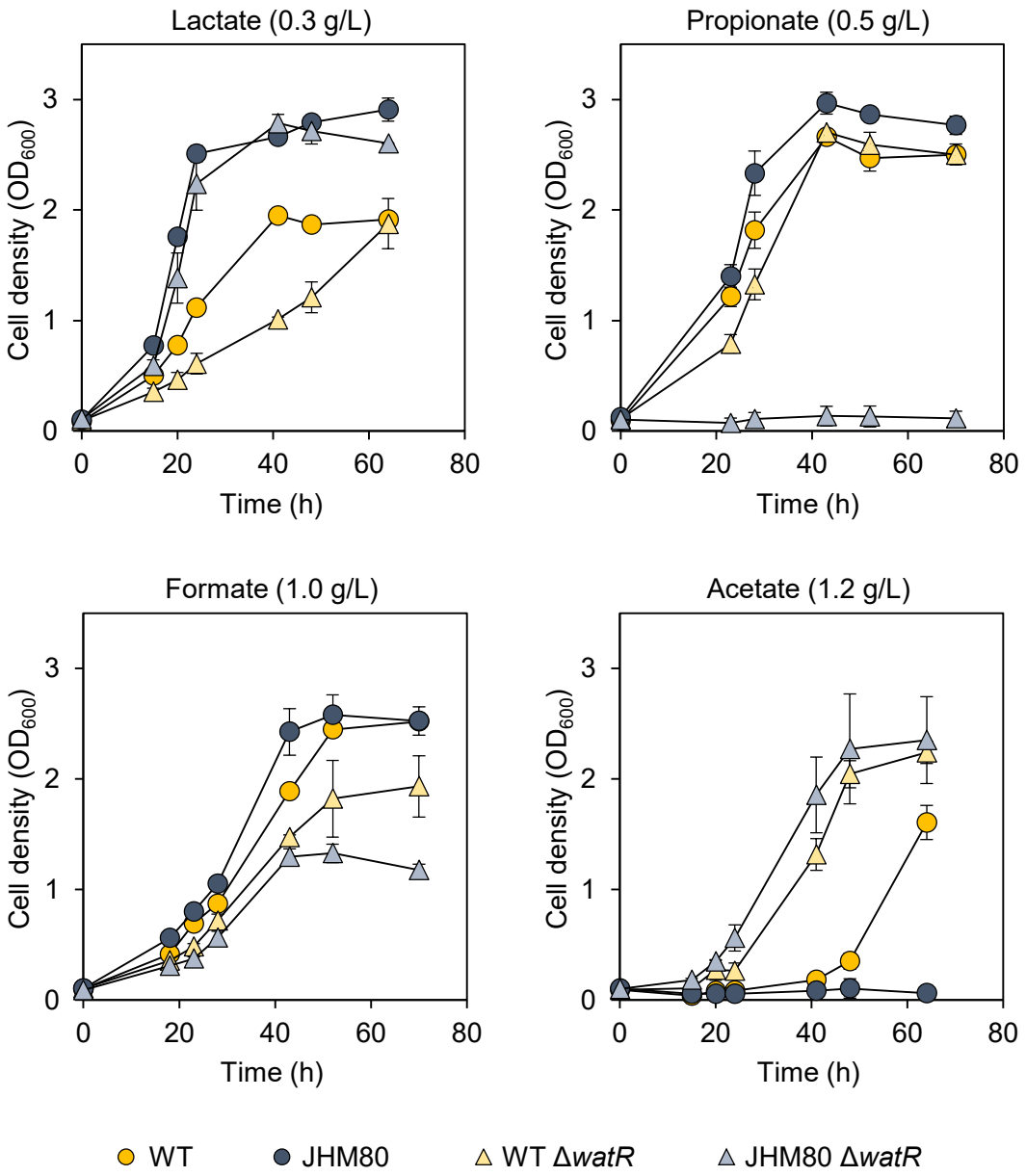

Figure A.1. Effects of watR deletion in the wild-type and JHM80 strains in weak organic acid tolerance.

Wild-type, JHM80, Δ*watR* (JHM15), and JHM80 Δ*watR* (JHM82) strains were grown in NMS media with 20% (v/v) methane containing 0.3 g/L lactate, 0.5 g/L propionate, 1.0 g/L formate, and 1.2 g/L of acetate with pH neutralization. Two independent experiments were averaged and plotted with standard deviations.

A.4. Determination of genome-wide binding sites and target genes of WatR

A.4.1. Genome-wide binding sites of WatR

To understand the role of WatR in weak acid tolerance, I identified genome-wide binding targets of WatR via ChIP-Seq analysis. To obtain reliable ChIP-seq signals, the *watR* gene in the genome of JHM80 strain was tagged with Flag (JHM80WF), resulting in overexpression of *watR*-Flag. The Flag tagging to WatR did not affect normal cell growth. I obtained 22 WatR binding peaks, among which 16 were located in the promoter regions of annotated genes (Table A.5). Conserved binding motifs, ATTGTT-[N]11-AACAA, were identified in the WatR-binding promoter regions (Figure A.2A), which is in agreement with the palindromic binding sites (T-[N]11-A) previously identified for LTTRs (16). Binding of WatR to some of the target promoters was also confirmed via ChIP-qPCR (Figure A.2B). The targets included functionally diverse genes encoding a sulfate/thiosulfate transporter and enzymes such as citrate synthase (*gltA1*), phosphomannomutase (*pmmM*), dethiobiotin synthase (*bioD*), and membrane-bound protease (*ftsH*).

A.4.2. WatR functions as a repressor of its expression and *gltA*

LTTRs are well known to autoregulate their expression (16). As previously reported, the expression of genes within the *watR* operon was upregulated in the JHM80 strain harboring a TT deletion in its own promoter (Figure A.2C). The deleted TT sequence is part of a putative -10 box, which overlaps with the predicted WatR binding sequence (Figure A.2D), suggesting that the TT deletion may reduce the binding of WatR to its promoter. This idea agrees with the fact that the *watR* promoter was not detected as a WatR binding site in our ChIP-seq experiment performed in the JHM80 strain background. Furthermore, I confirmed this hypothesis through an *in vitro* electrophoretic mobility shift assay with purified WatR protein (Figure A.2D). WatR exhibited a higher binding affinity to the wild-type watR promoter than the promoter harboring a TT deletion (Figure A.2D). These results suggest that WatR acts as a repressor of its expression. In the JHM80 strain, the TT deletion within the promoter region may prevent WatR binding, leading to derepression of the *watR* operon (Figure A.2E).

Among WatR target genes, *gltA1* encodes citrate synthase, which catalyzes the condensation between acetyl-CoA

and oxaloacetate in the first step of the TCA cycle. Overexpression of *gltA* has been shown to eliminate acetate production and redirect carbon flux toward the TCA cycle in other bacteria (27, 28). Therefore, I investigated whether *gltA1* expression levels are related to the WatR-dependent sensitivity to acetate stress. JHM80 exhibited lower *gltA1* expression than the wild type, which was restored following watR deletion, indicating that WatR represses *gltA1* (Figure A.2F). However, neither deletion nor overexpression of *gltA1* in the wild-type strain affected growth in the presence of acetate under my experimental conditions (data not shown).

Table A.5. WatR-binding sites determined by ChIP-seq

The nearest gene (AYM39_RS)	Promoter region	Name	Function	Reannotated gene
00605	+		hypothetical protein	*watS1*
02245				*watS2*
07485	+			*watS3*
19610	+			*watS4*
13560			glutathione S-transferase	*watS5*
17395	+	*watP*	hypothetical protein	
00705	+		FtsH/Yme1/Tma family protein	
02410	+		hypothetical protein	
02830			methionine adenosyltransferase	
02950	+		hypothetical protein	
04660	+	*pmmM*	phosphomannomutase	
06470	+		CRISPR_1	
06815			GDP-L-fucose synthase	
06860	+		hypothetical protein	
07170			aspartate kinase	
08175	+	*bioD*	dethiobiotin synthase	
09830	+		sulfate ABC transporter	
11160	+		hypothetical protein	
15960	+		hypothetical protein	
19040	+	*gltA1*	citrate (Si)-synthase	
19890			S41 family peptidase	
20750	+		hypothetical protein	

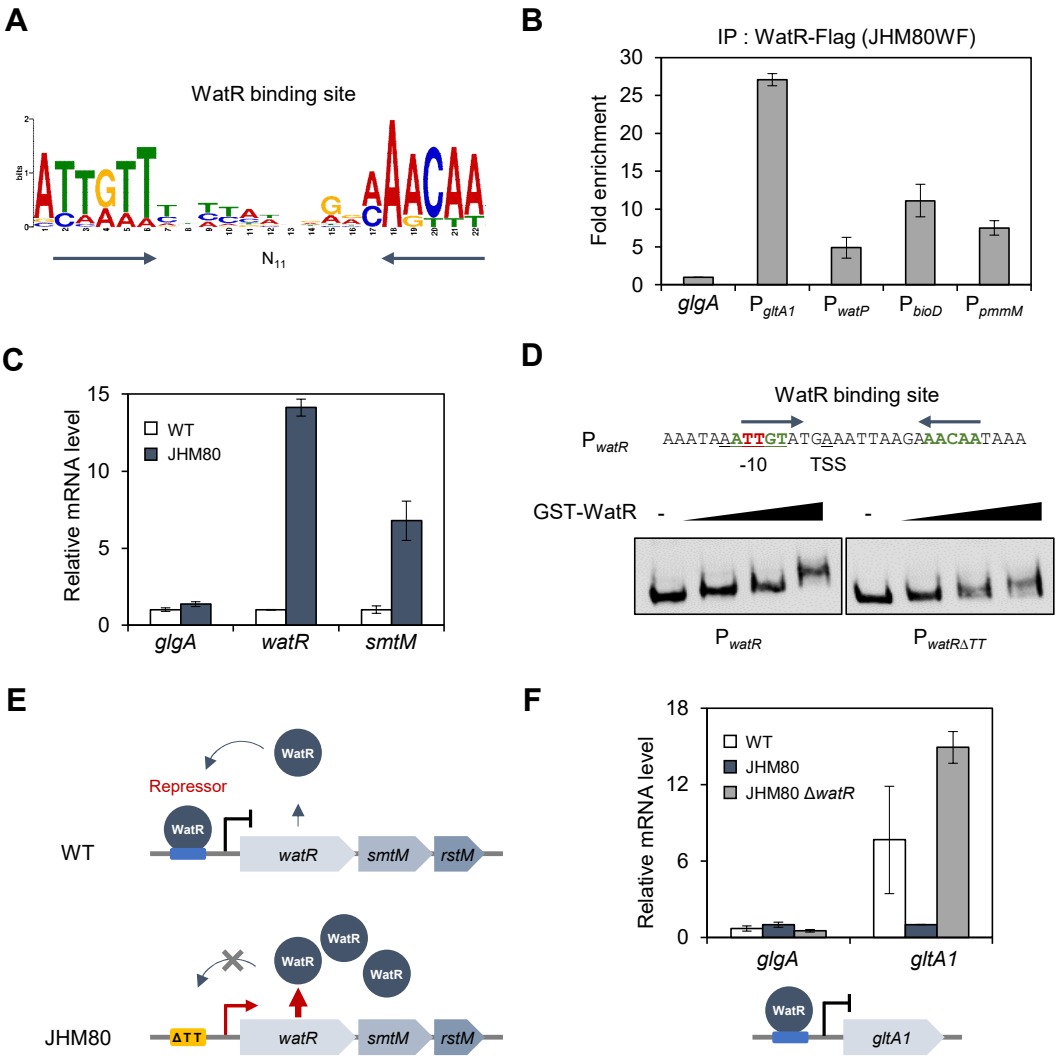

Figure A.2. Identification of genome-wide binding sites of WatR.

Figure A.2. Identification of genome-wide binding sites of WatR.

(A) Binding consensus sequence of WatR, discovered by ChIP-seq analysis.

(B) Confirmation of WatR binding to the promoters of four selected target genes identified by ChIP-seq analysis. Binding of WatR-Flag to the target promoters in JHM80WF strain was detected by ChIP with anti-Flag antibody and indicated as fold enrichment relative to the binding to a negative control (*glgA* ORF). Each value represents the average ± standard deviations from three independent experiments.

(C) Expression levels of *watR* operon genes (*watR* and *smtM*) in wild-type and JHM80 strains. The mRNA levels detected by qRT-PCR were normalized to that of *mxaF* gene and indicated as relative expression levels compared with those of wild type. The *glgA* gene was used as a negative control. Each value represents the average ± standard deviations from two independent experiments.

(D) Binding of WatR to the *watR* promoter detected through in vitro EMSA assay. EMSA assay was performed by incubating GST-WatR protein with biotin-labeled *watR* promoter probes with or without TT deletion. WatR binding sites (arrows), a putative -10 box, and transcription start site (TSS) are indicated. The deleted TT nucleotides in JHM80 are shown in red.

(E) Autoregulation of *watR* expression. In wild type, WatR binds to its own promoter, repressing the expression. In the JHM80 strain, the TT deletion in the promoter prevents WatR binding, resulting in derepression of the operon.

(F) Repression of *gltA1* by WatR. Expression levels of *gltA1* in the wild-type, JHM80, and JHM80 Δ*watR* strains were detected by qRT-PCR. The relative mRNA levels are indicated compared with those of JHM80. Each value represents the average ± standard deviations from two independent experiments.

A.4.3. WatR activates genes encoding an efflux pump involved in general weak organic acid tolerance

Among the WatR target genes, I identified a gene cluster encoding a resistance-nodulation-division (RND)-type efflux pump commonly found in gram-negative bacteria. The gene cluster includes an operon consisted of the AYM39_RS17395 gene (named *watP*), 17390 (named *watA*), and 17385 (named *watB*), as well as a divergently transcribed gene, AYM39_RS17405 (named *watO*) (Figure A.3A). Although *watP* has an unknown function, *watA*, *watB*, and *watO* were predicted to encode a membrane fusion protein, inner membrane protein, and outer membrane protein, respectively, forming a tripartite complex of the RND-type efflux pump (Figure A.3A). RND pumps are known to actively transport various antibiotics, organic substances, and metals (29, 30). Therefore, I hypothesized that this RND pump (named the WatABO pump) might be responsible for the WatR-dependent tolerance against various weak organic acids. The expression levels of these genes were higher in the JHM80 strain than in the wild type but restored when the *watR* gene was deleted in JHM80, suggesting that overexpressed WatR activates their transcription (Figure A.3B).

