Expanding the toolbox for genetic manipulations in *Methanosarcina acetivorans*

Ping Zhu

Aalto University

Aalto University publication series
Doctoral Theses 20/2025

Expanding the toolbox for genetic manipulations in *Methanosarcina acetivorans*

Ping Zhu

A doctoral thesis completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Chemical Engineering, at a public examination held at the lecture hall Otakaari 1, A123 A1 on 14th of February 2025 at 12 noon.

Aalto University
School of Chemical Engineering
Department of Bioproducts and Biosystems

Supervising professor
Professor Silvan Scheller, Aalto University, Finland

Thesis advisor
Professor Silvan Scheller, Aalto University, Finland

Preliminary examiners
Professor Yvonne Nygård, VTT Technical Research Centre of Finland, Finland
Professor Wenyu Gu, École Polytechnique Fédérale de Lausanne, Switzerland

Opponent
Professor Bastian Molitor, Leipzig University, Germany

Aalto University publication series
Doctoral Theses 20/2025

ISBN 978-952-64-2377-7 (paperback)
ISBN 978-952-64-2378-4 (PDF)
ISSN 1799-4934 (paperback)
ISSN 1799-4942 (PDF)

https://urn.fi/URN:ISBN:978-952-64-2378-4

PunaMusta Oy
Helsinki, 2025

Aalto-yliopisto

ABSTRACT

Author
Ping Zhu

Name of the doctoral thesis
Expanding the toolbox for genetic manipulations in *Methanosarcina acetivorans*

Article-based thesis

Number of pages 172

Keywords Genetic engineering, toolbox, CRISPR, promoter engineering, integration site, *Methanosarcina acetivorans*, Methanogen

Methanosarcina acetivorans is a model methane producer that can utilize many different one-carbon substrates. Genetic tools are needed to investigate its physiology and to extend the product scope beyond methane. The development of the current genetic toolbox in this archaeon, however, is far behind compared with those e.g. for acetogenic bacteria or *E. coli*. In this study, the genetic toolbox for *M. acetivorans* has been expanded for efficient genome edition and for controlling gene expression.

First, a CRISPR/Cas12a-based genome editing system (Publication **I**) was introduced for efficient marklerless genome editing in *M. acetivorans*. Different from the previously developed Cas9-editing system, the Cas12a recognizes distinct PAM sequence and demonstrates high efficiency in large DNA fragments deletion and alleviate the cloning efforts on multiplex genome editing. To facilitate a stable and reliable gene expression, four neutral integration sites in *M. acetivorans* genome were characterized (Publication **II**) for heterologous genes integration and metabolic pathway constructions. These neutral sites allow gene expression without disrupting cell viability, providing as non-essential loci for stable long-term genetic engineering.

Finally, a promoter-RBS library comparing 33 combinations was developed for fine-tuning gene expression in *M. acetivorans* (Publication **III**). By evaluating the native and engineered combinations, the library revealed diverse transcriptional regulation mechanisms mediated by RNA secondary structures. With a broad range of expression strength (ca. 140-fold), the library allows precise control of gene expression levels and enables fine regulation of metabolic flux for various synthetic biology applications.

The tools developed in the thesis enhance the capacity for genomic manipulation of *M. acetivorans* and facilitate the development of metabolic engineering for efficient methane production, carbon flux rewiring, and other biotechnological processes in this organism.

Acknowledgements

The research in the thesis was carried out from 2020-2024 in the Biochemistry group at the Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University. This work was financially supported by the Research Council of Finland and the Novo Nordisk Foundation. I am also deeply grateful to the Gasum Gas Fund from Gasum Oy for providing personal funding support in the summer of 2020.

First and foremost, I express my sincere gratitude to my supervisor, Prof. Silvan Scheller. Since joining the lab as a research assistant in 10. 2019, your guidance and encouragement have been pivotal in shaping my doctoral journey. Your passion for science set a high standard for me, and your dedication and enthusiasm for research have been truly inspiring. I will always remember your motivational words, "Let's focus on science", which consistently reminded me of the essence of this pursuit. I am profoundly grateful to my research advisor, Prof. Merja Penttilä, and to the secretary of the Doctoral Program Committee, Sirje Liukko. Your advice during our annual meetings helped me approach my research with a broader perspective and long-term vision. This thesis would not have been possible without your valuable insights and support.

I extend my thanks to the thesis pre-examiners, Prof. Yvonne Nygård and Prof. Wenyu Gu, for their thorough review of my research. Your valuable comments and suggestions greatly enhanced the coherence and clarity of this thesis, and I deeply appreciate your contributions.

I would like to thank my colleagues and collaborators, Dr. Ingemar von Ossowski, Dr. Norman Adlung, Maxime Laird, Dr. Vera D. Jäger, Tejas Somvanshi, Dr. Andrea Rodil Garcia, Shreyash Borkar, Mariana Molina Resendiz, Osama Mohamed, Dr. Thinh Nguyen, Dr. Sebastian Keller, Enrique de Dios Mateos, Dr. Yufang Tian, Heta Telimaa, Dr. Mai Tran, Feilong Li from Biochemistry group, and Dr. He Li, Pankaj Goyal, Dr. Yin Yin, Ruxia Fan, Mengjie Shen, Dr. Koskela Salla, Dr. Xiangze Kong from our School. Thank you for your companionship, insightful discussions, and the memorable experience we shared both inside and outside the lab. These memories are treasures that I will carry forward.

To my dear friends, Dr. Xiaoping Yue, Jiangang Zhang, Yining Zhang, Dr. Xuehua Wang, Yan Cai, Guiyue Tang, I am deeply grateful for your companionship over the years. The joyful moments and cherished memories we created together

have brought immense meaning to this journey. I also extend my thanks to JoJo and all the friends not mentioned here. Your thoughtful advice and inspiring ideas have been invaluable throughout this process.

A special thanks to Jianhui Feng, whose unwavering support and optimism have been a source of strength for me. This journey was like a roller coaster, and your constant encouragement made both the challenges and highlights even more enjoyable. Your patience and assistance in life and career mean a lot to me, and I feel fortunate to have you by my side. I would also like to thank my adorable Heisen. Your appearance brings sparkle to my life, and I wish you to be a happy and carefree cat forever.

Finally, to my beloved family, especially my wonderful grandparents, your unconditional love has been my foundation, providing me with the courage and resilience to pursue my passions. I cannot wait to reunite with you soon!

Kaisankuja 1, Espoo
January 2025
Ping Zhu

Contents

List of Publications

This doctoral thesis is based on the following publications:

I. Ping Zhu; Tejas Somvanshi; Jichen Bao; Silvan Scheller. 2023. CRISPR/Cas12a toolbox for genome editing in *Methanosarcina acetivorans*. Frontiers in Microbiology, volume 14, 1235616. ISSN 1664-302X. DOI: 10.3389/fmicb.2023.1235616.

II. Ping Zhu; Silvan Scheller. Characterization of neutral integration sites for gene expression in *Methanosarcina acetivorans*. (Submitted to Metabolic Engineering Communications)

III. Ping Zhu; Mariana Molina Resendiz; Ingemar von Ossowski; Silvan Scheller. 2024. A promoter–RBS library for fine-tuning gene expression in *Methanosarcina acetivorans*. Applied and Environmental Microbiology, volume 90, e01092-24. ISSN 1098-5336. DOI: 10.1128/aem.01092-24.

Author's Contribution

Publication I: CRISPR/Cas12a toolbox for genome editing in *Methanosarcina acetivorans*

PZ, JB, and SS conceived the research. PZ performed the experiments. PZ and TS analysed data. PZ wrote the manuscript. JB and SS supervised the research and critically revised the manuscript. SS responsible for the funding acquisition.

Publication II: Characterization of neutral integration sites for gene expression in *Methanosarcina acetivorans*

PZ and SS conceived the research. PZ performed the experiments, analysed data, and wrote the manuscript. SS supervised the research, critically reviewed the manuscript, and responsible for the funding acquisition.

Publication III: A promoter–RBS library for fine-tuning gene expression in *Methanosarcina acetivorans*

PZ and SS conceived the research. PZ and MMR performed the experiments. PZ analysed data and wrote the manuscript. IvO critically reviewed the manuscript and corrected the grammar. SS supervised the research and responsible for the funding acquisition.

List of Abbreviations

5´UTR	5´Untranslated region
8ADP	8-Aza-2,6-diaminopurine
Amp	Ampicillin
Bla	*β*-Lactamase
BRE	Transcription factor B-recognition element
CM	Cytoplasmic membrane
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
dCas9	dead Cas9
DOTAP	1,2-Dioleoyl-3-trimethylammonium-propane
DR	Direct repeat
DSB	Double strand breaks
dsDNA	Double-stranded DNA
GalNAc	*N*-acetyl-D-galactosamine
GFP	Green fluorescent protein
GGA	Golden Gate Assembly
GlcUA	*D*-Glucuronic acid
Hpt	Hypoxanthine phosphoribosyltransferase
HR	Homologous recombination
LacZ	*β*-Galactosidase
MC	Methanochondroitin
MCS	Multiple cloning site

MeOH	Methanol
ORF	Open reading frame
Ori	Origin of replication
Pac	Puromycin acetyltransferase
PAM	Protospacer adjacent motif
PEG	Polyethylene glycol
RBS	Ribosome binding site
RNAP	RNA polymerase
ssuC	Sulfonate ABC transporter
TBP	TATA binding protein
TFB	Transcription factor B
TMA	Trimethylamine
tracrRNA	Trans-activating crRNA
TSS	Transcriptional start site
UidA	β-Glucuronidase
WGS	Whole genome sequencing

1 Introduction

Methanogenic archaea are the essential methane producers on Earth, which can be found in anaerobic habitats such as soil sediments, landfills, paddy fields, and the rumen of livestock[1–4]. Their ability to produce methane and thrive in diverse environments makes them key players in the global carbon cycle and necessitates studying them for fundamental research. Furthermore, methanogenesis pathways can be used for biogas production, CO_2 fixation or for biotechnological one-carbon conversions.

Genetically modifying methanogens can provide valuable insights into their physiology, substrate-related energy conservation and metabolisms for fundamental understanding. Though the genomes of many methanogenic species have been sequenced over the years, only two genera, *Methanococcus* and *Methanosarcina*, have well-developed genetic systems owing to their early revealed genomes, versatile methanogenesis, and relatively manageable cultivating conditions[5,6]. More genetic tools are steadily studied considering the thriving development of biotechnologies in these organisms[7,8], and the inspiring research progress that have been made in model prokaryotic and eukaryotic systems.

In this thesis, three genetic tools were developed to expand the current toolbox for metabolic engineering in *Methanosarcina acetivorans*. Publication I introduced a Cas12a-mediated genome editing system, which enabled efficient markerless gene editing in this organism. In publication II, four genomic integration sites were characterized for heterologous genes integration, aiming to facilitate the construction of metabolic pathways in *M. acetivorans*. Publication III established a promoter-RBS library comprising 33 promoter-RBS combinations with the expression strength over 140-fold, which enabled precise gene expression control for rewiring metabolic pathways and regulating metabolic fluxes in the biotechnological applications in *M. acetivorans*.

1.1 Methanogenic archaea

Methanogens, mainly from the phylum *Euryarchaeota*, are classified into eight different orders based on the 16S rRNA sequence similarity, including *Methanobacteriales, Methanococcales, Methanomicrobiales, Methanosarcinales, Methanopyrales, Methanocellales, Methanomassiliicoccales, and Methanonatronarchaeales*[9–12]. Despite their phylogenic diversities, most

methanogens utilize at least one of the three conventional methanogenic pathways for methene production from various substrates. Most of the methanogens, except for the *Methanomassiliicoccales* order[13], can reduce CO2 and derive the required electrons through different types of hydrogenase to produce methane by hydrogenotrophic methanogenesis[14,15] (**Figure 1a**). Members of the *Methanosarcinales* order containing cytochromes are known for utilizing methylated compounds (i.e., methanol, methylamines, *etc.*) as energy and carbon source by substrate-specific methyltransferases (i.e., MTI, MTII in **Figure 1b**) in methylotrophic methanogenesis[16]. Additionally, the *Methanomasiliiicoccales* order, a novel group of methanogens, had been described as an obligate methyl-respiring member[17,18]. Only two genera from *Methanosarcinales* order, *Methanosarcina* and *Methanothrix* (also known as *Methanosaeta*), could decompose acetate to CO_2 and methane by aceticlastic methanogenesis (**Figure 1c**)[19–21].

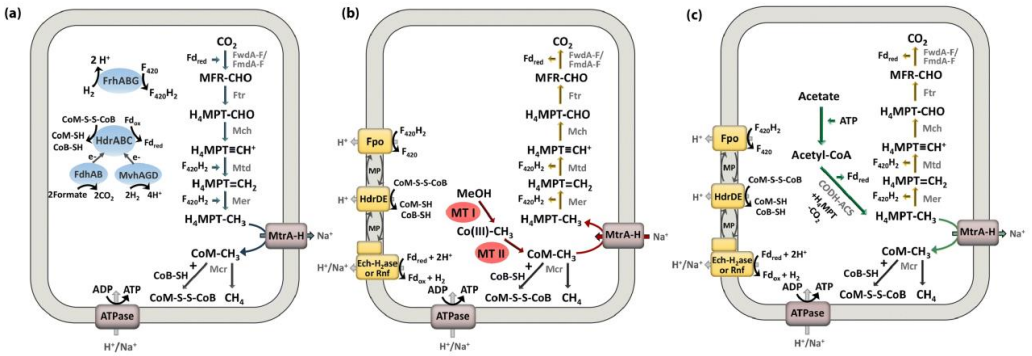

Figure 1 Three conventional methanogenesis pathways in methanogen. **(a)** Hydrogenotrophic methanogenesis, reducing CO_2 with hydrogenases to produce methane. FrhABG, coenzyme F420-reducing hydrogenase. HdrABC, soluble heterodisulfide reductase. FdhAB: formate dehydrogenase. MvhACD, F420-non-reducing hydrogenase. MtrA-H, tetrahydromethanopterin S-methyltransferase. ATPase: ATP synthase. **(b)** Methanogenetic methanogenesis reducing C1 compounds with the assistance of cytochromes for methane production. F420H2 dehydrogenase. HdrDE: membrane-bound heterodisulfide reductase. Ech-H2ase: energy-converting hydrogenase. MTI, MTII, substrate-specific methyltransferase. **(c)** Aceticlastic methanogenesis utilizing acetate with the assistance of cytochromes to produce methane. Adapted from Kurth *et al.*[22] under the Creative Commons CC BY license.

1.1.1 *Methanosarcinales*

Several species within the *Methanosarcinales* order (i.e., *M. acetivorans*, *Methanosarcina barkeri*, and *Methanosarcina mazei*, *etc.*) have been widely studied due to their versatile, cytochrome-containing methanogenesis and broad substrate utilization, which serving them promising models for studying energy conservation mechanisms[23,24] and potential chassis for other industrial

applications such as wastewater treatment, biofuels productions, and the generation of value-added biomaterials[7].

1.1.2 *Methanosarcina acetivorans*

As a representative organism in *Methanosarcina* genus, *M. acetivorans* was isolated from marine sediment and has been extensively studied thanks to its well-studied genomic information and the thriving genetic tools (discuss further in the section below) that have been established over the years[25,26]. *M. acetivorans* can utilize simple one carbon (i.e., MeOH, and methylated amines) through methylotrophic pathway, and two carbons such as acetate via aceticlastic pathway, as carbon and energy source for cell growth and methane production. Given its metabolic versatility, *M. acetivorans* has gained attention for various biotechnological applications. For instance, the expression of the esterase from *Pseudomonas* in this organism enabled the utilization of methyl acetate and methyl propionate through methylotrophic methanogenesis[27]. In another study, overexpression of a heterologous isoprene synthase in *M. acetivorans* successfully enabled the organism producing isoprene via mevalonate pathway with high carbon transferring efficiency[28]. These indicate the potential of utilizing this model methanogen as an industrial platform in the production of value-added products.

1.2 Genetic tools developed for *Methanosarcina* species

Exploring efficient genetic tools is crucial for investigating physiology, cellular biochemistry, and metabolic pathways of organisms. The development of genetic toolboxes for both prokaryotes and eukaryotes have greatly accelerated the fundamental research and applications in model organisms such as *Escherichia coli* and *Saccharomyces cerevisiae*[29–31]. Over the decades, various genetic tools for gene manipulation in methanogens[5], mainly in *Methanosarcina* and *Methanococcus* genera, have been studied, enhancing genetic research in these strains compared to other methanogenic species (**Figure 2**). In this section, the genetic tools that have been developed for metabolic engineering in *Methanosarcina* species are introduced, aiming to provide an overview of the current advancements in this area.

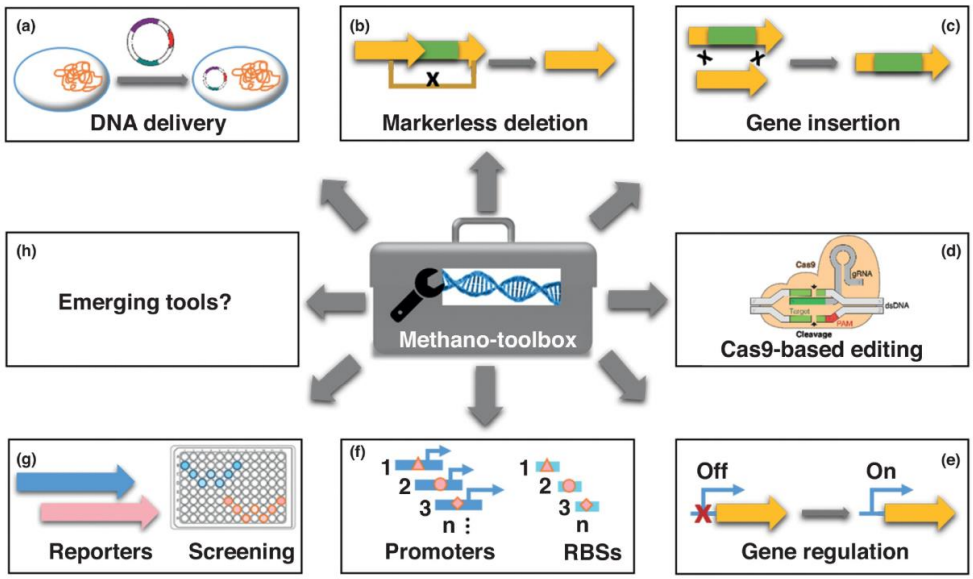

Figure 2 Genetic toolbox (Methano-toolbox) developed in methanogen. **(a)** DNA delivery methods. Homologous recombination-mediated **(b)** Markerless deletion and **(c)** Gene insertion. **(d)** CRISPR mediated genome editing (i.e., Cas9-based editing). **(e)** Gene regulation. **(f)** Promoters and ribosome binding sites (RBSs) engineering. **(g)** The exploration of reporter proteins and the cell-screening methods. **(h)** The exploration of the emerging tools. Reprinted with permission from Lyu and Whitman[32]. Copyright 2019 Elsevier Current Opinion in Biotechnology.

1.2.1 Shuttle vector and DNA delivery method

Shuttle vectors are plasmids containing genetic elements that allow them to propagate in two different hosts. The development of *E. coli* - *Methanosarcina* shuttle vectors has enabled genetic cloning to be conducted in *E. coli* host and transferred into *Methanosarcina* species for functional expression. To date, shuttle vectors for *Methanosarcina* have been derived from the pWM-series of plasmid such as pWM321[33], which comprising the *M. acetivorans* self-replicating plasmid pC2A[34], the puromycin acetyltransferase gene (*pac*) from *Streptomyces alboniger* as the selection marker for *Methanosarcina*, the origin of replication (*ori*) from plasmid R6K for replicating in *E. coli*, and the *β*-lactamase gene (*bla*) enables *E. coli* resistant to ampicillin. Subsequent shuttle vectors such as pJK-derivatives[35,36], pDN-CRISPR series[37,38], and the pMCp-CRISPR series[39] have been constructed upon this series and widely used in current genetic applications within *Methanosarcina* species. Recently, a simplified version of the vector pWM321 was established for an efficient and stable plasmid transformation in *M. mazei*[40]. Compared with the parental pWM321 (8,9 kb), the smaller vectors pRS1550 and pRS1595 disposed partial non-coding regions and the non-essential genes on the backbone (i.e., the *ssrA* gene encoding a site-specific recombinase, and the lacZ allowing the blue-white screening of *E. coli*), resulting in 35% reducing in size (ca.

5,8 kb) and allowing for shorter incubation time (ca. 5-7 days) to produce visible *M. mazei* transformants on solid media. This approach shed light on the construction of smaller vectors for time-saving transformations in other *Methanosarcina* species.

Self-replicating plasmids can get unstable in the environment without selection pressure, as they may be lost from the cell population. Considering this, chromosomal integration-based plasmids were developed, providing greater stability by ensuring the integrated genes being expressed as part of the genome. An example of this approach is the φC31 integrase-mediated site-specific recombination system[41], which has been implemented in engineered *M. acetivorans* and *M. barkeri* strains and the non-replicating plasmids (i.e., pJK027A, pAB79) possessing complementary recognizing elements could be integrated onto the specific site on the genome. Later, a simplified version of the suicide plasmids with the same integration system (i.e., pNB729 and pNB730) were constructed based on the commercially available *E. coli* plasmid pNEB193 for gene manipulation and reporter proteins expression in *M. acetivorans*[42,43].

