Doctoral Dissertation

Academic Year 2015

Multilateral understanding of hot springs by omics-based approaches

Graduate School of Media and Governance

Keio University

Shinnosuke Murakami

Title

Multilateral understanding of hot springs by omics-based approaches

Abstract

Hot springs are unique natural environments that have been used for recreational and/or therapeutic purposes (balneotherapy) since ancient times in various countries, especially, Japan. Japanese hot spring therapy, called “touji”, is world-renowned. Multilateral and comprehensive scientific study of hot springs and their benefits could contribute to making Japan a tourism-oriented country. As a first step, microbes living in hot springs and their genes were analyzed to understand the hot spring mircobiome. We detected 27 novel microorganisms and type-specific tRNA degradation. Hot spring water (HSW) consumption is one of the methods of balneotherapy. It has been reported that consumption of hydrogen carbonate- or sulfur-containing water may be prevent and/or improve type 2 diabetes. However, since the molecular mechanisms underlying the effects of balneotherapy have not been well elucidated, the physiological effects of HSW consumption were evaluated using omics-based approaches. In the HSW consumption periods, serum glycoalbumin levels, a glycemic control index, were significantly decreased. Metabolome analysis showed that concentrations of 19 blood metabolites including 4 glycolysis-related metabolites and 3 amino acids were significantly changed in the HSW consumption periods as compared with the tap water consumption periods, suggesting that HSW consumption may induce glycolysis and proteolysis alteration. Additionally, 8 families of gut microbiota were significantly changed, out of which lean-associated bacteria was significantly increased. Moreover, experiment on murine models was also conducted and these models may be useful for screening to evaluate the effectiveness of HSW consumption. The current research provides beneficial information for future studies investigating the molecular basis of balneotherapy. Taken together, our findings provide new insights into the microbial ecosystems in hot springs and the molecular mechanisms underlying the effects of balneotherapy. These findings contribute to the understanding of hot springs and their effect in improving human health.

Keywords: Hot spring, Environmental microbes, Balneotherapy, Type 2 diabetes, Gut microbiota

論文題目

オミクスアプローチによる温泉の多面的な理解

論文要旨

温泉は世界各国に存在する地球環境の一つであり、紀元前から保養や治療を目的に利用されてきた。現代でも日本をはじめとする各国において、娯楽や様々な疾患の治療を目的に利用されている。中でも日本独自の温泉療法である湯治の文化は世界に広く認知されている。そのため温泉に関して科学的かつ多面的に理解を深めることは、日本国が目指す観光立国を実現する上で重要な知見の一つになりうると考えられる。そこで本研究では、まず温泉そのものに対する理解を深めるため、温泉水中の微生物およびその遺伝子を網羅的に解析した。その結果、27 種類におよぶ新種の微生物の存在や tRNA 分子の切断に関する新規知見を得た。一方、このような温泉水を疾患の治療や予防目的で飲用する飲泉と呼ばれる温泉療法が存在し、炭酸水素塩や硫黄を含む温泉水の飲用は 2 型糖尿病の予防や治療効果が期待されている。しかし、飲泉によって有益な効果がもたらされるメカニズムはほとんど明らかにされていないため、本研究ではオミクスアプローチによって飲泉が生体へおよぼす影響を解析した。その結果、水道水飲用後と比べ、飲泉後は血糖状態を表す指標の一つであるグリコアルブミン値が有意に低下した。さらに血液のメタボローム解析の結果、19 種の代謝物質が温泉水飲用期間中に有意に増減しており、中には解糖系に関わる 4 種の代謝物質や 3 種のアミノ酸が含まれていた。したがって、体内における解糖系の亢進やタンパク質分解抑制が促されている可能性が示唆された。さらに、腸内細菌叢解析の結果、飲泉後は 8 種の細菌群が有意に増減しており、中でも抗肥満に関係していると考えられる細菌が有意に増加していた。また、本研究ではマウスを用いた飲泉試験も実施し、飲泉効果の有無をスクリーニングするための評価系として利用できる可能性が示唆された。これらの結果は、これまでそのメカニズムが明らかにされてこなかった飲泉による有益効果をオミクスアプローチによって解析したものであり、今後の温泉療法学の発展に寄与する知見をもたらしたものと考えられる。本研究は、温泉を取り巻く微生物生態系および温泉療法の分子メカニズムを解明する上での一助となり、温泉自体やその効能など、温泉の多面的な理解に貢献すると考えられる。

キーワード：温泉、環境微生物、温泉療法、2 型糖尿病、腸内細菌叢

Table of contents

Chapter 3

List of Tables

List of Figures

Chapter 1

Introduction

1.1 What are hot springs?

Hot springs are a unique natural environment that are present widely all over the world. Particularly in Japan, hot springs are familiar to most people as “onsen”. Generally, the term “hot spring” indicates warm water derived from natural springs or drilled wells, but the definition of the term is not yet unified, because the appropriate temperature for a spring to be considered “hot” is still under discussion [1]. Mainly, there are 2 definitions of “hot”: (1) at or above 36.7˚C (human body temperature) and (2) higher than local mean air temperature [1]. The Japanese Hot Spring Law employed the second definition, and spring is certified as “hot” if the temperature is at or above 25˚C (highest local mean air temperature in Japan when the Japanese Hot Spring Law was established). The Japanese Hot Spring Law has an additional definition; springs are also certified as “hot” if the concentrations of at least one of the mineral components meet the criteria defined by the Japanese Hot Spring Law, regardless of the temperature of the water (Table 1.1). Thus, it is not necessary for a Japanese hot springs to be warm. Similar to this, the definition of the hot spring varies across the world. In this dissertation, the term “hot spring” and “hot spring water (HSW)” were used based on the criteria defined by the Japanese Hot Spring Law.

The practice of bathing in natural resources such as hot springs extends across history. In Homeric times (8th century BC), bathing was primarily employed for

cleansing and hygiene [2]. By the time of Hippocrates (460-370 BC), bathing was used not only as a hygienic measure but also for improving health and healing, and it was considered beneficial for various diseases [2]. In ancient Rome, 3 types of baths were developed for home, private, and public use. The public baths were developed as "thermae" in huge edifices with a capacity for thousands of people [3]. Throughout the years, the implications of the baths gradually changed from medical treatment to relaxation and pleasure in Roman bathing culture [3], and these persist in the present times. "Touji", a type of Japanese hot spring therapy that became common in 14-15th century (Muromachi period in Japan) [4], is similar to "spa therapy" in European countries. However, a previous review indicates that there are some differences between Japanese and European bathing [4]. For example, Japanese baths were often used at 42-43˚C, but Europeans did not prefer this range of temperature. Additionally, acid springs (pH < 3) with strong antibacterial effects and aluminum ion-containing springs (more than 100 mg/kg water) do not exist in Europe.

As described above, hot springs are commonly used all over the world since over 2,000 years ago. Japan has the highest number of hot springs in the world [4]. There are over 27,000 active hot springs located over 3,000 hot spring areas, and hot spring bathing and/or hot spring therapy are still familiar to Japanese people. Additionally, Japanese hot spring culture is highly appreciated all over the world. A statistical survey conducted by the Japan Ministry of Land, Infrastructure, Transport and Tourism (http://www.mlit.go.jp/kankocho/siryou/archive/20110325.html) reported that 45% of foreign tourists were looking forward to visiting hot springs in their visit to

Japan; this percentage was the third highest, following Japanese food (64%) and shopping (52%). Therefore, Japanese hot springs have high potential as tourist attractions. Scientific research into this field will help improve the understanding of hot springs and enhance its importance worldwide, particularly, in Japan.

Table 1.1 Mineral concentrations required for hot spring certification based on the Japanese Hot Spring Law

Minerals	Formula	Criteria (per kg)
Total dissolved materials (exclude gas component)		> 1,000 mg
Carbon dioxide	CO_2	> 250 mg
Lithium ion	Li^+	> 1 mg
Strontium ion	St^{2+}	> 10 mg
Barium ion	Ba^{2+}	> 5 mg
Total iron ion	Fe^{2+} and Fe^{3+}	> 10 mg
Manganese ion	Mn^{2+}	> 10 mg
Hydrogen ion (proton)	H^+	> 1 mg
Bromide ion	Br^-	> 5 mg
Iodide ion	I^-	> 1 mg
Fluoride ion	F^-	> 2 mg
Hydrogen arsenate ion	$HAsO_4^{2-}$	> 1.3 mg
Metaarsenious acid	$HAsO_2$	> 1 mg
Total sulfur	HS^-, $S_2O_3^{2-}$, and H_2S	> 1 mg
Metaboric acid	HBO_2	> 5 mg
Metasilicic acid	H_2SiO_3	> 50 mg
Sodium bicarbonate	$NaHCO_3$	> 340 mg
Radon	Rn	> 2 nCi
Radium salt (as radium)	Ra	> 10 ng

1.2 Hot springs in environmental microbiology

Environmental microbiology is a field that explores the composition and physiology of microbes in the environment. Previous studies reported that the presence of hitherto uncultured microbes in the environment, and it was estimated that less than 1% of microbes are available by cultivation [5, 6]. These reports suggested that none of the environments on earth could be accurately reconstructed in experimental settings. However, Handelsman and colleagues proposed that analysis of the collective genomes derived from environmental microflora (metagenome) can yield fundamental knowledge and useful molecules, even if the microbes could not be cultured [7]. Therefore, environmental microbiological study using the metagenomic approach is very important in further understanding microorganisms.

Additionally, previous studies reported genome sequences of novel archaea that were reconstructed from metagenomic libraries without cultivation and isolation [8, 9]. Moreover, metatranscriptome analysis of environmental samples can also show which genes are actually important for survival in various environments [6]. Therefore, metagenome and metatranscriptome analyses of environmental samples could yield new insights into the microbial ecosystems, even though the archaea and/or bacteria are unculturable.

Recently, the distribution of archaea and bacteria in various environments has been represented using microbiome analysis [10, 11]. The studied environments include hot springs [12-15], oceans [16], hydrothermal vents [17], gold mines [9, 18], human skin [19, 20], human lacrimal fluids [21], and human gut [22]. In Japan, unique archaea

and bacteria have been isolated from hot spring environments: *Thermus thermophilus* was isolated from the Mine hot spring (Shizuoka, Japan) [23], *Sulfolobus* sp. NOB8H2 was isolated from the Noboribetsu hot spring (Hokkaido, Japan) [24], *Vulcanisaeta distributa* was isolated from the Owakudani hot spring (Kanagawa, Japan) [25], and *Sulfolobus tokodaii* was isolated from the Beppu hot spring (Oita, Japan) [26]. Examples from other countries include *Azoarcus taiwanensis*, isolated from a hot spring in the Yang-Ming Mountain (New Taipei City, Taiwan) [27], and *Thermus aquaticus* [28] and *Thermodesulfovibrio yellowstonii* [29], isolated from hot springs in the Yellowstone National Park (Wyoming, USA). Thus, many novel archaea and bacteria have been isolated from hot springs. Hot spring environments are generally thermal or hyperthermal and contain unique minerals at high concentrations. Additionally, the origins of hot springs are mainly located in the deep subsurface. Therefore, only unique specific microorganisms are expected to be adapted to survive in hot spring environments. One of the theories proposed that life originated in submarine hydrothermal vents [30]; therefore, microbes inhabiting hydrothermal environments such as hot springs might maintain some features of the ancient, original life forms.

Additionally, interesting and/or useful genes have been found in microbes isolated from hot springs. For instance, a unique DNA segregation machinery has been found from *Sulfolobus* sp. NOB8H2 [31], and a thermostable DNA polymerase was isolated from *T. aquaticus* [32] (it is commonly known as Taq DNA polymerase). The optimum temperature for growth of *T. aquaticus* is 70 ˚C [28], a characteristic necessary to thrive in thermal environments; this characteristic also enabled the organism to

synthesize products that are useful for humans.

Taken together, these reports suggest that the hot spring environments, especially in Japan, are rich sources of unique and interesting microbes and their genes. Therefore, further research involving hot spring environments is required to gain new insights into these microorganisms and their genes.

1.3 Hot springs in balneotherapy

Since ancient times, hot springs have been traditionally used as natural baths in various countries such as Greece, Rome, and Japan [2-4]. From the time of Hippocrates, it was considered that hot springs can exert beneficial effects over various diseases [2]. In Japan, hot spring therapy, called “touji”, was developed in the 14-15th century [4]. At present, the use of hot springs for medicinal purposes is defined as Balneotherapy [33]. Balneotherapy is the use of thermal and/or mineral water for treatment of human diseases by employing various methods such as bathing, drinking, mud therapy, and inhalation [34].

The effects of balneotherapy using bathing can be divided into 3 categories: mechanical, thermal, and chemical [33]. The mechanical effects include the displacement of fluids (blood and lymph fluid) induced by hydrostatic pressure, causing hemodilution and increased diuresis. These effects occur regardless of the type of hot spring. The thermal effects, include superficial vasodilation induced by bathing in hot water, and this effect is upregulated by the carbon dioxide contain in the HSW [35]. The chemical effects are attributable to the mineral components in the HSW, and these can be gained through both bathing and drinking. However, the essential mineral components and the ideal concentration of each mineral required for the beneficial effects remain unclear [33].

It has been previously reported that balneotherapy has beneficial effects for various diseases such as type 2 diabetes (T2D) [36-38], rheumatism [39], low back pain [34], fibromyalgia syndrome [40], atopic dermatitis [41], and cardiovascular disease [35,

42]. In particular, consumption of hydrogen carbonate- or sulfur-containing water has been reported to prevent or improve T2D [36-38].

Complementary and alternative medicine (CAM) is a group of diverse medical and health care systems, practices, and products that are not considered part of any current Western health care system [43, 44]. Balneotherapy, which is expected to have beneficial effects on various diseases, is also considered CAM [45], and the molecular mechanisms underlying the effects of balneotherapy have not been elucidated. Therefore, further investigations are necessary to determine the effectiveness, safety, standard procedures, and potential side effects of the various CAM methods [44].

1.4 Objectives and Summary

The aim of this study was to multilaterally understand hot springs based on omics-based approaches. As described above, hot springs have not only gained the attention of biological researches, but also of tourists. Scientific research on hot springs could promote development of tourism to these sights as well as advancing our understanding of their health benefits. One of the theories proposed that life originated in submarine hydrothermal vents; therefore, microbes living in hydrothermal environments such as hot springs might be important and interesting research targets for research into the origin of life. In addition, it has been known that traditional fermented foods such as yogurt, which contain probiotics, have beneficial effects on human health [46]. Because HSW has also been traditionally used for maintenance of human health, it suggests that microbes present in the HSW may contribute to its therapeutic effects. Therefore, obtaining the microbial profiles of HSW may be important for understanding its therapeutic potential. For this purpose, firstly, microbes living in the HSW and their genes were analyzed in Chapter 2. Additionally, the physiological effects of HSW consumption were evaluated in Chapter 3 to understand the molecular basis of balneotherapy.

In Chapter 2, microbiome analysis of the Yunohama hot spring (Yamagata, Japan) environment was conducted, and unique archaeal and bacterial species were found. In addition, metatranscriptome analysis of the Yunohama HSW revealed novel candidate small non-coding RNAs and type-specific tRNA degradation. These results suggest that hot spring environments are important and interesting research targets in

environmental microbiology and that further novel microbes and their genes will be revealed from hot spring environments in the future.

In Chapter 3, molecular level effects of HSW consumption were analyzed and alteration of the metabolic dynamics and gut microbiota were shown. Additionally, the murine model may be useful as screening tool to evaluate the effectiveness of HSW consumption. These results are expected to be important lines of evidence supporting the beneficial effects of balneotherapy, and this study could be the first step towards understanding molecular basis of balneotherapy.

Taken together, the findings in this dissertation provide new insights into microbial ecosystems in hot spring environments and the molecular mechanisms underlying the effects of balneotherapy. The results of this study can contribute to advancing the understanding of hot springs and their effectiveness in improving human health.

Chapter 2

Metatranscriptome analysis of microbes in an oceanfront deep subsurface hot spring reveals novel small RNAs and type-specific tRNA degradation

2.1 Introduction

There are huge numbers of microbes in the environment, but they are yet to be fully understood. Therefore, this study aimed to understand the population of microbes and their genes in a specified environment.

We selected an oceanfront deep subsurface hot spring, the Yunohama hot spring that is located Yamagata prefecture, Japan as the target environment in this study. Previously, unique bacteria, *Thermus thermophilus* [23] and archaea, *Vulcanisaeta distributa* [25] were found in Japanese hot springs, therefore further unique findings are expected by applying comprehensive microbiome analysis to Japanese hot spring environment.

This study focused on the small noncoding RNAs (sRNAs) as the target genes. sRNAs are known to be one of the major players in the regulation of gene expression in the prokaryote, in both bacteria and archaea [47-49]. A huge number of sRNA studies, particularly that of bacteria, have been undertaken in the last decade [48, 50]. For example, over 100 sRNAs from *Escherichia coli* have been registered in RegulonDB [51], which have been identified with several methods: high-throughput computational searches [52-55], shotgun cloning [56, 57], and tiling array analyses [58, 59]. In recent

years, a new powerful method, high-throughput sequencing technology, has been used to discover new sRNAs in bacteria and the results have shown that hundreds of undiscovered sRNA genes might still exist in the *E. coli* genome [60, 61]. It has been reported that the majority of bacterial sRNAs are regulators of gene expression in response to environmental stresses [62] and are typically involved in the modulation of translation efficiency and the stability of mRNA molecules [50].

Interestingly, many tRNA fragments are observed in the small RNA fractions prepared from *E. coli* [57, 61]. Developmentally regulated cleavage of tRNAs has also been observed in the bacterium *Streptomyces coelicolor* [63]. Plasmid-encoded enzymes, the colicins, are involved in the cleavage of the anticodons of specific tRNAs in *E. coli* [64, 65], resulting in the production of tRNA fragments. Similarly, it is reported that tRNA cleavage is reportedly a response to stresses in the eukaryote [66, 67], and cancer- or age-associated changes in fragmented tRNAs have been observed [68, 69]. Recently, rRNA fragmentation has become another topic of interest [70, 71], although the functions of fragmented rRNAs remain unclear. There are also accumulating number of papers that report the unique RNA regulatory sequences called "riboswitches" in prokaryotic cells [72]. These regulatory RNA sequences are predominantly located in the 5' untranslated regions of some mRNAs and control target translation [72]. Taken together, these findings confirm that sRNAs are important regulators in the prokaryote.

However, previous sRNA studies were usually conducted in model organisms, so the knowledge of environmental sRNAs derived from uncultured microbes is still

sparse. On the other hand, microbiome analysis is a powerful technique, which can establish population dynamics and phylogenies of uncultured microbes. However, the metagenomic approach does not identify the kinds of genes that are actually expressed in certain environments, which reflect the major functional activities and adaptations of the microorganisms that express them. Therefore, the metatranscriptome approach has become increasingly important in defining these overall environmental gene expression profiles [6, 73, 74] and in identifying both the precursor and processing status of such transcripts. Therefore, a combination of metagenome and metatranscriptome approaches would allow an overview of microbial community diversity and the dynamics of their gene expression to be established. Like other sRNA studies, recent reports have revealed the presence of large structured noncoding RNAs [75] and depth-specific novel sRNAs in the ocean [73], suggesting that these sRNAs play important roles in the adaptations required to live in this environment.

In the current study, we used both microbiome analysis and metatranscriptome approach to collect more information about sRNAs from microorganisms in hot spring water (HSW), which have minimal exposure to the external environment. The unique archaeal and bacterial species, and novel candidate sRNAs found in the environment, including huge amounts of tRNA fragments, are discussed.

2.2 Materials and Methods

2.2.1 Sample collection and DNA/RNA extraction

The sample was obtained from the Yunohama hot spring (Tsuruoka, Yamagata, Japan: 38°78′04.57″N, 139°75′21.18″E), which has a temperature of 57°C and a pH of 8.1, on 4 August, 2009. Unfiltered Yunohama HSW (20 L) was bottled from a tank into which the hot water had been pumped (by eight pumps) from depths ranging from 250 to 1,000 m for mixing and storage. For genomic DNA preparation, the sample water was concentrated by filtration using an UltraClean Water DNA Isolation Kit (MO BIO Laboratories), according to manufacturer's protocol. The genomic DNA (14.3 μg) was then extracted from 3.5 L of the sampled water using the same kit. For sRNA preparation, 3L of the sampled water was centrifuged (333 mL × 9 times) at 11,000×*g* for 10 min at 4°C. The pellet was stored at –80°C until analysis. The sRNA was prepared with the *mir*Vana miRNA Isolation Kit (Life Technologies). The stored pellet was initially resuspended in 1 mL of cold phosphate-buffered saline, to which 600 μL of lysis/binding buffer (*mir*Vana miRNA Isolation Kit) was added. Finally, the sRNA fraction (approximately < 200 bp in size; 650 ng) was isolated from 3 L of the sample water.

2.2.2 PCR amplification and determination of 16S rRNA gene sequences

16S rRNA genes were amplified from the genomic DNA isolated from the Yunohama sample, using an archaeal primer set described by Gantner *et al.* [76] (340F, 5'-CCCTAYGGGGYGCASCAG-3' and 1000R, 5'-GGCCATGCACYWCYTCTC-3'),

and a commonly used bacterial universal primer set described by Wilms *et al.* [77] (341F, 5'-CCTACGGGAGGCAGCAG-3' and 907R, 5'-CCGTCAATTCCTTTGAGTTT-3'). PCR was performed with TaKaRa Ex Taq DNA polymerase (Takara Bio Inc.). For the archaeal 16S rRNA genes, amplification proceeded with one denaturation step at 98°C for 2 min, followed by 40 cycles of95 °C for 30 sec, 57°C for 30 sec, and 72°C for 1 min 30 sec, with a final extension step at 72°C for 7 min. The bacterial 16S rRNA sequences were amplified with one denaturation step at 95°C for 5 min, followed by 30 cycles at 95°C for 30 sec, 50°C for 30 sec, and 72°C for 2 min, and with a final extension step at 72°C for 2 min. To amplify as many unknown archaeal species as possible, another set of primers (340F2, 5'-CCCTAYGGGGYGCASCAGGC-3' and 932R, 5'-GCYCYCCCGCCAATTCMTTTA-3') was designed for the archaeal 16S rRNA genes. The PCR amplification with these primers proceeded with one denaturation step at 95°C for 2 min, followed by 35 cycles of 95°C for 30 sec, 50°C for 15 sec, and 72°C for 1 min, with a final extension step at 72°C for 2 min.

The amplified PCR products (Fig. 2.1) were then cloned into the pGEM-T Easy vector (Promega Corporation). The resulting plasmids were transformed into competent *E. coli* DH5α cells (Takara Bio Inc.). The colonies were selected by blue-white screening on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside and isopropyl-β-D-thiogalactopyranoside plates. After the preparation of the plasmid DNAs, their inserts were sequenced on both the sense and antisense strands with an ABI PRISM 3100 DNA Sequencer (Life Technologies) using either the T7 primer

(5’-TAATACGACTCACTATAGGG-3’) or the SP6 primer (5’-ATTTAGGTGACACTATAGAA-3’).