When the above-mentioned genes were deleted in

JHM80, which can tolerate lactate at concentrations up to 8 g/L, the resulting strain exhibited severe lactate sensitivity even in the presence of 1 g/L lactate (Figure A.3C). This result suggested that the efflux pump plays a central role in the lactate tolerance of JHM80, possibly through pumping lactate out of cells. The WatABO-deficient JHM80 strain also exhibited sensitivity toward other weak organic acids including propionate, formate, and acetate, suggesting that the efflux pump works for a wide range of weak acids (Figure A.3C). Notably, the deletion of watR decreased tolerance against lactate, propionate, and formate but increased acetate tolerance (Figure A.1). Therefore, although the WatR-activated efflux pump can contribute to acetate efflux, other genes regulated by WatR seem to play more dominant roles in acetate tolerance.

A.4.4. Expression of WatR target genes is induced by acetate but not lactate

Since *watR* deletion increased acetate tolerance but decreased lactate tolerance, I investigated whether the expression of *watR* target genes is regulated by these acids. Wild-type and *watR* deletion strains were treated with 0.15 g/L lactate or 0.6 g/L acetate for 10 min, which did not affect cell growth (data not shown). Although WatR overexpression

increased lactate tolerance, the expression of WatR-repressed genes (*watR* and *gltA1*) and a WatR-activated gene (*watP*) was not considerably affected by lactate treatment (Figure A.4A). In contrast, acetate treatment induced the expression of *gltA1* and *watP* (Figure A.4B). The acetate-dependent induction of *watP* and *gltA1* was diminished via *watR* deletion, suggesting that their induction mainly depended on WatR activity, regardless of whether WatR functions as an activator or repressor for basal expression (Figure A.4B). In agreement with the WatR-dependent repression of *gltA1* (Figure A.2F), the *watR* deletion mutant exhibited higher basal expression levels of *gltA1* than the wild-type. In contrast, basal *watP* expression was not affected by *watR* deletion, suggesting that WatR activates *watP* only in the presence of acetate (Figure A.4B). However, the high basal expression level of *watP* in JHM80 suggests that high levels of WatR can activate the *watP* operon even in the absence of the inducer (Figure A.3B). The expression of *watR* gene was not induced by acetate (Figure A.3B), implying that acetate-dependent conformational changes in WatR may bring forth different effects depending on the target promoters.

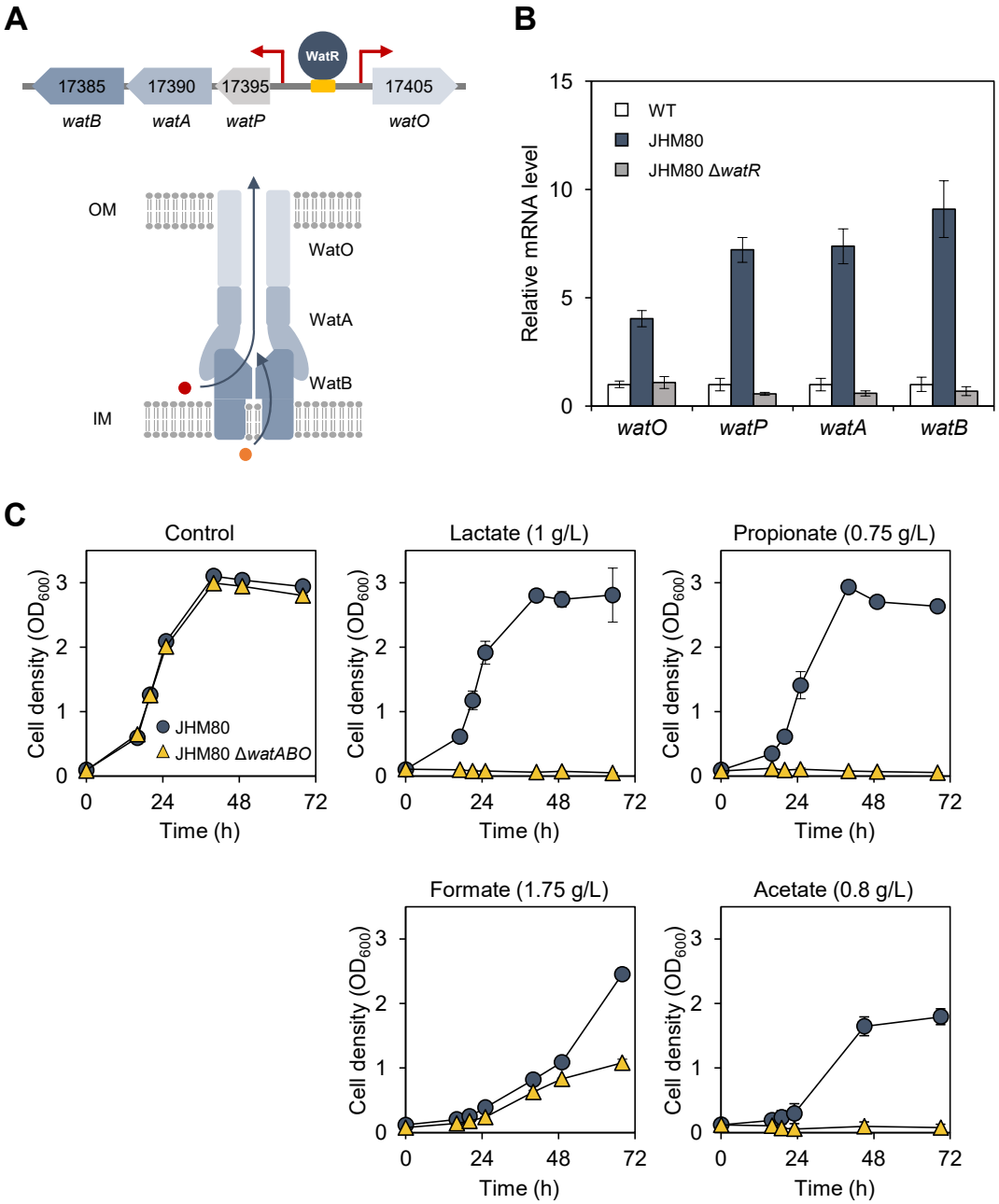

Figure A.3. WatR-dependent activation of genes encoding an RND-type efflux pump contributing to organic acid tolerance.

Figure A.3. WatR-dependent activation of genes encoding an RND-type efflux pump contributing to organic acid tolerance.

(A) Gene structure and putative functions of *watPAB* and *watO* genes regulated by WatR. Side view of the predicted WatABO efflux pump is shown. OM, outer membrane; IM, inner membrane.

(B) WatR-dependent activation of the efflux pump genes. Transcript levels were detected by qRT-PCR in the wild-type, JHM80, and JHM80 Δ*watR* strains and indicated as values relative to those of wild type.

(C) Effect of deleting the efflux pump genes on acid tolerance. The JHM80 strain and JHM80 strain lacking the WatABO efflux pump (JHM87) were grown in NMS media with 20% (v/v) methane without or with the indicated weak organic acids. Each value represents the average ± standard deviations from two independent experiments.

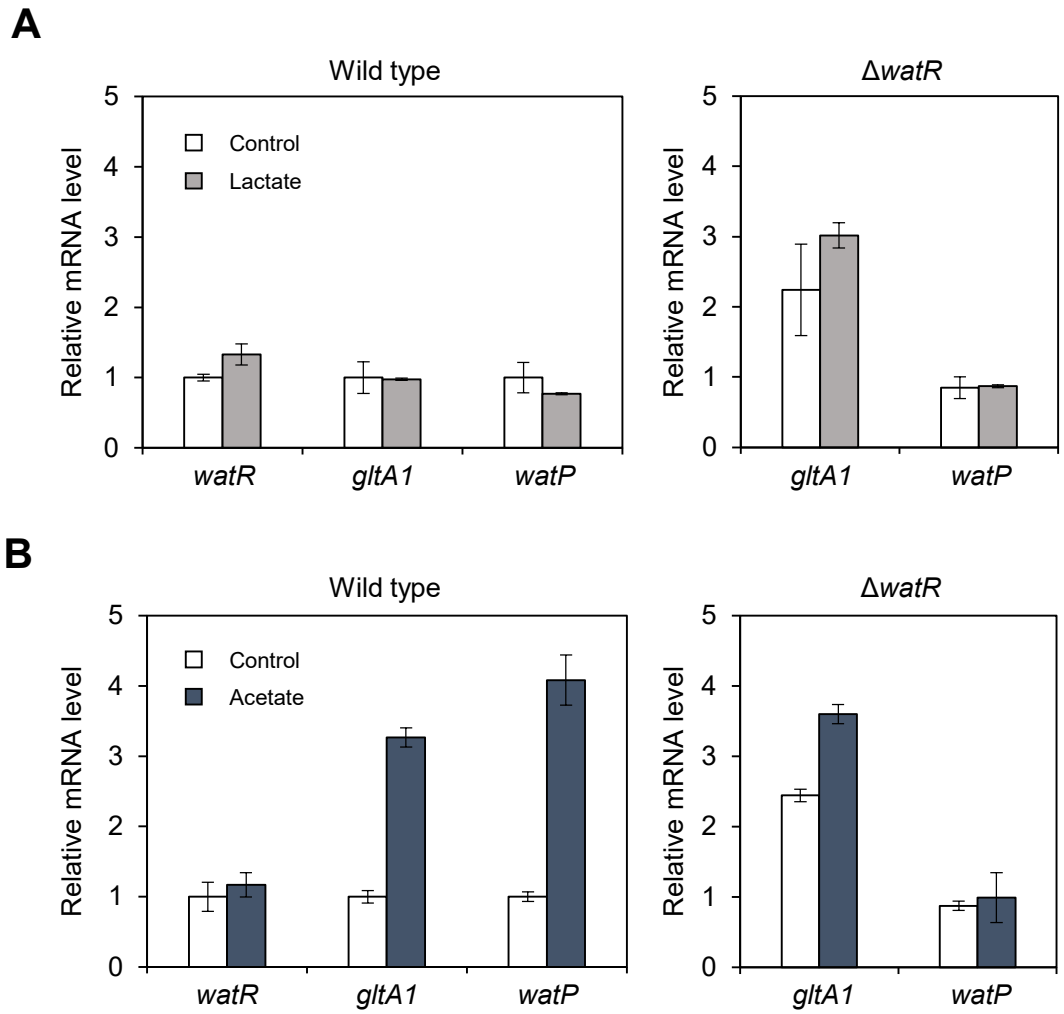

Figure A.4. Induction of WatR target genes by acetate but not lactate. The wild-type and ΔwatR strains were grown in NMS media with 20% (v/v) methane until early exponential phase and then treated with 0.15 g/L lactate (A) or 0.6 g/L acetate (B) for 10 min. Transcript levels were detected by qRT-PCR and indicated as values relative to those of untreated wild type. Each value represents the average ± standard deviations from two independent experiments.

A.5. Role of smORF-encoded peptides in acetate tolerance

A.5.1. WatR regulates the expression of previously unannotated small open reading frames (smORFs)

To further understand cellular responses against acetate stress, I analyzed changes in the transcriptome following acetate treatment in both wild-type and Δ*watR* strains using RNA-seq experiments. Acetate treatment resulted in the differential expression of 72 genes by ≥2-fold ($p < 0.05$), including 49 induced genes and 23 repressed genes (Table A.6, A.7 and A.8). The functional categories highly represented were membrane transporters in the induced genes and molecular chaperones in the repressed genes. Six of the induced genes exhibited at least two-fold reduced induction in Δ*watR*, suggesting WatR-dependent activation of these genes in response to acetate (Table A.6). These genes included three genes identified as direct WatR targets via ChIP-Seq analysis: AYM39_RS00605, 13560, and 17390 (*watA*). Consistent with the qRT-PCR experiments shown in Figure A.4B, the expression of *gltA1* also increased by acetate in a WatR-dependent manner. However, due to an induction fold (~1.8) lower than our filtration criteria, this gene was excluded from our initial selection. Two of the acetate repressed genes (AYM39_RS18690

and 18695) showed WatR-dependent repression, but WatR may regulate these gene indirectly because WatR binding to these genes was not detected (Table A.5 and Table A.8).

Unexpectedly, when I manually analyzed the RNA-seq peaks using an integrative genomics viewer (IGV), peaks assigned as AYM39_RS00605 and 13560 genes were not mapped to these ORFs but located in the intergenic regions where I could identify short unannotated ORF containing about 50 amino acids (Figure A.5A). The expression levels of these smORFs increased upon acetate stress in the wild type but not in the Δ *watR* strain. In addition, these genes exhibited higher basal expression levels in the ΔwatR strain. Therefore, the smORFs, named *watS1* and *watS5*, were repressed by WatR under normal conditions and activated upon acetate stress (Figure A.5A).

Based on these findings, I reexamined the 22 WatR-binding ChIP-seq peaks and RNA-seq data using the IGV browser, identifying three more WatR-regulated smORFs that had been matched with wrong ORFs or an intergenic region in our original ChIP-Seq analysis (Figure A.5A, Table A.6). The newly identified smORFs, named *watS2*, *watS3*, and *watS4*, also showed similar expression patterns to those of *watS1* and *watS5,* indicating the repression by WatR under normal conditions and activation upon acetate stress (Figure

A.5A). In agreement with the repressor-type regulation, the well-conserved WatR binding consensus sequences overlap with the predicted -35 box regions of these smORFs (Figure A.5B). Except for *watS5*, the polypeptides encoded by *watS1* to *watS4* showed highly homologous amino acid sequences containing a putative transmembrane domain (Figure A.5C).

In agreement with the qRT-PCR results shown in Figure A.4, the expression of genes encoding the RND-type efflux pump, *watPAB* and *watO*, was induced following acetate addition in the wild-type, but not the Δ*watR* strain (Figure A.5D). Promoter analysis revealed that a putative WatR binding site is located between the two expected -35 boxes of bidirectional genes (Figure A.5E).