DNA delivery efficiencies in methanogens are influenced by the composition of their cell membranes. In *Methanosarcina* species, the cell envelope comprises inner cytoplasmic membrane (CM), glycoprotein rich-surface layer (S-layer), and methanochondroitin (MC) composed of *N*-acetyl-*D*-galactosamine (GalNAc) and *D*-glucuronic acid (GlcUA) (**Figure 3a**). Under lab conditions where high salt concentrations (over 0.4 M) are used, *Methanosarcina* cells grow in a single-cell form without MC layer[44], which simplifies the DNA delivery due to the reduced membrane complexity. To introduce vectors into *Methanosarcina* cells, a highly efficient liposome-mediated transformation method had been developed[33]. The technique utilizes cationic liposome such as 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP) to combine the negatively charged DNA molecules and form a complex (**Figure 3b**). When mixed with protoplasts, the cationic headgroups of the complex interacts with anionic groups in cytoplastic membrane and gets into cells and releases DNA for editing. This method has been widely applied in genetic experiments including mutagenesis and the establishment of genetic tools in *M. acetivorans*[38,45,46]. Additionally, optimizations of this method enhanced its efficiency in mutants' construction in *M. mazei* and *M. barkeri*[47,48]. Polyethylene glycol (PEG)-mediated transformation is another highly efficient method for DNA delivery in methanogens. Similar to the liposome-mediated approach, PEG binds to the cytoplastic membrane in the presence of low concentrations of divalent cations (i.e., Mg^{2+}), enabling DNA entry into the cytoplast and release DNA. The method was initially described for plant genes transfer and was successfully attempted for *Methanococcus maripaludis* and *M. acetivorans*[49–51]. Natural transformation has been achieved in the engineered *M. maripaludis* JJ strain[52], however, this approach is unlikely to be applicable to *Methanosarcina* species, as

they lack functional type IV-like pili, which was reported as the portal for DNA uptake[53].

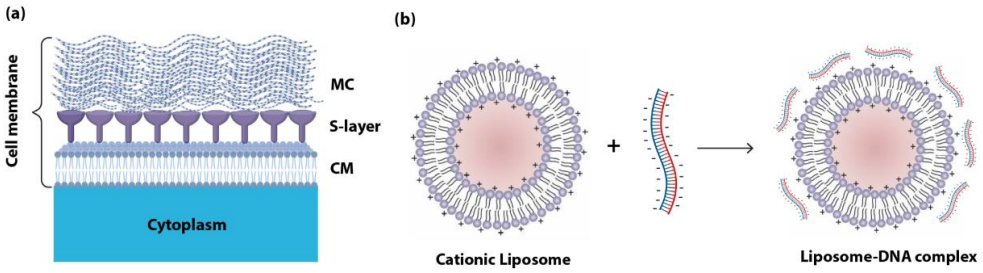

Figure 3 Cell membrane structure and DNA delivery in *Methanosarcina* species. **(a)** Cell membrane structure of *Methanosarcina*. CM, cytoplasmic membrane. S-layer, surface layer. MC, methanochondroitin. Adapted from Uali *et al.*[54] Copyright 2023 Taylor & Francis Group. **(b)** Schematic of cationic liposome-DNA complex in DNA delivery. Adapted from Luiz *et al.*[55] under the Creative Commons CC BY license.

1.2.2 Strategies of chromosomal mutagenesis

Chromosomal mutagenesis typically comprises gene deletions, gene insertions, and gene interference, each of which is crucial for investigating and manipulating biological features of organisms. Gene deletion is a fundamental technique for determining the essentiality and the function of specific genes within the genome. It is also an important step in metabolic engineering, where native genes are removed to rewire the metabolic pathways for better resource allocation and reducing byproduct production. Gene insertion, on the other hand, is a method for introducing novel functions into cells for exploring uncharacterized genes or constructing new pathways to produce valuable compounds in host organisms. The gene interference is a powerful tool in regulating gene expression mostly in a reversible manner. This approach is useful in exploring the native gene functions and uncovering the regulation mechanisms in host cells, where native genes or promoters are silenced to study the phenotypic changes. In metabolic engineering, this method enables fine-tuning gene expression level to control the metabolic fluxes, where essential genes that are impossible to delete can be downregulated to minimize the function without compromising viability.

Together, these approaches are crucial for optimizing engineered organisms for industrial applications. To achieve these goals, genetic tools based on methods like homologous recombination (HR) and integrase/transposase-mediated mutagenesis had been extensively explored in *Methanosarcina* species.

1.2.2.1 Homologous recombination-mediated mutagenesis

Homologous recombination is a DNA exchange process catalyzed by a group of dedicated enzymes. In *Methanosarcina*, the initial genetic editing attempt utilizing

this approach was carried out in *M. acetivorans* C2A, where the hypoxanthine phosphoribosyl transferase gene (*hpt*) was deleted by a "two-step markerless gene exchange" strategy to construct an engineered strain *M. acetivorans* WWM1[35]. Later, this standard and efficient method was applied and adapted for other species such as *M. mazei*[47].

Recent two-step markerless gene exchange mostly based on *Methanosarcina* Δ*hpt* mutants (i.e., *M. acetivorans* WWM1, and *M. barkeri* WWM85), where a vector containing DNA sequences homologous to regions upstream and downstream of the target site, and the *pac* and *hpt* genes, was integrated into genome (**Figure 4**). The *pac* and *hpt* genes confer resistance to puromycin and 8-aza-2,6-diaminopurine (8ADP) in the two rounds of recombination. In the first round of recombination, the vector integrates into the genome, creating merodiploid resistant to puromycin. Then, the merodiploid are cultivated without selective pressure for several generations to encourage the second round of recombination. The final desired mutants can be screened on the 8ADP containing-solid medium.

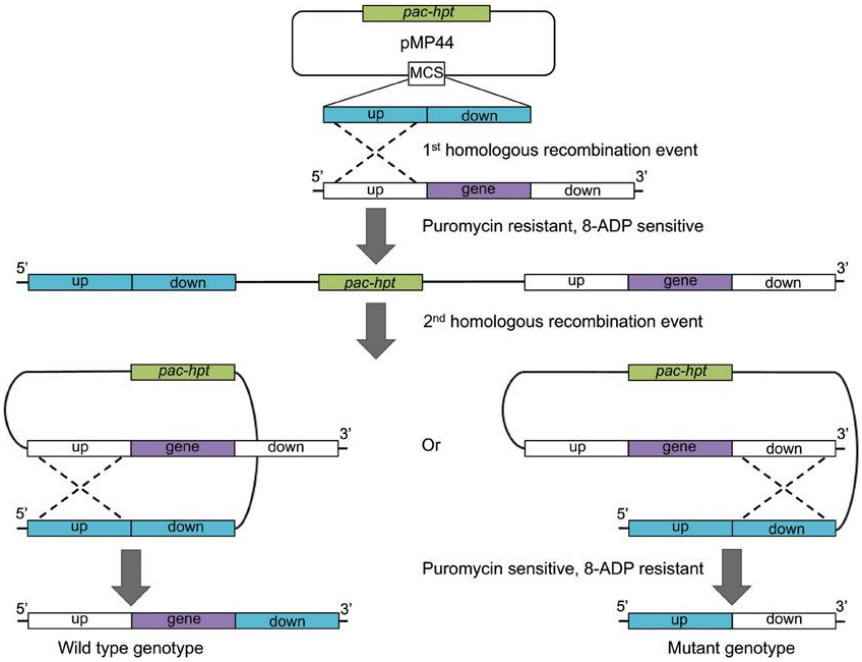

Figure 4 Scheme for the homologous recombination-mediated markerless gene deletion in *Methanosarcina* Δ*hpt* strains. pMP44[35], the non-replicating plasmid generating the desired genomic deletions by the upstream (up) and downstream (down) sequences. *pac-hpt*, the puromycin acetyltransferase gene *(pac)* and the hypoxanthine phosphoribosyl transferase gene *(hpt)* acting as the selective and counterselective marker in the 1st and 2nd round of homologous recombination event, individually. MCS, multiple cloning sites. Adapted from Kohler and Metcalf[6] under the Creative Commons CC BY license.

Similar with this strategy, a Flp recombinase-mediated two-step editing system adopted from yeast[56] was developed for gene disruption in *Methanosarcina* Δ*hpt* mutants, where a DNA fragment with the *pac* and *hpt* gene flanked with two 34-bp *Frt* (Flp recombination sites) was first integrated into the target genomic region. Subsequently, a non-replicating vector carrying *Flp* recombinase gene was introduced into cells to trigger the deletion of the region between the two *Frt* sites generating the desired strains[57,58].

1.2.2.2 φC31 integrase-mediated mutagenesis

Site-directed recombination mediated by recombinase and their recognizing DNA sites had been proven as an efficient approach for chromosomal mutagenesis. A well-established system is the *Streptomyces* bacteriophage φC31-mediated site-specific recombination that had been extensively studied in bacterial cells[59,60]. Unlike Flp recombinase, the φC31 integrase catalyses a unidirectional reaction, meaning that the enzyme cannot repetitively interact with its recognizing site (i.e., *attB* and *attP*) after the recombination, which stabilizes the integration over time. Inspired by this work in other organisms, the φC31 integrase-mediated site-integration system was first established in *M. acetivorans* and *M. barkeri*, which generated many representatively mutant strains and integrating vectors accordingly (i.e., *M. acetivorans* WWM73, *M. barkeri* WWM85, and plasmid pJK026A, *etc.*). Some of these vectors were later retrofitted with the commercial plasmid pNEB193 to generate a series of suicide plasmids such as pNB727 and pNB730 for gene (over)expression in *Methanosarcina* strains[41,43].

1.2.2.3 Transposon mutagenesis

Transposon mutagenesis is an efficient tool for uncovering gene functions of genes in less-characterized organisms. However, only a few host-independent transposons have been explored for methanogens over the years. The first transposon-mediated mutagenesis in methanogens involved generating mutations around the *nifH* gene in *Methanococcus maripaludis*[61]. Later, a modified system involving mariner mini-Himar1 elements and the corresponding Himar1 transposase was developed for random chromosomal mutagenesis in *M. acetivorans*[36]. This system was subsequently used to identify genes involved in formate catabolism (*in vitro*) and DNA uptake in the natural transformation (*in vivo*) in *M. maripaludis*[62,63].

1.2.2.4 CRISPR-mediated mutagenesis

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) systems have emerged as efficient and versatile tools for genome editing. Current CRISPR-

Cas systems are divided into two classes, where the Class 1 system requires multiple Cas-crRNA subunits to form a functional complex, while Class 2 operates with a single Cas-crRNA complex[64]. To date, six types of the CRISPR-Cas systems that have been identified (type I, IV, and III in Class 1, and type II, V, and VI in Class 2), and those from Class 2 had been mostly characterized given their simplicity and broad application potentials[65,66].

Among Class 2 types, the type II CRISPR-Cas9 system has been firstly characterized and applied for genome engineering in eukaryotic cells[67–69]. The Cas9 protein comprises a recognition (REC) lobe and a nuclease (NUC) lobe, which includes the RuvC, HNH endonuclease domains for DNA excision, as well as domains that interact with the protospacer adjacent motif (PAM) (**Figure 5a**). In the Cas9-mediated DNA cleavage, the Cas9 protein firstly recognizes the trans-activating crRNA (tracrRNA): crRNA (including protospacer and direct repeat) duplex, binds to the complimentary DNA strand after recognizing the guanine (G)-rich PAM sequence at the 3´end of the protospacer and then cleaves the double-stranded DNA (dsDNA) generating double strand breaks (DSB) with blunt ends (**Figure 5b**). Later, the type V CRISPR-Cas12a (formerly known as Cpf1) system requiring only RuvC domain for dsDNA excision has been reported for robust genome editing[70,71]. Different from Cas9-system, the Cas12a endonuclease recognizes thymine (T)-rich PAM sequence at the 5´end of the protospacer and only requires crRNA (without assistance of tracrRNA) to generate staggered cleavage sites[70,72–74].

Over the years, CRISPR-Cas systems have been extensively studied in eukaryotes and prokaryotes, whereas few research has been focused on the applications in archaeal domain. The first Cas9-mediated genetic tool was developed for genome editing in *M. acetivorans*[38]. Following this, the truncated Cas9 (dCa9)-mediated CRISPR interference (CRISPRi, elaborated in Section 1.2.3.3), was established for gene regulation in this organism[37]. Additionally, relevant Cas9-, and Cas12a-mediated genetic toolbox had been established for genome editing in another wildly studied *Methanococcus* species[75–77], which greatly advanced metabolic engineering applications in methanogens.

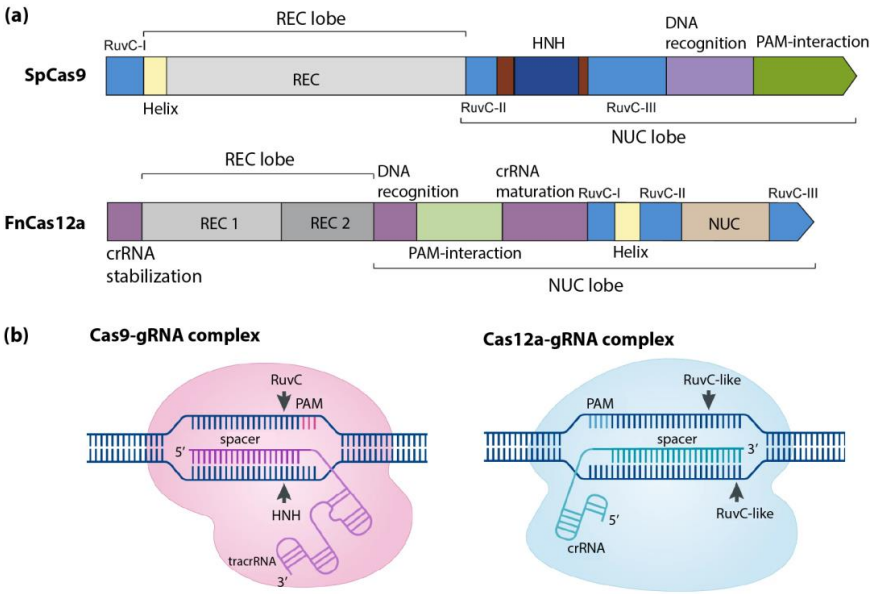

Figure 5 The comparison of the Cas9-, and Cas12a-mediated gene editing systems. **(a)** The compartment of the *Streptococcus pyogenes* (Sp) Cas9, and *Francisella novicida* (Fn) Cas12a endonuclease. REC lobe, recognition lobe. NUC lobe, nuclease lobe. RuvC, and HNH, endonuclease domains for DNA cleavage. PAM, protospacer adjacent motif. crRNA, CRISPR RNA including spacer and direct repeat sequence. Adapted from Paul and Montoya[78] under Creative Commons CC BY license. **(b)** The working mechanism of Cas9-, and Cas12a-gRNA complexes. Reprinted with permission from Pickar-Oliver and Gersbach[79]. Copyright 2019 Springer Nature Limited.

1.2.2.5 Chromosomal integration sites for efficient mutagenesis

Metabolic engineering often requires reprogramming intracellular pathways and introducing heterologous genes to improve biomass production. While gene expression can be achieved by using self-replicating plasmids, these constructs tend to become unstable when they are in large size hosting multiple genes. Additionally, the continuous use of antibiotics to maintain plasmid stability adds excessive labour and cost, making this approach less suitable for large-scale or long-term applications. Alternatively, integrating genes directly into the genome offers greater stability and reduced maintenance. Furthermore, the development of CRISPR-Cas genetic editing tools has made genomic gene integration more efficient and accessible in many organisms.

To identify suitable chromosomal integration sites for gene expression, several factors have been considered for future applications. First, integration sites should not disrupt auxotrophic genes, as this could impair cell growth. Meanwhile, these locations should maintain consistent gene expression levels. Finally, high integration efficiency is preferred, particularly when using CRISPR-Cas genetic tools to facilitate the process. To this end, studies exploring neutral integration sites for gene expression have been successfully evaluated in model organisms across

eukaryotes, such as *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Aspergillus oryzae*[80–82] as well as in prokaryotes like *E. coli* and *Bacillus subtilis* strains[83,84].

Systematic research into neutral integration sites for methanogenic archaea remains limited. Although some endogenous gene sites have been occasionally reported in biochemical research[85,86], their suitability as neutral integration sites for consistent gene expression and integration efficiency has yet to be studied.

1.2.3 Regulation of gene expression

Regulating gene expression in methanogens presents challenges due to their distinct biological characteristics and the relatively limited availability of genetic tools compared to bacteria and eukaryotes. Despite these hurdles, several strategies have been developed to regulate gene expression in methanogenic species, improving the ability to manipulate these organisms for research and biotechnological applications.

1.2.3.1 The transcriptional regulatory tools

Transcription is a process that involves RNA polymerase and associated factors for its initiation, elongation, and termination. Due to the different mechanisms of gene regulation, transcription factors and the corresponding promoter elements differ significantly across biological domains. The archaeal transcription initiates by the eukaryote-like RNA polymerase (RNAP)[87] binds to the two transcription factors, namely transcription factor B (TFB) and TATA binding protein (TBP) targeting the promoter elements TFB-recognition element (BRE), and the TATA box, respectively. The transcriptional start site (TSS) is the position where transcription begins, marked by the separation of the DNA strands to allow RNAP to bind to the template strand and initiate RNA synthesis[88,89] (**Figure 6a-c**).

Constitutive promoters drive continuous gene expression in host organisms. Regulation through these promoters has been achieved by engineering promoter motifs to generate variants with different expression strengths. In methanogens, promoter libraries with a range of strengths have been developed by introducing site-directed mutations to key elements such as the BRE, TATA-box, and ribosome binding site (RBS) in *Methanosarcina* and *Methanococcus* species[90,91]. To mitigate the risk of unwanted recombination caused by the high similarity of site-mutated promoters, later research shifted focus to identifying functional promoters from neighboring species[75,92].

Inducible promoters regulate gene expression by responding to specific stimuli, either activating or repressing gene expression. This dynamic control is valuable for studying gene function or producing proteins on demand. In methanogens, the tetracycline-regulated promoter system has been widely used for gene regulation. In this system, an engineered strain expresses the TetR protein, which binds to the

tetO operator-containing promoter to repress gene expression. Upon the addition of tetracycline, TetR binds to the substrate and releases the promoter, allowing transcription to proceed[41,93,94]. Due to the cost of induction in the TetR/*tetO* system, subsequent studies investigated nitrogen-dependent (*nif*) and temperature-dependent regulation of flagella (*fla*) promoters[95,96]. However, these promoters are not widely used due to their broad genomic effects. More recently, a phosphate-responsive promoter from the ATP-binding cassette transporter (*pst*) operon was evaluated in *M. maripaludis*[97]. The *Ppst* promoter and its derivatives have been reported as useful tools for Pi-induced heterologous gene expression in this organism.

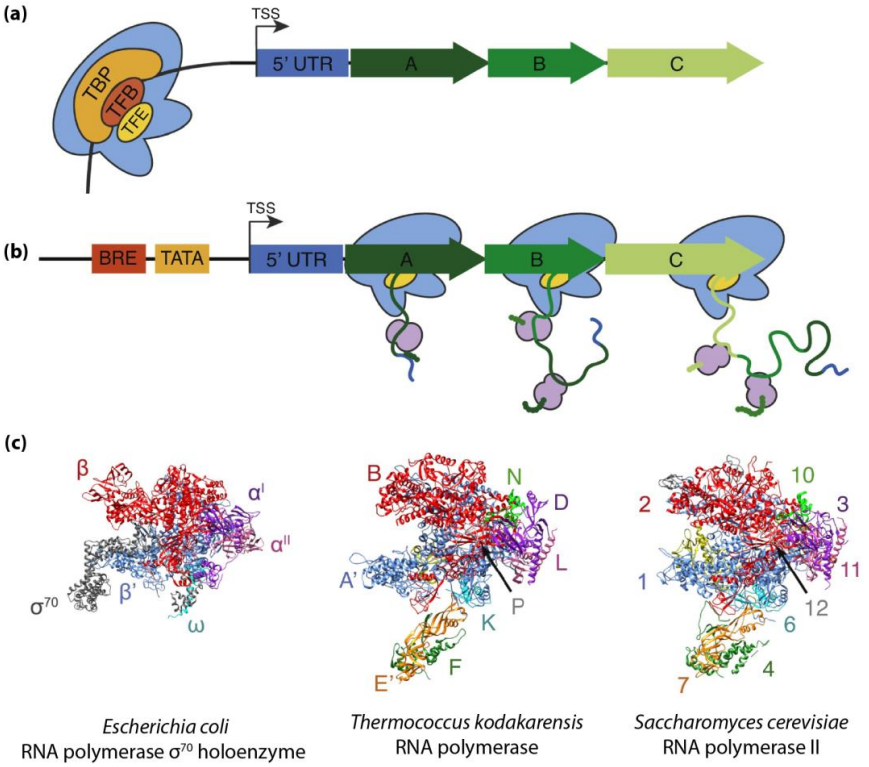

Figure 6 The transcription process in methanogenic archaea. **(a)** Transcription factors in methanogens. RNAP (light blue), RNA polymerase. TBP (orange), TATA binding protein. TFB (orange-red), transcription factor B. TFE (yellow), the non-essential transcription factor E enhances the RNAP binding efficiency. TSS (black arrow), transcriptional start site. 5´UTR (dark blue), 5´untranslated region in mRNA. A, B, and C, three coding sequences in a polycistronic mRNA. **(b)** Polycistronic mRNA structure. BRE, TFB-recognition element. TATA, TATA box. **(c)** RNA polymerase structure in bacteria (i.e., *Escherichia coli*), archaea (i.e., *Thermococcus kodakarensis*), and eukaryote (i.e., *Saccharomyces cerevisiae*). Reprinted with permission from Shalvarjian and Nayak[98]. Copyright 2021 Elsevier Ltd.