2.2.3 Construction of a cDNA library from the sRNA fraction and 454 pyrosequencing

The sRNAs were initially treated with tobacco acid pyrophosphatase (Nippon Gene), and then used to construct a cDNA library. The cDNA library was constructed with the Small RNA Cloning Kit (Takara Bio Inc.) with the following procedures: 1) 3’ adapters (5’-CATCGATCCTGCAGGCTAGAGAC-3’) and 5’ adapters (5’-AAAGATCCTGCAGGTGCGTCA-3’) were ligated with T4 RNA ligase; 2) the transcripts were then reverse transcribed to cDNAs with the reverse transcription primer (5’-GTCTCTAGCCTGCAGGATCGATG-3’, which is antisense sequence of the 3’ adapter); 3) the cDNAs were amplified with the F primer (the same DNA sequence as the 5′ adapter) and the reverse transcription primer for 15 cycles, followed by a second PCR for 18 cycles with the same primer set. The amplified cDNAs were separated on a 10% polyacrylamide gel, and the fragments were cut from the gel. The excised gel fragments were incubated at 37°C for 4 hour in a buffer containing 0.5 M ammonium acetate and 1 mM EDTA. The eluted cDNAs were purified by ethanol precipitation and dissolved in TE buffer. The final preparation of the sample was confirmed by 10-20% polyacrylamide gel electrophoresis (Fig. 2.2), and the samples were used for pyrosequencing with GS FLX (Roche Diagnostics).

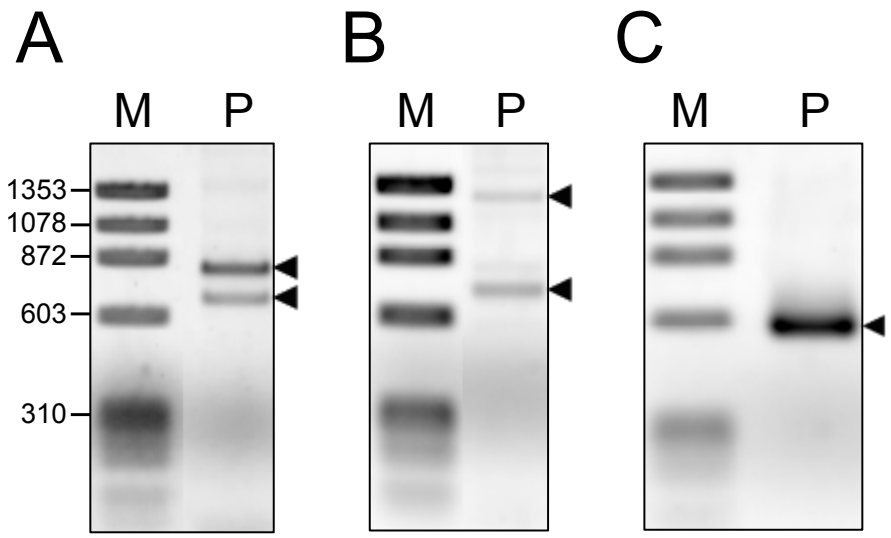


Figure 2.1 PCR amplification of microbial 16S rRNA genes

16S rRNA genes were amplified using either (A) a reported archaeal primer set (340F and 1000R), (B) the originally designed archaeal primer set (340F2 and 932R), or (C) the universal bacterial primer set (341F and 907R). Genomic DNA isolated from the Yunohama HSW was used as the template for all PCRs. The amplified PCR products (P) were visualized by ethidium bromide staining after 1.5% agarose gel electrophoresis with the DNA size marker φX174/*Hae*III digest (M). Arrowheads indicated the positions of specific PCR products.

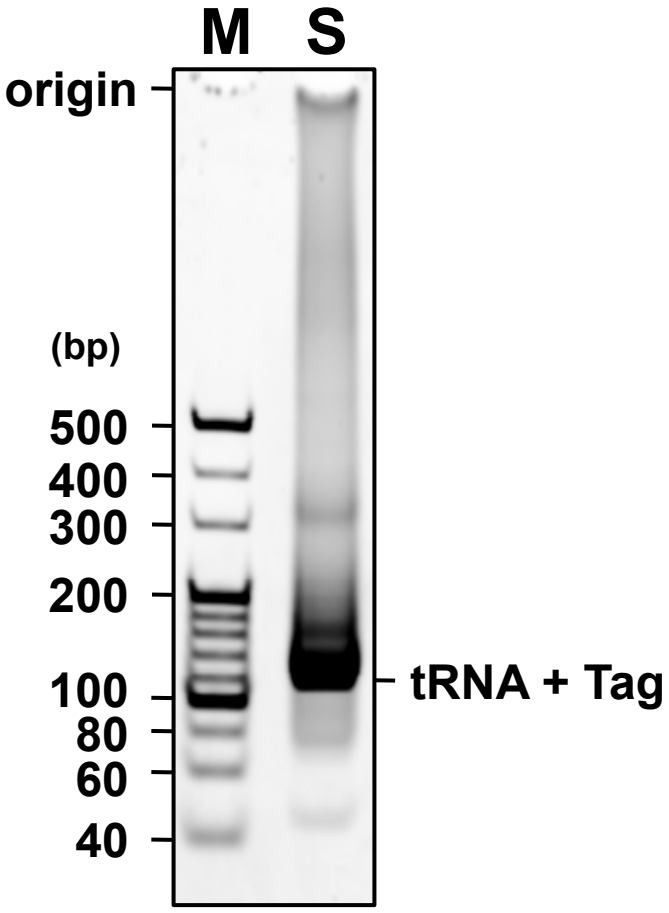


Figure 2.2 Gel electrophoresis of the cDNA library used for metatranscriptome analysis

A cDNA library of the Yunohama-derived small RNAs (< 200 nt) was constructed using the Small RNA Cloning Kit (Takara Bio Inc.) and was visualized by 10-20% polyacrylamide gel electrophoresis (S) together with a 20-bp ladder marker (M).

2.2.4 Bioinformatic analysis

The following software and databases were used for the analysis of the pyrosequencing data: tRNA sequences were detected with tRNAscan-SE [78], and the microbial taxon corresponding to each tRNA sequence was predicted based on the genomic tRNA database, GtRNAdb (http://gtrnadb.ucsc.edu/) [79]; 5S rRNA sequences were detected with the Basic Local Alignment Search Tool (BLAST) (ftp://ftp.ncbi.nih.gov/blast/executables/) against the 5S Ribosomal RNA Database [80]; fragment sequences of tRNA and 5S rRNA were detected with a BLAST search against the mature tRNA and 5S rRNA sequences detected in this study, respectively; 16S rRNA and 23S rRNA partial sequences were detected with a BLAST search against the Ribosomal Database Project (RDP) (http://rdp.cme.msu.edu/) [81] and the SILVA database [82]; and other noncoding RNAs were detected by comparison with the Rfam database (http://rfam.sanger.ac.uk/) [83]. We have summarized the numbers of redundant/nonredundant reads and the proportions of different types of sRNAs in Table 2.1. Sfold (http://sfold.wadsworth.org) [84, 85] was used to predict RNA secondary structures and to calculate the free energies of the candidate sRNAs.

For the phylogenetic analysis, representative 16S rRNA gene sequences (either 12 archaeal or 15 bacterial sequences) obtained from the Yunohama PCR clones were combined with known 16S rRNA gene sequences registered in the RDP (either 305 archaeal or 316 bacterial sequences), and then each phylogenetic tree was constructed using ClustalW 2.0 [86] with 100 bootstrap iterations. The tree was visualized with Interactive Tree Of Life (iTOL) [87].

2.2.5 Nucleotide sequence accession numbers

The pyrosequencing data have been submitted to the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/) under accession number DRA000364. The 16S rRNA gene sequences determined in this study have been deposited in the DDBJ database (http://getentry.ddbj.nig.ac.jp/) under accession numbers AB665420-AB665440, AB679512-AB679519, and AB683043 (Table 2.2). The small unique RNAs from Yunohama (SURFYs) were also deposited in the DDBJ database under accession numbers AB665588-AB665602 (Table 2.3).

Table 2.1 Number of reads and the percentage of each type of sRNA in the cDNA library

	Sequence type			
	Nonredundant		Redundant	
Small RNA type	No.	%	No.	%
tRNA				
Mature	8,069	40.2	43,092	67.1
Fragment	2,063	10.3	4,484	7.0
rRNA				
Mature 5S	532	2.7	2,036	3.2
5S fragment	320	1.6	673	1.0
16S fragment	3,499	17.4	5,640	8.8
23S fragment	3,348	16.7	5,133	8.0
Other noncoding RNA				
SRP bact[*1]	36	0.18	66	0.10
TPP[*2]	4	0.02	4	0.01
6S RNA	4	0.02	4	0.01
Alpha RBS[*3]	1	0.00	2	0.00
Unclassified				
SURFY (≥ 10 reads)	50	0.25	186	0.29
Others (< 10 reads)	2,131	10.6	2,874	4.5

[*1] Bacterial signal recognition particle RNA.

[*2] TPP riboswitch (THI element).

[*3] Alpha operon ribosome-binding site.

Table 2.2 Information and accession numbers for the 16S rRNA gene sequences determined in this study

	Clone #	Length (nt)	Number of related clones	Putative taxonomy	DDBJ/EMBL/ GenBank accessions
Archaea	A-110	784	53	ARMAN	AB679512
(340F-1000R)	A-115	646	14	Methanococcales	AB679514
	A-127	738	5	ARMAN	AB665425
	A-121	641	5	SAGMCG	AB679517
	A-120	645	3	Methanococcales	AB679516
	A-112	640	3	MCG	AB679513
	A-116	640	2	MCG	AB679515
	A-147	640	2	MCG	AB679518
Archaea	A-14	657	52	ARMAN	AB679519
(340F2-932R)	A-9	1187	6	HWCG I	AB665420
	A-52	611	3	ARMAN	AB683043
	A-32	507	1	AAG	AB665423
	A-54	513	1	FSCG	AB665424
	A-12	676	1	Marine Group	AB665421
	A-17	515	1	Thermoplasmatales	AB665422
Bacteria	B-73	548	48	β-Proteobacteria	AB665436
	B-65	548	20	β-Proteobacteria	AB665434
	B-35	549	7	β-Proteobacteria	AB665431
	B-10	542	5	Bacteroidetes	AB665427
	B-15	523	3	Aquificae	AB665428
	B-32	541	1	Bacteroidetes	AB665430
	B-38	542	1	Bacteroidetes	AB665432
	B-26	543	1	Bacteroidetes	AB665429
	B-71	541	1	Deinococcus-Thermus	AB665435
	B-3	551	1	δ-Proteobacteria	AB665426
	B-47	525	1	δ-Proteobacteria	AB665433
	B-75	550	1	δ-Proteobacteria	AB665437
	B-83	552	1	δ-Proteobacteria	AB665438
	B-96	550	1	γ-Proteobacteria	AB665440
	B-91	540	1	Unknown	AB665439

Table 2.3 Information of SURFYs

DDBJ/EMBL/ GenBank accessions	Name	Reads	Length (nt)	GC (%)	ΔG (kcal/mol)
AB665588	SURFY 01	19	38	60	–6.7
AB665589	02	17	101	57	–50.2
AB665590	03	14	68	58	–21.6
AB665591	04	14	27	51	–0.8
AB665592	05	13	62	59	–34.8
AB665593	06	12	43	48	–10.6
AB665594	07	12	46	58	–14.4
AB665595	08	12	71	67	–32.7
AB665596	09	11	75	56	–30.1
AB665597	10	11	38	60	–6.3
AB665598	11	11	31	54	–2.9
AB665599	12	10	98	61	–33.5
AB665600	13	10	49	42	–6.4
AB665601	14	10	62	61	–13.6
AB665602	15	10	85	76	–35.9

2.3 Results

2.3.1 Archaeal and bacterial communities in the Yunohama hot spring

To clarify the phylogeny of the microbes in the deep subsurface water of the Yunohama hot spring, 16S rRNA genes from both archaeal and bacterial species were amplified by PCR using domain-specific primer sets. Firstly, archaeal and bacterial species found in the Yunohama hot spring were summarized (Fig. 2.3).

Recently, Gantner *et al.* reported a primer set (340F and 1000R) for archaeal 16S rRNA gene amplification [76]. The authors validated these primers using three diverse environmental samples. Therefore, in the current study, this primer set was used to amplify the archaeal 16S rRNA genes from the genomic DNA prepared from the Yunohama HSW. Two bands were amplified (approximately 680 and 820 nucleotide [nt] in size; Fig. 2.1A). After the PCR products were subcloned, 87 clones were randomly selected, and their nucleotide sequences were determined. DNA sequencing and phylogenetic analysis of the archaeal 16S rRNA clones (Fig. 2.4) revealed that the longer band (58/87 clones) corresponded to only one taxon, the Archaeal Richmond Mine Acidophilic Nanoorganisms (ARMAN) [8], whereas the shorter band (29/87 clones) corresponded to the following three taxa: 1) Methanococcales, 2) a miscellaneous crenarchaeotic group (MCG) [18], and 3) the South African Gold Mine Crenarchaeotic Group (SAGMCG) [88]. Therefore, archaeal species belonging to ARMAN comprised 67% (58/87) of the archaeal clones derived from the Yunohama hot spring (Fig. 2.3A).

Interestingly, intervening repeat sequences were present in two distinct regions of the 16S rRNA gene sequences of the Yunohama ARMAN (Fig. 2.5A and 2.5B), whereas these repeat sequences were not observed in the 16S rRNA gene sequence of the candidate *Micrarchaeum acidiphilum* (ARMAN-2) [89], which is the species phylogenetically closest to the Yunohama ARMAN (Fig. 2.4). Moreover, the Yunohama ARMAN were classifiable into two groups based on the sequence of the second repeat region (Fig. 2.5B), either type A-110 (two repeats) or type A-127 (one repeat), although the sequences of the first repeat region (Fig. 2.5A) were identical in both types. The nucleotide sequences of these repeats are quite conserved (Fig. 2.5C), but the sequences did not have any sequence similarity with previously known sequences according to the BLASTn analysis against the National Center for Biotechnology Information (NCBI) nucleotide collection (nr/nt).

To amplify as many unknown archaeal species as possible, archaea-specific primer set (340F2 and 932R) was designed (Fig. 2.6 and 2.7), and it was also used for the PCR cloning of 16S rRNA genes from the sample of the Yunohama genomic DNA. The primer set was designed to detect uncultured archaea based on 307 recently available archaeal 16S rRNA gene sequences obtained from the RDP (August 2010). PCR amplification of the 16S rRNA genes from the Yunohama genomic DNA again produced two bands, approximately 700 and 1200 nt in size (Fig. 2.1B). In this case, most of the 700 nt band corresponded to ARMAN and only a single clone of the same size classified in the Marine Group [90] (Fig. 2.3B). However, the 1200 nt band (which included an approximately 680 nt insertion in the rRNA sequence; clone #A-9 in Table

2.2) corresponded to the Hot Water Crenarchaeotic Group (HWCG) I [18]. Other smaller clones that did not correspond to any major band were also isolated and belonged to the following three taxa: the Ancient Archaeal Group (AAG) [91], the Forest Soil Crenarchaeotic Group (FSCG) [92], and the Thermoplasmatales.

The bacterial species in this environment were analyzed using a bacteria-specific universal primer set [77] to amplify their 16S rRNA gene sequences. In total, 93 sequences derived from a single PCR band (approximately 600 nt in size; Fig. 2.1C) were identified. These were further classified into 15 nonredundant sequences and mapped to the bacterial phylogenetic tree (Fig. 2.8). A BLAST search showed that one sequence (B-91) was similar (83% sequence similarity) to the 16S rRNA gene sequence of an uncultured bacterium obtained from deep groundwater in a uranium mine in Japan [93]. Whereas, other 14 sequences showed 91-100% sequence similarity to 16S rRNA gene sequences from uncultured bacteria registered in the NCBI database, but all sequences were not identical to any 16S rRNA gene sequences derived from previously isolated bacteria. All these sequences belong to one of the followed six taxa: β-Proteobacteria, Bacteroidetes, Aquificae, Deinococcus-Thermus, δ-Proteobacteria, or γ-Proteobacteria, whereas B-91 remains unclassified, branching outside the domain Bacteria but lacking similarity to the archaea. Three species belonging to the phylum β-Proteobacteria comprised approximately 81% (75/93) of the bacterial clones derived from the Yunohama hot spring environment, and of these, B-35 was rooted at the deepest position in the currently known β-proteobacterial clade (Fig. 2.8).

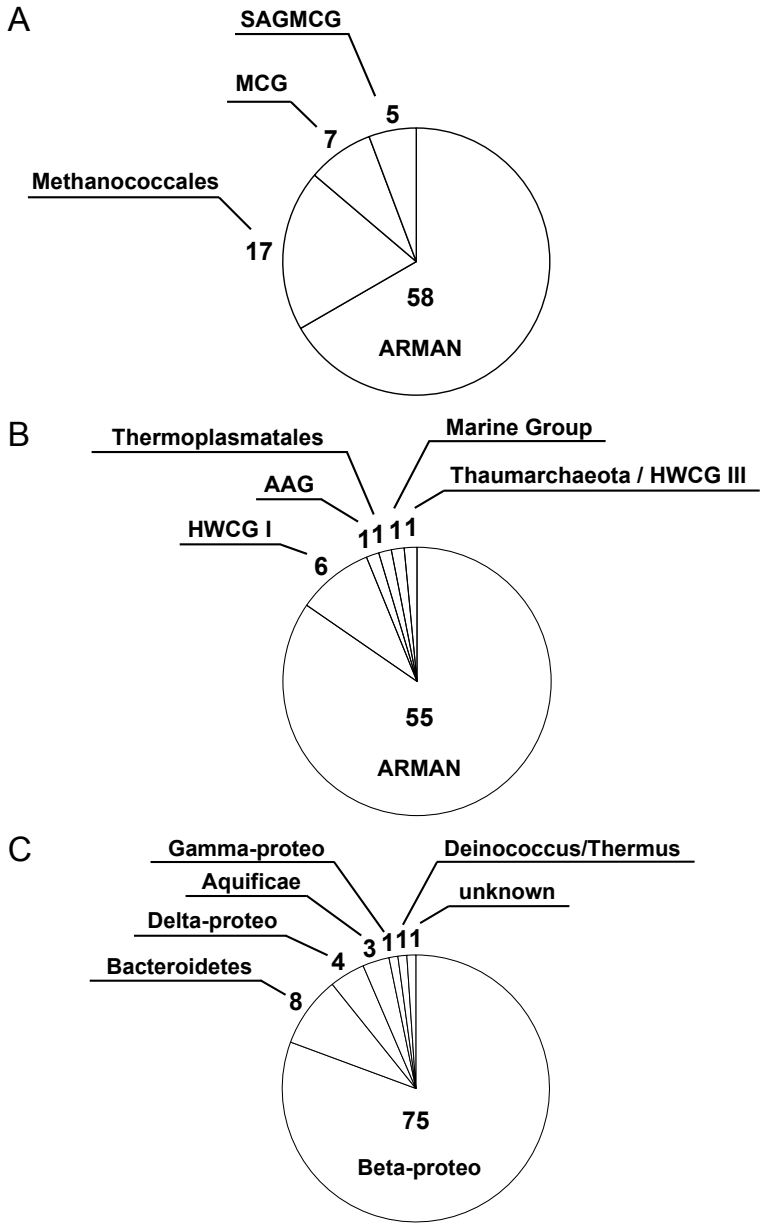


Figure 2.3 Archaeal and bacterial species found in the Yunohama hot spring

Pie charts represent the phylogenetically closest taxa and corresponding read numbers for archaeal (A-B) and bacterial (C) 16S rRNA gene sequences. Archaeal 16S rRNA genes were amplified by using 2 primer sets, 340F-1000R (A) and 340F2-932R (B), and their sequences were analyzed. The closest taxa were estimated based on the phylogenetic analysis. Beta-proteo, Delta-proteo, and Gamma-proteo indicate Betaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria, respectively.

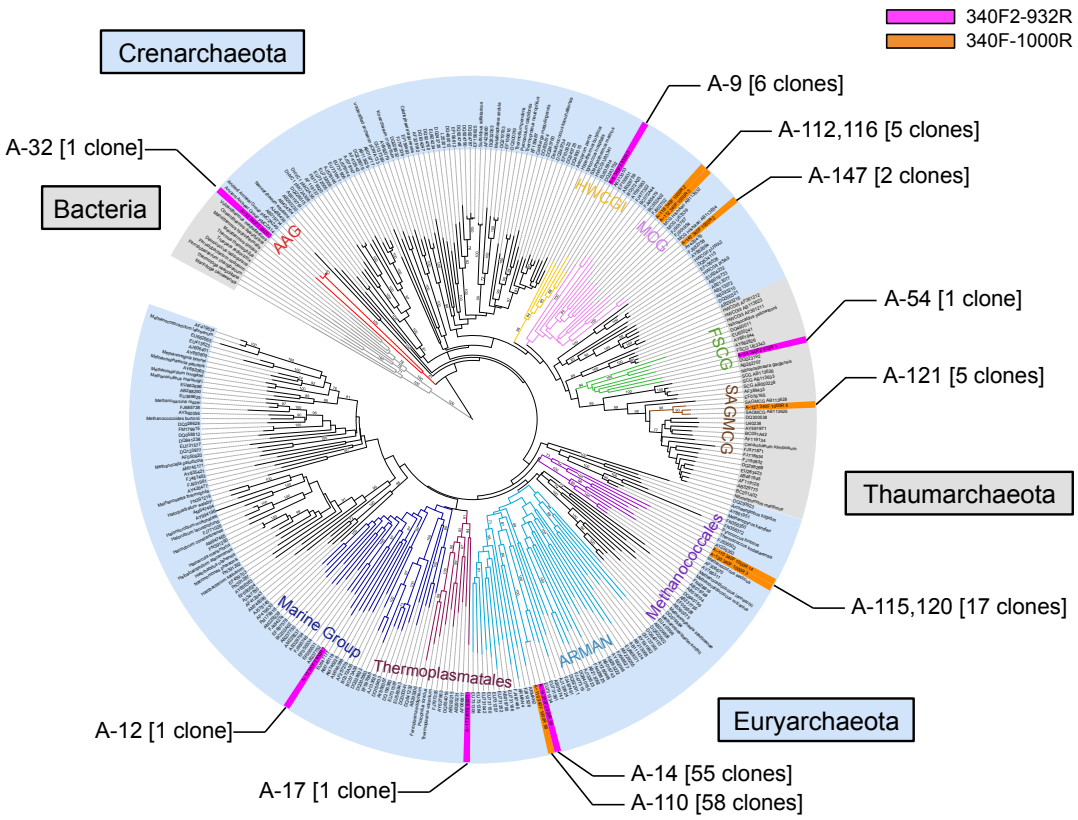


Figure 2.4 Phylogenetic positions of the archaeal 16S rRNA gene sequences isolated from the Yunohama HSW

An archaeal phylogenetic tree was constructed for 318 archaeal 16S rRNA gene sequences, including 13 nonredundant Yunohama-derived sequences, using a bacterial 16S rRNA as the outgroup (gray). The Yunohama-derived sequences were obtained by PCR cloning using either the recently reported archaeal primer set (orange) or the originally designed primer set (magenta). The phylogenetic tree was constructed based on the neighbor-joining method, with 100 bootstrap iterations (bootstrap values > 70 are shown on each branch). Branch colors represent the different phyla or groups containing sequences derived from the Yunohama hot spring. The numbers of sequenced reads are shown in brackets.