Table A.6. The list of genes induced in WatR-dependent manner upon acetate treatment in wild-type and Δ *watR* strains

gene number AYM39_RS	function	SmORF	log2(acetate/control)		log2(Fc WT/Fc Δ*watR*)	log2(Fc WT/Fc Δ*watR*) >1	WatR-dependent induction (manual curation)	Binding of WatR
			WT	Δ*watR*				
00605	hypothetical protein		1.126	-0.073	1.199	O		
		watS1	1.911	0.001	1.910		O	O
		watS2	3.232	0.122	3.110		O	O
		watS3	1.703	0.029	1.674		O	O
		watS4	2.279	0.333	1.946		O	O
13560	glutathione S-transferase		2.986	-0.037	3.023	O		
		watS5	4.541	0.170	4.371		O	O
17390	subunit of RND-type efflux pump (watA)		1.426	0.296	1.130	O	O	O
17385	subunit of RND-type efflux pump (watB)		1.218	0.255	0.963		O	O
17395	hypothetical protein (watP)		1.078	0.340	0.738		O	O
08535	hypothetical protein		3.681	-0.197	3.878	O	O	
07495	glutathione S-transferase		1.022	-0.265	1.287	O	O	
08545	ATP-grasp domain-containing protein		1.014	-0.077	1.091	O	O	

Table A.7. The list of genes induced in WatR-independent manner upon acetate treatment in wild-type and Δ*watR* strains

gene number AYM39_RS	function	log2(acetate/control)		log2(Fc WT/Fc Δ*watR*)
		WT	Δ*watR*	
09160	hypothetical protein	3.105	2.553	0.552
17295	MCE family protein	2.829	2.998	-0.169
17305	CHAT domain-containing protein	2.684	2.502	0.182
17300	caspase family protein	2.636	2.791	-0.155
09155	peptidase domain-containing ABC transporter	2.577	2.426	0.151
17290	lysozyme	2.493	2.645	-0.152
11750	hybrid non-ribosomal peptide synthetase/type I polyketide synthase	2.416	2.228	0.188
09145	hypothetical protein	2.358	2.860	-0.502
09135	hypothetical protein	2.354	2.256	0.098
09140	hypothetical protein	2.354	2.599	-0.245
09150	HlyD family efflux transporter periplasmic adaptor subunit	2.346	2.250	0.096
08110	hypothetical protein	2.165	2.244	-0.079
17310	hypothetical protein	1.981	1.833	0.148
09165	cell envelope integrity protein CreD	1.969	1.732	0.237
20175	hypothetical protein	1.849	1.967	-0.118
17320	phage tail sheath family protein	1.835	2.409	-0.574
11805	hypothetical protein	1.8	1.952	-0.152
06155	squalene/phytoene synthase family protein	1.76	1.698	0.062
17285	autotransporter outer membrane beta-barrel domain-containing protein	1.697	1.788	-0.091
18680	sulfate adenylyltransferase	1.675	3.117	-1.442
11745	non-ribosomal peptide synthetase	1.637	1.869	-0.232
20170	hypothetical protein	1.617	1.912	-0.295
04975	diguanylate cyclase	1.556	1.878	-0.322
18685	sulfate adenylyltransferase subunit CysD	1.517	1.896	-0.379
06150	squalene--hopene cyclase	1.501	1.631	-0.130
06145	phosphorylase	1.42	1.639	-0.219
11800	hypothetical protein	1.33	1.585	-0.255
13875	hypothetical protein	1.312	1.156	0.156
11810	hypothetical protein	1.234	1.169	0.065
18675	YchE family NAAT transporter	1.217	2.835	-1.618
23920	hypothetical protein	1.198	0.749	0.449
05670	hypothetical protein	1.165	1.327	-0.162
18670	hypothetical protein	1.162	2.721	-1.559
17085	hypothetical protein	1.127	0.765	0.362
06140	aspartate aminotransferase family protein	1.089	0.945	0.144
20150	HAD-IC family P-type ATPase	1.074	1.200	-0.126
15345	EamA family transporter	1.071	1.520	-0.449
17080	hypothetical protein	1.07	0.929	0.141
12350	ABC transporter ATP-binding protein	1.059	1.532	-0.473
12345	phosphoenolpyruvate carboxylase	1.024	1.645	-0.621
06135	copper resistance protein CopC	1.006	1.009	-0.003

Table A.8. The list of genes repressed upon acetate treatment in wild-type and Δ*watR* strains

gene number AYM39_RS	function	log2(acetate/control)		log2(Fc WT/Fc Δ*watR*)	WatR-dependent repression (RNA-seq)
		WT	Δ*watR*		
18695	NADH dehydrogenase	-2.713	-0.067	-2.646	O
18690	DUF2309 domain-containing protein	-1.345	0.649	-1.994	O
03490	chaperonin GroEL	-2.7	-2.446	-0.254	
17600	RNA-binding protein	-2.634	-3.664	1.030	
22490	ThiF family adenylyltransferase	-2.458	-1.946	-0.512	
15300	molecular chaperone HtpG	-2.423	-2.558	0.135	
03485	co-chaperone GroES	-2.319	-2.443	0.124	
22735	DUF1508 domain-containing protein	-2.169	-1.988	-0.181	
23475	hypothetical protein	-2.124	-1.978	-0.146	
14200	molecular chaperone DnaK	-1.858	-1.669	-0.189	
17605	cold-shock protein	-1.784	-2.422	0.638	
14205	nucleotide exchange factor GrpE	-1.761	-1.676	-0.085	
23170	type II toxin-antitoxin system RelE/ParE family toxin	-1.687	-1.593	-0.094	
14185	molecular chaperone DnaJ	-1.634	-1.554	-0.080	
14195	hypothetical protein	-1.617	-1.662	0.045	
14180	4-hydroxy-tetrahydrodipicolinate reductase	-1.591	-1.406	-0.185	
14190	hypothetical protein	-1.482	-1.667	0.185	
17595	DEAD/DEAH box helicase	-1.44	-2.657	1.217	
15020	ribonuclease HII	-1.248	-0.963	-0.285	
12585	RNA polymerase sigma factor RpoH	-1.152	-0.845	-0.307	
14210	heat-inducible transcriptional repressor HrcA	-1.123	-0.949	-0.174	
13175	ADP-glyceromanno-heptose 6-epimerase	-1.086	-1.231	0.145	
16265	lipopolysaccharide heptosyltransferase I	-1.033	-0.865	-0.168	

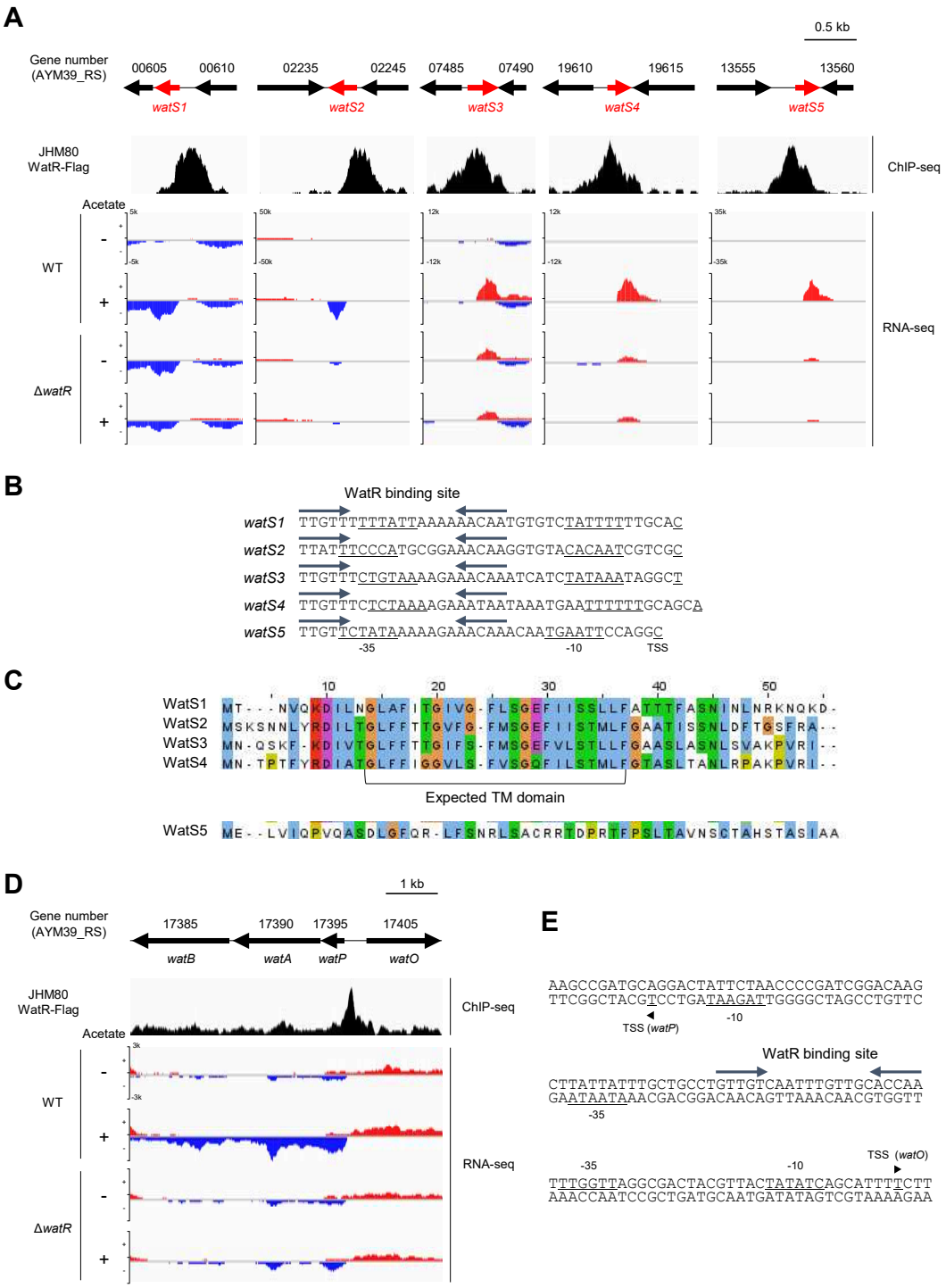

Figure A.5. WatR-dependent regulation of smORF genes upon acetate stress.

Figure A.5. WatR-dependent regulation of smORF genes upon acetate stress.

(A) WatR-dependent regulation of smORFgenes. Locations of five unannotated smORFs (*watS1-watS5*) are aligned with the WatR binding peaks detected via ChIP-Seq and transcript levels detected via RNA-seq using the IGV 2.3.72 program. RNA-seq analysis was performed in the wild-type and Δ*watR* strains with or without acetate treatment.

(B) The promoter sequences of *watS1-watS5* with their expected -35 box, -10 box, and TSS. The conserved WatR binding sites are shown as inverted arrows.

(C) The homology alignment of amino acid sequences of WatS1 to WatS5. A putative transmembrane domain region conserved in WatS1 to WatS4 is indicated.

(D) WatR-dependent regulation of WatABO efflux pump genes. The gene locations were aligned with the WatR binding peaks detected via ChIP-seq analysis and transcript levels detected by RNA-seq.

(E) The promoter sequences of the divergently transcribed *watPAB* and *watO* genes with their expected -35 box, -10 box, and TSS. The putative WatR binding sites are shown as inverted arrows.

A.5.2. WatR binding affinity to both activator and repressor target genes increased under acetate stress

Next, I investigated whether the binding affinity of WatR to its target genes changes under acetate stress. ChIP-qPCR analysis was performed in the wild-type strain overexpressing *watR*-Flag from the *EFTu* promoter (JHM16WF). The WatR target genes were induced at 0.6 g/L acetate in the wild-type strain (Figure A.4B). However, due to the higher expression levels of *watR* in the JHM16WF strain, the amount of acetate had to be increased to 3.0 g/L to observe the induction of *watP* and *watS1* (Figure A.6A). I performed ChIP experiments under the same acetate stress conditions inducing the expression of *watP* and *watS1* (Figure A.6B). WatR exhibited stronger binding to the promoter of repressor-type target *watS1* than to the activator-type target watP. However, irrespective of the regulation type, WatR binding affinity increased upon acetate stress (Figure A.6B).

The activity of LTTRs is typically regulated by conformational changes induced via effector binding to the C-terminal domain. Therefore, acetate itself or another metabolite generated upon acetate stress may act as a ligand regulating WatR activity. In the case of activator-type target genes, such as *watP*, the activator activity of WatR seems to

be enhanced by acetate, which involves an increase in DNA binding (Figure A.6C). In the case of repressor-type target genes, such as *watS1*, WatR acts as a repressor under normal conditions (Figure A.6D). Upon acetate stress, instead of WatR derepressing target genes by being released from the promoter, WatR seems to change to an activator, possibly by shifting binding sites in the promoter, thus exposing the RNA polymerase binding site (Figure A.6D). This hypothesis is supported by the RNA-seq data showing higher acetate-induced mRNA levels of smORF genes in the wild-type compared to the *watR* deletion mutant (Figure A.5A).