1.2.3.2 Translational regulatory tools

Translation is a process of synthesizing proteins from an mRNA template through the coordinated actions of ribosomes, tRNAs, and various translation factors. Like bacteria and eukaryotes, archaeal translation is divided into initiation, elongation, and termination. Among these stages, initiation is critical and varies across domains. In archaea, the translational machinery exhibits feature of both bacteria and eukaryotes. Specifically, the mRNA structure is bacteria-like, containing either RBS (Shine-Dalgarno sequence in bacteria) for small ribosomal subunit recognition or being leaderless, while the translation factors resemble those have been reported in eukaryotes[99].

In methanogens, translation-mediated gene regulation has been achieved through the construction of RBS libraries, either by discovering new RBS sequences with various strengths or through RBS engineering. Factors such as the RBS motif, the adjacent sequence, the distance between the RBS and the start codon, and the RBS-influenced mRNA secondary structure have been found to significantly affect protein expression levels in *M. acetivorans* and *M. maripaludis*[75,90–92]. Notably, unlike the leaderless 5´untranslated region (5´UTR)-mediated translation reported in halophilic archaea[100], methanogens such as *Methanosarcina* species typically possess long 5´UTR (i.e., 100-500 bp), suggesting the presence of additional transcriptional regulation[101]. Therefore, the impact of the 5´UTR RNA secondary structure resulting from RBS engineering should be considered during genetic engineering.

1.2.3.3 CRISPR interference (CRISPRi)-mediated gene regulation

CRISPRi is an emerging tool for gene regulation by utilizing catalytically inactive Cas proteins. Instead of inducing double-strand breaks in the genome, these deactivated Cas proteins bind to specific DNA sequences (e.g., promoter or coding regions), physically blocking RNA polymerase and preventing the initiation or elongation of transcription.

Mutations in the active sites of the RuvC (D10A) and HNH (H841A) domains of the Cas9 protein eliminate its endonuclease activity, generating a non-cleaving dead Cas9 (dCas9) for gene regulation[102]. dCas9 binds to either the upstream promoter or the open reading frame (ORF) and the transcription is repressed by blocking RNAP movement. Fusing the dCas9 to corresponding transcriptional repressor or activator had been reasoned as effective and stable[103,104]. Unlike Cas9, the Cas12a protein only requires its RuvC domain for cleaving the target strand and non-target strand[70]. Introducing site mutations in the RuvC domain (D917A, E1006A) generated the dead dCas12a (formerly known as dCpf1)[73,105] for transcription regulation. Additionally, owing to the inherent RNase activity for crRNA maturation, multiplex gene regulation generated by the dCas12a-crRNA array had been developed for pathway evolution in metabolic engineering

applications[106,107]. CRISPRi-mediated gene regulation in methanogenic species, primarily within the *Methanosarcina* and *Methanococcus* genera, was discussed earlier in Section 1.2.2.4.

1.2.4 Reporter proteins

Reporter proteins are essential for investigating gene expression, protein localization, and cellular processes. In archaea, however, selecting suitable reporter proteins becomes challenging due to their variations in codon usage, protein folding, and environmental conditions across domains. For example, although the green fluorescent protein (GFP) is widely used for real-time transient studies in bacteria and eukaryotes, its reliance on O_2 for fluorescence activation limits its application in anaerobic archaea. Despite these challenges, several reporter proteins have been identified for genetic studies in methanogenic archaea, including colorimetric reporter β-glucuronidase (UidA/GUS), β-galactosidase (LacZ), and β-lactamase (Bla), fluorescent reporter (e.g., mCherry), and fluorescence-activating protein such as FAST (**Table 1**). These discoveries pave the way for further exploration of suitable reporter proteins in anaerobic organisms, enhancing our understanding of their biological processes.

Table 1 The reporter system developed in methanogens.

Reporter	Representative genus	Pros/Cons	Example
UidA/GUS	*Methanosarcina, Methanococcus*	*Pros.* a. Non-toxic b. Quantitative measurement *Cons.* a. O_2 required for activation b. Endpoint assay, real-time measurement not applicable	108–110
LacZ	*Methanosarcina, Methanococcus*	*Pros.* a. Non-toxic b. Quantitative measurement *Cons.* a. Optimal pH 3.5–5.5, unclear stability b. O_2 required for activation c. Endpoint assay, real-time measurement not applicable	90,111
Bla	*Methanosarcina*	*Pros.* a. Stable in high-salt medium b. O_2 not required *Cons.* a. Substrate (nitrocefin) unstable	112

		b. Enzyme excretion required	
mCherry	*Methanococcus*	*Pros.* a. Non-toxic b. Quantitative measurement *Cons.* a. O_2 required for maturation b. Endpoint assay, real-time measurement not applicable	[113]
FAST	*Methanosarcina, Methanococcus*	*Pros.* a. Non-toxic b. O_2 not required c. Real-time measurement applicable *Cons.* a. Cellular autofluorescence b. Reversible, the stability of the fluorescence to be considered	[42,114]

2 Goal of the thesis

The aim of the thesis was to expand the genetic toolkit for *M. acetivorans* to facilitate the manipulations such as genome editing, gene expression, and metabolic pathways engineering in this organism.

The goal of Publication **I** was to establish an effective CRISPR-Cas system for genome editing in *M. acetivorans*. Though the CRISPR-Cas9 had been established, the system has its limitations such as recognizing only Guanine-rich PAM sequences, while *M. acetivorans* genome has high A-T content. Additionally, lower editing efficiency in large DNA fragments deletion was observed in that system, highlighting the needs for alternative genetic editing tools as complements to overcome these obstacles.

The second study was inspired during the construction of the CRISPR-Cas12a system, where a limited number of genomic sites were applicable when assessing Cas12a-mediated gene insertion efficiency. Thus, the aim of the Publication **II** was to identify additional neutral integration sites on *M. acetivorans* genome to facilitate heterologous gene expression in this organism. Besides, this study also developed a Golden Gate Assembly-mediated Cas12a-editing plasmid for the efficient construction of large shuttle vectors.

Publication **III** aimed to develop a tool that could fine-tune gene expression for rewiring metabolic fluxes in *M. acetivorans*. Instead of focusing on the commonly studied *mcr* promoters in *Methanosarcina* species, the potential of promoters from various methanogenic species were worth exploring. This approach revealed diverse regulatory mechanisms across methanogens and provided insights for future studies on the physiological functions among methanogenic genera.

3 Material and Methods

This section outlines the materials and methods used in the thesis, including the parental strains and vectors, cloning techniques, as well as analytical methods. Additional information can be found in Publication **I-III**.

3.1 Strains and Plasmids

The parental plasmids utilized in the thesis have been listed in **Table 2**. The derived constructs were prepared by using Golden Gate Assembly, Gibson Assembly, and T4 ligase-mediated ligation techniques. More details for plasmid construction and sequence data have been described in each publication.

Table 2 Parental plasmids used in this study.

Plasmid	Description	Source
pY016	Cas12a gene sequence source from *Lachnospiraceae bacterium* (*LbCas12a*)	[70]
pWM321	*E. coli-Methanosarcina* shuttle vector	[33]
pM001	pWM321-derived vector, *ColE1* origin of replication (high-copy-number from *E. coli*), *pac* and and *hpt* gene cassette	Publication **I**
pNB730	*pUC19 ori, bla*$^{+}$*, lacZ*$^{+}$, φC31 *attB*, PmcrB_*M. voltae pac (opt)*, *strep-his MCS his-strep* expression cassette	[43]

The parental strains used in the thesis are listed in **Table 3**. The parental strain WWM 73 was utilized as the wild-type (WT) strain in all the publications. The parental *E. coli* strains were obtained from the respective commercial sources. Detailed information and their derivatives could be found in the corresponding publications.

Table 3 Parental strains used in this study.

Strain	Properties	Source
E. coli		
DH10B	F– *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara-leu*)7697 *galU galK* λ– *rpsL* (*Str*^{R}) *nupG*	New England Biolabs

XL10-Gold	Tet^r Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [*F´ proAB lacI*^{q}*Z*Δ*M15* Tn*10* (Tet^r) Amy Cam^r]	Agilent Technologies
M. acetivorans		
WWM73	Δ*hpt*::P*mcrB-tetR*-φC31-*int*-attP	[41]

3.2 Plasmid transformation

Chemically competent *E. coli* cells were utilized for cloning shuttle vectors. PEG 4000-mediated transformation for DNA uptake in *M. acetivorans* was carried out following a modified version of the protocol from previous study[49] with modifications: a) 10 mL of *M. acetivorans* cells were grown to the mid-to-late exponential phase (OD600 = 0.6-1.0) and harvested by centrifugation at 3,200× *g* for 10 min. b) remove the supernatant and resuspend the cell pellet gently with 5 mL Transformation buffer (TB, pH 7.4) and then centrifuge at 3,200× *g* for 10 min. Meanwhile, mix proper amount of DNA with 500 µL TB-buffered PEG-4000 solution (pH 7.4). c) discard the supernatant and carefully resuspend cell pellet with 500 µL TB buffer and mix with the PEG-DNA mixture. d) incubate the mixture for 4 hours at room temperature in anaerobic chamber, after which inoculate the entire volume of mixture slowly into 10 mL HS/Na_2S ·$9H_2O$/substrate medium and incubate 12-16 hours (overnight) at 37 °C. e) Finally, plate appropriate volume of the overnight culture onto the solid HS medium containing antibiotic (i.e., 2 mg/mL puromycin), and incubate the plates at 37°C for 12–15 days.

3.3 Strain cultivation

The *E. coli* strains were cultivated in Luria-Bertani (LB) medium supplemented with 100 µg/mL ampicillin for antibiotic selection. The HS medium for single-cell *M. acetivorans* was prepared according to a previous protocol[44]. For cultivation, *M. acetivorans* strains were grown in HS medium with either 125 mM MeOH or 50 mM trimethylamine (TMA) as the sole carbon and energy source, with 0.4 mM $Na_2S·9H_2O$ as the sulfur source. Solid HS medium was prepared by adding 1.4% agar (Bacto™ Dehydrated, Fisher Scientific). The plates were incubated anaerobically in sealed jars under an N_2/CO_2/1% H_2S (75/20/5) atmosphere at 37°C for over 10 days to ensure adequate growth.

3.4 Genomic DNA extraction of *M. acetivorans* cells

The genomic DNA extracted for whole genome sequencing (WGS), as described in Publication I, was carried out using the Wizard® Genomic DNA Purification Kit (Promega). For downstream applications such as colony PCR and Sanger sequencing of PCR amplicons, a crude genome extraction method was employed.

In this method, 100 μL of the M. acetivorans culture was centrifuged at 20,000× g at room temperature (RT) for 10 min. After discarding the supernatant, the cell pellet was then resuspended with 300 μL ultrapure/Milli-Q water to lyse the cells via osmotic shock. The mixture was then boiled at 95 °C for 10 min to deactivate DNAase and RNAase activity. Following this, the lysate was centrifuged at 20,000× g at RT for 10 min, and the supernatant was collected as the genomic DNA template for subsequent PCR amplifications. Store the solution at -20°C for long-term usage.

3.5 Counterselection experiment for markerless genome editing

To achieve markerless genome editing in *M. acetivorans*, engineered strains containing Cas12a-editing plasmids were inoculated (0.5% inoculum, v/v) and cultivated for 5-10 generations in HS medium under non-selective conditions (without puromycin) to encourage plasmid loss. The cells were then plated onto solid HS medium containing 20 μg/mL 8ADP (instead of puromycin), and the plates were incubated at 37°C for 10–12 days. Colonies that grew on the counterselection medium, presumed to be plasmid-free, were screened via colony PCR using proper primers, and correct markerless mutants were further validated by Sanger sequencing of the engineered region(s). The primers for colony PCR and PCR amplifications were introduced in each publication.

3.6 UidA enzymatic assay

The UidA enzymatic assay was carried out in the Publication II and Publication III for identifying the integration efficiency and the strength of the promoter-RBS combinations. The details of assessing enzyme activity as well as the standard curve had been described in Publication III.

3.7 *In silico* design and analysis

The gRNAs used for Cas12a-mediated genome editing in Publication **I** and Publication **II** were designed by Benchling website (https://www.benchling.com/) using parameters by default. Sanger sequencing results in this thesis were aligned with SnapGene software. The assessment of the WGS data has been described in Publication I. The details for DNA sequence alignment and RNA secondary structure predictions were outlined in Publication **III**.

4 Results and Discussion

4.1 Construction of a Cas12a-mediated genome editing system

CRISPR-Cas genome editing systems have been widely established in prokaryote and eukaryote. However, the development of these tools has been far behind in archaea domain. In this chapter, a Cas12a-mediated system was developed for genome editing in *M. acetivorans*.

4.1.1 Plasmid construction

To develop the Cas12-mediated gene editing system, a commonly used shuttle vector pWM321[33] was modified. First, vector pM001 was created by replacing the *ori* R6K to the *ori* ColE1 and inserting the *hpt* gene to the downstream of *pac* under the control of *mcr* promoter from *Methanococcus voltae* (*Pmcr*) (**Figure 7a**). The *hpt* gene was used as a counterselectable marker for markerless editing in *Methanosarcina* Δ*hpt* hosts[41]. Next, the Cas12a gene from *Lachnospiraceae bacterium* (*LbCas12a*, simplified as *Cas12a* in the thesis) driven by promoter *PmcrB (tetO1)*[41] was inserted into the MCS region of pM001 to construct the Cas12a-expressing plasmid pMCp4 (**Figure 7b**). Later, a gRNA cassette comprising the promoter and terminator from the *M. acetivorans* methanol-specific methyltransferase (*PmtaC1* and *TmtaC1*) was then added to the downstream of the LbCas12a cassette in pMCp4 for constructing Cas12a-gRNA targeting system (**Figure 7c**). Finally, repair fragments were inserted into the downstream of the gRNA cassette to become the Cas12a-gRNA editing system (**Figure 7d**).

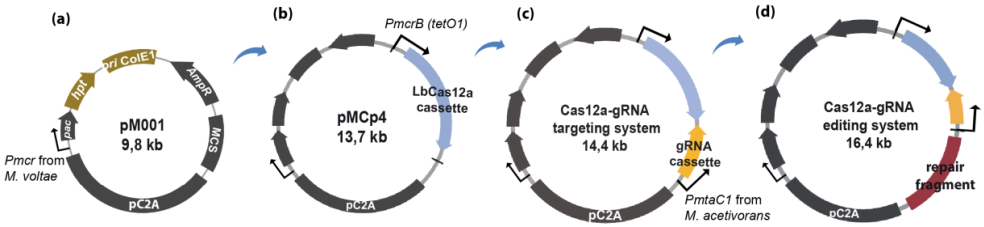

Figure 7 Schematic of the construction of Cas12a-mediated gene editing system. **(a)** pM001, the *E. coli-M. acetivorans* shuttle vector derived from the parental vector pWM321[33]. *hpt* (bronze arrow), hypoxanthine phosphoribosyl transferase gene counterselective marker. *ori* ColE1 (bronze fragment), the origin of replication in *E. coli*. **(b)** pMCp4, pM001-derived vector expressing LbCas12a protein in *M. acetivorans*. LbCas12a cassette (blue arrow), the *Cas12a* gene equipped with promoter *PmcrB (tetO1)*[41] and *mcr* terminator *Tmcr*. **(c)** Cas12a-gRNA targeting system, pMCp4-derived plasmid expressing Cas12a-gRNA complex generating genomic double-strand breaks. gRNA cassette (yellow arrow), gRNA comprising spacer and direct repeat sequences equipped with promoter *PmtaC1* and terminator *TmtaC1*. **(d)** Cas12a-gRNA editing system, pMCp4-derived plasmid expressing Cas12a-gRNA complex and the repair fragment (dark-red) for genome editing. Adapted from Publication **I** under the Creative Commons CC BY license.

To evaluate the impact of the Cas12a expression on cell growth, the plasmid pMCp4 was transformed into *M. acetivorans* through PEG4000-mediated transformation, given the unsuccessful trials of liposome-mediated plasmids transformation were observed at the beginning of the research. The empty vector pM001 was used as the positive control in comparison. Transforming 2 µg of plasmid DNA into *M. acetivorans*, the pMCp4 yielded 34,000 ± 19,000 CFUs of Pur^R transformants, which is comparable to the positive control ($p = 0.75$ in two-tailed *t-test*) (**Figure 8a**). Additionally, growth curves of the wild-type (WT) strain, the pM001-containing strain, and the pMCp4-containing strain were analysed, and no significant difference was observed (**Figure 8b**). These findings demonstrate the safety of utilizing Cas12a-mediated gene editing system in *M. acetivorans*.

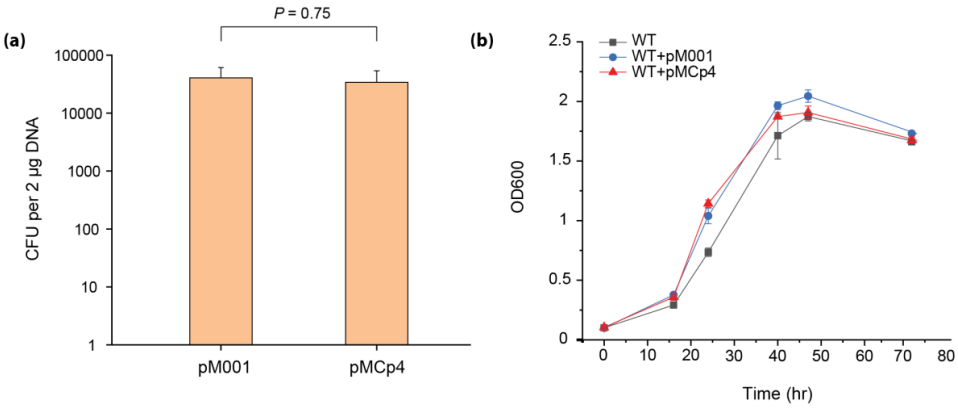

Figure 8 Growth effect of the Cas12a protein expression on *M. acetivorans* WWM73. **(a)** Plasmids transformation efficiency. pM001, the backbone of the shuttle vector served as the positive control. pMCp4, pM001-derived plasmid expressing Cas12a in *M. acetivorans*. **(b)** Growth profile of *M. acetivorans* and its derivatives. WT, wild-type WWM73 strain. Adapted from Publication **I** under the Creative Commons CC BY license.

4.1.2 Cas12a-mediated gene deletion

To assess the efficiency of the Cas12a-mediated gene deletion, the non-essential *ssuC* gene encoding a permease subunit of the sulfonate ABC transporter was targeted by a gRNA (namely g1RNA, 5´-TTGCGATTCCCTCAGCCATGCCC-3´). Deletions of varying lengths (i.e., 100 bp, 500 bp, 1000 bp, and 2000 bp) were generated by adjusting the distance of 1 kb in HR sequences (**Figure 9a**). 2 µg of DNA of each construct was transformed into *M. acetivorans*. After 15 days of incubation, highest yield of transformants (1,800 ± 600 CFUs) was observed with the smallest deletion (i.e., Δ100 bp). For deletions with 500 bp or larger, ca. 20% of the yield of Δ100 bp-edits were observed (**Figure 9b**). Ten Pur^R transformants of each deletion trial were randomly selected to confirm the correctness of the editing with the primers in Publication **I**, and over 80% of positive rate was observed (**Figure 9c-d**), indicating high editing specificity. Compared to the previously developed Cas9-based gene editing tool[38], our Cas12a-toolbox demonstrated higher efficiency in deleting larger DNA fragments (i.e., 1000 bp, and 2000 bp).

To investigate the upper limit of the gene deletion, a separate deletion trial was conducted. Here, the previously proved silent *frh* operon (ca. 3500 bp) encoding coenzyme F420 hydrogenase[109] was targeted by a single gRNA (g9RNA, 5´-CATCCCCGGGACGGAGACGAAGA-3´) (**Figure 9e**). Surprisingly, this attempt yielded a higher number of Pur^R transformants (2,800 ± 1,400 CFUs) than the 2000 bp-deletion trial, and with 100% deletion efficiency (**Figure 9f**). These results indicate that the Cas12a-mediated gene deletion performs efficiently especially in large fragments deletion, although the actual deletion efficiency may vary depending on the essentiality of the genes being targeted.