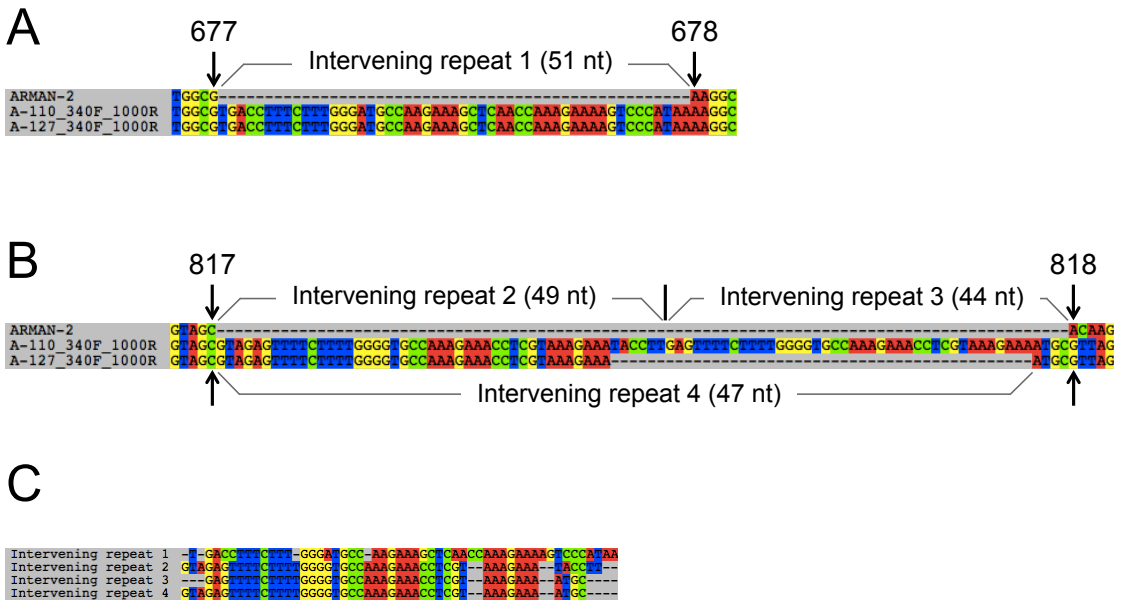


Figure 2.5 Comparison of the intervening repeat sequences found in the 16S rRNA genes of the Yunohama ARMANs

Similar intervening repeat sequences (#1, #2, #3, and #4) were found in two distinct regions (A or B) of the16S rRNA gene sequences from the Yunohama ARMAN. The arrows indicate the nucleotide positions in 16S rRNA gene sequence of ARMAN-2. Comparisons of each intervening repeat sequence (C). The sequences were aligned using ClustalW 2.0. Gaps (-) were inserted to maximize the numbers of nucleotide matches. The aligned sequences in C were also manually improved to maximize the number of nucleotide matches.

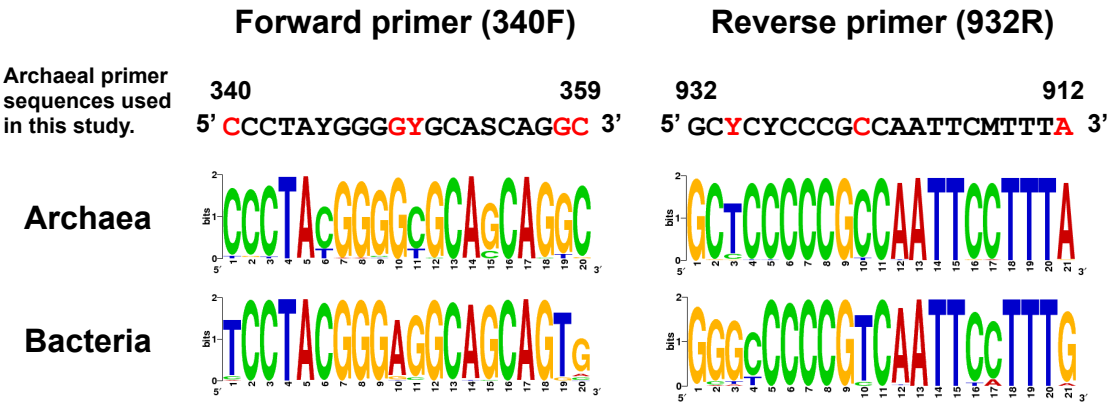


Figure 2.6 PCR primers designed to amplify the archaeal 16S rRNA genes in this study

Archaeal primers were designed based on archaeal and bacterial consensus sequences. The sequence logos were created by WebLogo using 65 archaeal (upper) and 100 bacterial rRNA gene sequences (bottom) obtained from the RDP. The nucleotide positions are numbered based on the 16S rRNA gene of *E. coli*. The nucleotides that differ from the bacterial consensus sequence in the designed archaeal primers are written in red at the top of the sequence logos.

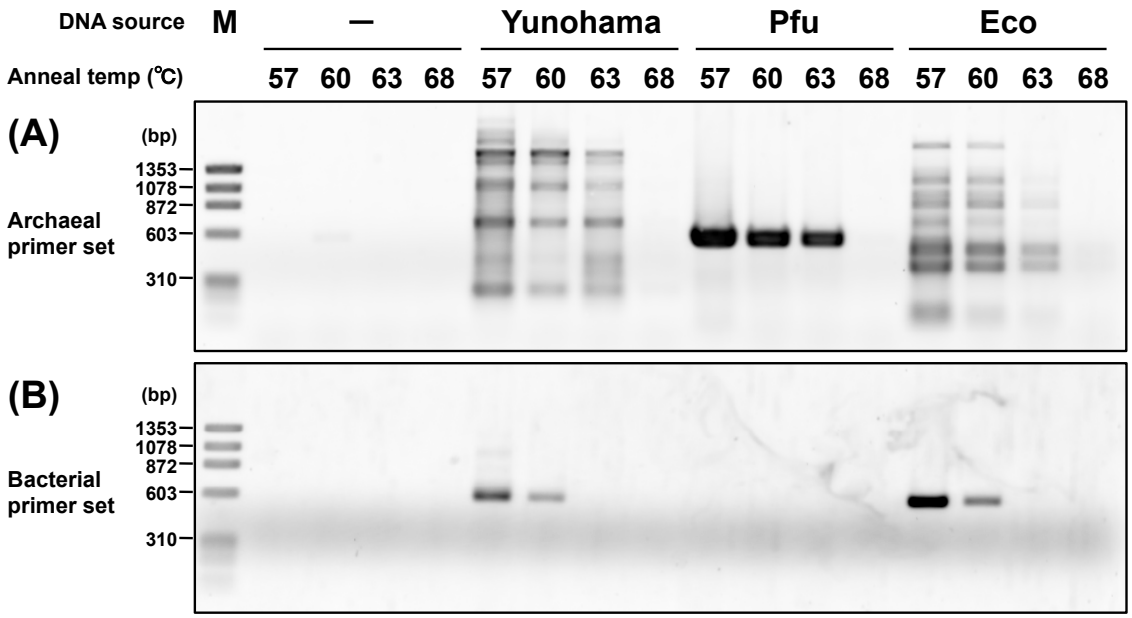


Figure 2.7 Test amplification of microbial rRNAs from different DNAs

(A) The original archaeal primer set (340F and 932R) and (B) the bacterial universal primer set (341F and 907R) were validated using three types of DNAs: blank (-), genomic DNA isolated from the Yunohama HSW (Yunohama), genomic DNA isolated from *Pyrococcus furiosus* (Pfu), and genomic DNA isolated from *E. coli* W3110 (Eco). PCR was performed at four annealing temperatures (57-68°C). The amplified products were visualized by ethidium bromide staining after 1.5% agarose gel electrophoresis with DNA size marker φX174/ *Hae*III digest (M).

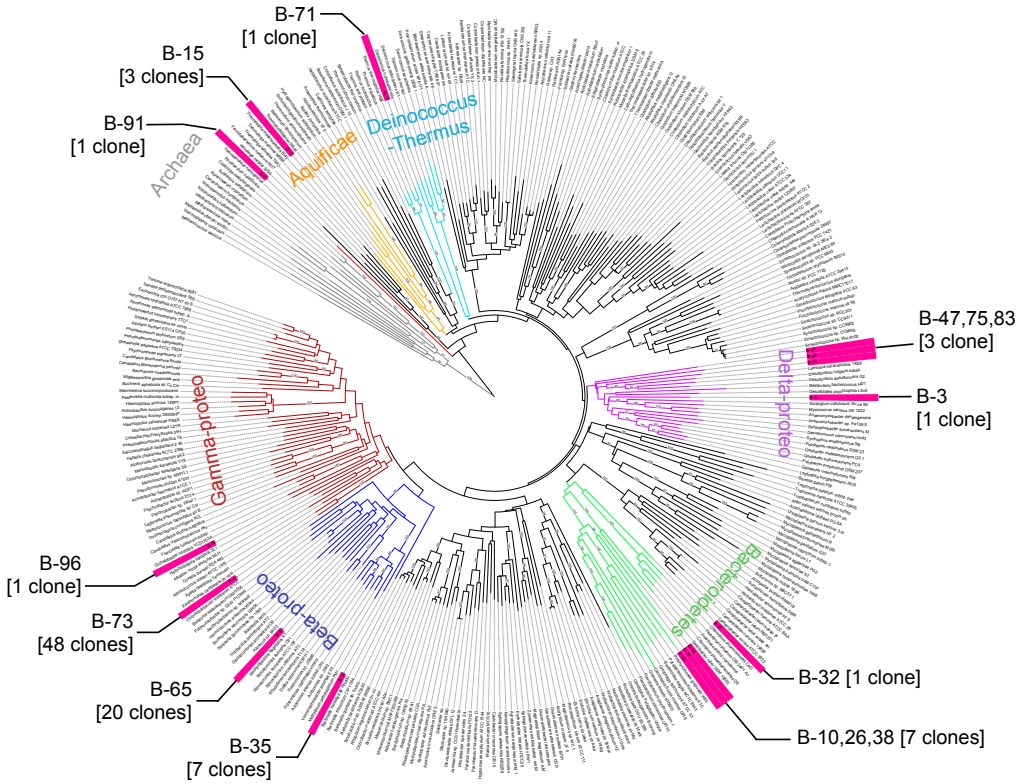


Figure 2.8 Phylogenetic positions of the bacterial 16S rRNA gene sequences isolated from the Yunohama HSW

A bacterial phylogenetic tree was constructed for 331 bacterial 16S rRNA sequences, including 15 nonredundant Yunohama-derived sequences (magenta), using an archaeal 16S rRNA as the outgroup (gray). The method of phylogenetic tree construction and other information are the same as those described in Figure 2.4. Beta-proteo, Delta-proteo, and Gamma-proteo indicate Betaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria, respectively.

2.3.2 Metatranscriptome analysis of candidate sRNAs from the Yunohama hot spring

To understand the sRNA population in the Yunohama hot spring and its sequence characteristics, a metatranscriptome analysis of the sRNA fraction (< 200 bp) was conducted using high-throughput sequencing technology. After 5' and the 3' RNA tags were added to the isolated RNAs for PCR amplification (cDNA library construction), a major band of approximately 120 nt was obtained with 10-20% polyacrylamide gel electrophoresis (Fig. 2.2). High-throughput sequencing the library then produced 86,236 sequence reads.

To characterize these reads, a nonredundant data set was firstly prepared by extracting the reliable reads and combining the redundant reads to yield 20,057 nonredundant reads, ranging in length from 11 to 197 nt (average length, 62 nt) (Fig. 2.9). Known noncoding RNAs were extracted with a series of filtering approaches based on homology searches and prediction software (Fig. 2.9). Of the final reads, 17,831 (88.9%) were either tRNAs, fragmented tRNAs, rRNAs, or fragmented rRNAs. Only 45 reads (0.2%) had certain sequence similarities to the following noncoding RNAs: bacterial signal recognition particle RNA [94], 6S RNA [95], and two known RNA elements, the thiamine pyrophosphate (TPP) riboswitch [96] and the alpha operon ribosome-binding site [97]. According to the BLAST search, archaeal-specific transcripts were not isolated from the Yunohama hot spring sample. Furthermore, as suggested by the tRNA sequence analysis described below, almost all the reads are

expected to have derived from the bacteria. The numbers of these noncoding RNAs are summarized in Table 2.1.

The remaining 2,181 reads (10.9%) were categorized as “unclassified”. These unclassified RNAs potentially include novel classes of sRNAs. To extract reliable candidate sRNAs, the extremely closely related sequences (with more than 95% sequence similarity) were combined, which further narrowed the sequence set to 1,972 reads, ranging in length from 11 nt to 176 nt (average length, 55 nt). A BLASTn analysis against the NCBI nucleotide collection (nr/nt) was conducted, which identified 354 reads (18%) with a certain degree of sequence similarity (with e-values ranging from 2e–4 to 1e–41) to known nucleotide sequences. Of these 354 reads, 266 (75%) were similar to known genomic sequences (mostly bacterial), whereas 88 (25%) showed similarity to environmental DNA sequences. In particular, 71 reads had a degree of sequence similarity to the following known genes: 5S rRNAs, 23S rRNAs, and protein-coding genes. Although rRNAs were identified in the previous steps (Fig. 2.9), a few sequences were still retained in the unclassified dataset because the numbers and kinds of environmental 5S and 23S rRNA gene sequences were incomplete in the databases.

The abundances of the unclassified RNAs showed that 1,574 reads (79.8%) were detected only once, whereas 398 reads (20.2%) were found more than twice (maximum 19 reads). The 15 unclassified RNAs that appeared 10 or more times were then further characterized as candidate sRNAs, because they were relatively highly abundant in the environment (Table 2.3). These 15 sequences were designated SURFYs.

Fourteen of the SURFYs had no sequence similarity to any known nucleotide sequences, whereas SURFY03 shared some degree of similarity (91% identity in 57 nt) to the gene encoding a prokaryotic elongation factor (EF-Tu) in the nitric bacterium *Nitrosococcus halophilus*. The secondary structures of these candidate sRNAs were predicted with Sfold and the free energies (ΔG) of the SURFYs (mean, –20 kcal/mol) were calculated (Table 2.3). The sequence lengths (average, 60 nt) and GC contents (average, 58%) of the SURFYs were also shown in Table 2.3. Because the secondary structures of noncoding RNAs are known to be important for their functions [98], the secondary structures of the four SURFYs with the lowest free energies are shown (Fig. 2.10).

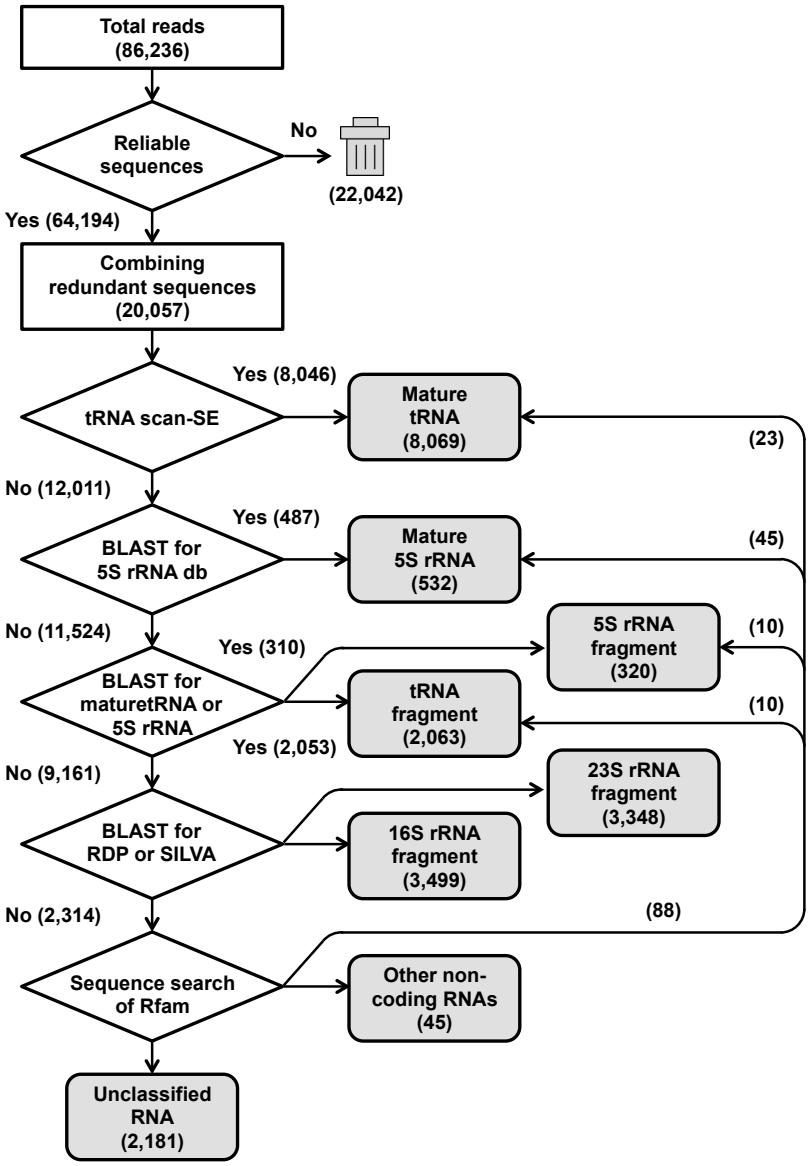


Figure 2.9 Computational scheme for the extraction of known and novel sRNAs

The whole scheme of the computational analysis used in this study is illustrated as a flowchart. The numbers indicate the read numbers at each step. Initially, a dataset containing 20,057 nonredundant reads was constructed from a total of 86,236 reads (raw data). These were classified in a stepwise manner as mature tRNA or rRNA (5S rRNA), fragmented tRNA or rRNA (5S, 16S, or 23S rRNAs), or noncoding RNA. The software tRNAscan-SE and BLAST (provided by NCBI), and the databases RDP, SILVA, and Rfam were used in our analysis.

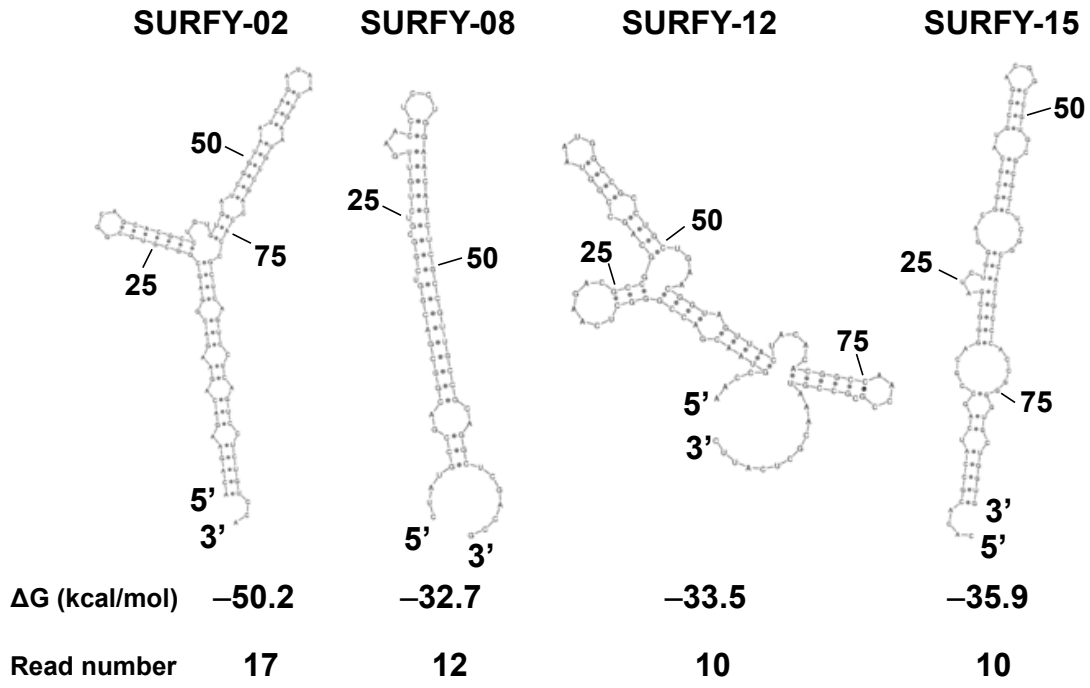


Figure 2.10 Four examples of abundant candidate sRNAs

The nucleotide sequences, predicted secondary structures, and free energies are shown. The secondary structures and free energies (ΔG) were calculated with Sfold. Detailed information about all candidate RNAs is given in Table 2.3.

2.3.3 Environmental tRNAs and their fragments

Of the sRNAs that were isolated from the microbes in the Yunohama hot spring, 74.1% (redundant reads) were either mature tRNA sequences or their fragments (Table 2.1). In total, 8,069 tRNAs (43,092 redundant reads) were classified into 20 groups, based on their corresponding amino acids. Neither selenocystenyl-tRNA nor pyrrolysyl-tRNA was detected in the sRNA library constructed from the Yunohama microbes. The microbial taxa corresponding to each tRNA group were predicted based on the genomic tRNA database (GtRNAdb), and most tRNAs were shown to be derived from bacterial species (Fig. 2.11). Approximately 80% of the mature tRNAs belonged to the β-Proteobacteria (Fig. 2.11), which is consistent with the rRNA gene sequence analysis of the bacterial clade (Fig. 2.3C).

We also found 2,063 tRNA fragments (4,484 redundant reads) with 100% nucleotide sequence identity to one (or more) of the sequenced the full-length tRNAs, suggesting that these fragments were produced through either tRNA processing or the nonspecific degradation of the tRNAs. There was no correlation between the abundances of full-length tRNA reads and the abundances of their corresponding tRNA fragment reads (Fig. 2.12), which indicates that these fragments are not the consequence of nonspecific degradation, but are instead generated by specific exonucleolytic/endonucleolytic degradation pathways. To verify this hypothesis, the position of the tRNA cleavage site in each of the 20 tRNA groups was analyzed and the three most frequently cleaved positions are represented in a piled bar graph (Fig. 2.13). Because most of the tRNA fragments were very short and parts of the tRNA sequences

are common to all tRNAs, many tRNA fragments could be mapped to several independent tRNAs. Therefore, the number of truncated positions was normalized by dividing the number of observed tRNA fragments by the number of corresponding mature tRNAs. These truncation positions are mainly located in the anticodon loop and the D-loop, although tRNA cleavage was also observed at several other positions, such as in the D-stem and variable region. Figure 2.13 shows that the truncation positions of the Lys-tRNAs in the Yunohama hot spring are concentrated in the D-loop. Furthermore, the frequent truncations at 15/16 were commonly seen in tRNA-Ile, tRNA-Lys, tRNA-Asp, tRNA-Asn, and tRNA-Val.