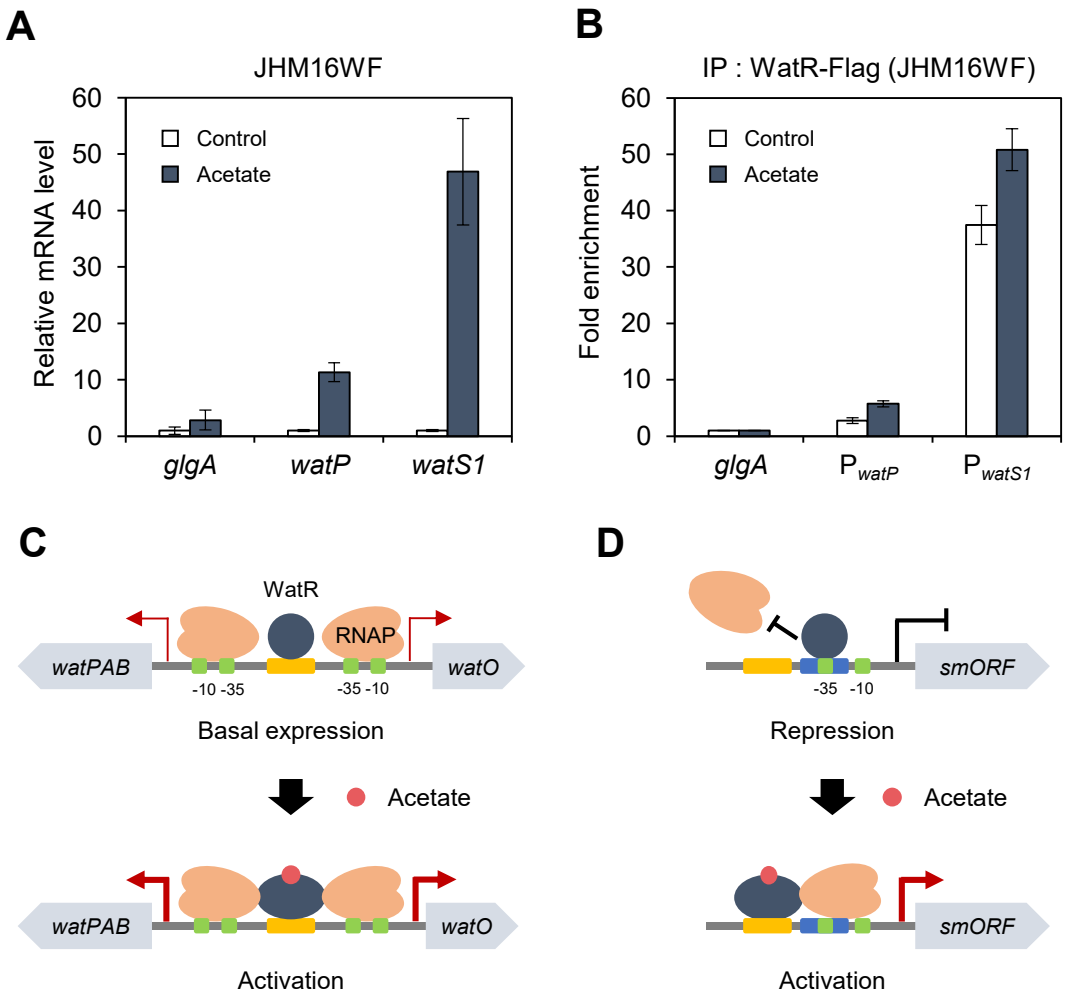

Figure A.6. Changes in DNA binding affinity of WatR upon acetate stress.

Figure A.6. Changes in DNA binding affinity of WatR upon acetate stress.

(A) Induction of *watP* and *watS1* gene expression by acetate. The JHM16WF strain expressing *watR*-Flag from the P_{EFTu} promoter was grown in NMS medium with 20% (v/v) methane and treated with 3.0 g/L of acetate for 10 min. The mRNA expression levels were detected by qRT-PCR and indicated as values relative to those of untreated control.

(B) Changes in WatR DNA binding upon acetate stress. The JHM16WF strain was grown in NMS medium with 20% (v/v) methane and treated with 3.0 g/L of acetate for 10 min. ChIP analysis was performed with anti-Flag antibody and WatR binding to the promoters was detected by qPCR. Each value represents the average ± SD of the relative fold enrichment, normalized to a negative control (*glgA* ORF).

(C) Model for the WatR-dependent transcriptional regulation of *watPAB* and *watO* genes. The WatR binding sites does not overlap with the RNA binding sites, enabling basal transcription. Upon acetate stress, WatR activates transcription, which involves increasing DNA binding affinity.

(D) Model for the WatR-dependent transcriptional regulation of smORFs. The WatR binding sites overlap with the RNA binding sites, repressing basal transcription. Upon acetate stress, WatR activates transcription possibly by shifting the binding site to expose the RNA polymerase binding site.

A.5.3. smORFs are responsible for acetate tolerance via efflux pump regulation

Since smORFs were identified as major targets regulated by WatR upon acetate stress, I next investigated the role of these smORFs in acetate tolerance. When the *watS1* and *watS5* genes were overexpressed under the control of the strong *mxaF* promoter, acetate tolerance increased compared with the control strain (Figure A.7A). Because *watS* genes were overexpressed by replacing the *fliE* ORF, the Δ*fliE* strain was used as a control. Overexpression of *watS1* was more effective than that of *watS5* in increasing acetate tolerance (Figure A.7A). Consistent with the high homology among WatS1 to WatS4 (Figure A.5C), strains overexpressing *watS2*, *watS3*, and *watS4* also showed higher acetate tolerance than that of *watS5*-overexpressing strain (Figure A.8). However, overexpression of *watS1* did not improve tolerance against lactate and propionate, suggesting that WatS1 function is specific to acetate (Figure A.7A). In line with the positive effect of *watS1* overexpression on acetate tolerance, deletion of *watS1* decreased tolerance to acetate only but not propionate and lactate (Figure A.7B). Therefore, I further investigated acetate tolerance mechanisms induced by *watS1*.

I first confirmed that the *watS1* encodes a protein. The

strain expressing the T7-tagged *watS1* (JHM161T) showed a band of the expected size (~7 kDa) in western blotting analysis (Data not shown). Therefore, *watS1* is expected to have a role as a smORF-encoded polypeptide (SEP). One of the known roles of SEPs is the regulation of membrane transporters. For example, AcrZ, a SEP in E. coli, binds to the AcrB subunit of an RND-type efflux pump, inducing conformational changes within the drug-binding pocket, which in turn affect the selectivity for transporting antibiotics (31). Since WatR activates the expression of the WatABO pump involved in organic acid tolerance, I hypothesized that the WatR-regulated SEPs might regulate efflux pump specificity. The presence of a transmembrane domain in WatS1 to WatS4 also supports our hypothesis. The deletion mutant for the efflux pump genes (Δ*watABO*: Δ*watPAB* and Δ*watO*) exhibited higher acetate sensitivity than the Δ*watS1* strain (Figure A.7C). However, additional deletion of *watS1* in the efflux pump deletion mutant did not further increase acetate sensitivity (Figure A.7C), suggesting that WatS1 and the WatABO pump might work in the same pathway. In addition, overexpression of the *watS1* gene increased acetate tolerance in the wild type (Figure A.7A) but could not rescue the acetate sensitivity of the Δ*watABO* strain (Figure A.7D), further supporting the hypothesis that

WatS1 might function through the WatABO pump.

Taken together, I propose a working model of WatS1-4 controlling the specificity of the WatABO efflux pump (Figure A.9). The efflux pump functions as a general transporter for several weak organic acids, including formate, acetate, lactate, and propionate under normal conditions. In response to acetate stress, the expression of the efflux pump genes and *watS1-4* is activated in a WatR-dependent manner. The SEPs WatS1-4 may then interact with the WatABO pump, shifting specificity toward acetate. This regulatory mechanism enables efficient cellular protection against acetate by switching the general weak organic acid efflux pump to an acetate-specific efflux pump in the presence of acetate.

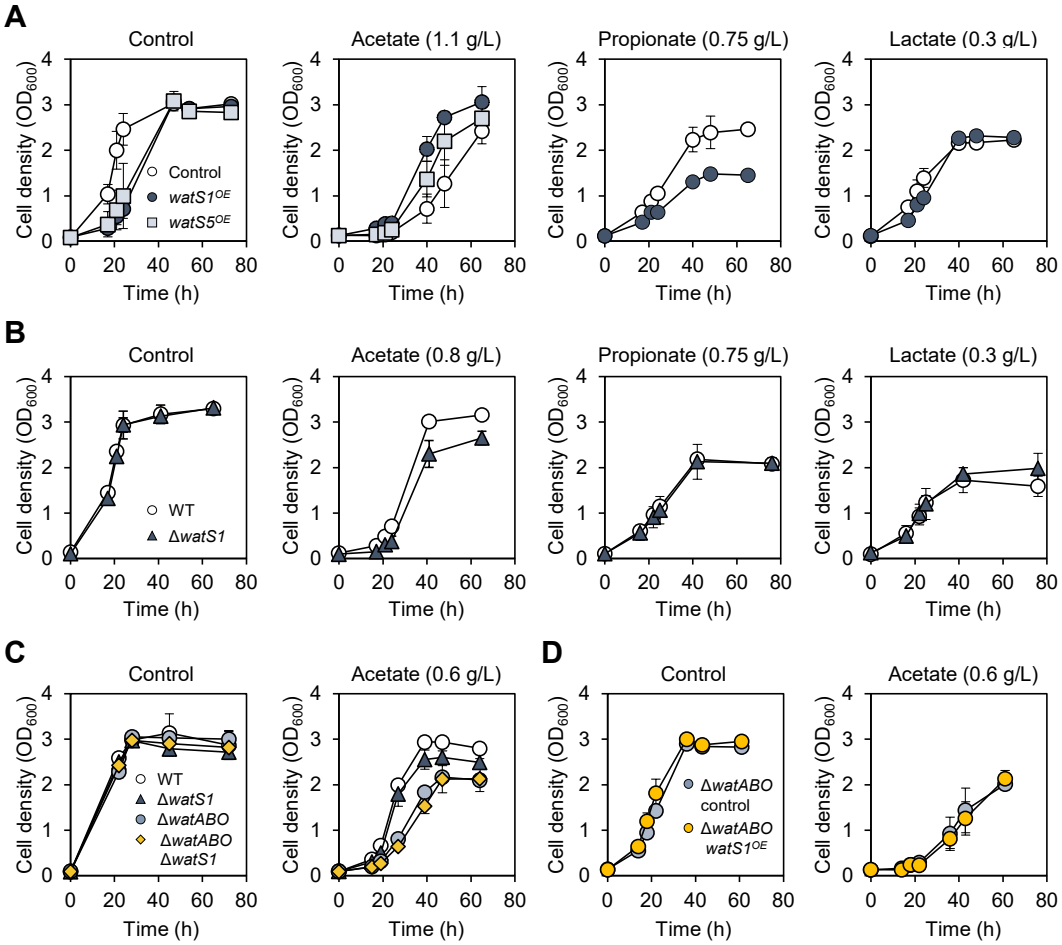

Figure A.7. WatS1 controls acetate tolerance only in the presence of WatABO pump.

Figure A.7. WatS1 controls acetate tolerance only in the presence of WatABO pump.

(A) Effect of overexpressing the *watS1* and *watS5* genes on acid tolerance. The Δ*fliE* control strain (JHM16) and strains replacing the *fliE* gene with *watS1* or *watS5* overexpression cassette (JHM161 and JHM165) were grown in NMS media with 20% (v/v) methane without or with weak organic acids as indicated. Each value represents the average ± standard deviations from two independent experiments.

(B) Effect of deleting the *watS1* gene on acid tolerance. The wild-type and watS1 deletion (JHM17) strains were grown in NMS media with 20% (v/v) methane without or with weak organic acids as indicated. Each value represents the average ± standard deviations from two independent experiments.

(C) Effect of deleting the *watS1* and WatABO pump genes. The wild-type, Δ*watS1*, Δ*watABO* (JHM18), and Δ*watS1* Δ*watABO* (JHM182) strains were grown in NMS media with 20% (v/v) methane without or with 0.6 g/L acetate. Each value represents the average ± standard deviations from two independent experiments.

(D) Effect of overexpressing the watS1 gene without WatABO pump genes. The control Δ*watABO* strain with *fliE* deletion (JHM181) and Δ*watABO* strain overexpressing *watS1* (JHM183) were grown in NMS media with 20% (v/v) methane without or with 0.6 g/L acetate. Each value represents the average ± standard deviations from two independent experiments.

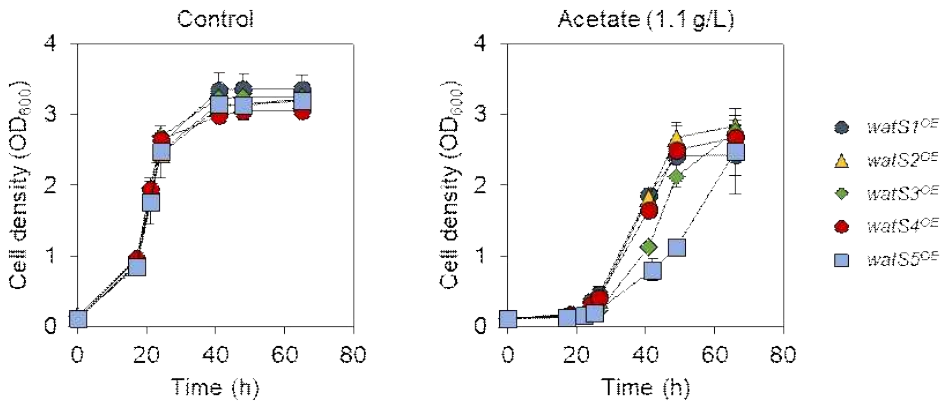

Figure A.8. Increase in acetate tolerance by overexpression of WatS1 to WatS5

All strains were grown in NMS media with 20% (v/v) methane with or without 1.1 g/L acetate. Each value represents the average ± standard deviations from two independent experiments.

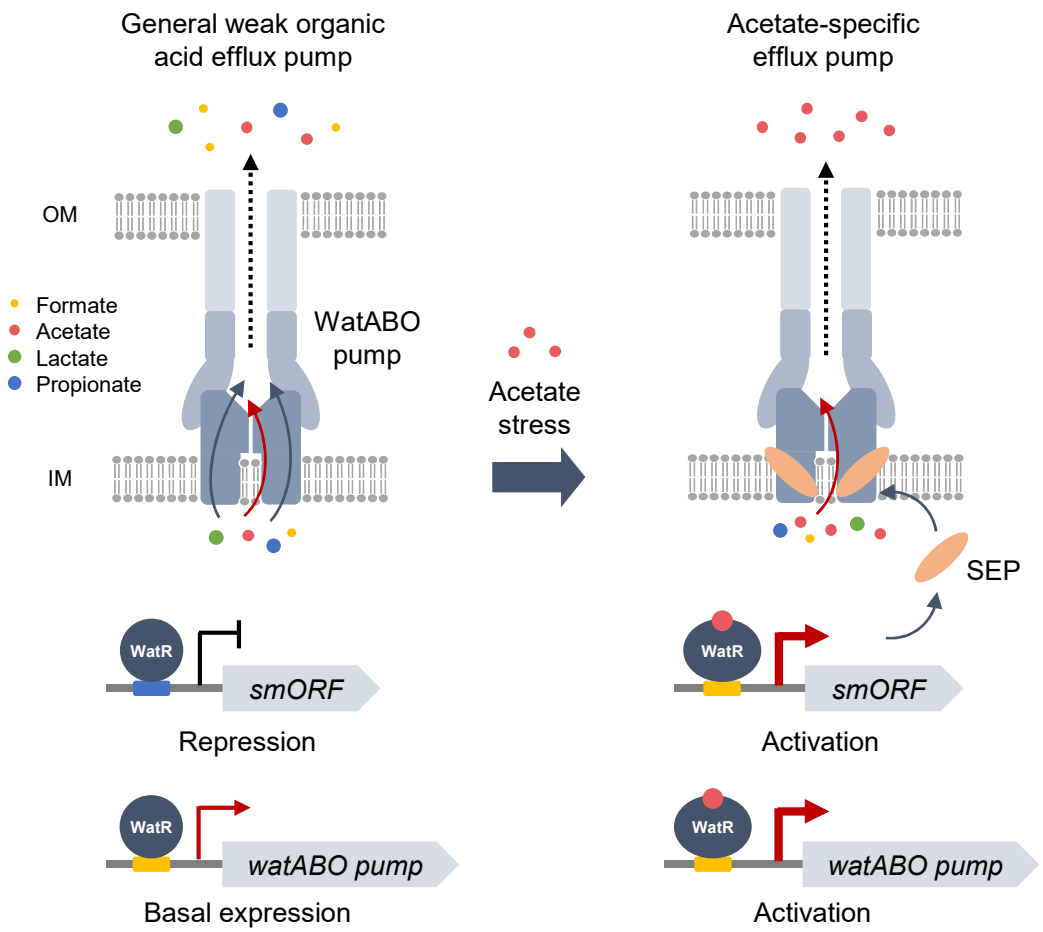

Figure A.9. Model for SEP-dependent regulation of the WatABO pump upon acetate stress.