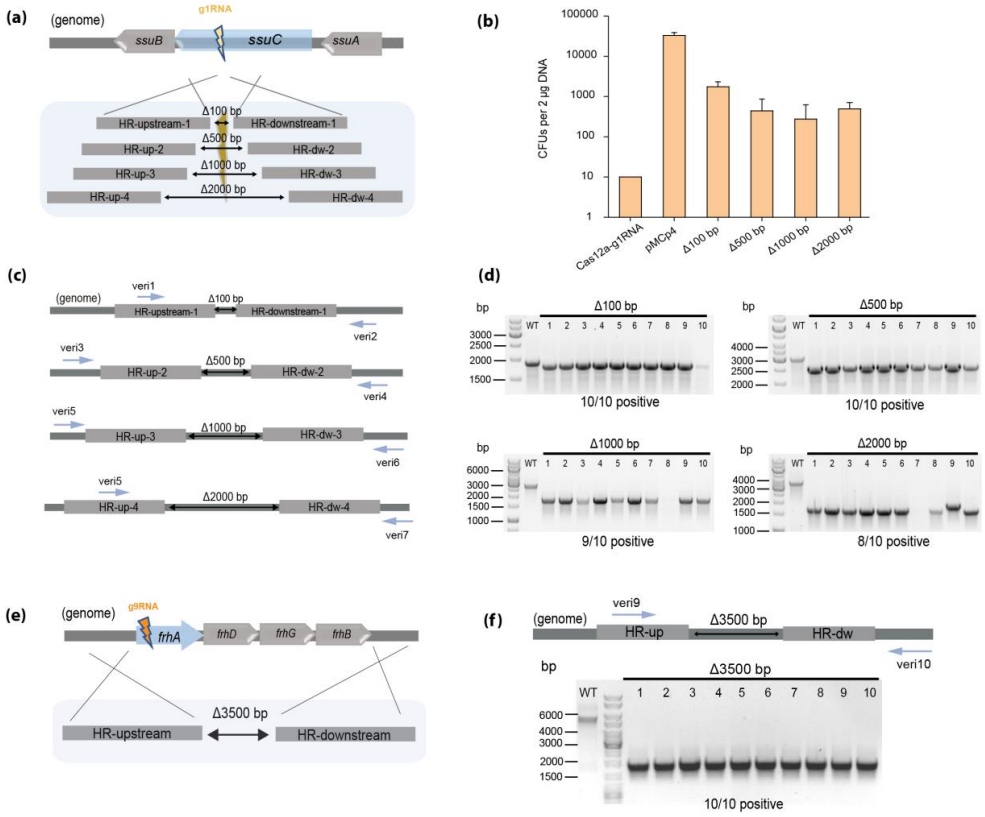

Figure 9 Cas12a-mediated gene deletion. **(a)** Schematic of generating varying-size deletions on the genomic *ssuC* locus. g1RNA, gRNA sequence targeting *ssuC*. Δ100 bp, Δ500 bp, Δ1000 bp, and Δ2000 bp, constructs generating 100 bp, 500 bp, 1000 bp, and 2000 bp on genome. **(b)** The transformation efficiency of the gene deletion with different size of DNA fragments. Cas12a-g1RNA, plasmid expressing Cas12-g1RNA complex targeting *ssuC* and generating DSB. pMCp4, plasmid expressing Cas12a serving as positive control. **(c)** Primers used for screening positive Pur^R transformants in *ssuC* gene deletion trials. veri1-veri7 are primers reported in Publication **I**. **(d)** The efficiency of Cas12a-mediated *ssuC* deletion. WT, wild-type strain served as the control. **(e)** Schematic of the deletion of *frh* operon. g9RNA, gRNA sequence targeting *frhA* locus. **(f)** The efficiency of Cas12a-mediated *frh* gene operon deletion. veri9 and veri10, primers used for PCR verification below, and the sequences were reported in Publication **I**. WT, wild-type strain served as the control. Adapted from Publication **I** under the Creative Commons CC BY license.

4.1.3 Cas12a-mediated gene insertion

Gene insertion is one of the requirements in genome engineering. To evaluate the efficiency of the Cas12a-mediated gene insertion, the *uidA* expression cassette comprising the *mcr* promoter *PmcrB* and terminator (*Tmcr*) from *M. barkeri*, the *uidA* gene from *E. coli* BL21 genome was inserted into *ssuC* region, where the gene was interrupted by the g1RNA in gene deletion trial (**Figure 10a**). The control plasmid pMCp4 yielded 117,000 ± 12,000 CFUs, while the insertion plasmid yielded 9,000 ± 2,000 CFUs of the Pur^R transformants. Verification of randomly

picked colonies exhibited 100% of insertion efficiency (20/20) (**Figure 10b**) with no heterozygous detected. Overall, the results proved that the Cas12a-based system enables efficient gene insertion.

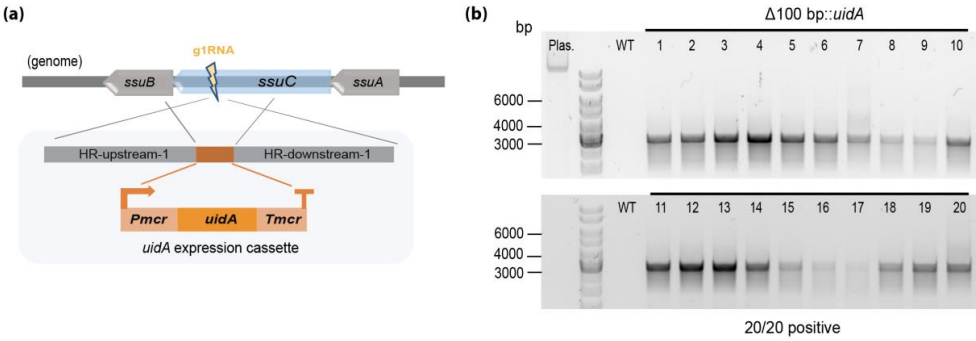

Figure 10 Cas12a-mediated gene insertion. **(a)** Schematic of inserting the *uidA* expression cassette (orange) into genomic *ssuC* locus. g1RNA (yellow sharp mark), the gRNA sequence targeting on the *ssuC*. HR-upstream-1 and HR-downstream-1, the upstream and downstream homologous recombination arms generating 100 bp-deletion in g1RNA-targeting region. **(b)** The insertion efficiency of the Pur^R transformants. Δ100 bp::*uidA*, WWM73 mutants with a 100 bp-deletion and *uidA* expression cassette insertion on *ssuC* site. Pls. and WT, plasmid and the wild-type strain served as the negative control. Adapted from Publication **I** under the Creative Commons CC BY license.

4.1.4 Cas12a-mediated multiplex gene editing

Previous research had demonstrated that the Cas12a endonuclease can drive multiple gRNAs (also called CRISPR array, crRNA array) in a single transcript in eukaryotic cells[115]. To evaluate the precision and effectiveness of the Cas12a/crRNA array-mediated multiplex genome engineering in *M. acetivorans*, two gRNA sequences were designed to simultaneously target the *ssuC* and *frhA* genes. First, a plasmid (pMCp3-g1g9–100 in Publication **I**) containing a cascade gRNA sequence targeting the two genes above was established (g1RNA, and g9RNA in **Figure 11a**) was constructed. The repairing fragments were designed to create approximately 100 bp deletions in each targeted region. Transformation of the plasmid generated 100% of deletion efficiency (8/8) (**Figure 11b**), although a relatively lower number of the Pur^R transformants were obtained (ca. 100 CFUs per 2 μg DNA). Later, another plasmid (pMCp3-g1-uid-g9-100 in Publication **I**) was constructed to achieve different types of edits simultaneously: a gene deletion in *ssuC* and a gene insertion in *frhA* loci. Transformation of this plasmid yielded six Pur^R colonies, with 5/6 were verified with desired edits, suggesting that the multiplex genome editing in both deletion and insertion are manageable by using the Cas12a-mediated editing system in this organism.

Cas12a-mediated multiplex genome editing has proven to be both efficient and straightforward in eukaryotes as the Cas12a protein alone can catalyse the maturation of pre-crRNA sequences. Regarding to the genome mutations, the

Cas12a had been found with better performance in deletions compared to insertions in eukaryotic systems[116]. In our study, the conclusion appears to be applicable in *M. acetivorans*, where higher efficiency and specificity observed in the double-site deletion trails. Despite of the lower specificity of the insertion in a multiplex editing, no evidence assures that Cas9 could be a more desired replacement in this scenario.

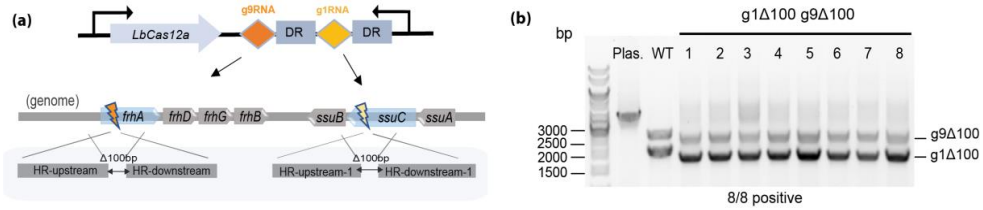

Figure 11 Cas12a/crRNA array-mediated simultaneous two-sites gene deletion and insertion. **(a)** Schematic of simultaneous 100 bp-deletions generated on *ssuC* and *frhA* loci. g1RNA (yellow) and g9RNA (orange) targeting *ssuC* and *frhA*, separately. DR, direct repeat sequence for gRNA maturation and recognition by Cas12a nuclease. **(b)** The double-site deletion efficiency. g1Δ100 g9Δ100, WWM73 mutants with 100 bp-deletions on both *ssuC* and *frhA* locus. Pls. and WT, plasmid and the wild-type strain served as control groups. Adapted from Publication **I** under the Creative Commons CC BY license.

4.1.5 The workflow of the Cas12a-mediated markerless genome editing and the occasional genome translocation concerns

To achieve a markerless genome editing in *M. acetivorans*, a workflow was generated based on the daily research progress (**Figure 12**). First, the Cas12a-gRNA editing plasmids were transformed into *M. acetivorans* using appropriate (e.g., PEG4000-, or liposome-mediated) transformation methods and the cells were plated on solid HS medium containing 2 μg/mL puromycin and incubated at 37°C for ca. 12 days. Next (Day 13-16), Pur^R transformants were selected from solid plates and cultivated in 1 mL HS media with required supplements (e.g., substrate, and sulfur source, *etc.*) for 2 days. Genomic DNA was then extracted from the cultures above, which served as the DNA template in colony PCR to identify positive clones. After that (Day 17-33), the positively edited strains were cultivated in HS media without antibiotic for 2-3 days (See Section 3.5) to occasionally lose the Cas12a-editing plasmids. Cultures were then plated on solid HS plates containing 20 μg/mL 8ADP and incubated at 37°C for ca. 13 days. In the following week (Day 34-40), transformants were picked from plates and cultivated to confirm the absence of the Cas12a-editing plasmids. Meanwhile, the edited genomic regions were PCR-amplified and sent for sequencing to assure the precise edition. When necessary, WGS needs to be performed to ensure the fidelity of the genome editing. The whole process of markerless editing can last ca. 40 days. Once confirmed, the engineered strains are ready for the downstream analyses.

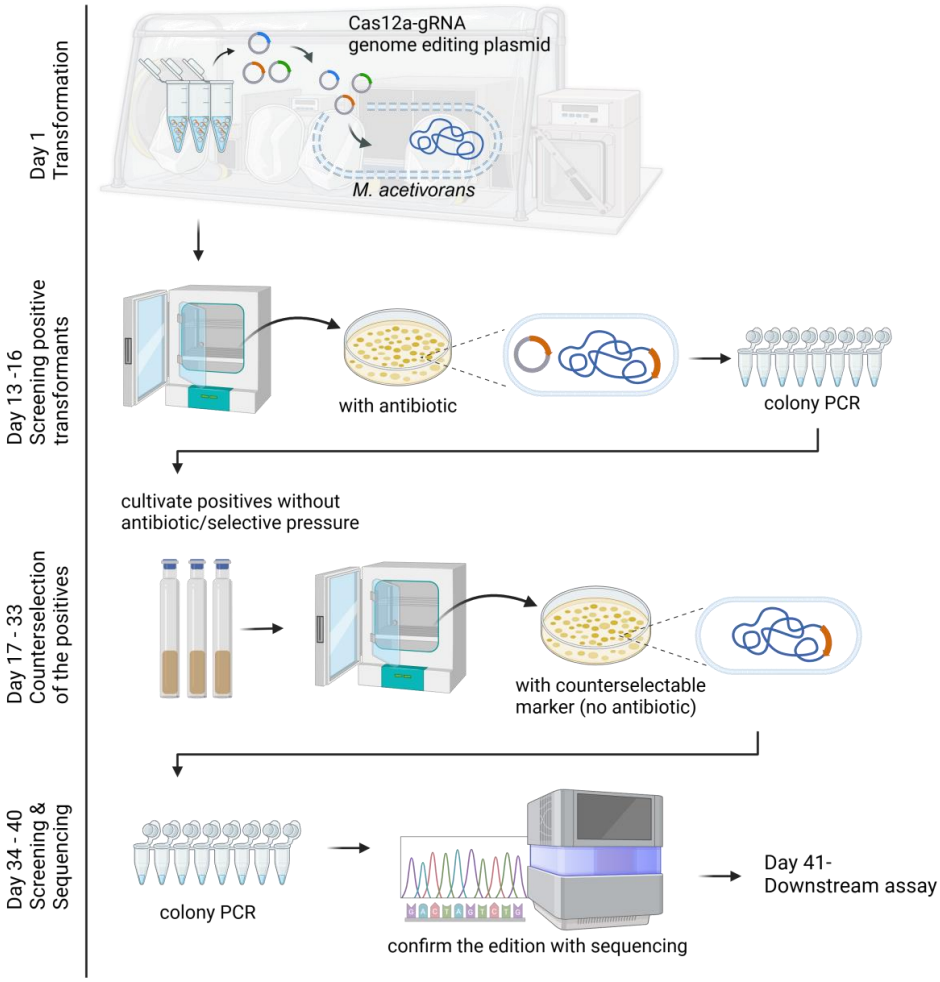

Figure 12 Streamline of a successful Cas12a-mediated markerless genome editing in *M. acetivorans*.

In this study, all the engineered strains were acquired by following the complete workflow, and the WGS was conducted to verify the accuracy of the genome edits. While the intended mutations were introduced in all the tested strains, genome translocations were observed in 3/5 of these mutants. To explain the phenomenon, several hypotheses were proposed and discussed in this study. Our hypothesis suggests that the translocations were triggered by transposable elements in the organism. In *Methanosarcina* species, over 90% of archaeal insertion sequences encoding the transposable elements were found to be active[117], indicating a high potential of genomic translocations occurring in this genus. Among these elements, the TnpB transposon, which typically functions alongside the TnpA transposase, has been identified as the evolutionary ancestor of CRISPR-Cas nucleases, such as the representative Cas12a, and Cas9[118–120]. Therefore, the observed genome

translocations could be the results of TnpA/TnpB complex-mediated dsDNA editing induced by the expression of heterologous Cas12a endonuclease. Furthermore, a recent study in *M. mazei* identified a Cas1solo-dependent translocation, where the native Cas1solo integrase/endonuclease actively translocated an external plasmid onto genome when the plasmid equipped with casposon-recognizing sites[121]. Additionally, stress-induced translocations mediated by this casposon have been demonstrated in this organism, suggesting that similar translocation events could occur in *M. acetivorans* when a Cas1solo-alike element has ever been present. Although attempts to identify such sites in *M. acetivorans* genome were unsuccessful, further investigation of the immunity-related proteins insights into the underlaying mechanisms of the translocation and contribute to assessing the safety of using CRISPR/Cas system for genome editing in *M. acetivorans*.

4.2 Exploration of neutral integration sites for heterologous gene expression

Metabolic engineering in an organism often requires multiple heterologous genes expression, and a collection of neutral integration sites provides choices for this process and facilitates the development of genetic editing. Unlike many studies where neutral integration sites were systematically investigated in bacteria and eukaryotes[80,81,122], few of them have been focusing on methanogenic species due to their lagged behind genetic studies. As a representative organism, *M. acetivorans* possesses the largest genome in archaea domain and holds high potential in biomass production through metabolic rewiring[7,25]. Though non-essential genes such as heterodisulfide reductase (*hdrD2*)[85], CO dehydrogenase (*cooS2*)[86], *ssuC* and monomethylamine methyltransferase isozymes (*mtmCB1*)[38] had been reported in previous physiological and genetic studies, few of them have been characterized as neutral sites for efficient genes integration and expression. In this section, four loci were systematically characterized as neutral integration sites that enable efficient and stable gene expression in *M. acetivorans*.

4.2.1 Screening for neutral integration sites and Cas12a-mediated gene disruption

Four sites had been selected for characterization (**Table 4**), including previously reported *ssuC*, *mtmC2* and the archaeal flagellum gene operons *fla2A* and *fla2B*, which were considered as silent considering no mobility had been detected in *Methanosarcina* genus under lab cultivating conditions[25].

To assess the editing ability of the screened sites, the Cas12a-mediated gene disruption was performed, where three gRNA sequences targeting on each locus were designed by Benchling platform (see Section 3.7) and their targeting

efficiencies were estimated by the Cas12a-gRNA targeting system (**Figure 13a**). The gRNA sequences achieving the highest targeting efficiency in each locus were selected for further genome editing trials (**Figure 13b**).

Table 4 The integration sites assessed in this section.

Gene locus	Location	Description	Source
ssuC	MA_RS00330	permease subunit of the sulfonate ABC transporter	[38], Publication **I**
mtmC2	MA_RS15570	methyltransferase cognate corrinoid protein	[123], Publication **II**
fla2AB	MA_RS16015	B-subunit of the type II flagellum *fla2A* cluster	Publication **II**
fla2BB	MA_RS16090	B-subunit of the type II flagellum *fla2B* cluster	Publication **II**

4.2.2 Editing efficiency of the integration sites

To evaluate the editing efficiency of each integration site, the *uidA* expression cassette (adopted from Publication **I**) flanked with 1 kb homologous repairing arms of each integration site was fused to the targeting plasmids, generating the Cas12a-gRNA editing constructs, namely Δ*ssuC*::*uidA*, Δ*mtmC2*::*uidA,* Δ*fla2AB*::*uidA,* and Δ*fla2BB*::*uidA*. It is worth noting that the construct Δ*ssuC*::*uidA* was adopted from Publication **I**, where the repairing arms generated 100 bp-deletion instead of 50 bp-deletions as other constructs did. The empty vector pMCp2-gX described in Publication **I** was used as positive control in transformation efficiency comparison.

To assess the editing efficiency, 2 μg of each construct was transformed into *M. acetivorans* WWM73 and the Pur^R transformants were counted from solid plates after 12-15 days of incubation. As a result, all the constructs yielded comparable number of Pur^R transformants to the control vector pMCp2-gX, except for the construct Δ*ssuC*::*uidA* with nine-fold less colonies observed (**Figure 13c**). Specifically, the pMCp2-gX, Δ*mtmC2*::*uidA,* Δ*fla2AB*::*uidA,* and Δ*fla2BB*::*uidA* yielded 16,000 ± 2,000 CFUs, 28,000 ± 2,000 CFUs, 33,000 ± 2,000 CFUs, and 33,000 ± 2,000 CFUs per 2 μg DNA, while the Δ*ssuC*::*uidA* yielded 2,500 ± 200 CFUs. The significant decrease of Pur^R transformants observed on the construct Δ*ssuC*::*uidA* might be due to its bigger deletion size generated on the genome, and similar results were observed in the previously developed Cas9-gene editing tool[38] and our Cas12a-study (Publication **I**). As the *ssuC* has been successfully utilized as an integration site for a pair of genes expression[38], our results suggested a high potential of using these sites for expressing heterologous genes in *M. acetivorans*.

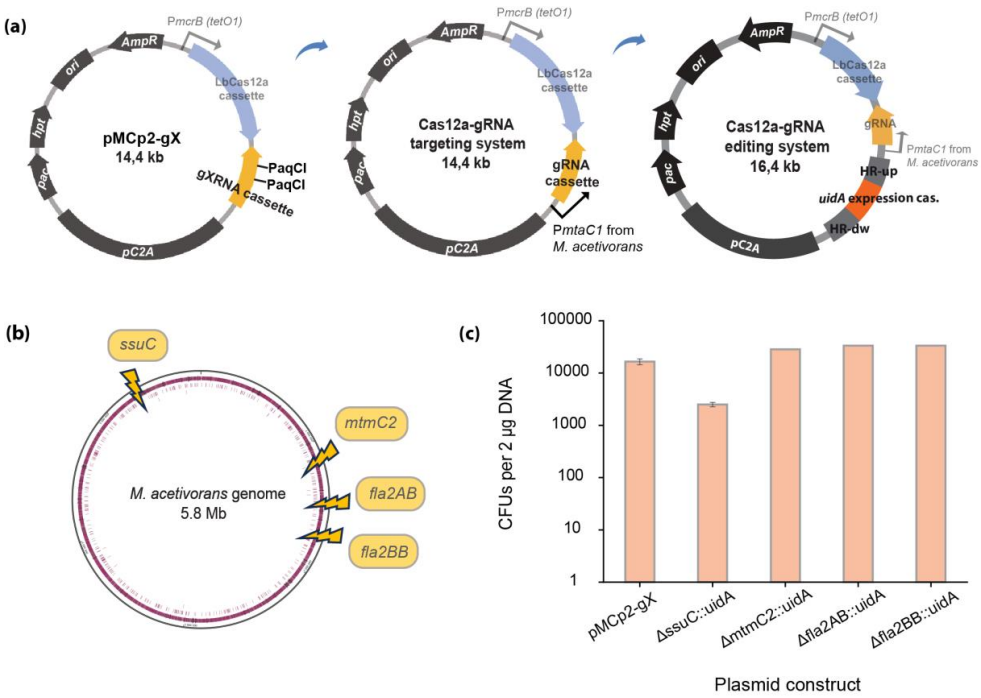

Figure 13 Cas12a-mediated *uidA* expression cassette insertion on the four integration sites. **(a)** Schematic of the genome editing plasmids construction. pMCp2-gX, created in Publication **I**, expresses Cas12a without editing ability served as positive control. Cas12a-gRNA targeting system, expresses Cas12a-gRNA complex generating double-strand-breaks on genome. Cas12a-gRNA editing system, expresses Cas12a-gRNA complex and repairing fragments (HR-up, *uidA* expression cas. and HR-dw) for genome editing. **(b)** The screened integration sites in this study. **(c)** The transformation efficiency of the genome editing plasmids. *ΔssuC::uidA*, *ΔmtmC2::uidA*, *Δfla2AB::uidA*, and *Δfla2BB::uidA*, the plasmid constructs generating disrupting *ssuC*, *mtmC2*, *fla2AB* and *fla2BB* operons and inserting the *uid* expression cassette on these loci, individually. Adapted from Publication **II**.