Next, the most frequently cleaved sites for each tRNA anticodon are listed in Figure 2.14. Precisely, tRNAs that are truncated at position 34/35 have a purine nucleotide (A or G) as the first base of their anticodons, whereas tRNAs that are cleaved at positions outside of the anticodon region have a pyrimidine nucleotide (C or U) as the first base of their anticodons (Fig. 2.14). This phenomenon was summarized in the Figure 2.15.

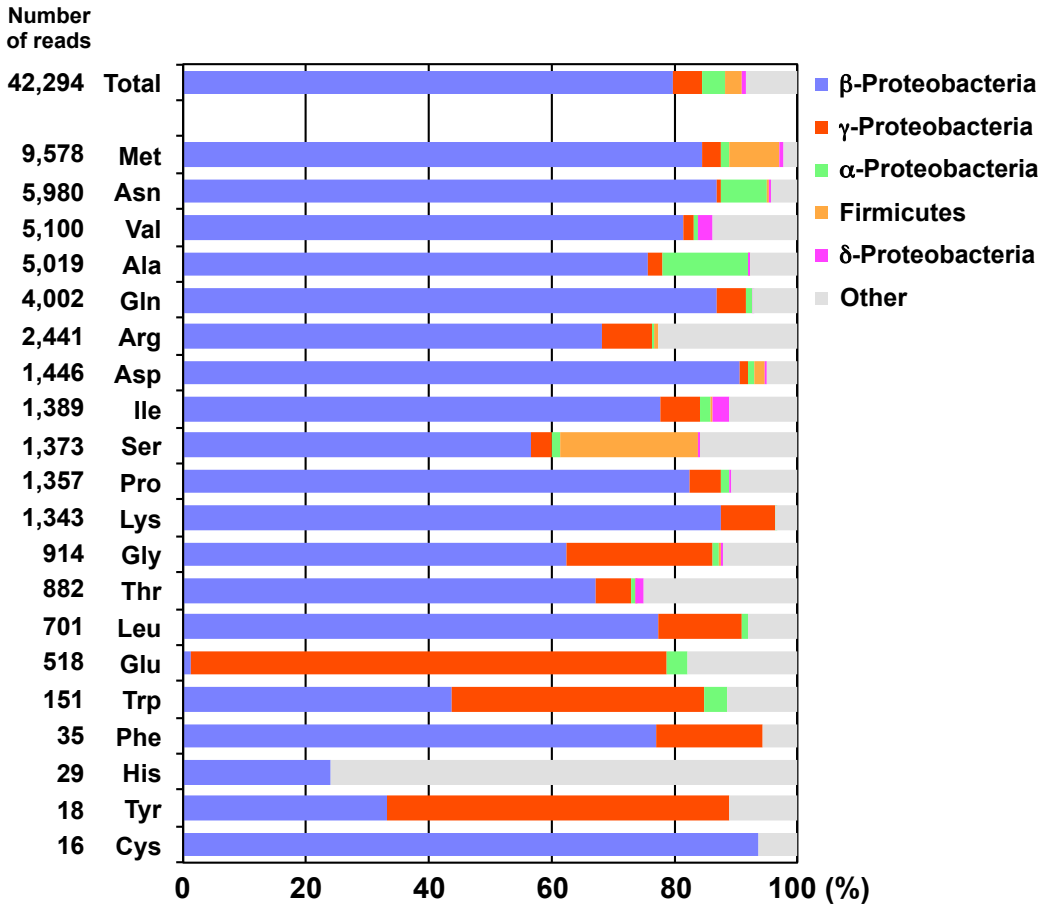


Figure 2.11 Numbers of environmental tRNAs and their corresponding taxa identified in this study

In total, 42,294 redundant tRNA sequences were classified based on the amino acids transferred by them. Suppressor tRNAs and pseudo tRNAs (tRNAscan-SE could not identify their anticodons) are not included. Colored bars indicate the percentage of each taxon identified with a BLAST search of the genomic tRNA database. The five most abundant taxa are colored: β-Proteobacteria (blue), γ-Proteobacteria (red), α-Proteobacteria (green), Firmicutes (orange), and δ-Proteobacteria (magenta); other minor bacterial taxa are colored gray.

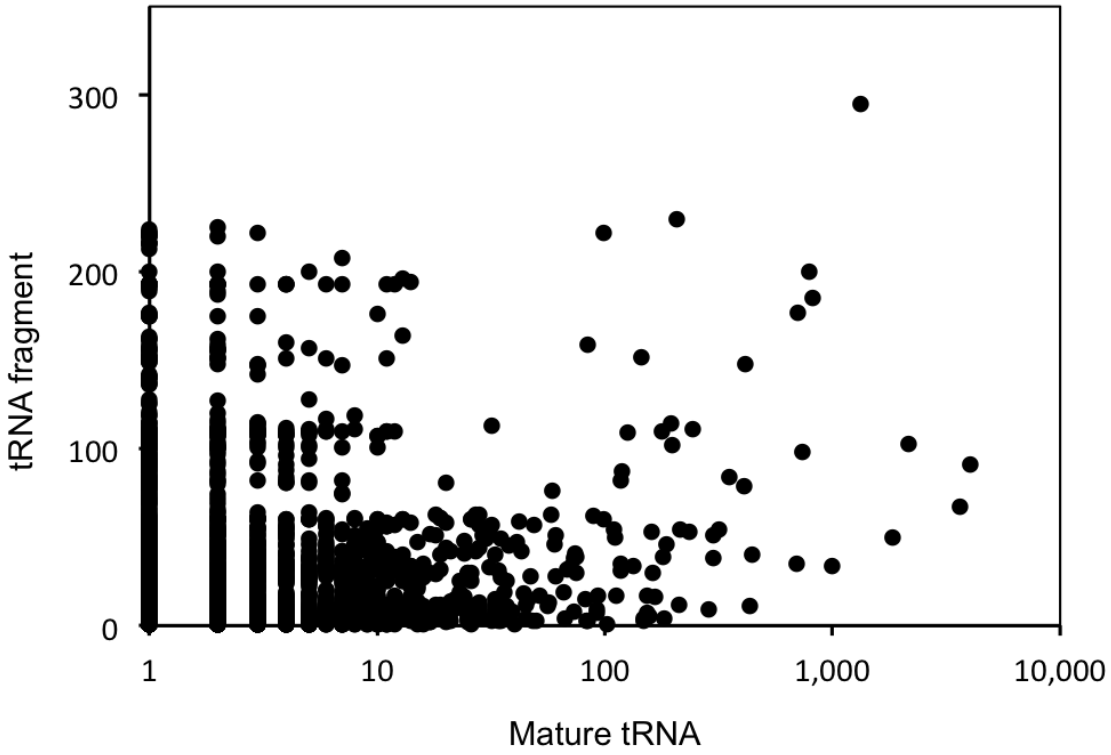
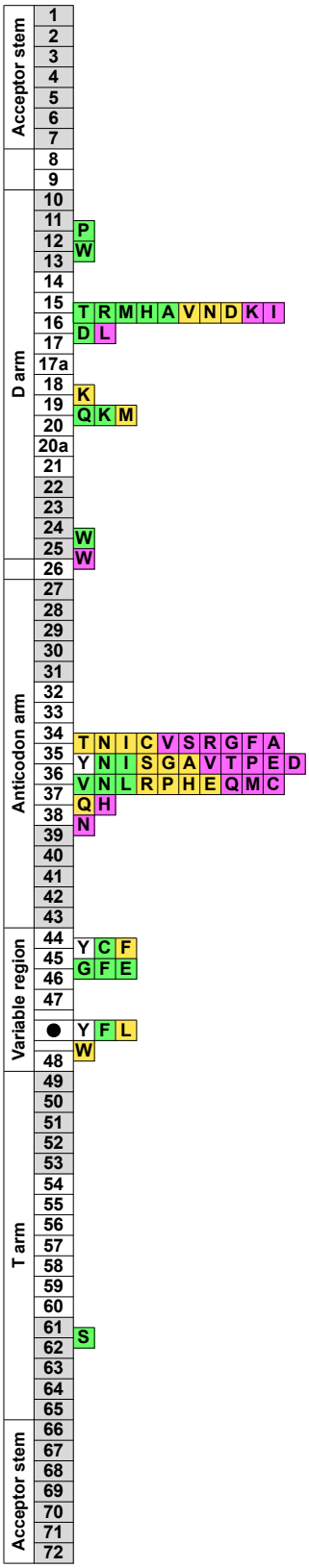


Figure 2.12 Correlation analysis of read numbers for the mature tRNAs and their fragments

A scatter plot of mature tRNAs (numbers of reads of mature tRNA sequences) and tRNA fragments (numbers of reads of partial tRNA sequences) is shown (n = 5,448). The correlation coefficient is 0.15.



(Legend on the next page)

Figure 2.13 Summary of the fragmented tRNAs isolated from the environment Numbered boxes (1-72) indicate the tRNA nucleotide positions based on a comprehensive report of tRNAs [99]. The gray boxes indicate the stem regions and the white boxes indicate the loop regions. The three most frequently cleaved positions are shown for each tRNA that transfers a specific amino acid (alphabetical box): most frequent (magenta), second most frequent (yellow), and third most frequent (green). For example, a magenta box with "A" located between nucleotides 34 and 35 means that tRNAs that transfer alanine were most frequently cleaved between nucleotides 34 and 35 in the anticodon loop. In the case of asparagine tRNA, tryptophan tRNA, and phenylalanine tRNA, the third most frequently cleaved sites were observed at two positions, so two green boxes occur at these positions. The variable stem/loop region is compressed into one box, marked with a black circle, between 47th and 48th nucleotide, and the cleaved positions observed in the variable region are mapped together on the black circle. Because there were few fragments of tyrosine tRNA, the three positions observed twice (most frequently) in the truncated fragments are mapped as white boxes.

1st base of anticodon		2nd base of anticodon								3rd base of anticodon
		A		G		T		C		
A										A
		AAG	-			ATG	37/38	ACG	34/35	G
										T
				AGC	35/36					C
G		GAA	34/35	GGA	35/36	GTA	-	GCA	36/37	A
		GAG	34/35	GGG	36/37	GTG	36/37	GCG	34/35	G
		GAT	15/16	GGT	35/36	GTT	38/39	GCT	34/35	T
		GAC	34/35	GGC	34/35	GTC	35/36	GCC	34/35	C
T				TGA	35/36	TTA	35/36			A
		TAG	V	TGG	36/37	TTG	36/37	TCG	37/38	G
				TGT	15/16	TTT	15/16	TCT	57/58	T
		TAC	15/16	TGC	35/36	TTC	35/36	TCC	35/36	C
C		CAA	-	CGA	V	CTA	V	CCA	25/26	A
		CAG	16/17	CGG	35/36	CTG	37/38	CCG	15/16	G
		CAT	36/37	CGT	35/36	CTT	19/20,57/58	CCT	15/16	T
		CAC	15/16	CGC	35/36	CTC	45/46	CCC	55/56	C

Figure 2.14 Most frequent cleavage sites for each tRNA in the anticodon table

The sites most frequently cleaved in the anticodons of the tRNAs are shown. "V" indicates variable stem/loop regions. Two types of cleavage sites are colored: 34/35 (blue) and outside of the anticodon region (magenta). No fragmented tRNAs with anticodons corresponding to blank cells (white) were isolated. Because the number of fragmented tRNAs with anticodons corresponding to "-" was low, the rankings of these cleavage sites were not calculated.

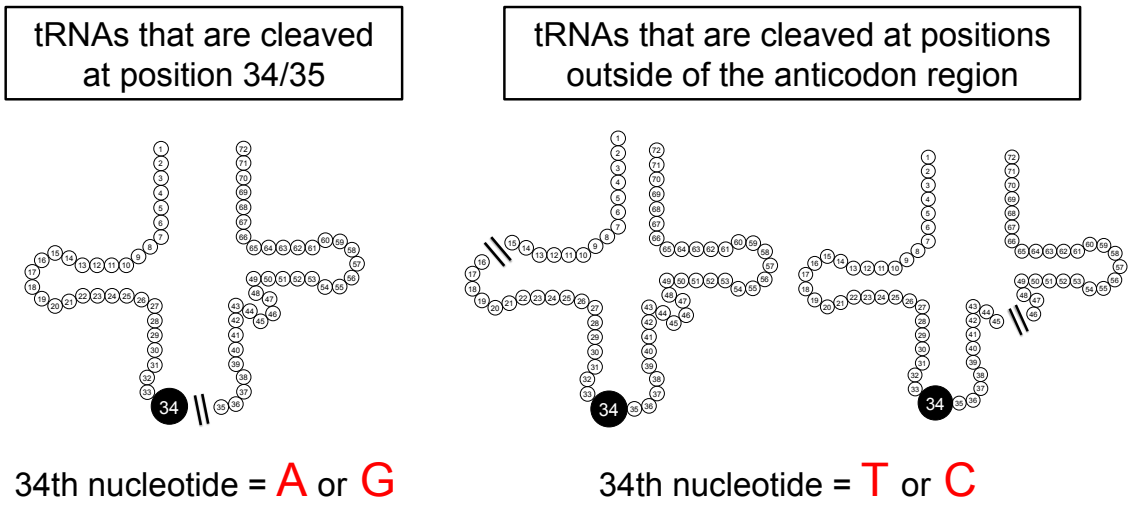


Figure 2.15 Schematic figure of the nucleotide preference for site-specific tRNA degradation

Preference of 34th nucleotide related to site-specific tRNA degradation were summarized. Circled numbers indicates nucleotide positions of canonical tRNA. Based on the results of this study, 34th nucleotide of tRNAs that are well cleaved at position 34/35 are surely A or G, whereas 34th nucleotide of tRNAs that are well cleaved at positions outside of the anticodon region are almost T or C.

2.4 Discussion

In the current study, 16S rRNA genes of archaea were amplified, and their nucleotide sequences were determined. The results indicate 12 types of archaea were possibly present in the Yunohama hot spring environment. Additionally, all of them were expected to be novel species because same nucleotide sequences of their 16S rRNA genes were not registered in NCBI database. In this study, in order to amplify 16S rRNA genes from as many archaeal species as possible, a novel universal primer set was created. 5 out of 12 archaeal 16S rRNA genes were amplified by the newly constructed primer set, suggesting that the new primer set designed in this study can potentially capture further diverse archaeal species.

According to the microbiome analysis of archaea, the most abundant species of Yunohama hot spring environment was closely related to ARMAN. Previously, ARMAN have only been found in acidic environments throughout the world [89]. As such, the Yunohama ARMAN could be the first example of ARMAN living in an alkaline environment (pH 8.1). In the 16S rRNA gene sequences of Yunohama ARMAN, there were intervening repeat sequences comparing with that of previously known ARMAN (candidate *Micrarchaeum acidiphilum*). Additionally, the intervening sequences were repeated once or twice in the 16S rRNA gene sequences of Yunohama ARMAN. This result suggests that the horizontal transfer of the intervening region could possibly have occurred in the deep subsurface hot spring environment.

Among the clones isolated using the originally designed primers, AAG is an interesting taxon, rooted at the deepest position on the archaeal tree, and is only found

in deep-sea hydrothermal vent environments [91]. Therefore, the discovery of archaeal species related to ARMAN and AAG extends the knowledge of the environments inhabitable by these taxa. Due to the fact that the Yunohama hot spring is located on the oceanfront, the detection of AAG, the Marine Group, and Thaumarchaeota members, these groups were originally found in oceanic environments, is reasonable with its geographic location. HWCG I, also known as *Caldiarchaeum subterraneum* [9], has been isolated from subsurface water in the Hishikari gold mine (Kagoshima, Japan), where the pH (pH 6.19-6.80) and temperature conditions (71.5-85.0°C) differ from those at the Yunohama hot spring. Therefore, the archaeal species with rRNA sequences highly similar to those of HWCG I are geographically interesting. The universal primer set for archaeal 16S rRNA gene amplification that was reported by Gantner *et al.* and used in this study also identified clones corresponding to MCG and SAGMCG, which were also observed at the Hishikari gold mine [18], suggesting that these three taxa might be common archaeal groups in subsurface environments in Japan.

In this study, 16S rRNA genes of bacteria were also amplified and their nucleotide sequences were determined. The results indicate 15 types of bacteria were possibly present in the Yunohama hot spring environment. Because nucleotide sequences of their 16S rRNA genes were not identical to 16S rRNA genes registered in NCBI database that were derived from previously isolated bacteria, all bacteria found in Yunohama hot spring may be novel species. Interestingly, phylogenetic analysis indicated that the clone B-91 was branched outside the domain bacteria but lacking similarity to the archaea. In addition, B-35 was rooted at the deepest position in the

currently known β-proteobacterial clade. These results suggest that the various types of microbes coexist, and a lot of unique archaea and bacteria are present in the Yunohama hot spring. Therefore, it is expected that further unique microbes will be found from the hot spring environments, thus the hot spring environments are potentially quite interesting and pose as important targets for environmental microbiological researches.

In the current study, we performed metatranscriptome analysis of small RNA fraction that was prepared from Yunohama HSW. Various types of tRNAs, rRNAs, and sRNA candidates were detected. Although the existence of unique archaeal species were indicated by microbiome analysis, transcripts that were expected to be derived from archaea were not observed. These results suggest that bacterial abundances are greater than that of archaea in the Yunohama hot spring environments and/or archaea living in the environments may exist in dormant states such as spores.

The relative abundances of each bacterial species that were estimated by compositions of mature tRNAs are nearly consistent with that of microbiome analysis, therefore transcripts might be derived from bacteria living in the Yunohama hot spring environment. However, the compositions of four tRNA groups (Glu, Trp, His, and Tyr) are not consistent with that of other types of tRNAs, but the reason remains unclear. Interestingly, metatranscriptome analysis revealed that there are many candidate sRNAs, SURFYs. Their high copy numbers and stable RNA structures suggest that SURFYs are stable in the Yunohama hot spring environment, probably reflecting their roles in specific biological functions.

A recent metatranscriptome analysis revealed unique microbial sRNAs in the ocean water from the HOT Station ALOHA site, Hawaii [73]. The authors identified some types of candidate sRNAs that were also found in the current study, such as signal recognition particle RNA and 6S RNA. Comparative genomics also revealed 104 structured RNAs, including candidate riboswitches, of the archaea and bacteria [100]. It has recently been reported that unique sRNAs existed in some environmental samples [73, 100]. Depth-dependent variations in putative sRNAs have been reported in a metatranscriptome analysis of an oceanic water sample and the authors suggested that these play potential roles in niche adaptation [73].

In this study, many fragmented RNAs especially tRNA fragments were also found by metatranscriptome analysis. As the characteristics of truncate positions of the tRNA fragments were analyzed, preferences of truncate positions were revealed in each type of tRNA. Previous studies of tRNA degradation have usually been performed in *E. coli*, and have demonstrated the specific degradation of only six types of tRNAs: Lys-tRNA is truncated between nucleotides 33 and 34 (33/34) by PrrC; tRNA-Tyr, tRNA-His, tRNA-Asn, and tRNA-Asp are truncated at 34/35 by colicin E5; and tRNA-Arg is truncated at 38/39 by colicin D [64]. In addition, it has been reported that the cleavage of tRNAs were developmentally regulated in the bacterium *S. coelicolor* [63]. However, previous knowledge of bacterial tRNA degradation is limited because only specific types of tRNAs were analyzed, and model organisms were used in these studies. In contrast, the current study has shown the truncate positions of all types of tRNAs, and analysis was performed across various types of microbes living in the

environment by using a metatranscriptome approach. According to the results, many types of tRNAs tended to truncate around anticodon region, but several types of tRNAs such as tRNA-Ile, tRNA-Lys, and tRNA-Phe were well truncated at D-loop region or variable region. Particularly, the truncation position at 15/16 is a novel cleavage site for bacterial tRNA degradation. Nucleotide positions 15 and 16 are known to be exposed in the tertiary tRNA structure. As such, an uncharacterized RNase may have access to this site for its specific cleavage.

Next, truncate positions were listed by anticodon sequences of tRNAs rather than charging amino acids of tRNAs. These results suggest that there is a nucleotide preference for site-specific degradation at the first base (34th nucleotide) of the anticodon. The 34th nucleotide of tRNAs that are well cleaved at position 34/35 are always purine nucleotides (A or G), whereas 34th nucleotide of tRNAs that are well cleaved at positions outside of the anticodon region are almost always pyrimidine nucleotide (T or C). These results suggest that microbes living in the Yunohama hot spring may express specific ribonucleases that can digest specific tRNAs by distinguishing the tRNA types and/or anticodon sequences.

In eukaryotes, tRNA cleavage is known to be a conserved response to some stresses [66], and several studies have reported position-specific tRNA degradation in some types of human cells [101, 102] and in yeast [103, 104]. Although the knowledge of bacterial tRNA degradation is previously still limited, more than 8,000 putative bacterial tRNAs and more than 2,000 fragments of them in the environment were analyzed in this study. The results suggest that this tRNA fragmentation was not the

result of nonspecific degradation. Because certain sRNAs are reported to occur in specific environments [73], many of the fragmented tRNAs isolated in the current study may have been induced by environmental factors, like heat or metal ions. Although the present study may be one of the earliest steps towards defining the unique characteristics of sRNAs in environments, further analyses are required to reveal the possible role(s) of these sRNAs in microbes. To this end, it should be very useful to collect as many sRNA samples as possible from various microbial habitats.

Chapter 3

The evaluation of the effects of hot spring water consumption on glycemic control based on metabolome and microbiome approaches

3.1 Introduction

Self-medication is an important approach to maintain and promote human health. There are many strategies of self-medication such as improving one's lifestyle and/or dietary habits, consumption of functional foods/beverages, and getting adequate exercise [105, 106]. HSW is traditionally used in public baths and balneotherapy in many countries. Balneotherapy is the use of thermal and/or mineral water for treatment of human health by employing various methods such as bathing, drinking, mud therapy, and inhalation [34]. According to previous studies, it has been reported that balneotherapy has beneficial effects for various diseases such as T2D [36-38], rheumatism [39], low back pain [34], and cardiovascular disease [42]. Especially, consumption of hydrogen carbonate- or sulfur-containing water has been reported to prevent or improve T2D [36-38]. However, as the reported benefits of the HSW were determined epidemiologically and clinically, the molecular mechanisms of the beneficial effects behind HSW remain unclear.

Metabolomics is a method of comprehensive measurement of metabolites, and it has been known to a powerful tool to gain new insights in various research fields such as metabolic syndrome [107], cancer [108, 109], chronic kidney disease [110], and

biomarker screening [111]. A recent study has reported that plasma metabolome profiles were altered between healthy people and T2D patients [112]. Moreover, recent studies reported that gut microbiota is involved in various types of diseases such as diabetes [113], obesity [114], inflammatory bowel disease [115], and autistic spectrum disorder [116]. There are several studies, which have been indicated that gut microbial composition and function in T2D patients are different from that of healthy subjects [113, 114, 117]. Therefore, preventive and/or therapeutic effects for T2D derived from HSW consumption are expected via influencing blood metabolites concentrations and gut microbial compositions.