Under normal conditions, the WatABO pump extrude a wide range of weak organic acids including acetate, formate, lactate, and propionate. Upon acetate stress, activated WatR induces transcription of WatABO pump and *watS* smORF genes. WatS SEP binds to the WatABO pump, increasing the specificity toward acetate for efficient removal of acetate from cells.

A.6. Discussion

A.6.1. Regulation of WatR activity under acetate stress

Understanding the acid tolerance mechanisms is critical for improving microbial production of useful organic acids. There is growing interest in utilizing methanotrophs for the bioconversion of methane into value-added chemicals, but little is known about their acid stress responses. In this appendix study, I elucidated the role of WatR, an LTTR, as a regulator of weak organic acid stress responses in *Methylomonas* sp. DH-1. Through the analysis of genome-wide binding targets of WatR and WatR-dependent transcriptional regulation, I identified that WatR functions both as a transcriptional repressor and an activator, with its activity regulated by acetate.

LTTR is among the largest families of bacterial regulators with diverse functions including stress response, biosynthesis, and biodegradation in response to various effector molecules binding to the C-terminal domain (16). In agreement with the classical regulatory model of LTTRs, WatR autoregulates its expression and activates target gene expression in response to acetate. WatR represses the basal expression of certain target genes, such as *watR*, *gltA*, and *smORFs* (*watS1*~*watS5*). Within their promoters, the WatR binding site overlaps with the RNA polymerase binding site; so,

access of the RNA polymerase is inhibited via WatR binding. In contrast, WatR overexpression activates the basal transcription of divergently transcribed genes encoding an RND-type efflux pump (*watPAB* operon and *watO*), where the WatR binding site does not overlap with the RNA polymerase binding site. In both cases, acetate treatment increased WatR binding to the target promoters and increased transcription. Therefore, conformational changes in WatR upon acetate stress may convert it to a transcriptional activator. In the case of WatR-repressed genes, rather than being derepressed via release of WatR from the promoter, acetate-dependent activation may shift the WatR binding site wherein WatR can activate transcription instead of preventing RNA polymerase binding. Such an effector-dependent transition from a repressor to an activator through changes in the binding sites has also been reported in other LTTRs (32). Acetate is known to directly regulate AlsR, an LTTR modulating acetoin production in Bacillus subtilis (33). In Saccharomyces cerevisiae, weak organic acids, including acetate, can bind directly bind to the Haa1 and War1 transcription factors involved in the cellular response against weak organic acids (34). Therefore, acetate may serve as a direct effector for WatR, but further studies are necessary to identify the specific effector molecules for WatR

regulation.

A.6.2. Roles of WatS SEPs as regulators of increasing acetate specificity of the WatABO efflux pump

Cells have evolved defense mechanisms against weak acids generated during normal cell growth or present in the environment. The major cellular defense mechanisms against weak acid stress include the export of excess cytosolic protons and acid anions through membrane transporters, restricted diffusion of weak acids by remodeling the cell wall and plasma membrane, and the metabolic conversion of weak acids (35-37). In this study, I demonstrate that WatR can affect tolerance against a wide range of weak organic acids, having a more specific role in response to acetate stress. I propose a novel defense mechanism against acetate stress: the SEP-mediated regulation of efflux pump specificity.

The WatR-activated WatABO pump was identified to work as an efflux pump for general weak organic acids, including formate, acetate, lactate, and propionate. Upon acetate stress, WatS SEPs may interact with the WatABO pump, changing its specificity toward acetate, leading to more efficient removal of acetate out of cells. The expression of *watS1* to *watS5* is repressed by WatR under normal conditions and

activated only upon acetate stress. In contrast, cells express basal or upregulated WatABO pump genes under normal and acetate conditions, respectively. This hypothesis is based on the well-established example of AcrB regulation by AcrZ, an SEP, in E. coli. AcrB is an inner membrane-binding component of an RND-type efflux pump exporting several antibiotics, organic solvents, and detergents (38-40). Genetic and cryo-EM-based structural studies revealed that AcrZ binding to AcrB leads to conformational changes in the drug-binding pocket of AcrB, altering specificity toward certain antibiotics, such as chloramphenicol (31, 38). In agreement with this working model, overexpression of *watS* SmORFs increased tolerance against acetate, but not other weak acids, only in the presence of the WatABO pump. I tried to detect a direct interaction between the WatS1 polypeptide and WatB inner membrane subunit through co-immunoprecipitation and split GFP assay. However, I was unsuccessful due to technical difficulties in tagging WatB without affecting cell growth.

The proposed role of WatR is in cellular defense against acetate stress, which contradicts the observed acetate-resistant phenotype of the *watR* deletion mutant. This inconsistency might be related to our experimental conditions of acetate stress. Since acetate was added at the beginning of the culture,

the basal expression levels of defense genes might affect the tolerance phenotypes I observed. Although WatR-dependent gene expression may play an important role in the cellular adaptation to dynamic changes in acetate levels, the high basal expression of *watS* genes and other WatR-repressed genes in the inoculum may be beneficial for survival of the Δ*watR* strain under our acid stress conditions. As the *watR* deletion decreased tolerance against formate, lactate, and propionate, tolerance against these weak acids seems more dependent on WatR target genes, such as WatABO pump genes, which are activated but not repressed by WatR. This requires further studies to understand the role of other WatR target genes in tolerance against different weak organic acids.

The identified role of WatR is similar to that of YdcI, an LTTR found in a wide range of gram-negative bacteria (41). Although WatR and YdcI have low sequence homology, YdcI is known to be involved in the acid stress response and pH homeostasis in Salmonella enterica serovar Typhimurium and E. coli (42, 43). In addition, both WatR and *E. coli* YdcI repress the expression of citrate synthase, and deletion of *watR* and *ydcI* both increased acetate tolerance (41). Although I could not observe the contribution of the citrate synthase gene (*gltA1*) in acetate tolerance under our experimental conditions, I

cannot rule out the possibility that the WatR-dependent activation of *gltA1* may contribute to alleviating acetate stress through the upregulation of acetate flux toward to the TCA cycle.

A.6.3. Emerging roles of smORFs in stress response

Comprehensive analysis of the ChIP-seq and RNA-seq data revealed five unannotated smORFs, *watS1* to *watS5*, as WatR target genes repressed under normal conditions and activated upon acetate stress. The smORFs are usually ignored in gene annotation programs, which use cut-off sizes of 50 and 100 amino acids for prokaryotes and eukaryotes, respectively (44) However, recent advances in genomics, proteomics, and bioinformatics have enabled the discovery of previously unannotated smORFs from bacteria to humans (45-49). smORFs and SEPs fall into two main functional categories: SEPs with their own function and upstream ORFs (uORFs) regulating the translation of a downstream gene in eukaryotes. In both cases, growing evidence supports the prevailing role of smORFs in stress responses. Translation of eukaryotic uORF prevents scanning and/or re-initiation at the downstream ORF, which can regulate stress-dependent translation of the downstream gene (50). In bacteria, the expression of smORFs is induced

under various stress conditions, including heat shock, cold shock, oxidative stress, low pH, and different nutrient conditions (51-53). To date, only a few functional SEPs have been characterized but these commonly regulate biological functions by modulating the activity or stability of other proteins or protein complexes (54). In addition, various functional SEPs have been identified as membrane proteins. Recent metagenomic analysis of the human microbiome revealed approximately 4,000 SEPs, about 30% of which are predicted to be secreted or membrane-bound (55). Like the proposed role of WatS in regulating the WatABO pump, several SEPs with one transmembrane domain are known to regulate membrane transporters in response to environmental signals, including nutrients and metal ions (56, 57). The SEP-dependent transporter regulation has also been reported in mammals. DWORF, an SEP localized within the sarcoplasmic reticulum membrane, interacts with the Ca2+-ATPase SERCA, thus increasing Ca2+ uptake (58).

Although the RND-type efflux pump is well known for transporting a broad range of chemicals, the SEP-dependent regulation of substrate specificity may provide an efficient and rapid cellular adaptation in response to environmental stress. The watP gene of unknown function in the *watPAB* operon is

also predicted to encode a relatively short protein of 89 amino acids, which has two transmembrane domains. Therefore, WatP might also act as a regulator or subunit of the WatABO pump. My study highlights the important roles of SEPs in fine-tuning stress responses by modulating specific interacting proteins. The *Methylomonas* sp. DH-1 genome contains at least 10 RND pump genes. This research is focused on acetate-responsive smORFs, but it would be interesting to determine whether SEPs regulate other RND pumps or transporters in response to different stress conditions. Understanding the SEP-dependent efflux mechanisms of various weak organic acids can contribute to the production of diverse organic acids in methanotrophs and other bacterial hosts that may share the same regulatory strategy of organic acid efflux.

A.7. Bibilography of appendix

1. Milkov, A.V., et al., Using global isotopic data to constrain the role of shale gas production in recent increases in atmospheric methane. Sci Rep, 2020. 10(1): p. 4199.
2. Schiermeier, Q., Global methane levels soar to record high. Nature, 2020.
3. Hur, D.H., et al., Selective bio-oxidation of propane to acetone using methane-oxidizing Methylomonas sp. DH-1. J Ind Microbiol Biotechnol, 2017. 44(7): p. 1097-1105.
4. Lee, J.K., et al., Efficient production of d-lactate from methane in a lactate-tolerant strain of Methylomonas sp. DH-1 generated by adaptive laboratory evolution.Biotechnol Biofuels, 2019. 12: p. 234.
5. Nguyen, T.T., et al., Bioconversion of methane to cadaverine and lysine using an engineered type II methanotroph, Methylosinus trichosporium OB3b. Green Chemistry, 2020. 22(22): p. 7803-7811.
6. Nguyen, D.T.N., et al., Metabolic engineering of the type I methanotroph

Methylomonas sp. DH-1 for production of succinate from methane. Metab Eng, 2019. 54: p. 170-179.
7. Pham, D.N., et al., Development of an engineered methanotroph-based microbial platform for biocatalytic conversion of methane to phytohormone for sustainable agriculture.Chemical Engineering Journal, 2022. 429: p. 132522.
8. Semrau, J.D., et al., Methanobactin and MmoD work in concert to act as the 'copper-switch' in methanotrophs. Environ Microbiol, 2013. 15(11): p. 3077-86.
9. Mustakhimov, II, et al., Identification and characterization of EctR1, a new transcriptional regulator of the ectoine biosynthesis genes in the halotolerant methanotroph Methylomicrobium alcaliphilum 20Z. J Bacteriol, 2010. 192(2): p. 410-7.
10. Sauer, M., et al., Microbial production of organic acids: expanding the markets. Trends Biotechnol, 2008. 26(2): p. 100-8.
11. Guan, N. and L. Liu, Microbial response to acid stress: mechanisms and applications. Appl Microbiol Biotechnol, 2020. 104(1): p. 51-65.
12. Mira, N.P. and M.C. Teixeira, Microbial mechanisms of tolerance to weak acid stress. Front Microbiol, 2013. 4: p. 416.
13. Carpenter, C.E. and J.R. Broadbent, External concentration of organic acid anions and pH: key independent variables for studying how organic acids inhibit growth of bacteria in mildly acidic foods.J Food Sci, 2009. 74(1): p. R12-5.
14. Bearson, S., B. Bearson, and J.W. Foster, Acid stress responses in enterobacteria. FEMS Microbiol Lett, 1997. 147(2): p. 173-80.
15. Hirshfield, I.N., S. Terzulli, and C. O'Byrne, Weak organic acids: a panoply of effects on bacteria. Sci Prog, 2003. 86(Pt 4): p. 245-69.
16. Maddocks, S.E. and P.C.F. Oyston, Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology (Reading), 2008. 154(Pt 12): p. 3609-3623.
17. Knapp, G.S. and J.C. Hu, Specificity of the E. coli LysR-type transcriptional regulators. PLoS One, 2010. 5(12): p. e15189.
18. Marx, C.J. and M.E. Lidstrom, Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. Biotechniques, 2002. 33(5): p. 1062-7.
19. Henard, C.A., et al., Bioconversion of methane to lactate by an obligate methanotrophic bacterium. Sci Rep, 2016. 6: p. 21585.
20. Bolger, A.M., M. Lohse, and B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 2014. 30(15): p. 2114-2120.
21. Kim, D., et al., Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol, 2019. 37(8): p. 907-915.
22. Liao, Y., G.K. Smyth, and W. Shi, The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res, 2013. 41(10): p. e108.
23. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 2014. 15(12): p. 550.