M. acetivorans has been found to possess an average of 17 genome copies per cell through the varying growth phase when methanol as the sole carbon and energy source, while 3-5 genome copies per cell has been observed when acetate as the substrate[124]. Thus, it is necessary to confirm the existence of the desired edition as well as possible heterozygous genotypes after each editing. With this attempt, ten Pur^R transformants from each edit were selected for colony PCR using two sets of primers (See Supplementary Materials in Publication **II**) to target the inserted *uidA* gene and the deleted region. The positive Pur^R transformants should contain the desired *uidA* gene without deleted region being detected. As a result, all the four sites exhibited high insertion efficiency, where 10/10 of the edits were detected with *uidA* inserted (**Figure 14a-d**). No heterozygous variants were observed in Δ*ssuC*::*uidA*, Δ*mtmC2*::*uidA,* and Δ*fla2BB*::*uidA* edits, except for Δ*fla2AB*::*uidA*, where 1/10 of the edits was found with unedited wild-type genome copies, which indicates a slightly lower editing efficiency on this site .

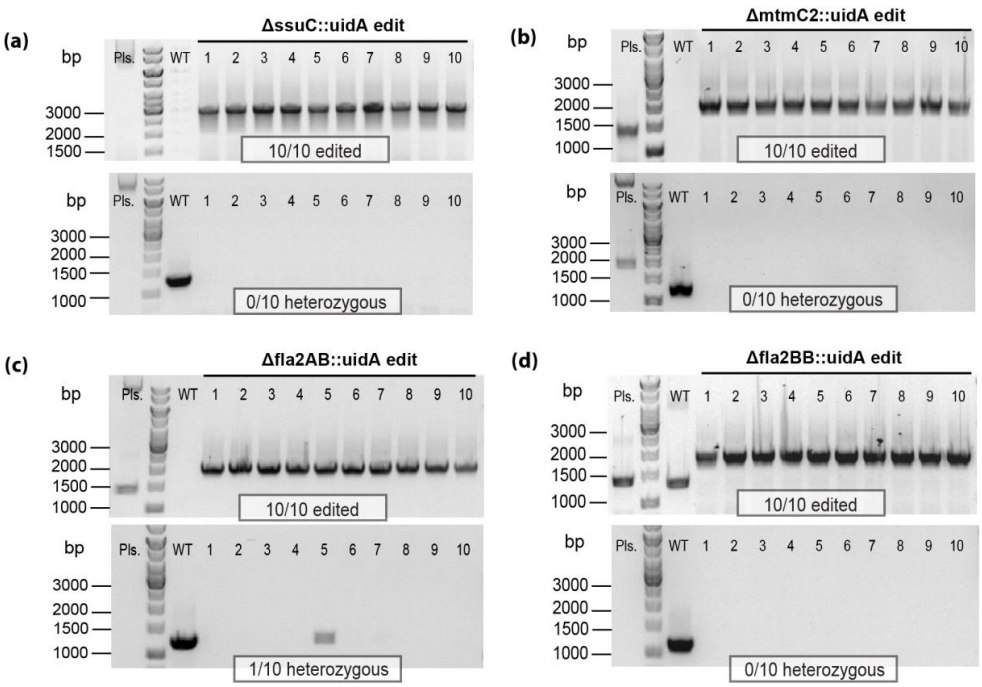

Figure 14 PCR verification of the gene insertion specificity on the four integration sites. **(a)** Δ*ssuC*::*uidA* edit, WWM73 mutants with *uidA* inserted in *ssuC* locus. **(b)** Δ*mtmC2*::*uidA* edit, WWM73 mutants with *uidA* inserted in *mtmC2* locus. **(c)** Δ*fla2AB*::*uidA* edit, WWM73 mutants with *uidA* inserted in *fla2AB* locus. **(d)** Δ*fla2BB*::*uidA* edit, WWM73 mutants with *uidA* inserted in *fla2BB* locus. In each figure: Upper panel, the existence of the *uidA* expression cassette. Bottom panel, the existence of unedited genome copies (heterozygous). Pls. and WT, the genome editing plasmid and the wild-type genome served as control. Adapted from Publication **II**.

4.2.3 Growth profile and the UidA expression of the engineered strains

The gene expression levels are known to be affected by the transcription and translation factors such as promoter, RBS, and upstream/downstream regulators. Effects of the chromosome position on gene expression had been extensively explored in bacteria and yeast[122,125]. Recent study revealed the correlations between archaeal chromosome conformation and gene expression[126]. Considering this, the effect of the integration site on gene expression level needs to be evaluated. Here, a counterselection experiment was performed to remove the Cas12a-editing plasmids and the positive edits with desired *uidA* insertion were identified by colony PCR and sequencing towards the edited genomic areas (**Figure 15a**). Finally, the screened markerless WWM73 edits, namely PZ51, PZ52, PZ53, and PZ54, were utilized for assessing the growth effect and the UidA enzyme activity. Finally, no cell impairment was observed among all the edits when they were cultivated with MeOH as the sole carbon source (**Figure 15b**), suggesting that editing on these

four sites causes no effect on cells phenotype. Meanwhile, comparable enzyme activities ($p > 0.1$ in two-tailed *t-test*) were detected among the edits (**Figure 15c**), which indicates that these locations do not interact in any genomic regulations under current lab cultivation conditions and can be utilized for heterologous genes expression.

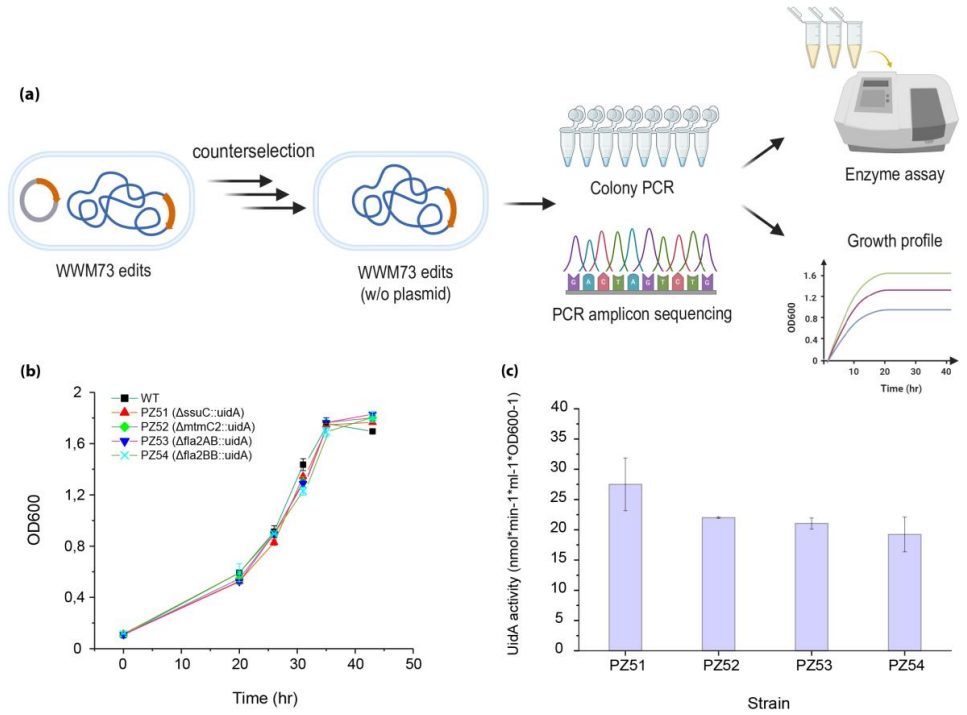

Figure 15 The growth profile and UidA expression level on the integration sites. **(a)** The workflow of preparing the engineered WWM73 edits before the assay. **(b)** Growth curve of the engineered strains (PZ51, PZ52, PZ53, and PZ54) under MeOH condition. WT, wild-type WWM73 strain. **(c)** The UidA activity of the engineered strains. Adapted from Publication **II**.

4.2.4 Golden Gate-based IS-USER construction for large plasmids cloning

In genetic studies of *M. acetivorans*, constructing shuttle vectors for engineering purposes has become an efficient manner. However, most of the shuttle vectors currently constructed in *M. acetivorans* are the derivatives of pWM321 (8,9 kb), where the *M. acetivorans* endogenous self-replicating plasmid pC2A combined with the cloning elements in *E. coli* (i.e., *ori*, selective marker, *etc.*) as well as the *pac* gene acting as the antibiotic resistant marker in *M. acetivorans*[33]. Due to the large size of the parent vector, most of the derivatives possesses over 10 kb in size, causing difficulties and low efficiency in genetic cloning process. In our research, similar problems arose in the construction of the Cas12a-mediated editing

plasmids, where the plasmid sizes can be over 18 kb and extremely low efficiency was observed in the cloning with either classical digestion-ligation or homologous recombination assembly. In contrast, a method called Golden Gate Assembly (GGA) with using the Type IIS restriction enzyme-T4 DNA ligase for seamless assembly has been observed with high efficiency and fidelity. To facilitate effortless genetic cloning for Cas12a-mediated gene editing, an GGA-based vector, IS-USER, has been developed.

This IS-USER vector was derived from the plasmid pMCp4 in Publication **I** where the restriction enzyme BsmBI- (5´-CGTCTC) and PaqCI- (5´-CACCTGC) recognition sites were removed by site-directed mutagenesis. Subsequently, a USER-gRNA cassette and USER-MCS DNA fragments were synthesized and inserted into the downstream of the Cas12a-epression cassette through digestion-ligation. Specifically, the USER-gRNA cassette comprises *mtr* promoter and mtaC1 terminator from *M. barkeri*, the DR sequence and an insertion fragment flanked with BsmBI digestion sites where the gRNA sequence(s) could be fused by GGA with BsmBI to construct the targeting plasmid named USER-Target (Step 1 in **Figure 16**). And the USER-MCS equipped with multiple enzyme digestion sites as well as PaqCI recognition sites, allowing the construction of the genome editing plasmid USER-Edit by GGA with PaqCI (or classical digestion-ligation when necessary) (Step 2 in **Figure 16**). With using this tool, users can easily manage plasmid cloning for Cass12-mediated genome targeting and editing in *M. acetivorans*.

In summary, four integration sites were characterized for heterologous gene expression in *M. acetivorans*. Demonstrated over 90% of editing efficiency and specificity, disrupting these sites had no impact on cell growth, and comparable gene expression levels have been managed, making them optimal targets for gene integration and expression in this organism. The adapted IS-USER vector not only simplifies the genetic manipulation process but also enhances the potential for metabolic engineering applications in this organism. Furthermore, this study shed light on future exploration of neutral integration sites in related methanogenic species.

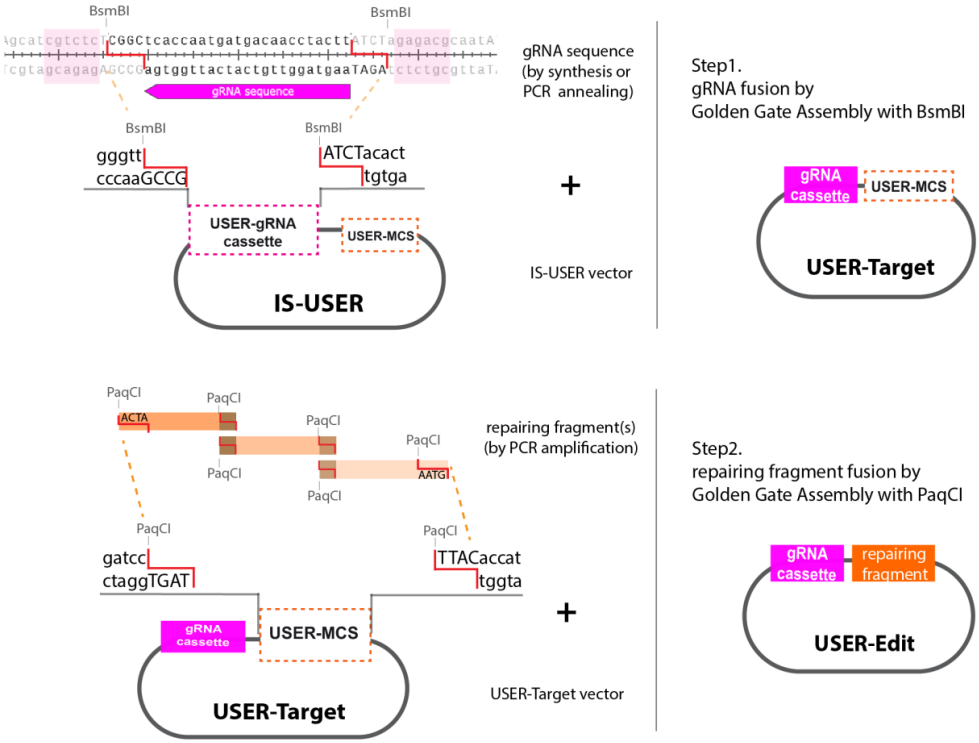

Figure 16 Schematic of using IS-USER vector for constructing Cas12a-medaited targeting and editing plasmids. **Step 1.** USER-Target plasmid cloning for Cas12a-medaited genome targeting. The gRNA sequence fused to the USER-gRNA cassette area on IS-USER by GGA with BsmBI. **Step 2.** USER-Edit plasmid cloning for Cas12a-medaited genome editing. The repairing fragments are assembled and fused to the USER-MCS area of the constructed USER-Target plasmid by GGA with PaqCI. Adapted from Publication **II**.

4.3 Development of a promoter-RBS library for fine-tuning gene expression

Promoter and RBS are the two essential elements in the regulation of genes expression. Over the decades, versatile promoter and RBS libraries have been extensively studied for precise control of gene expression in the model organisms in prokaryotes and eukaryotes[127–130], yet only few studies with various limitations had focused on engineering these elements for adjustable gene expression level in methanogenic species (i.e., *Methanosarcina*[90], *Methanococcus*[75,97], *etc.*). Until recently, a promoter library with a broad range of gene regulation was developed in *M. maripaludis*[91], inspiring the construction of such effective genetic tools in the neighbouring species. In this next section, a promoter-RBS library was established for fine-tuning gene expression in *M. acetivorans*.

4.3.1 Screening wild-type promoter-RBS combinations from other methanogens

Considering the potential instability caused by the repetitive use of the native promoter-RBS combinations in *M. acetivorans*, wild-type promoter-RBS combinations were selected from other methanogenic species and with ca. 300 bp-500 bp in length considering the existence of long 5´UTR in *Methanosarcina* species[101]. To access the strengths of the promoter-RBS combinations, the UidA was utilized as the reporter protein as it has been extensively employed in *M. acetivorans*[85,109,110,131–134]. The *mcr* terminator (*Tmcr*) from *M. barkeri* 227 was fused to the downstream of the *uidA* sequence to create the *uidA* cassette (**Figure 17a**). To construct the *uidA* expression cassette, the promoter-RBS sequences were amplified from methanogenic genomes accordingly and assembled to the *uidA* cassette. The vector pNB730[43] was adopted for integrating the cassette onto *M. acetivorans* WWM73 genome by φC31 integrase-mediated recombination.

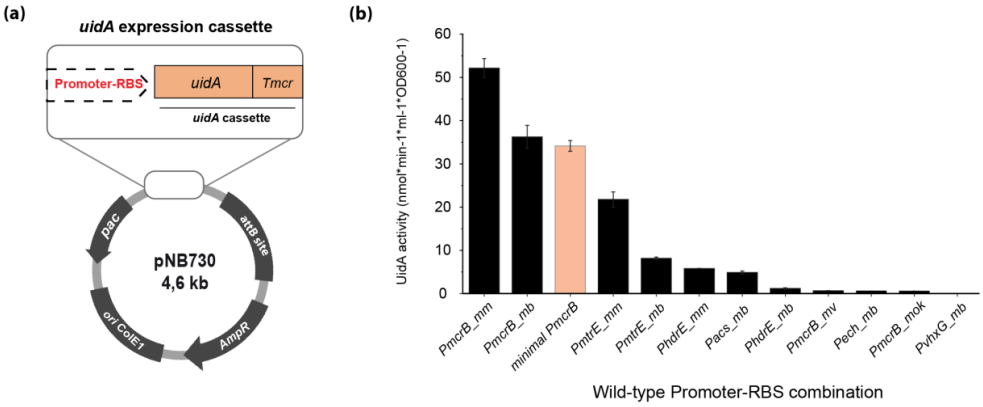

Figure 17 Characterization of wild-type promoter-RBS combinations from various methanogens. **(a)** Construction of the pNB730-*uidA* expression cassette plasmids. *pac*, puromycin acetyltransferase gene. *ori* ColE1, the origin of replication in *E. coli*. *AmpR*, ampicillin-resistant gene. *attB* site, recognition site of the *attP* on WWM73 genome. **(b)** Strength of the screened wild-type promoter-RBS combinations in MeOH. Adapted from Publication **III** under the Creative Commons CC BY license.

Eleven wild-type promoter-RBS combinations were first identified with MeOH as the carbon source. Generally, the promoter-RBS combinations from two *Methanosarcina* species, *M. mazei* and *M. barkeri*, exhibited higher strength than those from far relative *Methanococcus* species, which might be due to the similarity of the transcriptional and translational machinery (e.g., transcription factors, RNAP, and RBS, *etc.*) among the methanogenic species vary from the phylogenetic distance. Especially, the *mcr* promoter from *M. mazei* (*PmcrB_mm*) and *M. barkeri* (*PmcrB_mb*), showed significantly stronger activity than minimal *PmcrB*, which was used as the reference for strength comparison in this study (**Figure**

17b). The *PvhxG_mb* was identified with no activity, which is identical to a previous study[109].

4.3.2 Engineering wild-type combinations for strong-, and medium-strength variants

Although broad strength variance was observed among the wild-type promoter-RBS combinations, it is inadequate to utilize them for fine-tuning gene expression due to the limited number of the strong- and medium-strength combinations. To achieve this goal, three strategies were adopted in this study: 1) engineer the strongest *PmcrB_mm* for promoter-RBS combinations with higher strength than the parent, 2) hybrid the promoter/RBS of the strongest *PmcrB_mm* and those of the weak wild-type combinations to improve their strengths, and 3) explore more wild-type promoter-RBS combinations with medium strengths.

Protein overexpression can be useful in biochemical applications and relevant genetic systems have been widely explored in prokaryotes and eukaryotes. For instance, the T7 RNA polymerase-mediated gene overexpression system has been developed in *E. coli* and *B. subtilis*[135,136], and the inducible GAL1 promoter-mediated gene expression has been effectively applied in yeast[137]. In comparison, fewer genetic tools were developed for this purpose in the model organisms in archaea domain such as *Methanosarcina*. Previous studies had proved the 5´UTR engineering as effective on improving the promoter strength in *Bacillus* species[138]. Inspired by this strategy, here, six variants were created by engineering the 5´UTR (mainly the 30 bp upstream of the start codon) of the strongest *PmcrB_mm* according to the predicted RNA secondary structure and the free energy (**Figure 18a**), and their strengths were identified on MeOH-condition. Compared with the reference minimal *PmcrB*, four variants (*PUTR2*, *PUTR3*, *PUTR4*, and *PUTR5*) were identified as strong, and two variants (*PUTR1* and *PUTR6*) were categorized as medium, while no variants were observed with strength higher than the parent (**Figure 18b**). The failed attempt of seeking stronger promoter-RBS combination might be either due to the less mutants were assessed in this study or the existence of a dynamic balance between the gene expression and the cell fitness caused by the limited cellular resources that hindered proteins expression in this organism, which had been reported in previous research[139].