For this reason, we conducted clinical trial and animal experiment to elucidate the molecular basis of HSW consumption. In the clinical trial, blood metabolome analysis and gut microbiome analysis were conducted to elucidate the molecular basis of hydrogen carbonate-containing HSW (HSW1) consumption on human glycemic control. In the animal experiment, physiological effects on glycemic control derived from 8 types of HSW (HSW1-8) were evaluated and compared (HSW1 was used both clinical trial and animal experiment). Since previous studies have reported that consumption of hydrogen carbonate- or sulfur-containing water has potential to prevent or improve T2D [36-38], several hydrogen carbonate- and/or sulfur-containing HSW were used in the animal experiment. Additionally, several other types of HSW such as carbon dioxide- or chloride-containing HSW were also used in the animal experiment to compare the effects derived from consumption of the HSW.

In the clinical trial, we show that the serum glycoalbumin level, a glycemic control index, was slightly but significantly decreased, and 19 metabolites including glycolysis-related metabolites and 3 amino acids were significantly changed after consumption of HSW1. Additionally, microbiome analysis showed that relative compositions of 8 families were significantly changed after HSW1 consumption. Especially, relative composition of family Christensenellaceae that has been known as lean-associated bacteria was most significantly increased (in 75% of subjects) in the HSW1 consumption periods. Moreover, animal experiment suggest that consumption of 4 out of 8 types of HSW has the possible potential to improve glycemic control in high-fat diet-induced obese mice. This current study is an important research reporting the molecular basis of HSW consumption in the field of balneotherapy and CAM.

3.2 Materials and Methods

3.2.1 Tap water and HSW used in this study

In the current study, 2 types of tap water (TW and TWAF) and 8 types of HSW (HSW1-8) were used (Table 3.1) in total. Tap water used in the clinical trial was purchased from Nishikawa water treatment plant (Yamagata, Japan), and this commercially available tap water is designated as “TW” in this dissertation. Tap water used in the animal experiment was obtained from the water tap in the animal facility of National Institute of Technology, Tsuruoka College, and this water is designated as “TWAF” in this dissertation. 8 types of HSW (HSW1-8) were obtained from 3 different hot spring areas but their origins of hot springs are different. In the clinical trial, HSW1 was used because it has been orally reported by a local physician that the HSW have beneficial effects for glycemic control. In the animal experiment, all of 8 types of HSW including HSW1 were used. HSW1 and 2 were obtained from Nagayu hot spring area (Taketa-city, Oita, Japan), HSW3 and 4 were obtained from Hijiori hot spring area (Okura-village, Yamagata, Japan), and HSW5-8 were obtained from Shiobara hot spring area (Nasushiobara-city, Tochigi, Japan). Categories of each HSW were evaluated according to the results of mineral component analysis that were performed by administrators of each hot spring and determined by the Standard Methods of Analysis for Mineral Springs (published by the Ministry of the Environment, Japan).

Table 3.1 Tap water and HSW used in this study

Name of water	Use for[*1]	Origin of water[*2]	Category of hot springs	Temp (˚C)[*3]
TW	Human	Nishikawa water treatment plant.	Not hot spring (tap water)	-
TWAF	Mice	Water tap in the animal facility of National Institute of Technology, Tsuruoka College.	Not hot spring (tap water)	-
HSW1	Human and Mice	Nagayu	Magnesium • calcium • sodium - hydrogen carbonate springs	49.2
HSW2	Mice	Nagayu	Carbon dioxide-containing - magnesium • calcium • sodium - hydrogen carbonate springs	32.3
HSW3	Mice	Hijiori	Simple carbon dioxide springs	9.4
HSW4	Mice	Hijiori	Sodium - hydrogen carbonate • chloride springs	37.4
HSW5	Mice	Shiobara	Sodium - hydrogen carbonate • chloride springs	73.2
HSW6	Mice	Shiobara	Sodium - chloride springs	55.8
HSW7	Mice	Shiobara	Calcium • sodium - hydrogen carbonate • chloride springs	65.1
HSW8	Mice	Shiobara	Sulfur-containing - sodium - hydrogen carbonate • chloride springs	50.6

*1 This column indicates water used for clinical trial (Human) and/or mice experiment (Mice).

*2 “Nagayu” indicates Nagayu hot spring area (Taketa-city, Oita, Japan), “Hijiori” indicates Hijiori hot spring area (Okura-village, Yamagata, Japan), and “Shiobara” indicates Shiobara hot spring area (Nasushiobara-city, Tochigi, Japan).

*3 This column indicates temperature of the HSW when the water was obtained from the origin of spring.

3.2.2 Clinical trial

This clinical trial was approved by the ethical committees of the Japan Health and Research Institute, and Keio University Shonan Fujisawa Campus. All subjects were informed of the purpose of this study, and written consent was obtained from all subjects.

In this study, HSW (HSW1) consumption test was conducted amongst 19 healthy subjects (7 men and 12 women, ages from 26 to 59, 47 years old on the average). Initially, individual identification numbers were randomly assigned to 26 volunteers from N01 to N26, however N12 and N25 dropped out due to personal reasons. Additionally, N01, N02, N14, N18, and N21 were excluded from analysis because they forgot to drink TW and/or HSW1 at least once. HSW1 were collected in 500 ml plastic bottles and stored in refrigerator until consumption. HSW1 were consumed within 10 days from bottling to ensure freshness. TW bottles were also similarly stored in refrigerator until consumption. During the test, volunteers opened 1 plastic bottle of TW or HSW1 every day and drank the 500 ml of TW or HSW1 divided thrice daily (30-60 min before breakfast, lunch, and dinner). TW consumption period and HSW1 consumption period lasted for a week each, and this cycle was repeated twice. Volunteers were instructed to keep to their normal dietary habits during the test, but consumption of medicinal drugs was prohibited. Blood and fecal samples were collected on the first day of the test and last days of every week. Blood samplings that were collected from the same subject were performed at approximately same time

during the test. Body weight, body mass index (BMI), height, abdominal circumference, and blood pressure were measured on the first day of the test. Averages and standard deviations of each parameter during 19 volunteers were calculated and shown (Table 3.2). A schematic representation of the experimental design is as shown in Figure 3.1A. It is noted that fecal sample of week 0 could not collected from subject N22, therefore this subject was excluded from fecal microbiome and metabolome analysis.

3.2.3 Animal experiment

The animal experiment was performed using protocols approved by Animal Studies Committees of National Institute of Technology, Tsuruoka College.

5-weeks-old male C57BL/6JJcl mice were purchased from CLEA Japan Inc. and housed in the animal facility of National Institute of Technology, Tsuruoka College under a 12-hour light-dark cycle. Initially, mice were randomly grouped and fed a control diet (CE-2, CLEA Japan Inc.) and TWAF for 1 week to acclimatize to a new environment. After acclimatization, diet was replaced with high-fat diet (HFD32, CLEA Japan Inc.). In addition, drinking water was also replaced with each corresponding HSW in HSW1-8 consumption groups. All mice had *ad libitum* access to food and water. During weeks 15-17, oral glucose tolerance test (OGTT), intraperitoneal glucose tolerance test (IPGTT), and intraperitoneal insulin tolerance test (IPITT) were performed (Fig. 3.1B). Sample size was determined based on published studies [118] using similar assays as well as the previous experience.

Table 3.2 Anthropometric characteristics of volunteers

	Average	SD
Age	46.9	11.1
Weight (kg)	63.0	11.8
Height (cm)	165.0	10.1
BMI (kg/m^2)	23.0	3.2
Abdominal circumference (cm)	83.8	7.5
Blood pressure (top) (mmHg)	129.7	14.9
Blood pressure (bottom) (mmHg)	76.1	8.6

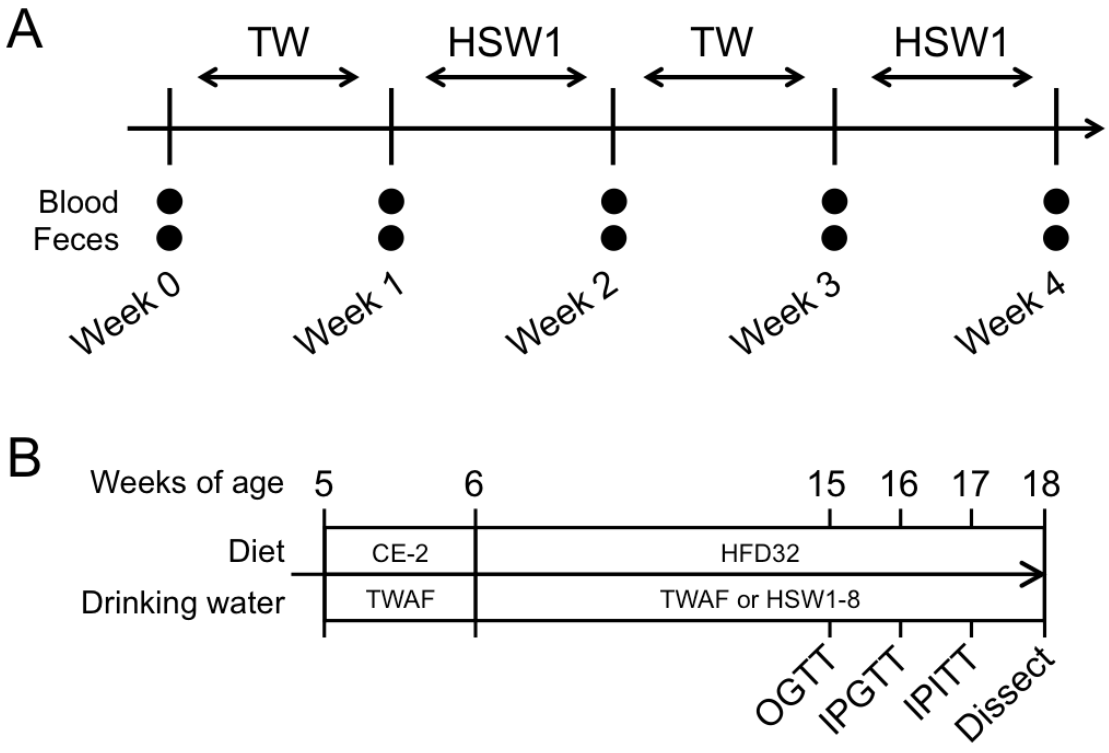


Figure 3.1 Schematic representation of the experimental design

(A) For clinical trial, HSW1 consumption test was performed over 4 weeks. TW consumption period and HSW1 consumption period lasted for a week each, and this cycle was repeated twice. Blood and fecal samples were collected on the first day of the test and last days of every week. (B) For mice experiment, C57BL/6JJcl male mice were bred from 5-weeks-old to 18-weeks-old. In the first week, mice were fed standard diet (CE-2) and TWAF. From the second week, mice were fed high-fat diet (HFD32) and TWAF or HSW until end of the experiment. OGTT, IPGTT, and IPITT were performed 15, 16, and 17-weeks-old, respectively. At 18-weeks-old, mice were dissected to collect cecal contents.

3.2.4 Clinical blood tests

For clinical trial, clinical blood tests were performed at each sampling point, including measurement of fasting plasma glucose, serum glucose, glycoalbumin, insulin, total cholesterol, high-density lipoprotein (HDL) cholesterol, low-density lipoprotein (LDL) cholesterol, triglyceride, urate, sodium, chloride, calcium, magnesium, and cortisol. The measurement of the concentrations of these parameters was outsourced to RINTEC Co., Ltd.

Plasma glucose measurement was performed using morning fasting blood samples obtained from only 8 volunteers. For these samples, the homeostasis model assessment ratio (HOMA-R) was calculated from plasma glucose and insulin levels.

3.2.5 Metabolome analysis

Metabolome analysis of blood and fecal samples obtained from clinical trial were conducted as described previously with some modifications [110]. In brief, to extract metabolites from blood, 400 μl of methanol including the internal standards (20 μM each of methionine sulfone and D-camphor-10-sulfonic acid (CSA)) was added to the 40 μl of blood samples. Next, this mixture was then mixed with 120 μl of ultrapure water and 400 μl of chloroform before centrifuging at 10,000×*g* for 3 min at 4°C. Subsequently, the aqueous layer was transferred to a centrifugal filter tube (UltrafreeMC-PLHCC 250/pk for Metabolome Analysis, Human Metabolome Technologies) to remove protein and lipid molecules. The filtrate was centrifugally

concentrated and dissolved in 20 μl of ultrapure water that contained reference compounds (200 μM each of 3-aminopyrrolidine and trimesic acid) immediately before capillary electrophoresis with electrospray ionization time-of-flight mass spectrometry (CE-TOFMS) analysis.

To extract metabolites from feces, samples were initially lyophilized by using VD-800R lyophilizer (TAITEC) for at least 18 hours. Freeze-dried feces were disrupted with 3.0 mm zirconia beads by vigorous shaking (1,500 rpm for 10 min) using Shake Master neo (Biomedical Science). 500 μl of methanol including the internal standards (20 μM each of methionine sulfone and CSA) was added to the 10 mg of disrupted feces. Samples were further disrupted with 0.1 mm zirconia/silica beads by vigorous shaking (1,500 rpm for 5 min) using Shake Master neo. Next, 200 μl of ultrapure water and 500 μl of chloroform were added before centrifuging at 4,600×*g* for 15 min at 20°C. Subsequently, the aqueous layer was transferred to a centrifugal filter tube (UltrafreeMC-PLHCC 250/pk for Metabolome Analysis, Human Metabolome Technologies) to remove protein and lipid molecules. The filtrate was centrifugally concentrated and dissolved in 100 μl of ultrapure water that contained reference compounds (200 μM each of 3-aminopyrrolidine and trimesic acid) immediately before CE-TOFMS analysis.

Additionally, metabolome analysis of 8 types of HSW used in this study was also performed. For measurements of the metabolites in the HSW, 50 μl of ultrapure water including the internal standards (4 mM each of methionine sulfone, CSA,

3-aminopyrrolidine, and trimesic acid) was added to the 1 ml of HSW. Samples were transferred to centrifugal filter tubes (UltrafreeMC-PLHCC 250/pk for Metabolome Analysis, Human Metabolome Technologies) to remove protein and lipid molecules. The filtrate was centrifugally concentrated and dissolved in 20 μl of ultrapure water immediately before CE-TOFMS analysis.

The measurement of extracted metabolites in both positive and negative modes was performed by CE-TOFMS. All CE-TOFMS experiments were performed using the Agilent capillary electrophoresis system (Agilent Technologies). Annotation tables were produced from measurement of standard compounds and were aligned with the datasets according to similar m/z value and normalized migration time. Then, peak areas were normalized against those of the internal standards methionine sulfone or CSA for cationic or anionic metabolites, respectively. Concentrations of each metabolite were calculated based on their relative peak areas and concentrations of standard compounds.

After statistical analysis, metabolite set enrichment analysis (MSEA) [119] was performed using the blood metabolites that were significantly different between TW and HSW1 consumption periods in the clinical trial.

3.2.6 Microbiome analysis

3.2.6.1 DNA isolation

DNA isolation using fecal samples collected from clinical trial and cecal samples

collected from animal experiment were performed as described previously with some modifications [120]. Briefly, fecal and cecal samples were initially lyophilized by using VD-800R lyophilizer (TAITEC) for at least 18 hours. Freeze-dried feces were disrupted with 3.0 mm zirconia beads by vigorous shaking (1,500 rpm for 10 min) using Shake Master neo (Biomedical Science). Fecal samples (10 mg) were suspended with DNA extraction buffer containing 200 µl of 10% (w/v) SDS/TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) solution, 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1), and 200 µl of 3 M sodium acetate. Feces in mixture buffer were further disrupted with 0.1 mm zirconia/silica beads by vigorous shaking (1,500 rpm for 5 min) using Shake Master neo. Freeze-dried cecal contents were manually disrupted. Cecal contents (10 mg) were suspended with DNA extraction buffer containing 200 µl of 1% (w/v) SDS/TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) solution, 400 µl of phenol/chloroform/isoamyl alcohol (25:24:1), and 200 µl of 3 M sodium acetate. Cecal contents in mixture buffer were disrupted with 3.0 mm zirconia beads and 0.1 mm zirconia/silica beads by vigorous shaking (1,500 rpm for 15 min) using Shake Master neo. After centrifugation at 17,800×*g* for 5 min at room temperature, bacterial genomic DNA was purified by the standard phenol/chloroform/isoamyl alcohol protocol. RNAs were removed from the sample by RNase A treatment, and then DNA samples were purified once more by the standard phenol/chloroform/isoamyl alcohol protocol.

3.2.6.2 16S rRNA gene sequencing

16S rRNA genes in the fecal DNA samples obtained from clinical trial and cecal DNA samples obtained from animal experiment were analyzed using the MiSeq sequencer (Illumina). The V1-V2 region of the 16S rRNA genes were amplified from the DNA isolated from feces or cecal contents using bacterial universal primer set 27Fmod (5'-AGRGTTTGATYMTGGCTCAG-3') and 338R (5'-TGCTGCCTCCCGTAGGAGT-3') [121]. PCR was performed with Tks Gflex DNA Polymerase (Takara Bio Inc.) and amplification proceeded with one denaturation step at 98°C for 1 min, followed by 20 cycles of 98°C for 10 sec, 55°C for 15 sec, and 68°C for 30 sec, with a final extension step at 68°C for 3 min. The amplified products were purified using Agencourt AMPure XP (Beckman Coulter), and then further amplified using forward primer (5'-AATGATACGGCGACCACCGAGATCTACAC-NNNNNNNN-TATGGTAATTGT-AGRGTTTGATYMTGGCTCAG-3') containing the P5 sequence, a unique 8-bp barcode sequence for each sample (indicated in N), Rd1 SP sequence, and 27Fmod primer, and reverse primer (5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNNNN-AGTCAGTCAGCC-TGCTGCCTCCCGAGGAGT-3') containing the P7 sequence, a unique 8-bp barcode sequence for each sample (indicated in N), Rd2 SP sequence, and 338R primer. After purification using Agencourt AMPure XP, mixed sample was prepared by pooling approximately equal amounts of PCR amplicons from each sample. Finally, MiSeq

sequencing was performed according to the manufacturer's instructions. In this study, 2×300 bp paired-end sequencing was employed.

3.2.6.3 Analysis of 16S rRNA gene sequences

Initially, to assemble the paired-end reads, fast length adjustment of short reads (FLASH) (v1.2.11) [122] was used. Assembled reads with an average Q-value less than 25 were filtered out using in-house script. 5,000 filter-passed reads were randomly selected from each sample and used further analysis. Reads were then processed using quantitative insights into microbial ecology (QIIME) (v1.8.0) pipeline [123]. Sequences were clustered into operational taxonomic units (OTUs) using 97% sequence similarity, and OTUs were assigned to taxonomy using RDP classifier. For OTUs corresponding to Clostridiales, *Clostridium* clusters were assigned according to the previous study [120, 124].

3.2.7 Glucose tolerance test and insulin tolerance test

In the animal experiment, glucose or insulin tolerance tests were performed as described previously with some modifications [118]. For glucose tolerance test, 1.5 mg of glucose per gram of body weight was orally or intraperitoneally administrated to mice after 15-hour fasting in the OGTT or IPGTT, respectively. In the IPITT, 0.75 mU of human insulin (Humulin N (Eli Lilly Japan)) per gram of body weight were intraperitoneally administrated to mice after 5-hour fasting. The glucose concentrations were measured

before administration and at 15, 30, 60, and 120 min after administration using Stat Strip XP3 (NIPRO CORPORATION), and plasma samples were collected from tail vein at the same time of the measurements of glucose concentrations.

3.2.8 ELISA assay of insulin and glucagon-like peptide-1

In the OGTT conducted in the animal experiment, plasma insulin and glucagon-like peptide-1 (GLP-1) levels were measured by ELISA assay. Plasma samples were collected from the mice in heparin-coated glass capillaries and mixed with 10 μl of DPP-IV inhibitor (DPP4-010, Merck Millipore). After centrifuging at 1,000×*g* for 15 min at 4°C, plasma samples were obtained. Plasma insulin levels were analyzed using insulin ELISA kit (MS303, Morinaga Institute of Biological Science) and plasma active GLP-1 levels were analyzed using GLP-1 (active) ELISA kit (AKMGP-011, Shibayagi Co., Ltd.).

3.2.9 Statistical analysis

For clinical trial, to analyze the intra-individual alterations, statistical evaluation between two groups was performed by Wilcoxon signed-rank test (non-parametric paired test) using the R package exactRankTests (available at https://cran.r-project.org/web/packages/exactRankTests/index.html) or XLSTAT (v2014.6.04) (Addinsoft). For multiple comparisons, the data were analyzed using Friedman's test and post hoc Nemenyi test using XLSTAT. For animal experiment,

statistical evaluation between TWAF consumption group and each HSW consumption group were performed by Dunnett test using the R package multcomp (available at https://cran.r-project.org/web/packages/multcomp/index.html). Orthogonal partial least squares discriminate analysis (OPLS-DA) on the mineral components data was conducted with the SIMCA-P+ software v12.0 (Umetrics Inc.). Kendall rank correlation coefficients were calculated by using the R software. All statements indicating significant differences show at least a 5% level of probability.

3.2.10 Nucleotide sequence accession number

The microbiome analysis data obtained from clinical trial have been deposited at the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/) under accession number DRA004008.

3.3 Results

3.3.1 Mineral components of the water used in this study

For clinical trial and murine experiment to evaluate the molecular basis of HSW consumption, 8 different types of HSW were collected and used. Mineral contents and pH of the tap water and HSW used in this study were measured (Table 3.3). In the almost of all HSW, concentrations of hydrogen carbonate, chloride, sulfate, sodium, magnesium, calcium, and potassium ions were higher as compared with tap water. For undissociated molecules, metasilicic acid, metaboric acid, and carbon dioxide were also observed in high concentrations in HSW. There are no notable differences between tap water used in clinical trial (TW) and animal experiment (TWAF). Only HSW8 contains sulfur (hydrogen sulfide ion and hydrogen sulfide). According to the results of metabolome analysis of HSW, no metabolites were detected.

3.3.2 Comparisons of clinical parameters between TW and HSW1 consumption periods in the clinical trial

In the clinical trial, intra-individual alterations of clinical parameters were firstly analyzed. For each clinical parameter, mean values from week 1 and 3 (TW); and week 2 and 4 (HSW1) were calculated and compared (Table 3.4). Among 15 clinical parameters, glycoalbumin level, a glycemic control index, was slightly but significantly decreased and calcium concentration was slightly but significantly increased in the HSW1 consumption periods as compared with the TW consumption periods. However,

there were no significant correlations (spearman correlation coefficient was 0.02, P = 0.83) between glycoalbumin levels and calcium concentrations (Fig. 3.2A). In this study, other biochemical parameters involved with glycemic control, hypercholesterolemia, hyperuricemia, mineral consumption, and stress were also measured, but consumption of HSW1 did not have any impact on the above-mentioned parameters apart from glycoalbumin and calcium levels (Table 3.4). Mean relative glycoalbumin levels of week 2 and 4 (HSW1) were significantly decreased as compared with that of week 1 and 3 (TW) (Fig. 3.2B). However, it was also noted that only 0.93% of relative glycoalbumin levels were decreased in average. Although glycoalbumin levels have been known to reflect blood glucose levels during short-term (at least the last 14 days) [125], blood glucose levels were not significantly decreased (Fig. 3.2C), but tended to be lowered by HSW1 consumption (P = 0.092). According to the weekly changes of relative glycoalbumin levels, the concentrations were slightly decreased but not significantly after the first TW consumption period (week 1) as compared with before the test (week 0) (Fig. 3.2D-E). Glycoalbumin levels were continuously decreased after the first HSW1 consumption period (week 2), and this reduction was statistically significant as compared with before the test. Subsequently, it slightly increased after the second TW consumption period (week 3) and then slightly decreased again after the second HSW1 consumption period (week 4). Values at week 3 and 4 were also significantly low as compared with before the test. Alterations of observed serum glycoalbumin levels were slight, and it was decreased 0.2% of total albumin in average

in the HSW1 consumption periods (Fig. 3.2F, Table 3.4). There are 3 subjects whose initial glycoalbumin levels exceeded the reference interval (Fig. 3.2F). The glycoalbumin level of subject with the highest value was not enough decreased in the HSW1 consumption periods. However, the levels of other 2 subjects were decreased below the upper reference value in weeks 2 and 4. Whereas, it was also noted that serum glycoalbumin level of 1 subject was decreased below the lower reference value after consumption of TW and HSW1. According to the observed blood glucose levels, blood glucose concentrations tended to decrease after the first HSW1 consumption period (week 2) as compared with the first TW consumption period (week 1), but there are no significant differences (Fig. 3.2G).