24. Grainger, D.C., et al., Genomic studies with Escherichia coli MelR protein: applications of chromatin immunoprecipitation and microarrays. J Bacteriol, 2004. 186(20): p. 6938-43.
25. Heinz, S., et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell, 2010. 38(4): p. 576-89.
26. Machanick, P. and T.L. Bailey, MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics, 2011. 27(12): p. 1696-7.
27. De Maeseneire, S.L., et al., Metabolic characterisation of E. coli citrate synthase and phosphoenolpyruvate carboxylase mutants in aerobic cultures. Biotechnol Lett, 2006. 28(23): p. 1945-53.
28. Zhu, N., et al., Engineering of acetate recycling and citrate synthase to improve aerobic succinate production in Corynebacterium glutamicum. PLoS One, 2013. 8(4): p. e60659.
29. Kim, E.H., et al., Switch or funnel: how RND-type transport systems control periplasmic metal homeostasis. J Bacteriol, 2011. 193(10): p. 2381-7.
30. Nikaido, H. and Y. Takatsuka, Mechanisms of RND multidrug efflux pumps. Biochim Biophys Acta, 2009. 1794(5): p. 769-81.
31. Du, D., et al., Interactions of a Bacterial RND Transporter with a Transmembrane Small Protein in a Lipid Environment. Structure, 2020. 28(6): p. 625-634 e6.
32. Oliver, P., et al., Molecular and structural considerations of TF-DNA binding for the generation of biologically meaningful and accurate phylogenetic footprinting analysis: the LysR-type transcriptional regulator family as a study model. BMC Genomics, 2016. 17: p. 686.
33. Fradrich, C., et al., The transcription factor AlsR binds and regulates the promoter of the alsSD operon responsible for acetoin formation in Bacillus subtilis. J Bacteriol, 2012. 194(5): p. 1100-12.
34. Kim, M.S., et al., Activation of Haa1 and War1 transcription factors by differential binding of weak acid anions in Saccharomyces cerevisiae. Nucleic Acids Res, 2019. 47(3): p. 1211-1224.
35. Holyoak, C.D., et al., The Saccharomyces cerevisiaeweak-acid-inducible ABC transporter Pdr12 transports fluorescein and preservative anions from the cytosol by an energy-dependent mechanism. J Bacteriol, 1999. 181(15): p. 4644-52.
36. Ter Beek, A., et al., Transcriptome analysis of sorbic acid-stressed Bacillus subtilis reveals a nutrient limitation response and indicates plasma membrane remodeling. J Bacteriol, 2008. 190(5): p. 1751-61.
37. De Mets, F., L. Van Melderen, and S. Gottesman, Regulation of acetate metabolism and coordination with the TCA cycle via a processed small RNA. Proc Natl Acad Sci U S A, 2019. 116(3): p. 1043-1052.
38. Hobbs, E.C., et al., Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. Proc Natl Acad Sci U S A, 2012. 109(41): p. 16696-701.
39. Tsukagoshi, N. and R. Aono, Entry into and release of solvents by

Escherichia coli in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump. Journal of Bacteriology, 2000. 182(17): p. 4803-4810.
40. Takatsuka, Y., C. Chen, and H. Nikaido, Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of Escherichia coli.Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(15): p. 6559-6565.
41. Gao, Y., et al., Systematic discovery of uncharacterized transcription factors in Escherichia coli K-12 MG1655. Nucleic Acids Res, 2018. 46(20): p. 10682-10696.
42. Jennings, M.E., et al., Characterization of the Salmonella enterica serovar Typhimurium ydcI gene, which encodes a conserved DNA binding protein required for full acid stress resistance. J Bacteriol, 2011. 193(9): p. 2208-17.
43. Solomon, L., et al., Bacterial genus-specific tolerance for YdcI expression. Curr Microbiol, 2014. 69(5): p. 640-8.
44. Storz, G., Y.I. Wolf, and K.S. Ramamurthi, Small proteins can no longer be ignored. Annual review of biochemistry, 2014. 83: p. 753.
45. Jaffe, J.D., H.C. Berg, and G.M. Church, Proteogenomic mapping as a complementary method to perform genome annotation. Proteomics, 2004. 4(1): p. 59-77.
46. Martinez, T.F., et al., Accurate annotation of human protein-coding small open reading frames. Nat Chem Biol, 2020. 16(4): p. 458-468.
47. Bartholomaus, A., et al., smORFer: a modular algorithm to detect small ORFs in prokaryotes. Nucleic Acids Res, 2021. 49(15): p. e89.
48. Stringer, A., et al., Identification of novel translated small ORFs in Escherichia coli using complementary ribosome profiling approaches. J Bacteriol, 2021: p. JB0035221.
49. Smith, J.E., et al., Translation of small open reading frames within unannotated RNA transcripts in Saccharomyces cerevisiae. Cell Rep, 2014. 7(6): p. 1858-66.
50. Zhang, H., Y. Wang, and J. Lu, Function and Evolution of Upstream ORFs in Eukaryotes. Trends Biochem Sci, 2019. 44(9): p. 782-794.
51. Hemm, M.R., et al., Small stress response proteins in Escherichia coli: proteins missed by classical proteomic studies. J Bacteriol, 2010. 192(1): p. 46-58.
52. Zhang, Y., et al., A Stress Response that Monitors and Regulates mRNA Structure Is Central to Cold Shock Adaptation. Mol Cell, 2018. 70(2): p. 274-286 e7.
53. Wadler, C.S. and C.K. Vanderpool, A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. Proc Natl Acad Sci U S A, 2007. 104(51): p. 20454-9.
54. Saghatelian, A. and J.P. Couso, Discovery and characterization of smORF-encoded bioactive polypeptides. Nat Chem Biol, 2015. 11(12): p. 909-16.
55. Sberro, H., et al., Large-Scale Analyses of Human Microbiomes Reveal

Thousands of Small, Novel Genes. Cell, 2019. 178(5): p. 1245-1259 e14.
56. Lippa, A.M. and M. Goulian, Feedback inhibition in the PhoQ/PhoP signaling system by a membrane peptide. PLoS Genet, 2009. 5(12): p. e1000788.
57. Lloyd, C.R., et al., The Small Protein SgrT Controls Transport Activity of the Glucose-Specific Phosphotransferase System. J Bacteriol, 2017. 199(11).
58. Nelson, B.R., et al., A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science, 2016. 351(6270): p. 271-5.

Bibliography of main text

1. De Deken RH. The Crabtree effect: a regulatory system in yeast. J Gen Microbiol. 1966;44(2):149-56.
2. Pfeiffer T, Morley A. An evolutionary perspective on the Crabtree effect. Front Mol Biosci. 2014;1:17.
3. Hill GE. Cellular respiration: the nexus of stress, condition, and ornamentation. Integr Comp Biol. 2014;54(4):645-57.
4. Dashko S, Zhou N, Compagno C, Piskur J. Why, when, and how did yeast evolve alcoholic fermentation? FEMS Yeast Res. 2014;14(6):826-32.
5. de Alteriis E, Carteni F, Parascandola P, Serpa J, Mazzoleni S. Revisiting the Crabtree/Warburg effect in a dynamic perspective: a fitness advantage against sugar-induced cell death. Cell Cycle. 2018;17(6):688-701.
6. Van Urk H, Voll WS, Scheffers WA, Van Dijken JP. Transient-state analysis of metabolic fluxes in crabtree-positive and crabtree-negative yeasts. Appl Environ Microbiol. 1990;56(1):281-7.
7. Bender T, Pena G, Martinou JC. Regulation of mitochondrial pyruvate uptake by alternative pyruvate carrier complexes. EMBO J. 2015;34(7):911-24.
8. de Smidt O, du Preez JC, Albertyn J. The alcohol dehydrogenases of Saccharomyces cerevisiae: a comprehensive review. FEMS Yeast Res. 2008;8(7):967-78.
9. de Smidt O, du Preez JC, Albertyn J. Molecular and physiological aspects of alcohol dehydrogenases in the ethanol metabolism of Saccharomyces cerevisiae. FEMS Yeast Res. 2012;12(1):33-47.
10. Remize F, Andrieu E, Dequin S. Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae: role of the cytosolic Mg(2+) and mitochondrial K(+) acetaldehyde dehydrogenases Ald6p and Ald4p in acetate formation during alcoholic fermentation. Appl Environ Microbiol. 2000;66(8):3151-9.
11. Shiba Y, Paradise EM, Kirby J, Ro DK, Keasling JD. Engineering of the pyruvate dehydrogenase bypass in Saccharomyces cerevisiae for high-level production of isoprenoids. Metab Eng. 2007;9(2):160-8.
12. Overkamp KM, Bakker BM, Kotter P, Luttik MA, Van Dijken JP, Pronk JT. Metabolic engineering of glycerol production in Saccharomyces cerevisiae. Appl Environ Microbiol. 2002;68(6):2814-21.
13. Nevoigt E, Stahl U. Osmoregulation and glycerol metabolism in the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev. 1997;21(3):231-41.
14. Alarcon DA, Nandi M, Carpena X, Fita I, Loewen PC. Structure of glycerol-3-phosphate dehydrogenase (GPD1) from Saccharomyces cerevisiae at 2.45 A resolution. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2012;68(Pt 11):1279-83.
15. Pahlman AK, Granath K, Ansell R, Hohmann S, Adler L. The yeast glycerol 3-phosphatases Gpp1p and Gpp2p are required for glycerol

biosynthesis and differentially involved in the cellular responses to osmotic, anaerobic, and oxidative stress. J Biol Chem. 2001;276(5):3555-63.
16. Athenstaedt K, Weys S, Paltauf F, Daum G. Redundant systems of phosphatidic acid biosynthesis via acylation of glycerol-3-phosphate or dihydroxyacetone phosphate in the yeast Saccharomyces cerevisiae. J Bacteriol. 1999;181(5):1458-63.
17. Galdieri L, Mehrotra S, Yu S, Vancura A. Transcriptional regulation in yeast during diauxic shift and stationary phase. OMICS. 2010;14(6):629-38.
18. Geistlinger L, Csaba G, Dirmeier S, Kuffner R, Zimmer R. A comprehensive gene regulatory network for the diauxic shift in Saccharomyces cerevisiae. Nucleic Acids Res. 2013;41(18):8452-63.
19. Walther K, Schuller HJ. Adr1 and Cat8 synergistically activate the glucose-regulated alcohol dehydrogenase gene ADH2 of the yeast Saccharomyces cerevisiae. Microbiology (Reading). 2001;147(Pt 8):2037-44.
20. Yamashita T, Nakamaru-Ogiso E, Miyoshi H, Matsuno-Yagi A, Yagi T. Roles of bound quinone in the single subunit NADH-quinone oxidoreductase (Ndi1) from Saccharomyces cerevisiae. J Biol Chem. 2007;282(9):6012-20.
21. Luttik MA, Overkamp KM, Kotter P, de Vries S, van Dijken JP, Pronk JT. The Saccharomyces cerevisiae NDE1 and NDE2 genes encode separate mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH. J Biol Chem. 1998;273(38):24529-34.
22. Rigoulet M, Aguilaniu H, Averet N, Bunoust O, Camougrand N, Grandier-Vazeille X, et al. Organization and regulation of the cytosolic NADH metabolism in the yeast Saccharomyces cerevisiae. Mol Cell Biochem. 2004;256-257(1-2):73-81.
23. Borst P. The malate-aspartate shuttle (Borst cycle): How it started and developed into a major metabolic pathway. IUBMB Life. 2020;72(11):2241-59.
24. Maestre O, Garcia-Martinez T, Peinado RA, Mauricio JC. Effects of ADH2 overexpression in Saccharomyces bayanus during alcoholic fermentation. Appl Environ Microbiol. 2008;74(3):702-7.
25. Larsson C, Pahlman IL, Ansell R, Rigoulet M, Adler L, Gustafsson L. The importance of the glycerol 3-phosphate shuttle during aerobic growth of Saccharomyces cerevisiae. Yeast. 1998;14(4):347-57.
26. Jung JY, Kim TY, Ng CY, Oh MK. Characterization of GCY1 in Saccharomyces cerevisiae by metabolic profiling. J Appl Microbiol. 2012;113(6):1468-78.
27. Pavlik P, Simon M, Schuster T, Ruis H. The glycerol kinase (GUT1) gene of Saccharomyces cerevisiae: cloning and characterization. Curr Genet. 1993;24(1-2):21-5.
28. Ronnow B, Kielland-Brandt MC. GUT2, a gene for mitochondrial glycerol 3-phosphate dehydrogenase of Saccharomyces cerevisiae. Yeast. 1993;9(10):1121-30.
29. Pahlman IL, Larsson C, Averet N, Bunoust O, Boubekeur S, Gustafsson L, et al. Kinetic regulation of the mitochondrial glycerol-3-phosphate dehydrogenase by the external NADH dehydrogenase in Saccharomyces

cerevisiae. J Biol Chem. 2002;277(31):27991-5.
30. Boumans H, Grivell LA, Berden JA. The respiratory chain in yeast behaves as a single functional unit. J Biol Chem. 1998;273(9):4872-7.
31. Mileykovskaya E, Penczek PA, Fang J, Mallampalli VK, Sparagna GC, Dowhan W. Arrangement of the respiratory chain complexes in Saccharomyces cerevisiae supercomplex III2IV2 revealed by single particle cryo-electron microscopy. J Biol Chem. 2012;287(27):23095-103.
32. Heinemeyer J, Braun HP, Boekema EJ, Kouril R. A structural model of the cytochrome C reductase/oxidase supercomplex from yeast mitochondria. J Biol Chem. 2007;282(16):12240-8.
33. Kim JH, Roy A, Jouandot D, 2nd, Cho KH. The glucose signaling network in yeast. Biochim Biophys Acta. 2013;1830(11):5204-10.
34. Brondijk TH, Konings WN, Poolman B. Regulation of maltose transport in Saccharomyces cerevisiae. Arch Microbiol. 2001;176(1-2):96-105.
35. Marques WL, Raghavendran V, Stambuk BU, Gombert AK. Sucrose and Saccharomyces cerevisiae: a relationship most sweet. FEMS Yeast Res. 2016;16(1):fov107.
36. 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(11):6465-72.
37. Weinhandl K, Winkler M, Glieder A, Camattari A. Carbon source dependent promoters in yeasts. Microb Cell Fact. 2014;13:5.
38. Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP, Boles E. Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in Saccharomyces cerevisiae. FEBS Lett. 1999;464(3):123-8.
39. Kruckeberg AL. The hexose transporter family of Saccharomyces cerevisiae. Arch Microbiol. 1996;166(5):283-92.
40. Jordan P, Choe JY, Boles E, Oreb M. Hxt13, Hxt15, Hxt16 and Hxt17 from Saccharomyces cerevisiae represent a novel type of polyol transporters. Sci Rep. 2016;6:23502.
41. Elbing K, Stahlberg A, Hohmann S, Gustafsson L. Transcriptional responses to glucose at different glycolytic rates in Saccharomyces cerevisiae. Eur J Biochem. 2004;271(23-24):4855-64.
42. Reifenberger E, Boles E, Ciriacy M. Kinetic characterization of individual hexose transporters of Saccharomyces cerevisiae and their relation to the triggering mechanisms of glucose repression. Eur J Biochem. 1997;245(2):324-33.
43. Ozcan S, Johnston M. Function and regulation of yeast hexose transporters. Microbiol Mol Biol Rev. 1999;63(3):554-69.
44. Sanz P, Alms GR, Haystead TA, Carlson M. Regulatory interactions between the Reg1-Glc7 protein phosphatase and the Snf1 protein kinase. Mol Cell Biol. 2000;20(4):1321-8.
45. Hedbacker K, Carlson M. SNF1/AMPK pathways in yeast. Front Biosci. 2008;13:2408-20.