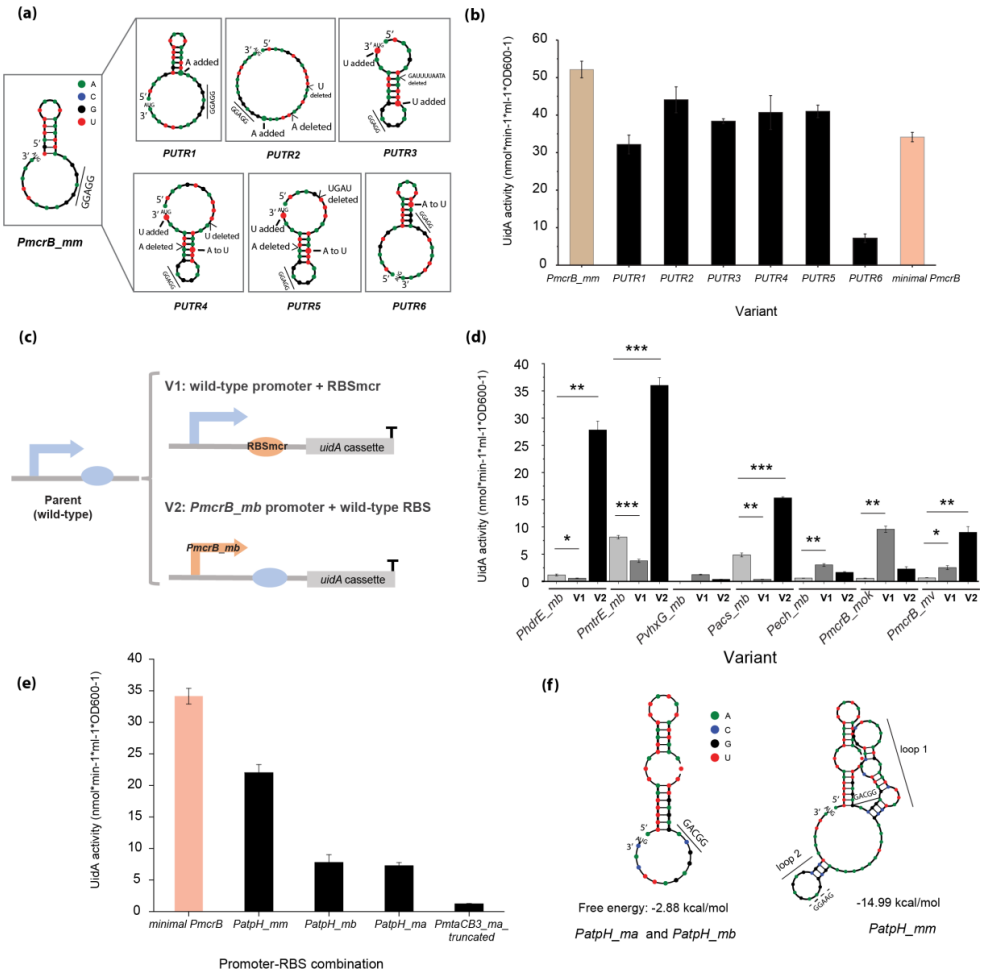

Figure 18 Strategies for expanding the promoter-RBS library for strong- and medium-strength variants. **(a)** Strategy one: 5´UTR engineering of the *PmcrB_mm*. GGAGG, putative *mcr* RBS in *Methanosarcina*. **(b)** The strength of the *PmcrB_mm* variants on MeOH. **(c)** Strategy two: promoter and RBS hybridization. V1 and V2, variant 1 and variant 2. RBSmcr, consensus *mcr* RBS sequence. *PmcrB_mb*, *mcr* promoter from *M. barkeri*. **(d)** The strength of the hybridized variants on MeOH. **(e)** Strategy three: rational evaluation of the wild-type promoter-RBS combinations with medium-strength. **(f)** RNA secondary structure of the *atp* promoter from *M. acetivorans* (*PatpH_ma*), *M. barkeri* (*PatpH_mb*), and *M. mazei* (*PatpH_mm*). GACGG, putative consensus RBS. GGAAG, the second putative RBS. Adapted from Publication **III** under the Creative Commons CC BY license.

To generate additional promoter-RBS combinations with medium-strength, the promoter and RBS region of the seven weak wild-type combinations (i.e., *PhdrE_mb*, *PmtrE_mb*, *PvhxG_mb, Pacs_mb, Pech_mb, PmcrB_mok*, and *PmcrB_mv*) were substituted with the one of strong *PmcrB_mm* (**Figure 18c**), creating fourteen derivatives. Notably, these modifications revealed significant differences in strength. For instance, those V2 comprising parental RBS and *PmcrB*

promoter derived from *PhdrE_mb*, *PmtrE_mb*, *Pacs_mb,* and *PmcrB_mv*, exhibited significant strength increase ($p < 0.05$ in two-tail *t-test*) (**Figure 18d**). These results suggested that their native RBS regions possess high strength, and the low activities of their parental combinations might be due to the wild-type promoter-involved repression in *M. acetivorans*. For better description of the variants, the engineered V2 are named as *PhdrE_mb*-2, *PmtrE_mb*-2, *Pacs_mb*-2, and *PmcrB_mv*-2 in **Figure 19b**. Additionally, the strength of the V1 of *PmcrB_mok* (later named as *PmcrB_mok*-1) showed significant increase when parental RBS was replaced with the strong RBSmcr, which indicates native RBS in the wild-type combination is weak. Interestingly, the *PvhxG_mb*, which was detected with no activity in the previous study, exhibited tracing strengths in both V1- and V2-generated derivatives. Sequence analysis of *PvhxG_mb* revealed the absence of a putative *Methanosarcina* RBS sequence (GGAGG), leading us to speculate that the lack of detectable activity may be due to the loss of RBS region, which negatively impact the gene translation. Overall, the promoter/RBS hybridization strategy successfully collected four variants (i.e., *PhdrE_mb*-2, *Pacs_mb*-2, *PmcrB_mok*-1*, and PmcrB_mv*-2) with medium strength.

To identify additional medium-strength promoter-RBS combinations, various wild-type and derived combinations were explored based on previous studies. For instance, the native promoter-RBS combinations of the *atp* operons encoding archaeal AOA1ATP synthase complex in *M. acetivorans*, *M. barkeri*, and *M. mazei* (namely *PatpH_ma*, *PatpH_mb*, and *PatpH_mm*, respectively) were assessed according to the quantitative transcript analysis of the *atp* operon in *M. acetivorans*[140], where over a 10-fold transcription level was detected compared to the native *hdrED* operon and the close relative *PhdrE_mb* was identified as weak combination in **Figure 17b**. Additionally, a truncated promoter-RBS from the *mtaCB3* operon (namely *PmtaCB3_ma_truncated*) encoding a methyltransferase in *M. acetivorans* was evaluated, as a previous study demonstrated improved performance under both MeOH and TMA growth conditions after truncation[132]. As expected, all screened promoter-RBS combinations exhibited lower strength than the standard *minimal PmcrB* (**Figure 18e**). Specifically, *PatpH_mb*, and *PatpH_ma* exhibited medium strength, whereas the *PmtaCB3_ma_truncated* showed ca. 25-fold lower activity than the standard and thus as a weak combination. Further assessment of the *PmtaCB3_ma_truncated* under TMA condition revealed comparable strength to that under MeOH, which is consistent with previous findings[132]. Notably, *PatpH_mm* displayed over three times the strength of those from neighbouring species. To further explain the results, the secondary structures of the 5´UTR regions of the *PatpH_mm, PatpH_mb*, and *PatpH_ma* were compared (**Figure 18f,** Publication **III**). As a result, an additional 57-bp sequence at the end of the *PatpH_mm* was observed, causing two additional hairpin loops and two putative RBS core regions, which may contribute to the increased mRNA stability and translation efficiency. Such effects had been revealed and discussed in

previous studies in prokaryotes and eukaryotes that the positions and shapes of the 5´UTR stem loops largely influence the stability and efficiency of mRNA structures[141–143]. Considering that non-native promoter-RBS combinations are preferred in our study, the medium strength variants *PatpH_mb* and *PatpH_mm* were collected for further use.

4.3.3 Substrate effect of the promoter-RBS combinations

Through all the trials above, 33 promoter-RBS combinations were collected for fine-tuning gene expression in *M. acetivorans*. To further investigate the substrate effects on these combinations, the MeOH-grown cells were then subjected to a different methylated environment where TMA was used as the sole carbon source. To obtain TMA-adapted cells, the MeOH-cultures were transferred from MeOH- to TMA-HS media and grown for at least 20 generations (over three transfers with 1% inoculum). Prior to the enzyme assay, all the edits were confirmed by Sanger sequencing the corresponding genomic regions to ensure the correct sequences (**Figure 19a**). Meanwhile, an attempt of exploring the effect of the promoter-RBS combinations on acetate, known as a C2 substrate for *M. acetivorans*, was unsuccessful due to the laboriously long shifting time from MeOH cultures. Thus, only methylated environment was discussed in our study.

The promoter-RBS combinations were classified as strong (> 100%), medium (20-100%), or weak (< 20%) based on their strength relative to the minimal *PmcrB* under MeOH conditions (**Figure 19b**), and their performance was further assessed in TMA conditions. Most of the promoter-RBS combinations exhibited comparable and increased expression strength in TMA-grown conditions, except for *PmtrE_mm*, *PhdrE_mm*, and *Pacs_mb*, which showed at least a two-fold decrease when substrate changed to TMA. In contrast, the *PmcrB_mok* showed a ca. five-fold increase in strength in TMA ($P < 0.001$ in two-tailed *t*-test), suggesting the existence of the substrate-related regulations in *M. acetivorans*.

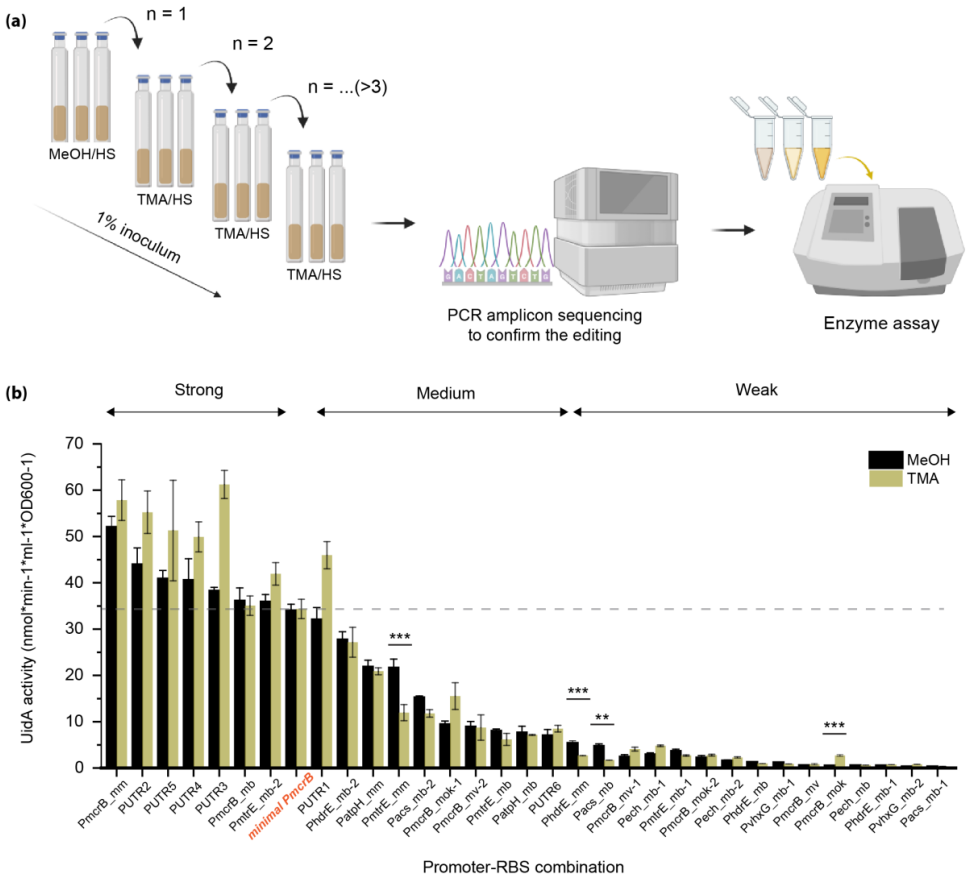

Figure 19 Characterization of the 33 promoter-RBS combinations in methylated conditions. **(a)** Schematic of shifting the carbon sauce of the strains from MeOH to TMA for UidA enzyme assay. **(b)** The strength of the promoter-RBS combinations under MeOH and TMA. Adapted from Publication **III** under the Creative Commons CC BY license.

4.3.4 Accessing the stability of the promoter-RBS combinations in different growth stage

To assess the impact of the promoter-RBS combinations on cell growth, growth curves of the relative strains were compared to the control stain (**Figure 20a-c**, namely PZ0 in Publication **III**) where the *uidA* cassette without any promoter-RBS sequence was integrated into genome by vector pNB730. MeOH served as the sole carbon source. All strains demonstrated comparable growth speed, reaching an OD600 of over 1.7 after 40 hours cultivation. In addition, the expression strengths of the promoter-RBS combinations were evaluated across various growth stages (e.g., the early-exponential (OD600 = 0.3-0.5), mid-exponential (OD600 = 0.5-0.8), and late-to-early stationary phases (OD600 = 0.8-1.2)), as the gene expression levels were found as coordination of the RNA polymerase and ribosome availability

in growing cells[144]. Overall, 23/33 of the promoter-RBS combinations displayed highest activity during the exponential phases, whereas 10/33 showed increasing activity along with the growth, suggesting that the RNA secondary structures of these combinations may be more stable in *M. acetivorans* (**Figure 20d**). Notably, though the expression strength of all the wild-type strong combinations decreased in late-exponential phase, the engineered ones maintained more consistent strength, which indicates that the 5´UTR engineering may contribute to the stability of the RNA secondary structures.

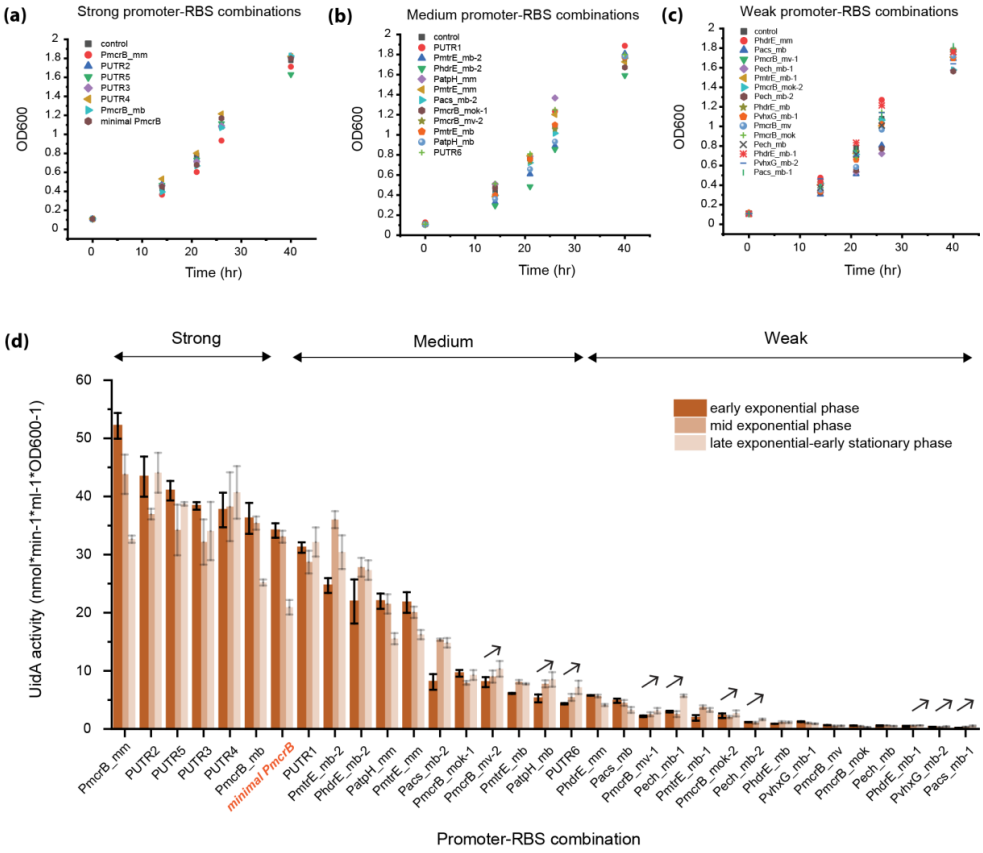

Figure 20 The growth effect of promoter-RBS combinations and their stabilities in different cell growth stages. The growth effect of the **(a)** Strong-, **(b)** Medium-, and **(c)** Weak promoter-RBS combinations on *M. acetivorans* when MeOH used as the carbon source. **(d)** The strength of promoter-RBS combinations in different cell growth stages. early exponential phase, OD600 = 0.3-0.5. mid exponential phase, OD600 = 0.5-0.8. late-to-early stationary phases, OD600 = 0.8-1.2. Adapted from Publication **III** under the Creative Commons CC BY license.

Taken together, this section provides a promoter-RBS library comprising 33 promoter-RBS combinations with expression strength ca. 140-fold of variance for fine-tuning gene expression in *M. acetivorans*. Unlike the libraries derived from

site-mutated promoters, the promoter-RBS sequences in this library share less gene similarity and thus lowers the risk of unintended homologous recombination events during genetic engineering. Additionally, the utilization of the UidA enzyme as the standardized reporter for comparison, our genetic tool provides a framework for the developing similar promoter libraries in other non-model methanogenic organisms.

5 Conclusions and perspectives

5.1 Conclusions

Genetic manipulations of methanogens are essential for studying their physiology and understanding their functions in natural evolution. Unlike the genetic manipulation methods that have been well-established in prokaryotes and eukaryotes, progress in genetic tools development for methanogens still faces a few challenges due to their demanding cultivation conditions, the relatively slow growth rate, as well as the lag-behind genomic studies[5,6]. Over the decades, extensive efforts have been made to expand the genetic toolbox for methanogens, yet progress remains limited. To bridge this gap, three additional genetic tools have been established in this thesis to enhance the current genetic toolbox for *M. acetivorans*. These tools enable precise manipulations for advancing metabolic engineering and understanding the diversities of methanogenesis in this organism.

First, a Cas12a-mediated genome editing system has been developed and assessed for gene deletions, gene insertions and multiplex gene editing. Compared to the previously developed Cas9-editing system, the Cas12a-system demonstrates advantages in deleting large DNA fragments (i.e., over 1,000 bp) and performing multiplex genes deletions using Cas12a-crRNA array. This improvement allows future studies to select either of these CRISPR/Cas systems for genome editing based on their specific needs. Additionally, a streamlined markerless genome editing procedure has been described for homologous recombination-mediated genomic manipulations in *M. acetivorans* Δ*hpt* strains, which offers insights for the metabolic engineering applications in this organism.

Stable heterologous gene expression requires carefully selected genomic integration sites. Towards this goal, four non-essential genomic loci were characterized and confirmed as suitable integration sites. These sites support efficient gene insertion onto the genome, enable genes to be functionally expressed and do not impair cell growth, holding a high potential for future metabolic pathways construction. Furthermore, an efficient GGA-based IS-UER vector was established to alleviate the cloning efforts in large plasmids construction. This vector can serve both as a backbone for Cas12a-mediated genome editing and as a shuttle vector for other genetic applications.

Lastly, a promoter-RBS library was constructed for fine-tuning gene expression in *M. acetivorans*. The library comprises 33 promoter-RBS combinations which

includes both wild-type combinations from the neighbouring species and the engineered edits. Unlike libraries characterized based on episomal plasmids that may introduce instability and inaccuracy in strength, the combinations in this thesis were evaluated by genomic integration, ensuring accurate strength measurements for genome-level engineering. Moreover, exploration of neighbouring instead of native promoter-RBS combinations reduces the unexpected instability caused by the high-similarity in gene sequences. And the engineered combinations with varying RNA secondary structures revealed promiscuous mechanisms of transcriptional regulation. Overall, this library offers a valuable tool for regulating metabolic flux, as well as for the development of synthetic biology in *Methanosarcina spp*.

5.2 Perspectives

Possessing versatile metabolic pathways, *M. acetivorans* has gained interest as the chassis for generating value-added product owing to its CO_2-reducing potentials. Genetic tools such as CRISPR/Cas systems, as a result, are essential in exploring the metabolic versatility[145], utilizing unconventional substrates[27,146], increasing methane yield[147] and producing high-value products in this organism[7,28]. Despite the Cas12a-genome editing system developed in this thesis, future research can also focus on optimizing such tools based on the current progress. For instance, the genome translocation issue observed in the Cas12a-mediated edits (discussed in Section 4.1.5) deserves further investigation. With several hypotheses proposed, future studies could explore the mechanisms of archaeal transposase-transposon complexes, such as TnpA/TnpB complexes, as well as the unidentified casposon-like elements in *Methanosarcina* species. Such research could pave the way for the development of novel genome editors, which further expands the genetic toolbox. Additionally, since the Cas12a protein facilitates to maturation of pre-crRNA, the multiplex gene repression achieved by dCas12a-crRNA array has proven to be efficient in bacterial transcriptional repression[148,149]. Such dCas12a-mediated gene interference in *M. acetivorans* can be further investigated.

The identification of integration sites for gene expression in this thesis not only provided four neutral genomic sites for future genetic engineering, but also offered a general strategy for identifying potential integration sites in other archaeal species. Despite these sites were successfully edited by the Cas12a-mediated genome insertion, constructing the Cas12a-editing systems proved to be a bottleneck in the research. The challenge came from the large size (ca. 14kb) of the empty vector, which led to low assembly efficiency. To address the issue, an IS-USER vector was designed for future plasmid construction, where the digestion sites for the Type IIS restriction enzymes were modified, allowing the Golden Gate Assembly method to be used in fusing the gRNA sequences and the interested gene operons. Future CRISPR/Cas-mediated genome editing attempts can benefit from

this construction to prepare the Cas12a-editing system with less labour and more efficiency.

During the development of the promoter-RBS library, exploration of promoter-RBS combinations in *M. acetivorans* through the hybridization of promoter and RBS regions generated various combinations with unexpected strengths. Though potential transcriptional regulatory factors were discussed in Publication **III**, the precise mechanisms underlying these regulation effects remain unclear. Previous research has identified regulatory mechanisms in methyltransferase operons[123,132]. The approach like promoter truncation could be possibly applied to the promoter-RBS combinations in this thesis to uncover the potential regulators (repressors and/or activators) and their mechanisms in *M. acetivorans*. This could be achieved by engineering the full-length promoter sequences and mutating the 5´UTR regions. Studies like these could further contribute to the development of inducible promoters in *Methanosarcina* species.