Taken together, glycoalbumin levels were at least statistically significantly decreased after HSW1 consumption therefore HSW1 consumption may have the possible benefit in glycemic control. However, the amounts of changes of glycoalbumin levels were slight and improvement of glycemic control could not completely supported by the alterations of blood glucose levels. Since the current study targeted healthy subjects, alterations of glycoalbumin and glucose levels might be slight. Therefore, it is important to perform additional studies among diabetes patients to further evaluate the beneficial effects of HSW1 consumption for glycemic control and/or T2D.

Table 3.3 Mineral contents and pH of the water used in this study

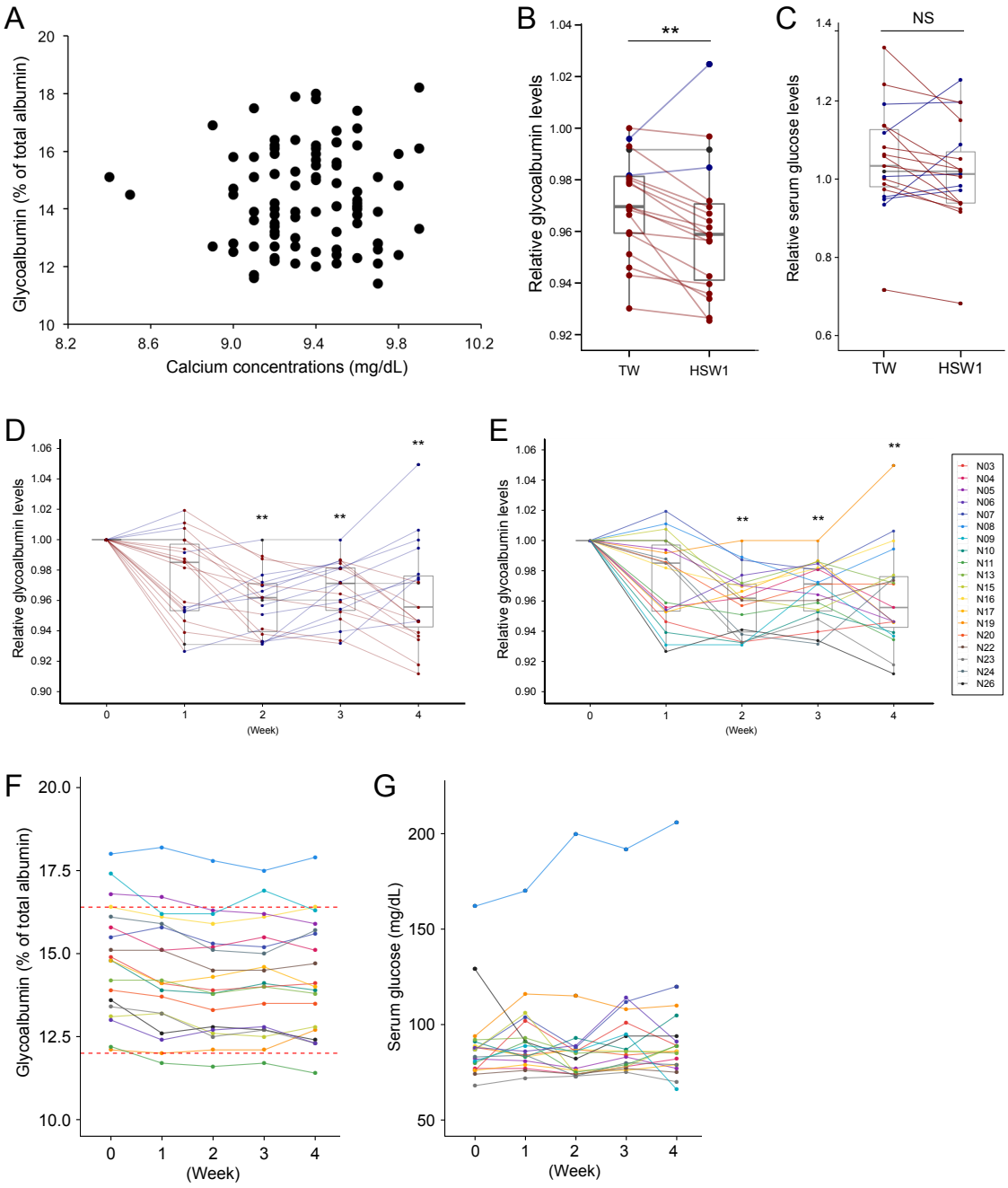
		Concentrations (mg/kg)									
Minerals	Formula	TW	TWAF	HSW1	HSW2	HSW3	HSW4	HSW5	HSW6	HSW7	HSW8
Hydrogen carbonate ion	HCO_3^-	28.0	27.5	2485	1232	20.7	3340	449	386	458	1084
Chloride ion	Cl^-	11.0	11.4	182	78.3	6.3	679	478	2239	343	906
Sulfate ion	SO_4^{2-}	6.9	6.9	355	185	11.6	0.1	28.7	54.3	46.1	65.3
Carbonate ion	CO_3^{2-}	-	-	1.2	-	-	0.9	-	-	-	0.3
Nitrate ion	NO_3^-	0.7	0.7	1.2	-	1.6	-	-	-	-	-
Fluoride ion	F^-	-	-	0.3	0.2	-	0.7	0.2	0.4	0.5	0.5
Hydrogen sulfide ion	HS^-	-	-	-	-	-	-	-	-	-	22.1
Iodide ion	I^-	-	-	-	0.3	-	0.3	0.6	2.8	0.8	0.8
Bromide ion	Br^-	-	-	-	-	-	1.7	0.7	1.0	0.5	2.7
Hydrogenphosphate ion	HPO_4^{2-}	-	-	-	-	-	0.6	-	-	-	-
Dihydrogenphosphate ion	$H_2PO_4^-$	-	-	-	-	-	-	-	-	-	3.2
Thiosulfate ion	$S_2O_3^{2-}$	-	-	-	-	-	-	-	-	-	-
Hydrogensulfate	HSO_4^-	-	-	-	-	-	-	-	-	-	-
Sodium ion	Na^+	10.0	10.3	412	214	7.4	1657	327	1204	286	860
Magnesium ion	Mg^{2+}	1.9	1.9	291	136	3.0	4.8	27.4	37.6	17.0	19.7
Calcium ion	Ca^{2+}	6.1	6.1	177	102	2.8	28	88.1	264	78.7	92.7
Potassium ion	K^+	-	-	80.0	35.3	1.3	8.9	44.9	107	29.1	61.4
Aluminum ion	Al^{3+}	0.2	0.2	0.6	0.2	1.0	0.4	0.6	1.0	0.6	-
Manganese ion	Mn^{2+}	-	-	0.4	0.6	0.1	0.3	0.2	1.5	0.3	0.2
Ferrous ion	Fe^{2+}	-	-	2.3	3.1	-	0.6	0.9	3.7	0.4	-
Ammonium ion	NH_4^+	-	-	1.9	1.0	0.1	2.9	0.1	0.4	0.6	0.5
Ferric ion	Fe^{3+}	-	-	-	-	-	-	-	-	-	-
Metasilicic acid	H_2SiO_3	10.0	10.1	207	164	32.2	24.6	235	92.7	181	177
Metaboric acid	HBO_2	0.8	0.8	6.2	3.5	1.2	35.5	34.5	126	28.0	106
Carbon dioxide	CO_2	0.9	0.9	161	411	241	284	46.6	102	31.2	178
Hydrogen sulfide	H_2S	-	-	-	-	-	-	-	-	-	65.7
Metaarsenious acid	$HAsO_2$	-	-	-	-	-	-	0.1	-	0.1	-
Sulfuric acid	H_2SO_4	-	-	-	-	-	-	-	-	-	-
Mercury	Hg	-	-	-	-	-	-	-	-	-	-
Copper	Cu	-	-	-	-	-	-	-	-	-	-
Zinc	Zn	-	-	-	-	-	-	-	-	-	-
Lead	Pb	-	-	-	-	-	-	-	-	-	-
Cadmium	Cd	-	-	-	-	-	-	-	-	-	-
pH		7.58	7.58	7.07	6.28	4.72	6.82	6.63	6.18	6.78	6.58

"-" indicates the value was under detection limit.

Table 3.4 Results of clinical blood test

Tests	TW		HSW1		*P* value	
	Average	SD	Average	SD		
Fasting blood glucose (plasma) (mg/dL)	91.2	4.3	88.6	4.7	0.156	
Blood glucose (serum) (mg/dL)	93.5	23.7	91.4	29.0	0.092	
Glycoalbumin (% of total albumin)	14.4	1.7	14.2	1.7	0.002	**
Insulin (μU/ml)	19.3	28.0	18.8	31.0	0.294	
HOMA-R	1.2	0.5	1.2	0.4	1.000	
Total cholesterol (mg/dL)	200.6	26.0	202.4	30.3	0.914	
HDL cholesterol (mg/dL)	60.9	14.9	60.1	14.1	0.685	
LDL cholesterol (mg/dL)	112.0	22.5	112.3	24.1	1.000	
Triglycerides (mg/dL)	151.2	142.1	180.7	159.0	0.123	
Urate (mg/dL)	5.2	1.2	5.2	1.2	0.396	
Na (mEq/L)	139.8	1.3	140.3	1.2	0.172	
Cl (mEq/L)	102.6	1.7	102.9	1.9	0.457	
Ca (mg/dL)	9.3	0.2	9.4	0.2	0.021	*
Mg (mg/dL)	2.3	0.1	2.3	0.1	0.518	
Cortisol (μg/dL)	10.1	4.0	10.0	3.6	0.623	

*$P < 0.05$; **$P < 0.005$.



(Legend on the next page)

Figure 3.2 Comparisons of clinical parameters during the test

(A) Scatter plot between calcium concentrations (x) and glycoalbumin levels (y). (B-C) Glycoalbumin levels (B) and serum glucose levels (C) were expressed as a relative value to week 0. Individual data of mean relative glycoalbumin levels from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in dot plots overlaid on box plots. Plots corresponding to same individuals were connected with red, blue, or gray lines when the values were decreased, increased, or not changed in the HSW1 consumption periods as compared with the TW consumption periods, respectively. Plots were also colored in the same color as their lines. (D-E) Individual weekly relative glycoalbumin levels were shown in dot plots overlaid on box plots. Significances between week 0 and each week were shown at top of the graph. Plots corresponding to same individuals were connected with red, blue, or gray lines when the values were decreased, increased, or not changed as compared with one week before. Plots were also colored in the same color as their lines (D). Same data were also color-coded by individual (E). (F-G) Individual observed glycoalbumin levels (F) and serum glucose levels (G) were shown. Reference interval of serum glycoalbumin level [126] was indicated as red dotted line (F). $^{**}P < 0.005$; NS, not significant.

3.3.3 HSW1 consumption-related changes of physiological metabolism

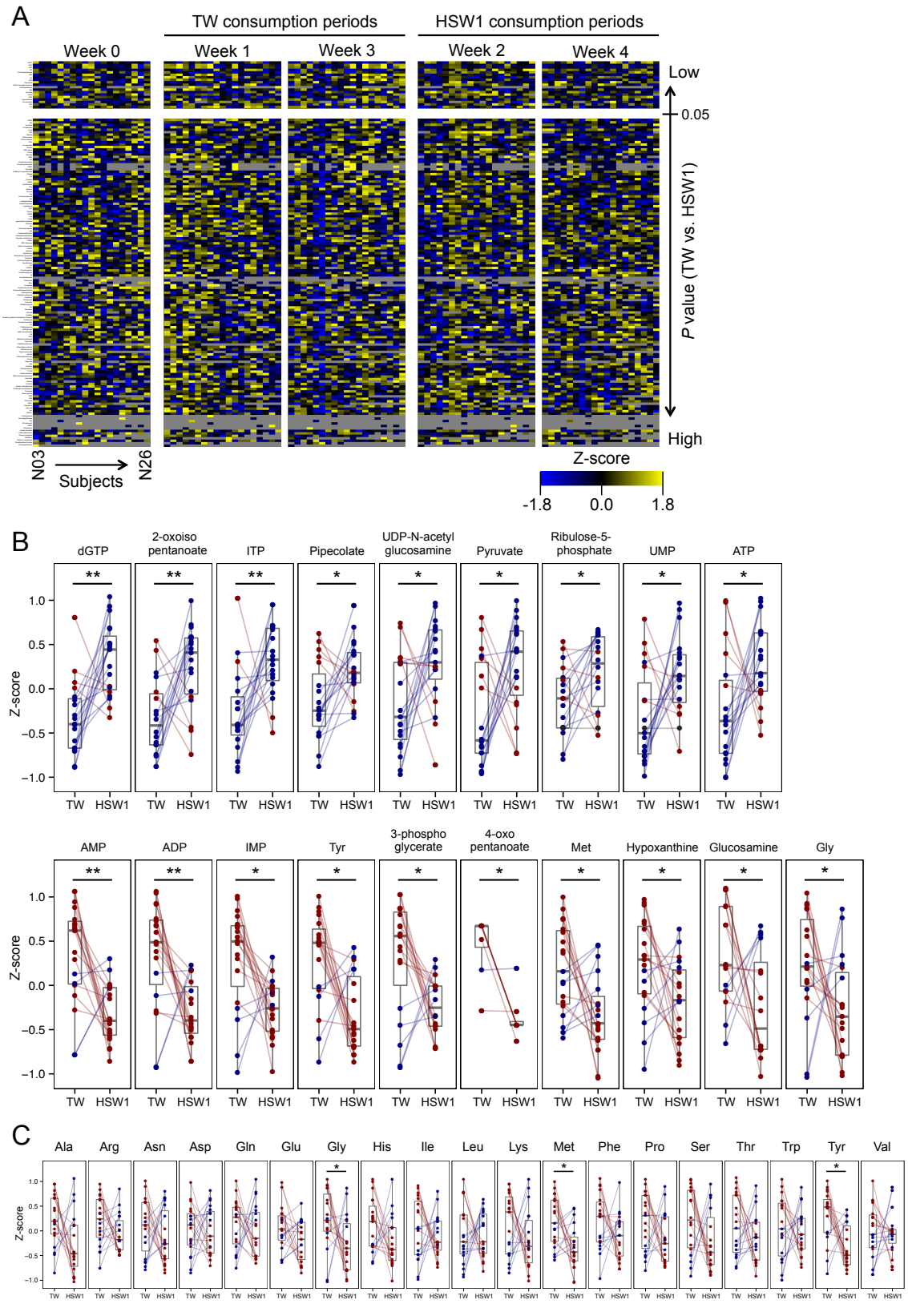
To evaluate the effect of HSW1 consumption on glycoalbumin reduction, blood metabolome analysis was performed using CE-TOFMS. A total of 152 metabolites were detected from blood samples at least from 1 subject and 1 time point, and concentrations of these metabolites were compared within subject. Over 85% of metabolites were not significantly changed in the HSW1 consumption periods (Fig. 3.3A), and it was expected that the concentrations of most metabolites remained consistent due to physiological homeostasis. However, the concentrations of 19 metabolites were significantly different between TW and HSW1 consumption periods (Fig. 3.3A, B). As compared with the TW consumption periods, 9 metabolites were significantly increased and 10 metabolites were significantly decreased in the HSW1 consumption periods (Fig. 3.3B).

Metabolites that may be related to glycemic control such as glycolysis-related metabolites (3-phosphoglyceric acid, pyruvate, ATP, and ADP), amino acids (tyrosine, methionine, and glycine), and UDP-N-acetylglucosamine were included in significantly changed metabolites. In addition, 3 amino acids were significantly decreased but almost of all amino acids were also lowered in the HSW1 consumption periods (Fig. 3.3C). Since the 4 glycolysis-related metabolites were significantly changed in the HSW1 consumption periods, relative concentrations of the metabolites corresponding to glycolysis and citric acid cycle were also represented (Fig. 3.4). In the glycolysis, 3-phosphoglycerate and ADP were significantly decreased, and fructose 1,6-bisphosphatase and phosphoenolpyruvic acid were tended to decrease. Whereas,

pyruvate and ATP were significantly increased, and lactate was tended to increase in the HSW1 consumption periods. In the citric acid cycle, citrate, *cis*-aconitate, *iso*-citrate, and malate were tended to increase, but there are no significant differences. These results suggest that HSW1 consumption might be induced glycolysis enhancement, but not citric acid cycle.

Additionally, MSEA was performed using the 19 metabolites that were significantly changed between TW and HSW1 consumption periods (Table 3.5). According to the result, purine metabolism, nitrogen metabolism (ammonia recycling, urea cycle, and methionine metabolism), and central metabolic pathways (glycolysis, gluconeogenesis, and citric acid cycle) were expected to be mainly altered after HSW1 consumption. This result supported hypothetical enhancement of glycolysis, therefore glycoalbumin reduction may be result of glycolysis enhancement. However, relationships among purine metabolism, nitrogen metabolism, and glycoalbumin reduction were still not clear.

On the other hand, it was observed that 3-hydroxybutyrate, one of the indexes of diabetic ketoacidosis, was not altered by HSW1 consumption (Fig. 3.5).



(Legend on the next page)

Figure 3.3 Comparisons of blood metabolites between TW and HSW1 consumption periods

(A) Relative concentrations of each metabolite in blood were transformed to Z-score by subjects and shown as heat maps using blue-black-yellow scheme. Gray color indicates that the metabolites were not detected in the sample. Total 152 metabolites were arranged in increasing order of *P* value that calculated by Wilcoxon signed-rank test between TW and HSW1 consumption periods. (B) Mean relative concentrations of each metabolite (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in dot plots overlaid on box plots in the same manner as Figure 3.2B. Only 19 metabolites that their concentrations were significantly increased (upper 9 metabolites) or decreased (lower 10 metabolites) in the HSW1 consumption periods (shown in upper block of panel A) were represented. (C) Mean relative concentrations of each amino acid (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in the same manner as Figure 3.2B. Relative concentrations of standard amino acids apart from cysteine were shown because cysteine cannot be detected by CE-TOFMS measurement. $*P < 0.05$; $**P < 0.005$.

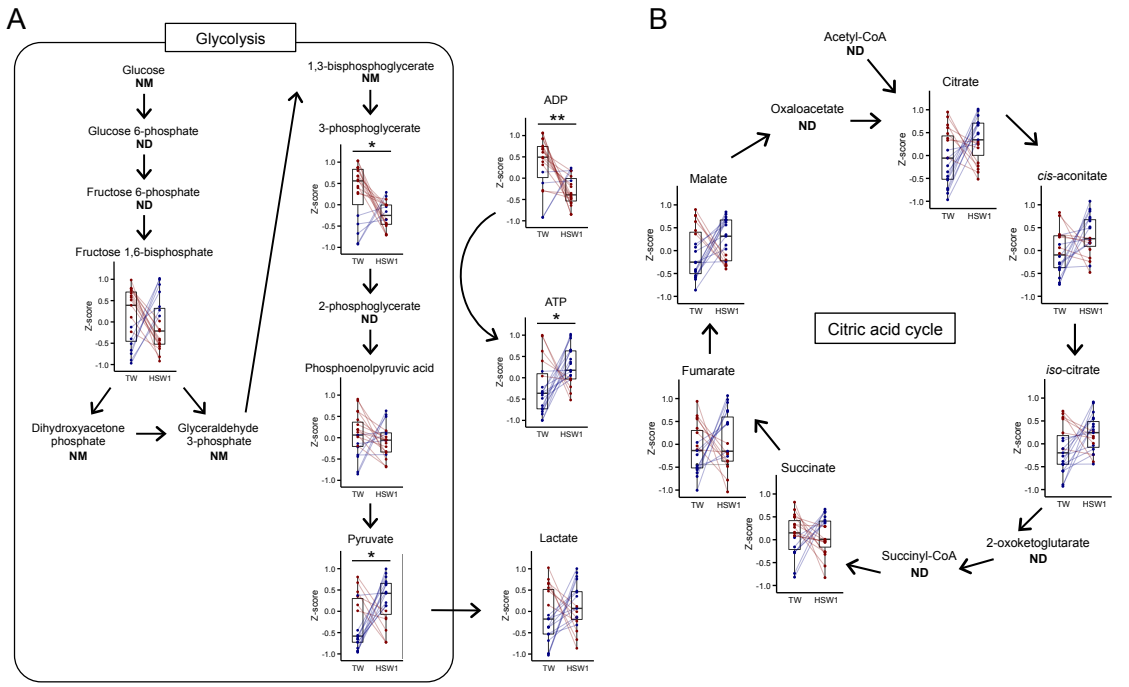


Figure 3.4 Comparisons of metabolites that related to glycolysis and citric acid cycle between TW and HSW1 consumption periods

Mean relative concentrations of each metabolite (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in dot plots overlaid on box plots in the same manner as Figure 3.2B. Metabolites that are related to glycolysis (A) and citric acid cycle (B) were represented. ND, not detected; NM, not measured; **P* < 0.05; ***P* < 0.005.

Table 3.5 Results of MSEA: List of metabolite sets/pathways that are significantly different between TW and HSW1 consumption periods

Pathway	Total[*1]	Hits[*2]	Expect[*3]	Fold change (hits/expect)	*P* value	FDR[*4]
Purine metabolism	45	7	0.98	7.1	< 0.001	0.001
Ammonia recycling	18	4	0.39	10.2	< 0.001	0.014
Urea cycle	20	4	0.44	9.2	< 0.001	0.014
RNA transcription	9	3	0.20	15.3	< 0.001	0.014
Intracellular signaling through prostacyclin receptor and prostacyclin	6	2	0.13	15.3	0.006	0.103
Glycolysis	21	3	0.46	6.6	0.009	0.121
Citric acid cycle	23	3	0.50	6.0	0.012	0.133
Methionine metabolism	24	3	0.52	5.7	0.013	0.133
Gluconeogenesis	27	3	0.59	5.1	0.018	0.163
Amino sugar metabolism	15	2	0.33	6.1	0.040	0.290
Mitochondrial electron transport chain	15	2	0.33	6.1	0.040	0.290

[*1] Total numbers of metabolites that corresponded in each pathway.

[*2] Observed numbers of metabolites that derived from given dataset in each pathway.