46. Fernandez-Garcia P, Pelaez R, Herrero P, Moreno F. Phosphorylation of yeast hexokinase 2 regulates its nucleocytoplasmic shuttling. J Biol Chem. 2012;287(50):42151-64.
47. Ahuatzi D, Riera A, Pela Ez R, Herrero P, Moreno F. Hxk2 regulates the phosphorylation state of Mig1 and therefore its nucleocytoplasmic distribution. J Biol Chem. 2007;282(7):4485-93.
48. Vallier LG, Carlson M. Synergistic release from glucose repression by mig1 and ssn mutations in Saccharomyces cerevisiae. Genetics. 1994;137(1):49-54.
49. Nehlin JO, Carlberg M, Ronne H. Control of yeast GAL genes by MIG1 repressor: a transcriptional cascade in the glucose response. EMBO J. 1991;10(11):3373-7.
50. Klein CJ, Olsson L, Ronnow B, Mikkelsen JD, Nielsen J. Alleviation of glucose repression of maltose metabolism by MIG1 disruption in Saccharomyces cerevisiae. Appl Environ Microbiol. 1996;62(12):4441-9.
51. Westholm JO, Nordberg N, Muren E, Ameur A, Komorowski J, Ronne H. Combinatorial control of gene expression by the three yeast repressors Mig1, Mig2 and Mig3. BMC Genomics. 2008;9:601.
52. Christensen TS, Oliveira AP, Nielsen J. Reconstruction and logical modeling of glucose repression signaling pathways in Saccharomyces cerevisiae. BMC Syst Biol. 2009;3:7.
53. Kim JH, Rodriguez R. Glucose regulation of the paralogous glucose sensing receptors Rgt2 and Snf3 of the yeast Saccharomyces cerevisiae. Biochim Biophys Acta Gen Subj. 2021;1865(6):129881.
54. Ozcan S, Dover J, Johnston M. Glucose sensing and signaling by two glucose receptors in the yeast Saccharomyces cerevisiae. EMBO J. 1998;17(9):2566-73.
55. Kayikci O, Nielsen J. Glucose repression in Saccharomyces cerevisiae. FEMS Yeast Res. 2015;15(6).
56. Pasula S, Chakraborty S, Choi JH, Kim JH. Role of casein kinase 1 in the glucose sensor-mediated signaling pathway in yeast. BMC Cell Biol. 2010;11:17.
57. Flick KM, Spielewoy N, Kalashnikova TI, Guaderrama M, Zhu Q, Chang HC, et al. Grr1-dependent inactivation of Mth1 mediates glucose-induced dissociation of Rgt1 from HXT gene promoters. Mol Biol Cell. 2003;14(8):3230-41.
58. Schmidt MC, McCartney RR, Zhang X, Tillman TS, Solimeo H, Wolfl S, et al. Std1 and Mth1 proteins interact with the glucose sensors to control glucose-regulated gene expression in Saccharomyces cerevisiae. Mol Cell Biol. 1999;19(7):4561-71.
59. Roy A, Jouandot D, 2nd, Cho KH, Kim JH. Understanding the mechanism of glucose-induced relief of Rgt1-mediated repression in yeast. FEBS Open Bio. 2014;4:105-11.
60. Li S, Li Y, Rushing BR, Harris SE, McRitchie SL, Dominguez D, et al. Multi-Omics Analysis of Multiple Glucose-Sensing Receptor Systems in

Yeast. Biomolecules. 2022;12(2).
61. Conrad M, Schothorst J, Kankipati HN, Van Zeebroeck G, Rubio-Texeira M, Thevelein JM. Nutrient sensing and signaling in the yeast Saccharomyces cerevisiae. FEMS Microbiol Rev. 2014;38(2):254-99.
62. Creamer DR, Hubbard SJ, Ashe MP, Grant CM. Yeast Protein Kinase A Isoforms: A Means of Encoding Specificity in the Response to Diverse Stress Conditions? Biomolecules. 2022;12(7).
63. Broach JR. Nutritional control of growth and development in yeast. Genetics. 2012;192(1):73-105.
64. Kim JH, Johnston M. Two glucose-sensing pathways converge on Rgt1 to regulate expression of glucose transporter genes in Saccharomyces cerevisiae. J Biol Chem. 2006;281(36):26144-9.
65. Wilson WA, Roach PJ. Nutrient-regulated protein kinases in budding yeast. Cell. 2002;111(2):155-8.
66. Muller D, Exler S, Aguilera-Vazquez L, Guerrero-Martin E, Reuss M. Cyclic AMP mediates the cell cycle dynamics of energy metabolism in Saccharomyces cerevisiae. Yeast. 2003;20(4):351-67.
67. Plank M. Interaction of TOR and PKA Signaling in S. cerevisiae. Biomolecules. 2022;12(2).
68. Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. Cell. 2006;124(3):471-84.
69. Santangelo GM. Glucose signaling in Saccharomyces cerevisiae. Microbiol Mol Biol Rev. 2006;70(1):253-82.
70. Papamichos-Chronakis M, Gligoris T, Tzamarias D. The Snf1 kinase controls glucose repression in yeast by modulating interactions between the Mig1 repressor and the Cyc8-Tup1 co-repressor. EMBO Rep. 2004;5(4):368-72.
71. Treitel MA, Kuchin S, Carlson M. Snf1 protein kinase regulates phosphorylation of the Mig1 repressor in Saccharomyces cerevisiae. Mol Cell Biol. 1998;18(11):6273-80.
72. Young ET, Dombek KM, Tachibana C, Ideker T. Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. J Biol Chem. 2003;278(28):26146-58.
73. Lesage P, Yang X, Carlson M. Yeast SNF1 protein kinase interacts with SIP4, a C6 zinc cluster transcriptional activator: a new role for SNF1 in the glucose response. Mol Cell Biol. 1996;16(5):1921-8.
74. Mehlgarten C, Krijger JJ, Lemnian I, Gohr A, Kasper L, Diesing AK, et al. Divergent Evolution of the Transcriptional Network Controlled by Snf1-Interacting Protein Sip4 in Budding Yeasts. PLoS One. 2015;10(10):e0139464.
75. Vincent O, Carlson M. Sip4, a Snf1 kinase-dependent transcriptional activator, binds to the carbon source-responsive element of gluconeogenic genes. EMBO J. 1998;17(23):7002-8.
76. Simon M, Adam G, Rapatz W, Spevak W, Ruis H. The Saccharomyces-Cerevisiae Adr1 Gene Is a Positive Regulator of Transcription of Genes Encoding Peroxisomal Proteins. Molecular and Cellular Biology.

1991;11(2):699-704.
77. Turcotte B, Liang XB, Robert F, Soontorngun N. Transcriptional regulation of nonfermentable carbon utilization in budding yeast. FEMS Yeast Res. 2010;10(1):2-13.
78. Bourgarel D, Nguyen CC, Bolotin-Fukuhara M. HAP4, the glucose-repressed regulated subunit of the HAP transcriptional complex involved in the fermentation-respiration shift, has a functional homologue in the respiratory yeast Kluyveromyces lactis. Mol Microbiol. 1999;31(4):1205-15.
79. Liu Z, Butow RA. A transcriptional switch in the expression of yeast tricarboxylic acid cycle genes in response to a reduction or loss of respiratory function. Mol Cell Biol. 1999;19(10):6720-8.
80. Lee P, Kim MS, Paik SM, Choi SH, Cho BR, Hahn JS. Rim15-dependent activation of Hsf1 and Msn2/4 transcription factors by direct phosphorylation in Saccharomyces cerevisiae. FEBS Lett. 2013;587(22):3648-55.
81. Bisschops MM, Zwartjens P, Keuter SG, Pronk JT, Daran-Lapujade P. To divide or not to divide: a key role of Rim15 in calorie-restricted yeast cultures. Biochim Biophys Acta. 2014;1843(5):1020-30.
82. Swinnen E, Wanke V, Roosen J, Smets B, Dubouloz F, Pedruzzi I, et al. Rim15 and the crossroads of nutrient signalling pathways in Saccharomyces cerevisiae. Cell Div. 2006;1:3.
83. Orzechowski Westholm J, Tronnersjo S, Nordberg N, Olsson I, Komorowski J, Ronne H. Gis1 and Rph1 regulate glycerol and acetate metabolism in glucose depleted yeast cells. PLoS One. 2012;7(2):e31577.
84. Nasaruddin ML, Tajul Arifin K. Application of Metabolomics in the Study of Starvation-Induced Autophagy in Saccharomyces cerevisiae: A Scoping Review. J Fungi (Basel). 2021;7(11).
85. Wu J, Zhang N, Hayes A, Panoutsopoulou K, Oliver SG. Global analysis of nutrient control of gene expression in Saccharomyces cerevisiae during growth and starvation. Proc Natl Acad Sci U S A. 2004;101(9):3148-53.
86. Holland P, Bergenholm D, Borlin CS, Liu G, Nielsen J. Predictive models of eukaryotic transcriptional regulation reveals changes in transcription factor roles and promoter usage between metabolic conditions. Nucleic Acids Res. 2019;47(10):4986-5000.
87. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature. 1997;387(6634):708-13.
88. Solis-Escalante D, Kuijpers NG, Barrajon-Simancas N, van den Broek M, Pronk JT, Daran JM, et al. A Minimal Set of Glycolytic Genes Reveals Strong Redundancies in Saccharomyces cerevisiae Central Metabolism. Eukaryot Cell. 2015;14(8):804-16.
89. Doorenbosch T, Mager WH, Planta RJ. Multifunctional DNA-binding proteins in yeast. Gene Expr. 1992;2(3):193-201.
90. Challal D, Barucco M, Kubik S, Feuerbach F, Candelli T, Geoffroy H, et al. General Regulatory Factors Control the Fidelity of Transcription by Restricting Non-coding and Ectopic Initiation. Mol Cell. 2018;72(6):955-69 e7.
91. Moehle CM, Hinnebusch AG. Association of RAP1 binding sites with

stringent control of ribosomal protein gene transcription in Saccharomyces cerevisiae. Mol Cell Biol. 1991;11(5):2723-35.
92. Wu ACK, Van Werven FJ. Transcribe this way: Rap1 confers promoter directionality by repressing divergent transcription. Transcription. 2019;10(3):164-70.
93. Klein C, Struhl K. Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. Mol Cell Biol. 1994;14(3):1920-8.
94. Baker HV. GCR1 of Saccharomyces cerevisiae encodes a DNA binding protein whose binding is abolished by mutations in the CTTCC sequence motif. Proc Natl Acad Sci U S A. 1991;88(21):9443-7.
95. Huie MA, Scott EW, Drazinic CM, Lopez MC, Hornstra IK, Yang TP, et al. Characterization of the DNA-binding activity of GCR1: in vivo evidence for two GCR1-binding sites in the upstream activating sequence of TPI of Saccharomyces cerevisiae. Mol Cell Biol. 1992;12(6):2690-700.
96. Tornow J, Zeng X, Gao W, Santangelo GM. GCR1, a transcriptional activator in Saccharomyces cerevisiae, complexes with RAP1 and can function without its DNA binding domain. EMBO J. 1993;12(6):2431-7.
97. Zeng X, Deminoff SJ, Santangelo GM. Specialized Rap1p/Gcr1p transcriptional activation through Gcr1p DNA contacts requires Gcr2p, as does hyperphosphorylation of Gcr1p. Genetics. 1997;147(2):493-505.
98. Barbara KE, Willis KA, Haley TM, Deminoff SJ, Santangelo GM. Coiled coil structures and transcription: an analysis of the S. cerevisiae coilome. Mol Genet Genomics. 2007;278(2):135-47.
99. Deminoff SJ, Santangelo GM. Rap1p requires Gcr1p and Gcr2p homodimers to activate ribosomal protein and glycolytic genes, respectively. Genetics. 2001;158(1):133-43.
100. Barbara KE, Haley TM, Willis KA, Santangelo GM. The transcription factor Gcr1 stimulates cell growth by participating in nutrient-responsive gene expression on a global level. Mol Genet Genomics. 2007;277(2):171-88.
101. Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF. Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae. PLoS Genet. 2014;10(4):e1004249.
102. Hossain MA, Claggett JM, Edwards SR, Shi A, Pennebaker SL, Cheng MY, et al. Posttranscriptional Regulation of Gcr1 Expression and Activity Is Crucial for Metabolic Adjustment in Response to Glucose Availability. Mol Cell. 2016;62(3):346-58.
103. Lee SY, Kim HU, Chae TU, Cho JS, Kim JW, Shin JH, et al. A comprehensive metabolic map for production of bio-based chemicals. Nat Catal. 2019;2(1):18-33.
104. Stephanopoulos G, Sinskey AJ. Metabolic engineering--methodologies and future prospects. Trends Biotechnol. 1993;11(9):392-6.
105. Cho JS, Kim GB, Eun H, Moon CW, Lee SY. Designing Microbial Cell Factories for the Production of Chemicals. JACS Au. 2022;2(8):1781-99.
106. Ida K, Ishii J, Matsuda F, Kondo T, Kondo A. Eliminating the