References

1. Asakawa, S., Akagawa-Matsushita, M., Morii, H., Koga, Y., & Hayano, K. (1995). Characterization of *Methanosarcina mazei* TMA isolated from a paddy field soil. *Current Microbiology*, *31*(1), 34–38. https://doi.org/10.1007/BF00294631

2. Dang, Q., Zhao, X., Li, Y., & Xi, B. (2023). Revisiting the biological pathway for methanogenesis in landfill from metagenomic perspective—A case study of county-level sanitary landfill of domestic waste in North China plain. *Environmental Research*, *222*, 115185. https://doi.org/10.1016/j.envres.2022.115185

3. Hook, S. E., Wright, A.-D. G., & McBride, B. W. (2010). Methanogens: Methane Producers of the Rumen and Mitigation Strategies. *Archaea*, *2010*, 945785. https://doi.org/20160811101719

4. Li, D., Ni, H., Jiao, S., Lu, Y., Zhou, J., Sun, B., & Liang, Y. (2021). Coexistence patterns of soil methanogens are closely tied to methane generation and community assembly in rice paddies. *Microbiome*, *9*(1), 20. https://doi.org/10.1186/s40168-020-00978-8

5. Myers, T., & Dykstra, C. M. (2024). Teaching old dogs new tricks: Genetic engineering methanogens. *Applied and Environmental Microbiology*, *0*(0), e02247-23. https://doi.org/10.1128/aem.02247-23

6. Kohler, P. R. A., & Metcalf, W. W. (2012). Genetic manipulation of *Methanosarcina* spp. *Frontiers in Microbiology*, *3*(JUL), 1–9. https://doi.org/10.3389/fmicb.2012.00259

7. Carr, S., & Buan, N. R. (2022). Insights into the biotechnology potential of *Methanosarcina*. *Frontiers in Microbiology*, *13*. https://www.frontiersin.org/articles/10.3389/fmicb.2022.1034674

8. Goyal, N., Zhou, Z., & Karimi, I. A. (2016). Metabolic processes of *Methanococcus maripaludis* and potential applications. *Microbial Cell Factories*, *15*(1), Article 1. https://doi.org/10.1186/s12934-016-0500-0

9. Iino, T., Tamaki, H., Tamazawa, S., Ueno, Y., Ohkuma, M., Suzuki, K., Igarashi, Y., & Haruta, S. (2013). *Candidatus* Methanogranum caenicola: A Novel Methanogen from the Anaerobic Digested Sludge, and Proposal of *Methanomassiliicoccaceae* fam. nov. and *Methanomassiliicoccales* ord. nov., for a Methanogenic Lineage of the Class *Thermoplasmata*. *Microbes and Environments*, *28*(2), 244–250. https://doi.org/10.1264/jsme2.ME12189

10. Liu, Y., & Whitman, W. B. (2008). Metabolic, Phylogenetic, and Ecological Diversity of the Methanogenic Archaea. *Annals of the New York Academy of Sciences*, *1125*(1), 171–189. https://doi.org/10.1196/annals.1419.019

11. Sakai, S., Imachi, H., Hanada, S., Ohashi, A., Harada, H., & Kamagata, Y. (2008). *Methanocella paludicola* gen. Nov., sp. Nov., a methane-producing archaeon, the first isolate of the lineage 'Rice Cluster I', and proposal of the new archaeal order *Methanocellales* ord. Nov. *International Journal of*

Systematic and Evolutionary Microbiology, *58*(4), 929–936. https://doi.org/10.1099/ijs.0.65571-0

12. Sorokin, D. Y., Merkel, A. Y., Abbas, B., Makarova, K. S., Rijpstra, W. I. C., Koenen, M., Sinninghe Damsté, J. S., Galinski, E. A., Koonin, E. V., & van Loosdrecht, M. C. M. (2018). *Methanonatronarchaeum thermophilum* gen. Nov., sp. Nov. And "*Candidatus* Methanohalarchaeum thermophilum", extremely halo(natrono)philic methyl-reducing methanogens from hypersaline lakes comprising a new euryarchaeal class *Methanonatronarchaeia* classis nov. *International Journal of Systematic and Evolutionary Microbiology*, *68*(7), 2199–2208. https://doi.org/10.1099/ijsem.0.002810

13. Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., & Tyson, G. W. (2019). An evolving view of methane metabolism in the Archaea. *Nature Reviews Microbiology*, *17*(4), 219–232. https://doi.org/10.1038/s41579-018-0136-7

14. Kaster, A.-K., Moll, J., Parey, K., & Thauer, R. K. (2011). Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proceedings of the National Academy of Sciences*, *108*(7), 2981–2986. https://doi.org/10.1073/pnas.1016761108

15. Thauer, R. K., Kaster, A., Seedorf, H., & Buckel, W. (2008). Methanogenic Archaea: Ecologically relevant differences in energy conservation. 6(august), 579–591. https://doi.org/10.1038/nrmicro1931

16. Mand, T. D., Kulkarni, G., & Metcalf, W. W. (2018). Genetic, biochemical, and molecular characterization of *Methanosarcina barkeri* mutants lacking three distinct classes of hydrogenase. *Journal of Bacteriology*, *200*(20). https://doi.org/10.1128/JB.00342-18

17. Kröninger, L., Gottschling, J., & Deppenmeier, U. (2017). Growth Characteristics of *Methanomassiliicoccus luminyensis* and Expression of Methyltransferase Encoding Genes. *Archaea*, *2017*, 2756573. https://doi.org/10.1155/2017/2756573

18. Poulsen, M., Schwab, C., Borg Jensen, B., Engberg, R. M., Spang, A., Canibe, N., Højberg, O., Milinovich, G., Fragner, L., Schleper, C., Weckwerth, W., Lund, P., Schramm, A., & Urich, T. (2013). Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nature Communications*, *4*(1), 1428. https://doi.org/10.1038/ncomms2432

19. Carr, S. A., Schubotz, F., Dunbar, R. B., Mills, C. T., Dias, R., Summons, R. E., & Mandernack, K. W. (2018). Acetoclastic Methanosaeta are dominant methanogens in organic-rich Antarctic marine sediments. *The ISME Journal*, *12*(2), 330–342. https://doi.org/10.1038/ismej.2017.150

20. Teh, Y. L., & Zinder, S. (1992). Acetyl-coenzyme A synthetase in the thermophilic, acetate-utilizing methanogen Methanothrix sp. Strain CALS-1. *FEMS Microbiology Letters*, *98*(1–3), 1–7. https://doi.org/10.1111/j.1574-6968.1992.tb05481.x

21. Welte, C., & Deppenmeier, U. (2014). Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochimica et Biophysica Acta - Bioenergetics*, *1837*(7), 1130–1147. https://doi.org/10.1016/j.bbabio.2013.12.002

22. Kurth, J. M., Op den Camp, H. J. M., & Welte, C. U. (2020). Several ways one goal—Methanogenesis from unconventional substrates. *Applied Microbiology and Biotechnology*, *104*(16), 6839–6854. https://doi.org/10.1007/s00253-020-10724-7

23. Deppenmeier, U. (2004). The Membrane-Bound Electron Transport System of *Methanosarcina* Species. *Journal of Bioenergetics and Biomembranes*, *36*(1), Article 1. https://doi.org/10.1023/B:JOBB.0000019598.64642.97

24. Mand, T. D., & Metcalf, W. W. (2019). Energy Conservation and Hydrogenase Function in Methanogenic Archaea, in Particular the Genus Methanosarcina. https://doi.org/10.1128/MMBR

25. Galagan, J. E., Nusbaum, C., Roy, A., Endrizzi, M. G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., … Birren, B. (2002). The Genome of *M. acetivorans* Reveals Extensive Metabolic and Physiological Diversity. *Genome Research*, *12*(4), 532–542. https://doi.org/10.1101/gr.223902

26. Sowers, K. R., Baron, S. F., & Ferry, J. G. (1984). *Methanosarcina acetivorans* sp. Nov., an Acetotrophic Methane-Producing Bacterium Isolated from Marine Sediments. *Applied and Environmental Microbiology*, *47*(5), 971–978. https://doi.org/10.1128/aem.47.5.971-978.1984

27. Lessner, D. J., Lhu, L., Wahal, C. S., & Ferry, J. G. (2010). An Engineered Methanogenic Pathway Derived from the Domains Bacteria and Archaea. *mBio*, *1*(5), e00243-10. https://doi.org/10.1128/mBio.00243-10

28. Aldridge, J., Carr, S., Weber, K. A., & Buan, N. R. (2021). Anaerobic Production of Isoprene by Engineered *Methanosarcina* Species Archaea. *Applied and Environmental Microbiology*, *87*(6), Article 6. https://doi.org/10.1128/AEM.02417-20

29. Besada-Lombana, P. B., McTaggart, T. L., & Da Silva, N. A. (2018). Molecular tools for pathway engineering in *Saccharomyces cerevisiae*. *Current Opinion in Biotechnology*, *53*, 39–49. https://doi.org/10.1016/j.copbio.2017.12.002

30. Han, T., Nazarbekov, A., Zou, X., & Lee, S. Y. (2023). Recent advances in systems metabolic engineering. *Current Opinion in Biotechnology*, *84*, 103004. https://doi.org/10.1016/j.copbio.2023.103004

31. Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology*, *5*, 172. https://doi.org/10.3389/fmicb.2014.00172

32. Lyu, Z., & Whitman, W. B. (2019). Transplanting the pathway engineering toolbox to methanogens. *Current Opinion in Biotechnology*, *59*, 46–54. https://doi.org/10.1016/j.copbio.2019.02.009

33. Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R., & Wolfe, R. S. (1997). A genetic system for Archaea of the genus *Methanosarcina*: Liposome-mediated transformation and construction of shuttle vectors. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(6), 2626–2631. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC20139/

34. Sowers, K. R., & Gunsalus, R. P. (1988). Plasmid DNA from the acetotrophic methanogen *Methanosarcina acetivorans*. *Journal of Bacteriology*, *170*(10), 4979–4982. https://doi.org/10.1128/jb.170.10.4979-4982.1988

35. Pritchett, M. A., Zhang, J. K., & Metcalf, W. W. (2004). Development of a Markerless Genetic Exchange Method for *Methanosarcina acetivorans* C2A and Its Use in Construction of New Genetic Tools for Methanogenic Archaea. *Applied and Environmental Microbiology*, *70*(3), 1425–1433. https://doi.org/10.1128/AEM.70.3.1425-1433.2004

36. Zhang, J. K., Pritchett, M. A., Lampe, D. J., Robertson, H. M., & Metcalf, W. W. (2000). In vivo transposon mutagenesis of the methanogenic archaeon *Methanosarcina acetivorans* C2A using a modified version of the insect mariner-family transposable element Himar1. *Proceedings of the National Academy of Sciences*, *97*(17), 9665–9670. https://doi.org/10.1073/pnas.160272597

37. Dhamad, A. E., & Lessner, D. J. (2020). A CRISPRi-dCas9 System for Archaea and Its Use To Examine Gene Function during Nitrogen Fixation by *Methanosarcina acetivorans*. *Applied and Environmental Microbiology*, *86*(21), Article 21. https://doi.org/10.1128/AEM.01402-20

38. Nayak, D. D., & Metcalf, W. W. (2017). Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(11), 2976–2981. https://doi.org/10.1073/pnas.1618596114

39. Zhu, P., Somvanshi, T., Bao, J., & Scheller, S. (2023). CRISPR/Cas12a toolbox for genome editing in *Methanosarcina acetivorans*. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1235616

40. Thomsen, J., & Schmitz, R. A. (2022). Generating a Small Shuttle Vector for Effective Genetic Engineering of *Methanosarcina mazei* Allowed First Insights in Plasmid Replication Mechanism in the Methanoarchaeon. *International Journal of Molecular Sciences*, *23*(19), 11910. https://doi.org/10.3390/ijms231911910

41. Guss, A. M., Rother, M., Zhang, J. K., Kulkkarni, G., & Metcalf, W. W. (2008). New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for *Methanosarcina* species. *Archaea*, *2*(3), 193–203. https://doi.org/10.1155/2008/534081

42. Adlung, N., & Scheller, S. (2023). Application of the Fluorescence-Activating and Absorption-Shifting Tag (FAST) for Flow Cytometry in Methanogenic Archaea. *Applied and Environmental Microbiology*, *0*(0), e01786-22. https://doi.org/10.1128/aem.01786-22

43. Shea, M. T., Walter, M. E., Duszenko, N., Ducluzeau, A.-L., Aldridge, J., King, S. K., & Buan, N. R. (2016). pNEB193-derived suicide plasmids for gene deletion and protein expression in the methane-producing archaeon, *Methanosarcina acetivorans*. *Plasmid*, *84–85*, 27–35. https://doi.org/10.1016/j.plasmid.2016.02.003

44. Sowers, K. R., Boone, J. E., & Gunsalus, R. P. (1993). Disaggregation of *Methanosarcina spp*. And Growth as Single Cells at Elevated Osmolarity. *Applied and Environmental Microbiology*, *59*(11), 3832–3839. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC182538/

45. Chadwick, G. L., Joiner, A. M. N., Ramesh, S., Mitchell, D. A., & Nayak, D. D. (2023). McrD binds asymmetrically to methyl-coenzyme M reductase improving active-site accessibility during assembly. *Proceedings of the National Academy of Sciences*, *120*(25), e2302815120. https://doi.org/10.1073/pnas.2302815120

46. Zhang, J. K., White, A. K., Kuettner, H. C., Boccazzi, P., & Metcalf, W. W. (2002). Directed Mutagenesis and Plasmid-Based Complementation in the Methanogenic Archaeon *Methanosarcina acetivorans* C2A Demonstrated by Genetic Analysis of Proline Biosynthesis. *Journal of Bacteriology*, *184*(5), 1449–1454. https://doi.org/10.1128/JB.184.5.1449-1454.2002

47. Ehlers, C., Weidenbach, K., Veit, K., Deppenmeier, U., Metcalf, W. W., & Schmitz, R. A. (2005). Development of genetic methods and construction of a chromosomal glnK1 mutant in *Methanosarcina mazei* strain Gö1.

Molecular Genetics and Genomics, *273*(4), 290–298. https://doi.org/10.1007/s00438-005-1128-7

48. Nayak, D. D., & Metcalf, W. W. (2018). Genetic techniques for studies of methyl-coenzyme M reductase from *Methanosarcina acetivorans* C2A. In F. Armstrong (Ed.), *Methods in Enzymology* (Vol. 613, pp. 325–347). Academic Press. https://doi.org/10.1016/bs.mie.2018.10.012

49. Oelgeschläger, E., & Rother, M. (2009). In vivo role of three fused corrinoid/methyl transfer proteins in *Methanosarcina acetivorans*. *Molecular Microbiology*, *72*(5), 1260–1272. https://doi.org/10.1111/j.1365-2958.2009.06723.x

50. Paszkowski, J., Shillito, R. D., Saul, M., Mandák, V., Hohn, T., Hohn, B., & Potrykus, I. (1984). Direct gene transfer to plants. *The EMBO Journal*, *3*(12), 2717–2722. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC557758/

51. Tumbula, D. L., Makula, R. A., & Whitman, W. B. (1994). Transformation of *Methanococcus maripaludis* and identification of a Pst I-like restriction system. *FEMS Microbiology Letters*, *121*(3), 309–314. https://doi.org/10.1111/j.1574-6968.1994.tb07118.x

52. Fonseca, D. R., Halim, M. F. A., Holten, M. P., & Costa, K. C. (2020). Type IV-like pili facilitate transformation in naturally competent archaea. *Journal of Bacteriology*, *202*(21), 1–12. https://doi.org/10.1128/JB.00355-20

53. Craig, L., Forest, K. T., & Maier, B. (2019). Type IV pili: Dynamics, biophysics and functional consequences. *Nature Reviews Microbiology*, *17*(7), Article 7. https://doi.org/10.1038/s41579-019-0195-4

54. Uali, A. S., Lam, T. Y. C., Huang, X., Wu, Z., Shih, H. J., Tan, G.-Y., & Lee, P.-H. (2024). Role and potential of the semi-classical/-quantum mechanism of the extracellular environment and cell envelope in Direct Interspecies Electron Transfer (DIET)-driven biomethanation. *Critical Reviews in Environmental Science and Technology*, *54*(7), 581–601. https://doi.org/10.1080/10643389.2023.2261813

55. Luiz, M. T., Dutra, J. A. P., Tofani, L. B., de Araújo, J. T. C., Di Filippo, L. D., Marchetti, J. M., & Chorilli, M. (2022). Targeted Liposomes: A Nonviral Gene Delivery System for Cancer Therapy. *Pharmaceutics*, *14*(4), 821. https://doi.org/10.3390/pharmaceutics14040821

56. Kopke, K., Hoff, B., & Kück, U. (2010). Application of the *Saccharomyces cerevisiae* FLP/FRT Recombination System in Filamentous Fungi for Marker Recycling and Construction of Knockout Strains Devoid of Heterologous Genes. *Applied and Environmental Microbiology*, *76*(14), 4664–4674. https://doi.org/10.1128/AEM.00670-10

57. Rother, M., & Metcalf, W. W. (2005). Genetic technologies for Archaea. *Current Opinion in Microbiology*, *8*(6), 745–751. https://doi.org/10.1016/j.mib.2005.10.010

58. Welander, P. V., & Metcalf, W. W. (2008). Mutagenesis of the C1 Oxidation Pathway in *Methanosarcina barkeri*: New Insights into the Mtr/Mer Bypass Pathway. *Journal of Bacteriology*, *190*(6), 1928–1936. https://doi.org/10.1128/jb.01424-07

59. Kuhstoss, S., & Rao, R. N. (1991). Analysis of the integration function of the *Streptomycete bacteriophage* φC31. *Journal of Molecular Biology*, *222*(4), 897–908. https://doi.org/10.1016/0022-2836(91)90584-S

60. Thorpe, H. M., & Smith, M. C. M. (1998). In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase

family. *Proceedings of the National Academy of Sciences*, *95*(10), Article 10. https://doi.org/10.1073/pnas.95.10.5505

61. Blank, C. E., Kessler, P. S., & Leigh, J. A. (1995). Genetics in methanogens: Transposon insertion mutagenesis of a *Methanococcus maripaludis* nifH gene. *Journal of Bacteriology*, *177*(20), 5773–5777. https://doi.org/10.1128/jb.177.20.5773-5777.1995

62. Fonseca, D. R., Loppnow, M. B., Day, L. A., Kelsey, E. L., Abdul Halim, M. F., & Costa, K. C. (2023). Random transposon mutagenesis identifies genes essential for transformation in *Methanococcus maripaludis*. *Molecular Genetics and Genomics*, *298*(3), 537–548. https://doi.org/10.1007/s00438-023-01994-7

63. Sattler, C., Wolf, S., Fersch, J., Goetz, S., & Rother, M. (2013). Random mutagenesis identifies factors involved in formate-dependent growth of the methanogenic archaeon *Methanococcus maripaludis*. *Molecular Genetics and Genomics*, *288*(9), 413–424. https://doi.org/10.1007/s00438-013-0756-6

64. Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., … Koonin, E. V. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, *13*(11), 722–736. https://doi.org/10.1038/nrmicro3569

65. Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, *37*, 67–78. https://doi.org/10.1016/j.mib.2017.05.008

66. Tang, Y., & Fu, Y. (2018). Class 2 CRISPR/Cas: An expanding biotechnology toolbox for and beyond genome editing. *Cell & Bioscience*, *8*(1), 59. https://doi.org/10.1186/s13578-018-0255-x

67. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. https://doi.org/10.1126/science.1231143

68. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

69. Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., & Doudna, J. A. (2014). Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science*, *343*(6176), 1247997. https://doi.org/10.1126/science.1247997

70. Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. *Cell*, *163*(3), 759–771. https://doi.org/10.1016/j.cell.2015.09.038

71. Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., Winblad, N., Choudhury, S. R., Abudayyeh, O. O., Gootenberg, J. S., Wu, W. Y., Scott, D. A., Severinov, K., van der Oost, J., & Zhang, F. (2017). Multiplex gene editing by CRISPR–Cpf1 using a single

crRNA array. *Nature Biotechnology*, *35*(1), Article 1. https://doi.org/10.1038/nbt.3737

72. Dong, D., Ren, K., Qiu, X., Zheng, J., Guo, M., Guan, X., Liu, H., Li, N., Zhang, B., Yang, D., Ma, C., Wang, S., Wu, D., Ma, Y., Fan, S., Wang, J., Gao, N., & Huang, Z. (2016). The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature*, *532*(7600), 522–526. https://doi.org/10.1038/nature17944

73. Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., & Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, *532*(7600), Article 7600. https://doi.org/10.1038/nature17945

74. Swarts, D. C., van der Oost, J., & Jinek, M. (2017). Structural basis for guide RNA processing and seed-dependent DNA targeting and cleavage by CRISPR-Cas12a. *Molecular Cell*, *66*(2), 221-233.e4. https://doi.org/10.1016/j.molcel.2017.03.016

75. Bao, J., de Dios Mateos, E., & Scheller, S. (2022). Efficient CRISPR/Cas12a-Based Genome-Editing Toolbox for Metabolic Engineering in *Methanococcus maripaludis*. *ACS Synthetic Biology*, *11*(7), 2496–2503. https://doi.org/10.1021/acssynbio.2c00137

76. Du, Q., Wei, Y., Zhang, L., Ren, D., Gao, J., Dong, X., Bai, L., & Li, J. (2024). An improved CRISPR and CRISPR interference (CRISPRi) toolkit for engineering the model methanogenic archaeon *Methanococcus maripaludis*. *Microbial Cell Factories*, *23*(1), 239. https://doi.org/10.1186/s12934-024-02492-0

77. Li, J., Zhang, L., Xu, Q., Zhang, W., Li, Z., Chen, L., & Dong, X. (2022). CRISPR-Cas9 Toolkit for Genome Editing in an Autotrophic CO2-Fixing Methanogenic Archaeon. *Microbiology Spectrum*, *10*(4), e01165-22. https://doi.org/10.1128/spectrum.01165-22