[*3] Expected observed numbers of metabolites that calculated by given dataset in each pathway.

[*4] False discovery rate (FDR) according to the Benjamini and Hochberg method that was provided by MSEA software.

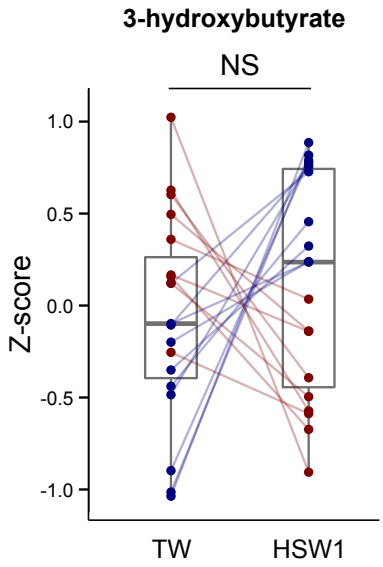


Figure 3.5 Relative concentrations of 3-hydroxybutyrate in blood

Mean relative concentrations of 3-hydroxybutyrate in blood (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in dot plots overlaid on box plots in the same manner as Figure 3.2B. NS, not significant.

3.3.4 HSW1 consumption-related changes of intestinal environment

To investigate the alteration of gut microbiota composition during the test, microbiome analysis was conducted using fecal samples that were collected weekly during the study. A total of 7,075 OTUs were constructed from 16S rRNA gene sequences derived from 19 subjects. These OTUs corresponded to 62 families and whole structures of fecal microbiota are shown (Fig. 3.6A). To investigate the intra-individual changes of gut microbiota during the test, relative abundances of each microbial taxon were compared between TW and HSW1 consumption periods within subjects. From the results, it was observed that over 85% of families were not altered similarly to blood metabolites, but relative abundances of 8 families especially Christensenellaceae were significantly different between TW and HSW1 consumption periods (Fig. 3.6B). As species level analysis, relative abundances of each OTU were compared between TW and HSW1 consumption periods, 23 OTUs were significantly altered (Fig. 3.7). One OTU corresponding to Bifidobacteriales was significantly decreased (Fig. 3.7A), but 9 OTUs corresponding to Bacteroidales were significantly increased (Fig. 3.7B) in the HSW1 consumption periods as compared with the TW consumption periods. In the family Clostridiales, 7 OTUs were significantly decreased whereas 5 OTUs were significantly increased (Fig. 3.7C) in the HSW1 consumption periods as compared with the TW consumption periods. These results suggest that consumption of HSW1 has a potential to change some gut microbial compositions.

Recently, gut microbiota-derived metabolites such as short-chain fatty acids (SCFAs), and bile acids have received a lot of attention. Therefore relative

concentrations of SCFAs, medium-chain fatty acids (MCFAs), and bile acids were compared between TW and HSW1 consumption periods. As a result, butyrate, pentanoate, decanoate, dodecanoate, and deoxycholate were significantly increased in the HSW1 consumption periods (Fig. 3.8).

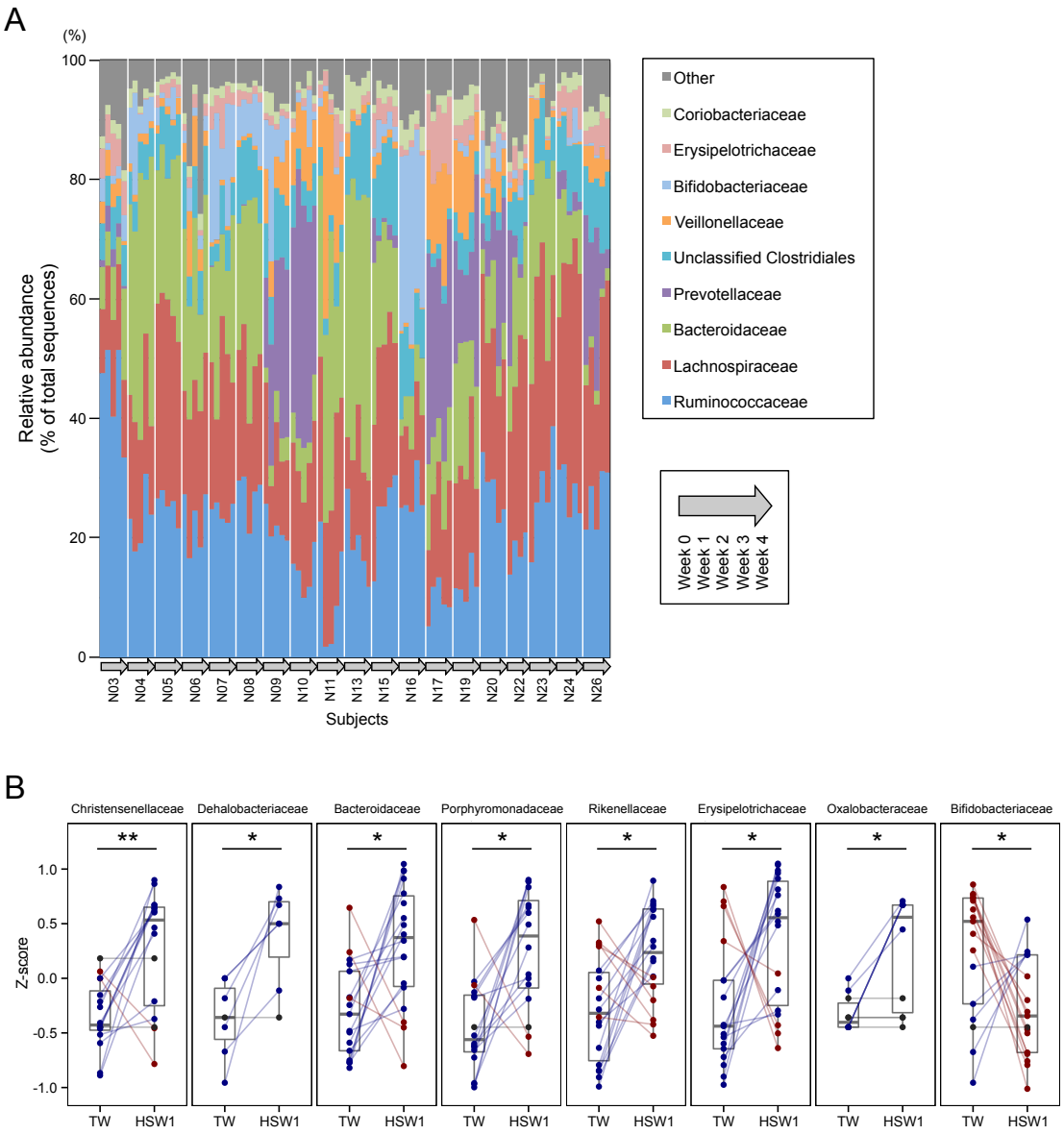


Figure 3.6 Comparisons of fecal microbiota compositions between TW and HSW1 consumption periods

(A) Family level compositions of fecal microbiota during the test. (B) Mean relative abundances of each taxa (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown in dot plots overlaid on box plots in the same manner as Figure 3.2B. Only 8 families had compositions that were significantly different between TW and HSW1 consumption periods were demonstrated. *$P < 0.05$; **$P < 0.005$.

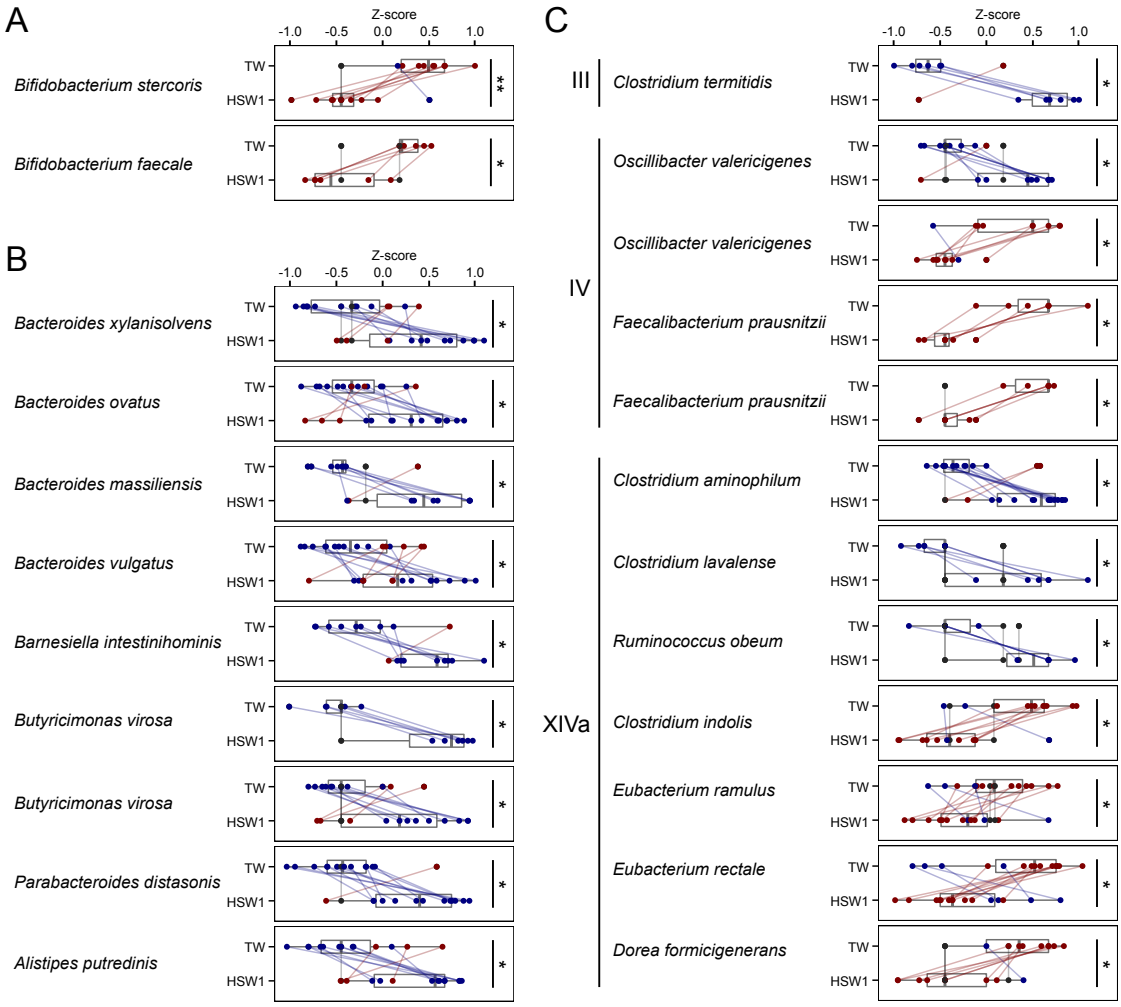


Figure 3.7 Comparisons of relative abundances of each OTU between TW and HSW1 consumption periods

Mean relative abundances of each OTU (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown. Significantly different OTUs between TW and HSW1 consumption periods that were assigned as Bifidobacteriales (A), Bacteroidales (B), and Clostridiales (C) were represented in dot plots overlaid on box plots in the same manner as Figure 3.2B. The closest taxon of each OTU that were determined by BLASTn analysis against the database "16S ribosomal RNA sequences (Bacteria and Archaea)" registered in NCBI was shown at the left of the plot. For only OTUs corresponded to Clostridiales (C), *Clostridium* clusters were also represented. *$P < 0.05$; **$P < 0.005$.

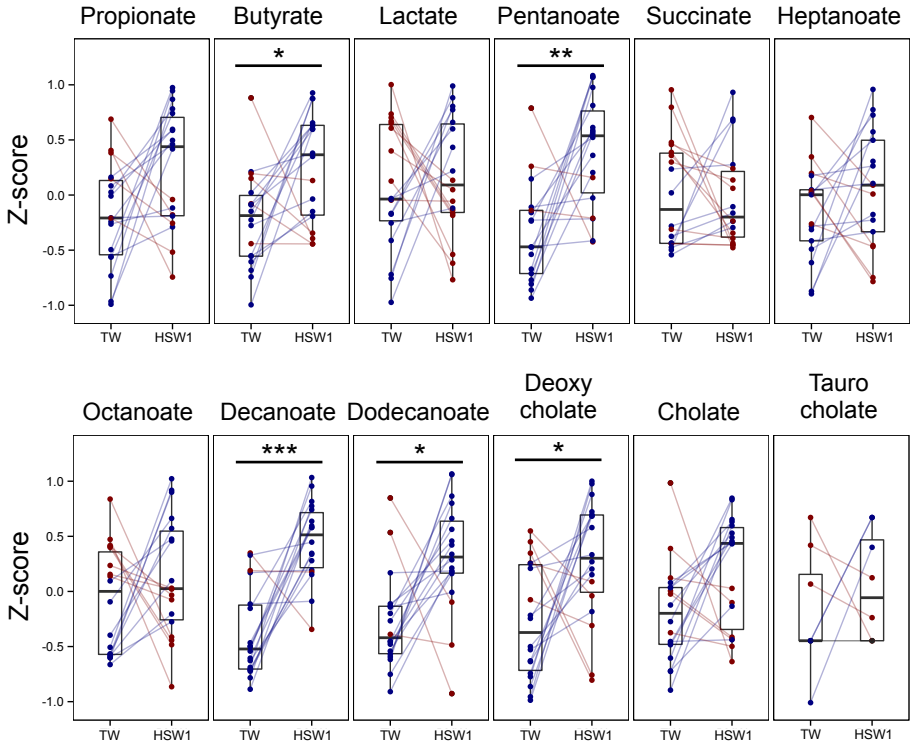


Figure 3.8 Comparisons of short- and medium-chain fatty acids and bile acids detected from fecal samples

Mean relative concentrations of short- and medium-chain fatty acids and bile acids (Z-score) from week 1 and 3 (TW); and week 2 and 4 (HSW1) were shown. Data were represented in dot plots overlaid on box plots in the same manner as Figure 3.2B. **P* < 0.05; ***P* < 0.005; ****P* < 0.0005.

3.3.5 Evaluation of the physiological effects of HSW consumption in animal experiment

According to the clinical trial, HSW1 consumption might have beneficial effects in glycemic control therefore it was consequently investigated in high-fat diet-induced obese mice. In addition, the physiological effects derived from consumption of various types of HSW were also investigated. Firstly, body weights of all groups were not different during the test (Fig. 3.9A). Although the body weights were not changed, food intakes of HSW4, 7, and 8 consumption groups, and water intakes of HSW1, 2, 4, 7, and 8 groups were significantly increased as compared with TWAF consumption group (Fig. 3.9B-E). Next, OGTT was performed to assess the glycemic status (Fig. 3.10). Fasting plasma glucose levels (measured before glucose administration in OGTT) of HSW3 and 8 consumption groups were significantly decreased. Additionally, plasma glucose level of HSW1, 3, 7, and 8 consumption groups at 15 min after glucose administration were significantly decreased as compared with TWAF consumption group. Therefore these 4 types of HSW were defined as “effective HSW”. As a result of OGTT, plasma glucose levels were significantly decreased at only 15 min after glucose administration in the effective HSW consumption groups. We suspected that this phenomenon might have occurred due to increment of insulin level via increased incretin secretion such as GLP-1. However, plasma GLP-1 (active form) and insulin levels were not improved in the effective HSW consumption groups (Fig. 3.11). Instead, insulin levels at 15 and 60 minutes after glucose administration in HSW8 consumption group were significantly decreased. To evaluate the insulin resistance, IPGTT and

IPITT were also performed but significant improvements were not observed between TWAF and each effective HSW consumption group. In fact, plasma glucose levels were significantly increased in HSW3 and 7 consumption groups in IPGTT and IPITT, respectively (Fig. 3.12).

Next, OPLS-DA was performed using concentrations of minerals in each HSW to seek the components that were contributed to improve glucose tolerance (Fig. 3.13). The resulting S-plot suggests that sodium, chloride, and hydrogen carbonate ions were contributed to discriminate non-effective HSW, but no components that were informative to discriminate effective HSW were found, although the cross-validated predictive ability and the total explained variance were quite low.

Finally, alterations of gut microbiota compositions derived from HSW consumption were investigated. For this purpose, microbiome analysis was conducted using the cecal contents that were collected from the mice of each HSW consumption group. Family level compositions of cecal microbiota were not significantly changed aside from HSW5 and 7 consumption groups (Fig. 3.14). Relative abundance of Lactobacillaceae was significantly decreased in HSW5 consumption group, and these of unclassified Bacteroidales and S24-7 were significantly increased in HSW7 consumption group as compared with TWAF consumption group (Fig. 3.14B-D). In addition, some OTUs were significantly changed between TWAF and each HSW consumption group (Table 3.6 and 3.7). Subsequently, the correlation between mineral concentrations in the drinking water and relative abundances of each OTU were analyzed. As a result, there are 71 OTUs that have more than 0.7 or less than –0.7 of

Kendall rank correlation coefficients (Fig. 3.15). Almost of all the OTUs were corresponded to Bacteroidales or Clostridiales, but the correlation patterns were not discriminated at order level taxonomy.

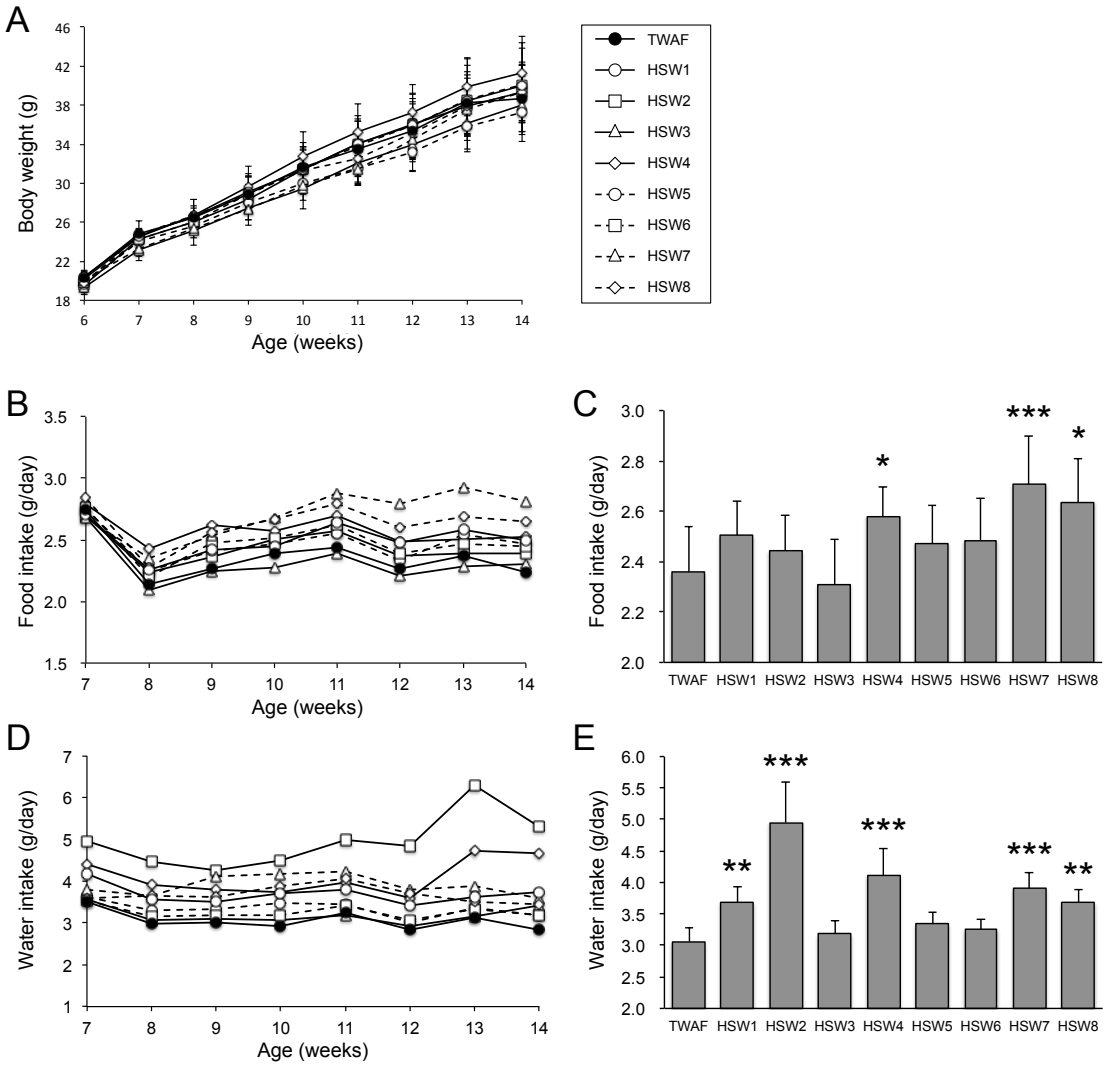


Figure 3.9 Body weights, food intakes, and water intakes during the experiment

Body weights (A), food intakes (B-C), and water intakes (D-E) of TWAF and HSW1-8 consumption groups are represented. Body weights are shown as mean ± SD (n = 4-5) (A). Food and water intakes were calculated by each cage (each group), and data are represented as gram per day per mouse. Weekly data (B, D) and mean ± SD during the experiment (C, E) are shown. *$P < 0.05$; **$P < 0.005$; ***$P < 0.0005$.