isoleucine biosynthetic pathway to reduce competitive carbon outflow during isobutanol production by Saccharomyces cerevisiae. Microb Cell Fact. 2015;14:62.
107. Tajima Y, Yamamoto Y, Fukui K, Nishio Y, Hashiguchi K, Usuda Y, et al. Effects of eliminating pyruvate node pathways and of coexpression of heterogeneous carboxylation enzymes on succinate production by Enterobacter aerogenes. Appl Environ Microbiol. 2015;81(3):929-37.
108. Paalman JWG, Verwaal R, Slofstra JH, Verkleij AJ, Boonstra J, Verrips CT. Trehalose and glycogen accumulation is related to the duration of the G(1) phase of Saccharomyces cerevisiae. Fems Yeast Research. 2003;3(3):261-8.
109. Vicente RL, Spina L, Gomez JPL, Dejean S, Parrou JL, Francois JM. Trehalose-6-phosphate promotes fermentation and glucose repression in Saccharomyces cerevisiae. Microb Cell. 2018;5(10):444-59.
110. Bertels LK, Fernandez Murillo L, Heinisch JJ. The Pentose Phosphate Pathway in Yeasts-More Than a Poor Cousin of Glycolysis. Biomolecules. 2021;11(5).
111. Lao-Martil D, Verhagen KJA, Schmitz JPJ, Teusink B, Wahl SA, van Riel NAW. Kinetic Modeling of Saccharomyces cerevisiae Central Carbon Metabolism: Achievements, Limitations, and Opportunities. Metabolites. 2022;12(1).
112. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nature Biotechnology. 2010;28(3):245-8.
113. Plaza J, Bogaerts P. Dynamic flux balance analysis for predicting biomass growth and ethanol production in yeast fed-batch cultures. Ifac Papersonline. 2018;51(2):631-6.
114. Kwak S, Yun EJ, Lane S, Oh EJ, Kim KH, Jin YS. Redirection of the Glycolytic Flux Enhances Isoprenoid Production in Saccharomyces cerevisiae. Biotechnol J. 2020;15(2):e1900173.
115. Liu Q, Yu T, Li X, Chen Y, Campbell K, Nielsen J, et al. Rewiring carbon metabolism in yeast for high level production of aromatic chemicals. Nat Commun. 2019;10(1):4976.
116. Jin C, Kim S, Moon S, Jin H, Hahn JS. Efficient production of shinorine, a natural sunscreen material, from glucose and xylose by deleting HXK2 encoding hexokinase in Saccharomyces cerevisiae. FEMS Yeast Res. 2021;21(7).
117. Lee SW, Oh MK. Improved production of N-acetylglucosamine in Saccharomyces cerevisiae by reducing glycolytic flux. Biotechnol Bioeng. 2016;113(11):2524-8.
118. Baek SH, Kwon EY, Bae SJ, Cho BR, Kim SY, Hahn JS. Improvement of d-Lactic Acid Production in Saccharomyces cerevisiae Under Acidic Conditions by Evolutionary and Rational Metabolic Engineering. Biotechnol J. 2017;12(10).
119. Bae SJ, Kim S, Hahn JS. Efficient production of acetoin in Saccharomyces cerevisiae by disruption of 2,3-butanediol dehydrogenase and

expression of NADH oxidase. Sci Rep. 2016;6:27667.
120. Kim S, Hahn JS. Efficient production of 2,3-butanediol in Saccharomyces cerevisiae by eliminating ethanol and glycerol production and redox rebalancing. Metab Eng. 2015;31:94-101.
121. Kim S, Hahn JS. Synthetic scaffold based on a cohesin-dockerin interaction for improved production of 2,3-butanediol in Saccharomyces cerevisiae. J Biotechnol. 2014;192 Pt A:192-6.
122. Ghosh A, Zhao H, Price ND. Genome-scale consequences of cofactor balancing in engineered pentose utilization pathways in Saccharomyces cerevisiae. PLoS One. 2011;6(11):e27316.
123. Hjersted JL, Henson MA. Steady-state and dynamic flux balance analysis of ethanol production by Saccharomyces cerevisiae. IET Syst Biol. 2009;3(3):167-79.
124. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, et al. Global analysis of protein expression in yeast. Nature. 2003;425(6959):737-41.
125. Bae SJ, Kim S, Park HJ, Kim J, Jin H, Kim BG, et al. High-yield production of (R)-acetoin in Saccharomyces cerevisiae by deleting genes for NAD(P)H-dependent ketone reductases producing meso-2,3-butanediol and 2,3-dimethylglycerate. Metab Eng. 2021;66:68-78.
126. Kim D, Song JY, Hahn JS. Improvement of glucose uptake rate and production of target chemicals by overexpressing hexose transporters and transcriptional activator Gcr1 in Saccharomyces cerevisiae. Appl Environ Microbiol. 2015;81(24):8392-401.
127. Hong J, Park SH, Kim S, Kim SW, Hahn JS. Efficient production of lycopene in Saccharomyces cerevisiae by enzyme engineering and increasing membrane flexibility and NAPDH production. Appl Microbiol Biotechnol. 2019;103(1):211-23.
128. Baek SH, Kwon EY, Kim YH, Hahn JS. Metabolic engineering and adaptive evolution for efficient production of D-lactic acid in Saccharomyces cerevisiae. Appl Microbiol Biotechnol. 2016;100(6):2737-48.
129. Noguchi C, Garabedian MV, Malik M, Noguchi E. A vector system for genomic FLAG epitope-tagging in Schizosaccharomyces pombe. Biotechnol J. 2008;3(9-10):1280-5.
130. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989;122(1):19-27.
131. Mumberg D, Muller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 1995;156(1):119-22.
132. Gasch AP. Yeast genomic expression studies using DNA microarrays. Methods Enzymol. 2002;350:393-414.
133. 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. Nat Protoc. 2012;7(3):562-78.

134. 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(4):R36.
135. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28(5):511-5.
136. Andrews S. FastQC: a quality control tool for high throughput sequence data 2010 [Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
137. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
138. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576-89.
139. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.
140. Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, Davidson I, et al. seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. 2011;39(6):e35.
141. Baker HV. Glycolytic gene expression in Saccharomyces cerevisiae: nucleotide sequence of GCR1, null mutants, and evidence for expression. Mol Cell Biol. 1986;6(11):3774-84.
142. Tornow J, Santangelo G. The GCR1 gene of Saccharomyces cerevisiae is a split gene with an unusually long intron. Genetics. 1994;138(3):973-4.
143. Turkel S, Liao XB, Farabaugh PJ. GCR1-dependent transcriptional activation of yeast retrotransposon Ty2-917. Yeast. 1997;13(10):917-30.
144. Dudley AM, Gansheroff LJ, Winston F. Specific components of the SAGA complex are required for Gcn4- and Gcr1-mediated activation of the his4-912delta promoter in Saccharomyces cerevisiae. Genetics. 1999;151(4):1365-78.
145. Lopez MC, Baker HV. Understanding the growth phenotype of the yeast gcr1 mutant in terms of global genomic expression patterns. J Bacteriol. 2000;182(17):4970-8.
146. Taylor R, Kebaara BW, Nazarenus T, Jones A, Yamanaka R, Uhrenholdt R, et al. Gene set coregulated by the Saccharomyces cerevisiae nonsense-mediated mRNA decay pathway. Eukaryot Cell. 2005;4(12):2066-77.
147. Gonzalez CI, Bhattacharya A, Wang W, Peltz SW. Nonsense-mediated mRNA decay in Saccharomyces cerevisiae. Gene. 2001;274(1-2):15-25.
148. Solinger JA, Pascolini D, Heyer WD. Active-site mutations in the Xrn1p exoribonuclease of Saccharomyces cerevisiae reveal a specific role in meiosis. Mol Cell Biol. 1999;19(9):5930-42.
149. Blasco-Moreno B, de Campos-Mata L, Bottcher R, Garcia-Martinez J, Jungfleisch J, Nedialkova DD, et al. The exonuclease Xrn1 activates

transcription and translation of mRNAs encoding membrane proteins. Nat Commun. 2019;10(1):1298.
150. Zeng AP, Byun TG, Posten C, Deckwer WD. Use of Respiratory Quotient as a Control Parameter for Optimum Oxygen-Supply and Scale-up of 2,3-Butanediol Production under Microaerobic Conditions. Biotechnology and Bioengineering. 1994;44(9):1107-14.
151. Parenteau J, Durand M, Veronneau S, Lacombe AA, Morin G, Guerin V, et al. Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol Biol Cell. 2008;19(5):1932-41.
152. Hooks KB, Naseeb S, Parker S, Griffiths-Jones S, Delneri D. Novel Intronic RNA Structures Contribute to Maintenance of Phenotype in Saccharomyces cerevisiae. Genetics. 2016;203(3):1469-81.
153. Juneau K, Nislow C, Davis RW. Alternative splicing of PTC7 in Saccharomyces cerevisiae determines protein localization. Genetics. 2009;183(1):185-94.
154. Grund SE, Fischer T, Cabal GG, Antunez O, Perez-Ortin JE, Hurt E. The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. J Cell Biol. 2008;182(5):897-910.
155. Lipinski KA, Kaniak-Golik A, Golik P. Maintenance and expression of the S. cerevisiae mitochondrial genome--from genetics to evolution and systems biology. Biochim Biophys Acta. 2010;1797(6-7):1086-98.
156. Celik A, He F, Jacobson A. NMD monitors translational fidelity 24/7. Curr Genet. 2017;63(6):1007-10.
157. Shimizu K, Matsuoka Y. Regulation of glycolytic flux and overflow metabolism depending on the source of energy generation for energy demand. Biotechnol Adv. 2019;37(2):284-305.
158. van den Brink J, Canelas AB, van Gulik WM, Pronk JT, Heijnen JJ, de Winde JH, et al. Dynamics of glycolytic regulation during adaptation of Saccharomyces cerevisiae to fermentative metabolism. Appl Environ Microbiol. 2008;74(18):5710-23.
159. Huberts DH, Niebel B, Heinemann M. A flux-sensing mechanism could regulate the switch between respiration and fermentation. FEMS Yeast Res. 2012;12(2):118-28.
160. Overal GB, Teusink B, Bruggeman FJ, Hulshof J, Planque R. Understanding start-up problems in yeast glycolysis. Math Biosci. 2018;299:117-26.
161. Wu P, Wang G, Wang G, Borresen BT, Liu H, Zhang J. Butanol production under microaerobic conditions with a symbiotic system of Clostridium acetobutylicum and Bacillus cereus. Microb Cell Fact. 2016;15:8.
162. Costenoble R, Valadi H, Gustafsson L, Niklasson C, Franzen CJ. Microaerobic glycerol formation in Saccharomyces cerevisiae. Yeast. 2000;16(16):1483-95.

국문초록

크랩트리 양성인 효모 *Saccharomyces cerevisiae*는 산소의 유무에 관계 없이 포도당을 해당과정과 에탄올 발효를 통해 빠르게 분해하는 대사적 특징을 가진다. Gcr1은 해당과정에 관여하는 유전자들을 조절하는 가장 중요한 전사조절인자로, 최근들어 $Gcr1^U$와 $Gcr1^S$의 두 가지 동위체가 존재함이 밝혀졌다.

GCR1 mRNA는 복잡한 스플라이싱 기작을 통해 최소 7개의 서로 다른 스플라이싱 mRNA 결과물을 형성한다. 그러나 그 중 두 개만 정상적인 단백질인 $Gcr1^A$와 $Gcr1^S$로 번역되며 나머지는 모두 제거된다. 스플라이싱 되지 않은 *GCR1*의 mRNA는 인트론의 중간에 위치한 개시코돈(ATG)을 이용하여 또 다른 단백질 형태인 $Gcr1^U$를 만든다. 본 연구에서는 세 형태의 단백질들 중 가장 주요하게 생성되는 $Gcr1^U$와 $Gcr1^S$의 차등적 역할을 규명하고, 이를 젖산 생산에 활용하였다.

첫째, CRISPR/Cas9을 기반으로 한 유전자 조작 기술을 이용하여 $Gcr1^U$ 또는 $Gcr1^S$만 따로 생성하는 균주를 제작하였다. 두 균주는 성장 및 타겟 유전자에서 차이를 거의 보이지 않았으나, 두 단백질의 프로모터에의 결합은 서로 다른 양상을 보였다. 또한, Gcr1의 단백질 도메인 혹은 co-activator *GCR2* 유전자 결손 연구를 통해 $Gcr1^U$ 단백질은 주로 co-activator인 Gcr2 단백질과의 결합하여 타겟 유전자들을 활성화시키는 반면, $Gcr1^S$ 단백질은 Gcr2 단백질 없이도 이량체 형성을 통해 활성화된다는 것을 밝혔다.

둘째, $Gcr1^U$의 N 말단에만 존재하는 55개 아미노산으로 이루어진 USS 도메인이 $Gcr1^U$의 이량체 형성을 막는다는 사실을 밝혔다. 이 도메인은 $Gcr1^S$의 이량체 형성 또한 트랜스-액팅 기작을 통해 저해한다. 이로 인해 형성된 $Gcr1^S$의 단량체는 *ALD4* (mitochondrial aldehyde dehydrogenase) 유전자의 프로모터에 직접 결합하여 발효에서 호흡으로의 대사 전환을 억제하지만, *ALD4* 유전자의 과발현시 이러한 결함이 해

소되었다. 이러한 표현형은 *GCR1*^{S}의 mRNA가 과발현되었을 때도 관찰되었으며, 이는 Gcr1^{S} 단백질의 발현양을 적정 수준으로 유지하는 것이 *S. cerevisiae*의 호흡에 중요함을 시사한다.

마지막으로, Gcr1^{S}의 비정상적 발현 조절로 인한 호흡 대사과정의 결함이 피루브산을 전구체로 하여 NADH를 조효소로 사용하는 물질 생산량 증가에 활용될 수 있음을 밝혔다. 호기성 조건에서 *GCR1*^{WT}을 과발현 한 경우, 과도한 해당과정의 강화로 인해 젖산 생산이 감소하였다. 반면 *GCR1*^{S}의 과발현시 호흡과정으로 가는 탄소 흐름의 억제를 통해 젖산 생산을 증가시킬 수 있었다. 본 연구는 적절한 Gcr1 동형체의 발현 조절을 통해 포도당 대사에 있어 발효와 호흡 경로의 분배를 조절함으로써 대사 물질 생산에 활용할 수 있음을 시사한다.

주요어 : *Saccharomyces cerevisiae*, 대체 접합, CRISPR/Cas9, 이조성 성장 전환, Gcr1, Gcr2, 해당과정, 해당과정 흐름 조절, 세포 호흡

학번 : 2016-21056

감사의 글