78. Paul, B., & Montoya, G. (2020). CRISPR-Cas12a: Functional overview and applications. *Biomedical Journal*, *43*(1), Article 1. https://doi.org/10.1016/j.bj.2019.10.005

79. Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, *20*(8), 490–507. https://doi.org/10.1038/s41580-019-0131-5

80. Maini Rekdal, V., van der Luijt, C. R. B., Chen, Y., Kakumanu, R., Baidoo, E. E. K., Petzold, C. J., Cruz-Morales, P., & Keasling, J. D. (2024). Edible mycelium bioengineered for enhanced nutritional value and sensory appeal using a modular synthetic biology toolkit. *Nature Communications*, *15*(1), 2099. https://doi.org/10.1038/s41467-024-46314-8

81. Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N. J., Keasling, J. D., & Mukhopadhyay, A. (2017). A Cas9-based toolkit to program gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, *45*(1), 496–508. https://doi.org/10.1093/nar/gkw1023

82. Ruan, S., Yang, Y., Zhang, X., Luo, G., Lin, Y., & Liang, S. (2024). Screening and characterization of integration sites based on CRISPR-Cpf1 in *Pichia pastoris*. *Synthetic and Systems Biotechnology*, *9*(4), 759–765. https://doi.org/10.1016/j.synbio.2024.06.002

83. Carruthers, D. N., Saleski, T. E., Scholz, S. A., & Lin, X. N. (2020). Random Chromosomal Integration and Screening Yields *E. coli* K-12 Derivatives Capable of Efficient Sucrose Utilization. *ACS Synthetic Biology*, *9*(12), 3311–3321. https://doi.org/10.1021/acssynbio.0c00392

84. Härtl, B., Wehrl, W., Wiegert, T., Homuth, G., & Schumann, W. (2001). Development of a New Integration Site within theBacillus subtilis Chromosome and Construction of Compatible Expression Cassettes. *Journal of Bacteriology*, *183*(8), 2696–2699. https://doi.org/10.1128/jb.183.8.2696-2699.2001

85. Buan, N. R., & Metcalf, W. W. (2010). Methanogenesis by *Methanosarcina acetivorans* involves two structurally and functionally distinct classes of heterodisulfide reductase. *Molecular Microbiology*, *75*(4), 843–853. https://doi.org/10.1111/j.1365-2958.2009.06990.x

86. Rother, M., Oelgeschläger, E., & W. Metcalf, W. (2007). Genetic and proteomic analyses of CO utilization by *Methanosarcina acetivorans*. *Archives of Microbiology*, *188*(5), 463–472. https://doi.org/10.1007/s00203-007-0266-1

87. Hirata, A., Klein, B. J., & Murakami, K. S. (2008). The X-ray crystal structure of RNA polymerase from Archaea. *Nature*, *451*(7180), 851–854. https://doi.org/10.1038/nature06530

88. Blombach, F., Matelska, D., Fouqueau, T., Cackett, G., & Werner, F. (2019). Key Concepts and Challenges in Archaeal Transcription. *Journal of Molecular Biology*, *431*(20), Article 20. https://doi.org/10.1016/j.jmb.2019.06.020

89. Dexl, S., Reichelt, R., Kraatz, K., Schulz, S., Grohmann, D., Bartlett, M., & Thomm, M. (2018). Displacement of the transcription factor B reader domain during transcription initiation. *Nucleic Acids Research*, *46*(19), 10066–10081. https://doi.org/10.1093/nar/gky699

90. Karim, A. A., Gestaut, D. R., Fincker, M., Ruth, J. C., Holmes, E. C., Sheu, W., & Spormann, A. M. (2018). Fine-Tuned Protein Production in *Methanosarcina acetivorans* C2A. *ACS Synthetic Biology*, *7*(8), 1874–1885. https://doi.org/10.1021/acssynbio.8b00062

91. Xu, Q., Du, Q., Gao, J., Chen, L., Dong, X., & Li, J. (2023). A robust genetic toolbox for fine-tuning gene expression in the CO2-Fixing methanogenic archaeon *Methanococcus maripaludis*. *Metabolic Engineering*, *79*, 130–145. https://doi.org/10.1016/j.ymben.2023.07.007

92. Zhu, P., Molina Resendiz, M., von Ossowski, I., & Scheller, S. (2024). A promoter–RBS library for fine-tuning gene expression in *Methanosarcina acetivorans*. *Applied and Environmental Microbiology*, *0*(0), e01092-24. https://doi.org/10.1128/aem.01092-24

93. Opulencia, R. B., Bose, A., & Metcalf, W. W. (2009). Physiology and Posttranscriptional Regulation of Methanol:Coenzyme M Methyltransferase Isozymes in *Methanosarcina acetivorans* C2A §. *JOURNAL OF BACTERIOLOGY*, *191*, 6928–6935. https://doi.org/10.1128/JB.00947-09

94. Yue, L., Li, J., Zhang, B., Qi, L., Li, Z., Zhao, F., Li, L., Zheng, X., & Dong, X. (2020). The conserved ribonuclease aCPSF1 triggers genome-wide transcription termination of Archaea via a 3′-end cleavage mode. *Nucleic Acids Research*, *48*(17), 9589–9605. https://doi.org/10.1093/nar/gkaa702

95. Ding, Y., Nash, J., Berezuk, A., Khursigara, C. M., Langelaan, D. N., Smith, S. P., & Jarrell, K. F. (2016). Identification of the first transcriptional activator of an archaellum operon in a euryarchaeon. *Molecular Microbiology*, *102*(1), 54–70. https://doi.org/10.1111/mmi.13444

96. Lie, T. J., Wood, G. E., & Leigh, J. A. (2005). Regulation of nif Expression in *Methanococcus maripaludis*: ROLES OF THE EURYARCHAEAL REPRESSOR NrpR, 2-OXOGLUTARATE, AND TWO OPERATORS.

Journal of Biological Chemistry, *280*(7), 5236–5241. https://doi.org/10.1074/jbc.M411778200

97. Akinyemi, T. S., Shao, N., Lyu, Z., Drake, I. J., Liu, Y., & Whitman, W. B. (2021). Tuning Gene Expression by Phosphate in the Methanogenic Archaeon *Methanococcus maripaludis*. *ACS Synthetic Biology*, *10*(11), 3028–3039. https://doi.org/10.1021/acssynbio.1c00322

98. Shalvarjian, K. E., & Nayak, D. D. (2021). Transcriptional regulation of methanogenic metabolism in archaea. *Current Opinion in Microbiology*, *60*, 8–15. https://doi.org/10.1016/j.mib.2021.01.005

99. Schmitt, E., Coureux, P.-D., Kazan, R., Bourgeois, G., Lazennec-Schurdevin, C., & Mechulam, Y. (2020). Recent Advances in Archaeal Translation Initiation. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.584152

100. Brenneis, M., Hering, O., Lange, C., & Soppa, J. (2007). Experimental Characterization of Cis-Acting Elements Important for Translation and Transcription in Halophilic Archaea. *PLOS Genetics*, *3*(12), e229. https://doi.org/10.1371/journal.pgen.0030229

101. Jäger, D., Sharma, C. M., Thomsen, J., Ehlers, C., Vogel, J., & Schmitz, R. A. (2009). Deep sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. *Proceedings of the National Academy of Sciences*, *106*(51), 21878–21882. https://doi.org/10.1073/pnas.0909051106

102. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

103. Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., & Marraffini, L. A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Research*, *41*(15), 7429–7437. https://doi.org/10.1093/nar/gkt520

104. Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, *154*(2), 442–451. https://doi.org/10.1016/j.cell.2013.06.044

105. Leenay, R. T., Maksimchuk, K. R., Slotkowski, R. A., Agrawal, R. N., Gomaa, A. A., Briner, A. E., Barrangou, R., & Beisel, C. L. (2016). Identifying and Visualizing Functional PAM Diversity across CRISPR-Cas Systems. *Molecular Cell*, *62*(1), 137–147. https://doi.org/10.1016/j.molcel.2016.02.031

106. Wu, Y., Li, Y., Jin, K., Zhang, L., Li, J., Liu, Y., Du, G., Lv, X., Chen, J., Ledesma-Amaro, R., & Liu, L. (2023). CRISPR–dCas12a-mediated genetic circuit cascades for multiplexed pathway optimization. *Nature Chemical Biology*, *19*(3), 367–377. https://doi.org/10.1038/s41589-022-01230-0

107. Zhang, X., Wang, J., Cheng, Q., Zheng, X., Zhao, G., & Wang, J. (2017). Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discovery*, *3*(1), 1–9. https://doi.org/10.1038/celldisc.2017.18

108. Beneke, S., Bestgen, H., & Klein, A. (1995). Use of the *Escherichia coli* uidA gene as a reporter in *Methanococcus voltae* for the analysis of the regulatory function of the intergenic region between the operons encoding selenium-free hydrogenases. *Molecular and General Genetics MGG*, *248*(2), 225–228. https://doi.org/10.1007/BF02190804

109. Guss, A. M., Kulkarni, G., & Metcalf, W. W. (2009). Differences in Hydrogenase Gene Expression between *Methanosarcina acetivorans* and *Methanosarcina barkeri*. *Journal of Bacteriology*, *191*(8), 2826–2833. https://doi.org/10.1128/JB.00563-08

110. Rother, M., Boccazzi, P., Bose, A., Pritchett, M. A., & Metcalf, W. W. (2005). Methanol-Dependent Gene Expression Demonstrates that Methyl-Coenzyme M Reductase Is Essential in *Methanosarcina acetivorans* C2A and Allows Isolation of Mutants with Defects in Regulation of the Methanol Utilization Pathway. *Journal of Bacteriology*, *187*(16), 5552–5559. https://doi.org/10.1128/JB.187.16.5552-5559.2005

111. Cohen-Kupiec, R., Blank, C., & Leigh, J. A. (1997). Transcriptional regulation in archaea: In vivo demonstration of a repressor binding site in a methanogen. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(4), Article 4. https://doi.org/10.1073/pnas.94.4.1316

112. Demolli, S., Geist, M. M., Weigand, J. E., Matschiavelli, N., Suess, B., & Rother, M. (2014). Development of *β*-Lactamase as a Tool for Monitoring Conditional Gene Expression by a Tetracycline-Riboswitch in *Methanosarcina acetivorans*. *Archaea*, *2014*(1), 725610. https://doi.org/10.1155/2014/725610

113. Lyu, Z., Shao, N., Chou, C.-W., Shi, H., Patel, R., Duin, E. C., & Whitman, W. B. (2020). Posttranslational Methylation of Arginine in Methyl Coenzyme M Reductase Has a Profound Impact on both Methanogenesis and Growth of *Methanococcus maripaludis*. *Journal of Bacteriology*, *202*(3), 10.1128/jb.00654-19. https://doi.org/10.1128/jb.00654-19

114. Hernandez, E., & Costa, K. C. (2022). The Fluorescence-Activating and Absorption-Shifting Tag (FAST) Enables Live-Cell Fluorescence Imaging of *Methanococcus maripaludis*. *Journal of Bacteriology*, *204*(7), e00120-22. https://doi.org/10.1128/jb.00120-22

115. Kim, H. K., Song, M., Lee, J., Menon, A. V., Jung, S., Kang, Y.-M., Choi, J. W., Woo, E., Koh, H. C., Nam, J.-W., & Kim, H. (2017). In vivo high-throughput profiling of CRISPR–Cpf1 activity. *Nature Methods*, *14*(2), 153–159. https://doi.org/10.1038/nmeth.4104

116. Endo, A., Masafumi, M., Kaya, H., & Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from Francisella novicida. *Scientific Reports*, *6*, 38169. https://doi.org/10.1038/srep38169

117. Filée, J., Siguier, P., & Chandler, M. (2007). Insertion Sequence Diversity in Archaea. *Microbiology and Molecular Biology Reviews*, *71*(1), 121–157. https://doi.org/10.1128/MMBR.00031-06

118. Altae-Tran, H., Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Kannan, S., Zhang, F., & Koonin, E. V. (2023). Diversity, evolution, and classification of the RNA-guided nucleases TnpB and Cas12. *Proceedings of the National Academy of Sciences*, *120*(48), e2308224120. https://doi.org/10.1073/pnas.2308224120

119. Kapitonov, V. V., Makarova, K. S., & Koonin, E. V. (2016). ISC, a Novel Group of Bacterial and Archaeal DNA Transposons That Encode Cas9 Homologs. *Journal of Bacteriology*, *198*(5), 797–807. https://doi.org/10.1128/jb.00783-15

120. Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., Kazlauskas, D., Venclovas, Č., & Siksnys, V. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease.

Nature, *599*(7886), 692–696. https://doi.org/10.1038/s41586-021-04058-1

121. Gehlert, F. O., Nickel, L., Vakirlis, N., Hammerschmidt, K., Vargas Gebauer, H. I., Kießling, C., Kupczok, A., & Schmitz, R. A. (2023). Active in vivo translocation of the *Methanosarcina mazei* Gö1 Casposon. *Nucleic Acids Research*, *51*(13), Article 13. https://doi.org/10.1093/nar/gkad474

122. Bryant, J. A., Sellars, L. E., Busby, S. J. W., & Lee, D. J. (2014). Chromosome position effects on gene expression in Escherichia coli K-12. *Nucleic Acids Research*, *42*(18), 11383–11392. https://doi.org/10.1093/nar/gku828

123. Nayak, D. D., & Metcalf, W. W. (2019). Methylamine-specific methyltransferase paralogs in *Methanosarcina* are functionally distinct despite frequent gene conversion. *ISME Journal*, *13*(9), 2173–2182. https://doi.org/10.1038/s41396-019-0428-6

124. Hildenbrand, C., Stock, T., Lange, C., Rother, M., & Soppa, J. (2011). Genome copy numbers and gene conversion in methanogenic archaea. *Journal of Bacteriology*, *193*(3), 734–743. https://doi.org/10.1128/JB.01016-10

125. Bai Flagfeldt, D., Siewers, V., Huang, L., & Nielsen, J. (2009). Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast*, *26*(10), 545–551. https://doi.org/10.1002/yea.1705

126. Takemata, N., Samson, R. Y., & Bell, S. D. (2019). Physical and Functional Compartmentalization of Archaeal Chromosomes. *Cell*, *179*(1), 165-179.e18. https://doi.org/10.1016/j.cell.2019.08.036

127. Alper, H., Fischer, C., Nevoigt, E., & Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(36), Article 36. https://doi.org/10.1073/pnas.0504604102

128. Peng, B., Wood, R. J., Nielsen, L. K., & Vickers, C. E. (2018). An Expanded Heterologous GAL Promoter Collection for Diauxie-Inducible Expression in *Saccharomyces cerevisiae*. *ACS Synthetic Biology*, *7*(2), 748–751. https://doi.org/10.1021/acssynbio.7b00355

129. Zhou, S., Lyu, Y., Li, H., Koffas, M. A. G., & Zhou, J. (2019). Fine-tuning the (2S)-naringenin synthetic pathway using an iterative high-throughput balancing strategy. *Biotechnology and Bioengineering*, *116*(6), 1392–1404. https://doi.org/10.1002/bit.26941

130. Johnson, A. O., Gonzalez-Villanueva, M., Tee, K. L., & Wong, T. S. (2018). An Engineered Constitutive Promoter Set with Broad Activity Range for Cupriavidus necator H16. *ACS Synthetic Biology*, *7*(8), 1918–1928. https://doi.org/10.1021/acssynbio.8b00136

131. Bose, A., Pritchett, M. A., Rother, M., & Metcalf, W. W. (2006). Differential regulation of the three methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Journal of Bacteriology*, *188*(20), 7274–7283. https://doi.org/10.1128/JB.00535-06

132. Bose, A., & Metcalf, W. W. (2008). Distinct regulators control the expression of methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Molecular Microbiology*, *67*(3), 649–661. https://doi.org/10.1111/j.1365-2958.2007.06075.x

133. Reichlen, M. J., Vepachedu, V. R., Murakami, K. S., & Ferry, J. G. (2012). MreA functions in the global regulation of methanogenic pathways in *Methanosarcina acetivorans*. *mBio*, *3*(4). https://doi.org/10.1128/mBio.00189-12

134. Matschiavelli, N., Oelgeschläger, E., Cocchiararo, B., Finke, J., & Rother, M. (2012). Function and Regulation of Isoforms of Carbon Monoxide Dehydrogenase/Acetyl Coenzyme A Synthase in *Methanosarcina acetivorans*. *Journal of Bacteriology*, *194*(19), 5377–5387. https://doi.org/10.1128/jb.00881-12

135. Rosenberg, A. H., Lade, B. N., Dao-shan, C., Lin, S.-W., Dunn, J. J., & Studier, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*, *56*(1), 125–135. https://doi.org/10.1016/0378-1119(87)90165-X

136. Ye, J., Li, Y., Bai, Y., Zhang, T., Jiang, W., Shi, T., Wu, Z., & Zhang, Y.-H. P. J. (2022). A facile and robust T7-promoter-based high-expression of heterologous proteins in *Bacillus subtilis*. *Bioresources and Bioprocessing*, *9*(1), 56. https://doi.org/10.1186/s40643-022-00540-4

137. Johnston, M., Flick, J. S., & Pexton, T. (1994). Multiple Mechanisms Provide Rapid and Stringent Glucose Repression of GAL Gene Expression in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *14*(6), 3834–3841. https://doi.org/10.1128/mcb.14.6.3834-3841.1994

138. Xiao, J., Peng, B., Su, Z., Liu, A., Hu, Y., Nomura, C. T., Chen, S., & Wang, Q. (2020). Facilitating Protein Expression with Portable 5′-UTR Secondary Structures in *Bacillus licheniformis*. *ACS Synthetic Biology*, *9*(5), 1051–1058. https://doi.org/10.1021/acssynbio.9b00355

139. Qi, L., Yue, L., Feng, D., Qi, F., Li, J., & Dong, X. (2017). Genome-wide mRNA processing in methanogenic archaea reveals post-transcriptional regulation of ribosomal protein synthesis. *Nucleic Acids Research*, *45*(12), 7285–7298. https://doi.org/10.1093/nar/gkx454

140. Rohlin, L., & Gunsalus, R. P. (2010). Carbon-dependent control of electron transfer and central carbon pathway genes for methane biosynthesis in the Archaean, *Methanosarcina acetivorans* strain C2A. *BMC Microbiology*, *10*, 62. https://doi.org/10.1186/1471-2180-10-62

141. Carrier, T. A., & Keasling, J. D. (1997). Mechanistic Modeling of Prokaryotic mRNA Decay. *Journal of Theoretical Biology*, *189*(2), 195–209. https://doi.org/10.1006/jtbi.1997.0509

142. Carrier, T. A., & Keasling, J. D. (1999). Library of Synthetic 5′ Secondary Structures To Manipulate mRNA Stability in Escherichia coli. *Biotechnology Progress*, *15*(1), 58–64. https://doi.org/10.1021/bp9801143

143. Jia, L., Mao, Y., Ji, Q., Dersh, D., Yewdell, J. W., & Qian, S.-B. (2020). Decoding mRNA translatability and stability from the 5′ UTR. *Nature Structural & Molecular Biology*, *27*(9), 814–821. https://doi.org/10.1038/s41594-020-0465-x

144. Santos-Navarro, F. N., Vignoni, A., Boada, Y., & Picó, J. (2021). RBS and Promoter Strengths Determine the Cell-Growth-Dependent Protein Mass Fractions and Their Optimal Synthesis Rates. *ACS Synthetic Biology*, *10*(12), Article 12. https://doi.org/10.1021/acssynbio.1c00131

145. Scheller, S., & Rother, M. (2022). Deconstructing *Methanosarcina acetivorans* into an acetogenic archaeon. 1–7. https://doi.org/10.1073/pnas.2113853119/-/DCSupplemental.Published

146. Ma, J.-Y., Yan, Z., Sun, X.-D., Jiang, Y.-Q., Duan, J.-L., Feng, L.-J., Zhu, F.-P., Liu, X.-Y., Xia, P.-F., & Yuan, X.-Z. (2024). A hybrid photocatalytic system enables direct glucose utilization for methanogenesis. *Proceedings of the National Academy of Sciences*, *121*(4), e2317058121. https://doi.org/10.1073/pnas.2317058121

147. Catlett, J. L., Ortiz, A. M., & Buan, N. R. (2015). Rerouting Cellular Electron Flux To Increase the Rate of Biological Methane Production. *Applied and Environmental Microbiology*, *81*(19), Article 19. https://doi.org/10.1128/AEM.01162-15

148. Li, M., Chen, J., Wang, Y., Liu, J., Huang, J., Chen, N., Zheng, P., & Sun, J. (2020). Efficient Multiplex Gene Repression by CRISPR-dCpf1 in Corynebacterium glutamicum. *Frontiers in Bioengineering and Biotechnology*, *8*, 357. https://doi.org/10.3389/fbioe.2020.00357

149. Liu, W., Tang, D., Wang, H., Lian, J., Huang, L., & Xu, Z. (2019). Combined genome editing and transcriptional repression for metabolic pathway engineering in *Corynebacterium glutamicum* using a catalytically active Cas12a. *Applied Microbiology and Biotechnology*, *103*(21), 8911–8922. https://doi.org/10.1007/s00253-019-10118-4

Business, Economy

Art, Design, Architecture

Science, Technology

Crossover

Doctoral Theses

Aalto DT 20/2025

ISBN 978-952-64-2377-7
ISBN 978-952-64-2378-4 (pdf)

Aalto University
School of Chemical Engineering
Department of Bioproducts and Biosystems
aalto.fi