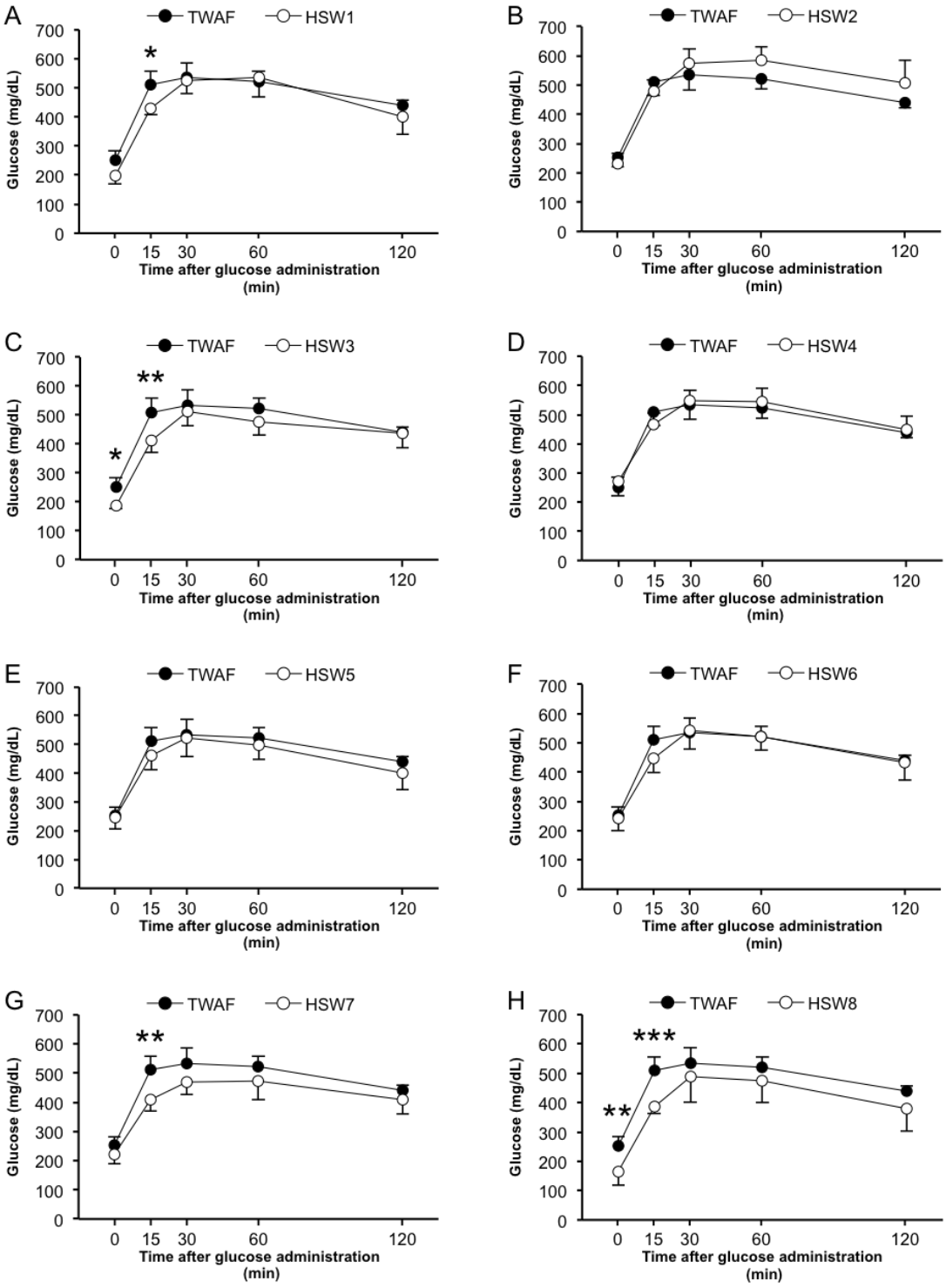


Figure 3.10 OGTT of TWAF and HSW1-8 consumption mice

(A-H) Blood glucose levels of TWAF and HSW1-8 consumption mice during OGTT were represented. Data were represented as mean ± SD (n = 4-5). *P < 0.05; **P < 0.005; ***P < 0.0005.

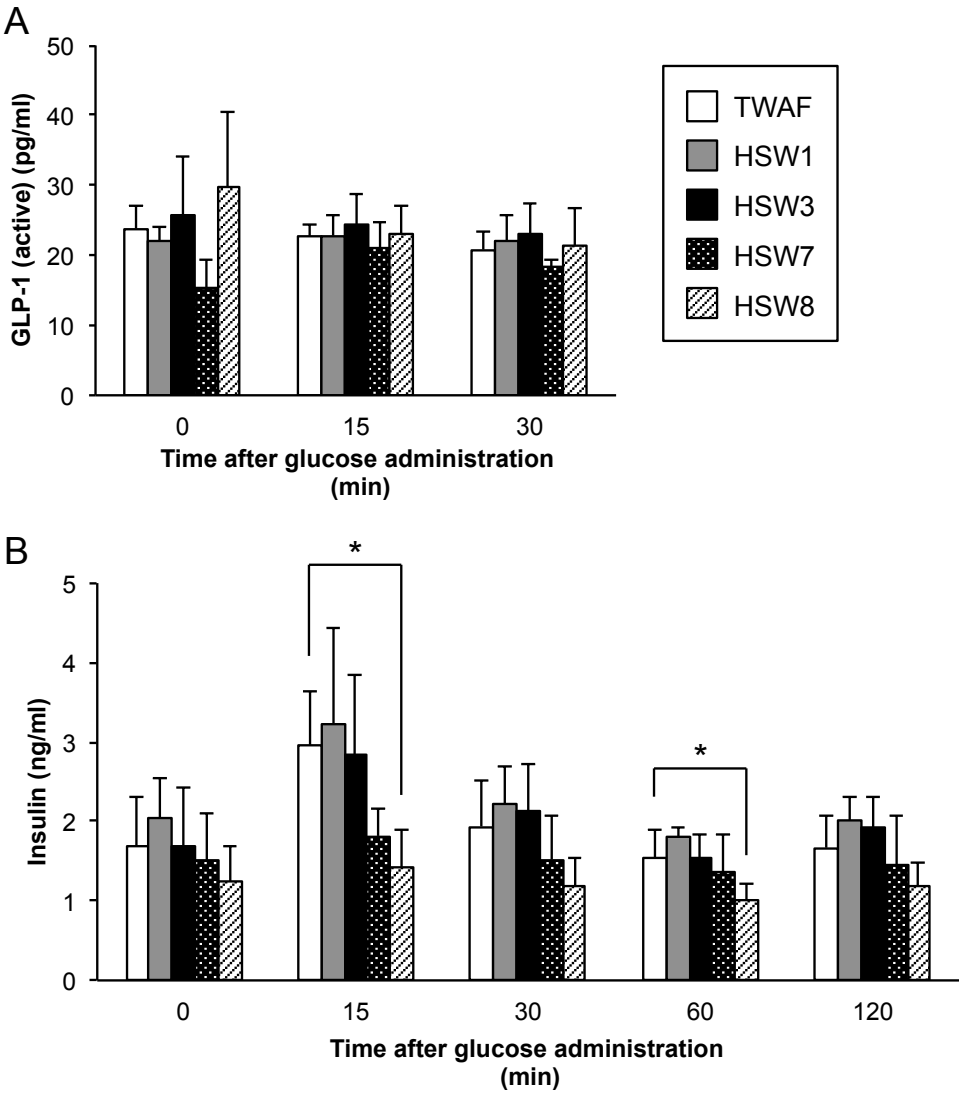


Figure 3.11 GLP-1 and insulin levels during OGTT

Plasma GLP-1 (active) (A) and insulin (B) levels of mice during OGTT were represented as mean ± SD (n = 4-5) of TWAF and effective HSW consumption mice. *$P < 0.05$.

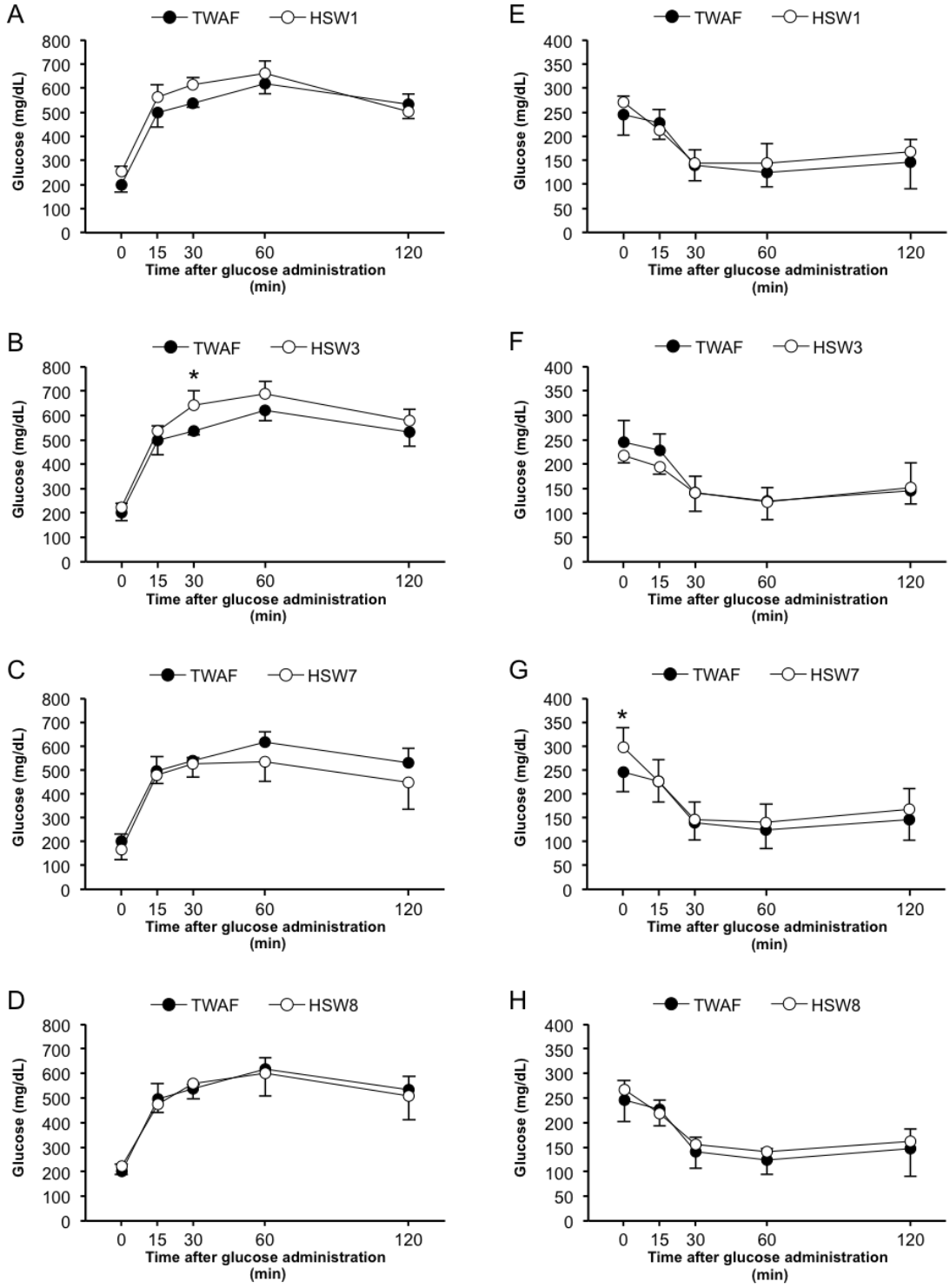


Figure 3.12 IPGTT and IPITT of TWAF and effective HSW consumption mice

Blood glucose levels of TWAF and effective HSW consumption mice during IPGTT (A-D) and IPITT (E-H) were represented. Data were represented as mean ± SD (n = 4-5). *$P < 0.05$.

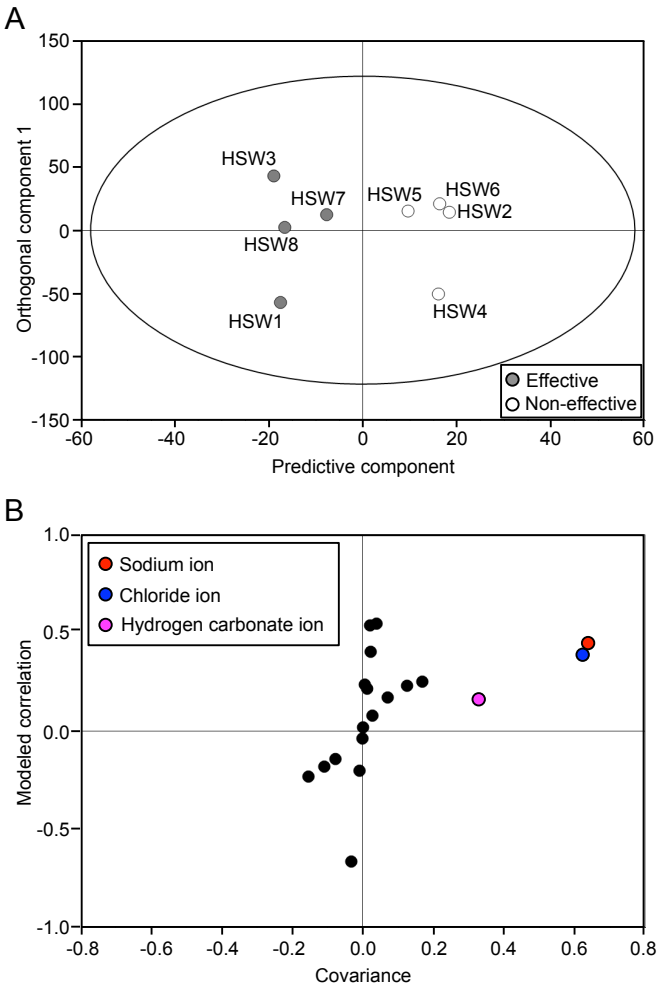
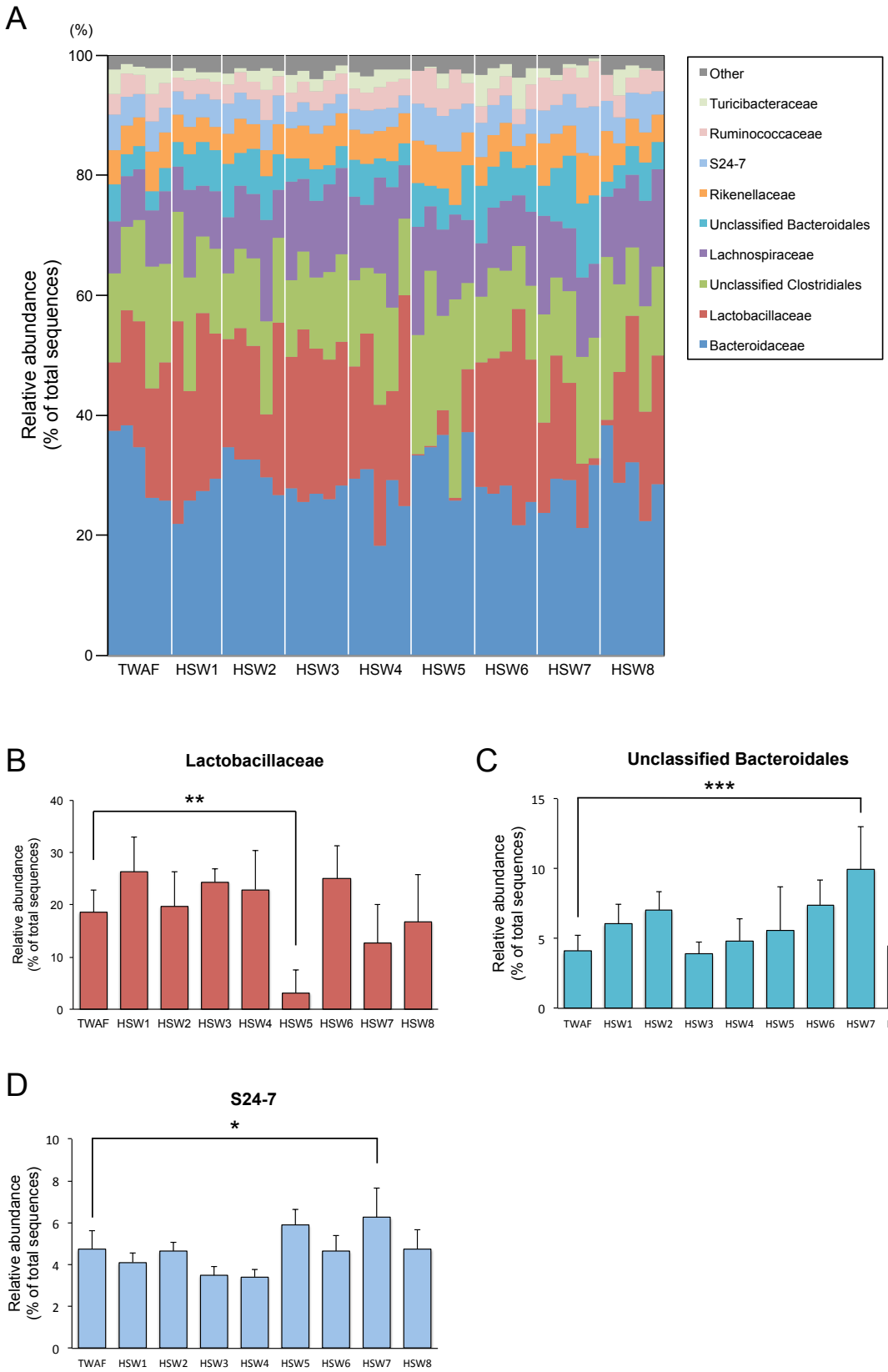


Figure 3.13 OPLS-DA on the mineral components of each HSW

(A) Cross-validated score plots from OPLS-DA of mineral components data of each HSW. The model resulted in one predictive and one orthogonal (1+5) components with the cross-validated predictive ability Q2(Y) = –1.81 and the total explained variance R2(X) = 0.09. The ellipse denotes the 95% significance limit of the model, as defined by Hotelling's t-test. (B) S-plot for predictive component from OPLS-DA of mineral components data of each HSW. The S-plot visualizes the variable influence in a model and corresponds to combining the contribution or magnitude (covariance) with the effect and reliability (correlation) for the model variables with respect to model component scores. The highlighted signals (|covariance| > 0.2) in the S-plot have high contribution and reliability for class separation between effective and non-effective HSW.



(Legend on the next page)

Figure 3.14 Comparisons of cecal microbiota compositions

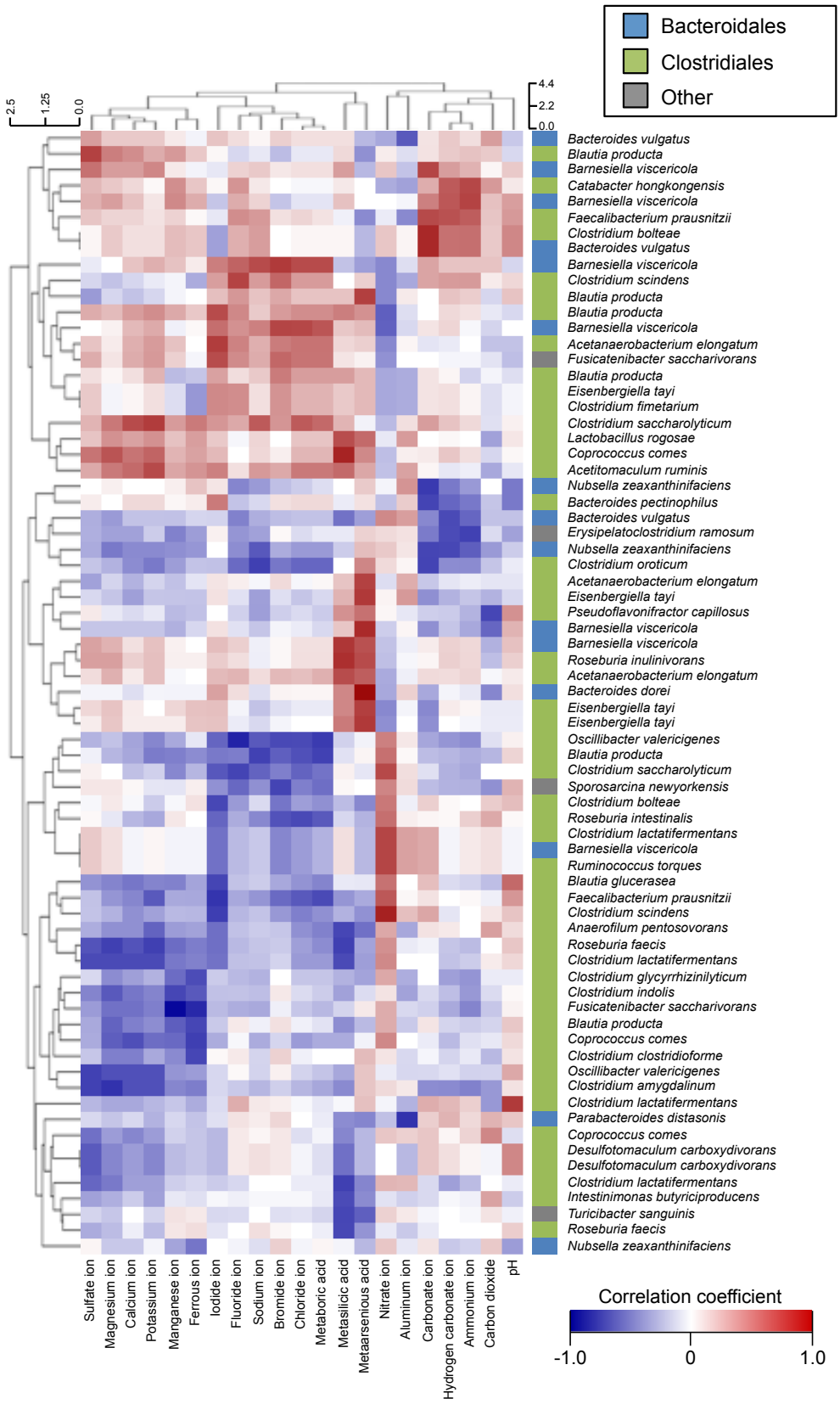
(A) Whole family level compositions of mice cecal microbiota in each mouse were shown. (B-D) Relative abundances of families that the compositions were significantly different between TWAF consumption group and each HSW consumption groups were individually shown; Lactobacillaceae (B), unclassified Bacteroidales (C), and S24-7 (D) (n = 4-5). *$P < 0.05$; **$P < 0.005$; ***$P < 0.0005$.

Table 3.6 Numbers of OTUs that significantly increased as compared with TWAF consumption group

Corresponding taxa (family level)	HSW1	HSW2	HSW3	HSW4	HSW5	HSW6	HSW7	HSW8
Unclassified Clostridiales	4	2	3	7	11	3	5	3
Lachnospiraceae	1	0	4	2	8	1	1	1
Ruminococcaceae	0	0	0	1	9	0	5	0
S24-7	3	1	0	0	0	0	2	0
Enterobacteriaceae	2	1	2	0	0	0	0	0
Lactobacillaceae	0	0	0	1	0	0	0	0
Bacteroidaceae	0	0	0	0	0	0	1	0
Streptococcaceae	0	0	0	0	1	0	0	0
Unclassified Bacteroidales	0	0	0	0	0	0	1	0
Unclassified	0	0	0	0	1	0	0	0

Table 3.7 Numbers of OTUs that significantly decreased as compared with TWAF consumption group

Corresponding taxa (family level)	HSW1	HSW2	HSW3	HSW4	HSW5	HSW6	HSW7	HSW8
Unclassified Clostridiales	1	6	2	1	6	4	5	5
Lachnospiraceae	4	3	2	3	4	5	2	3
Ruminococcaceae	1	0	1	1	1	0	1	1
Lactobacillaceae	0	0	0	0	2	0	1	0
Bacteroidaceae	0	1	0	1	0	1	0	0



(Legend on the next page)

Figure 3.15 Correlations between mineral concentrations and abundances of OTUs
Kendall rank correlation coefficients between mineral concentrations of each mineral water, and mean relative abundances of each OTU were shown as heat map using blue-white-red scheme. In the heat map, OTUs that have more than 0.7 or less than –0.7 of Kendall rank correlation coefficients were shown. The closest taxon of each OTU that were determined by BLASTn analysis against the database "16S ribosomal RNA sequences (Bacteria and Archaea)" registered in NCBI and shown at the right of the heat map. Order level taxonomy of each OTU that were assigned by RDP classifier was also shown as blue (Bacteroidales), green (Clostridiales), and gray (other) at the right of the heat map. Minerals and OTUs were clustered by hierarchical clustering (Euclidean distance) and tree was represented at the top and left of the heat map.

3.4 Discussion

In the present study, physiological effects especially on glycemic control derived from HSW consumption were investigated. Firstly, it was shown that the relative serum glycoalbumin level was slightly but significantly decreased after HSW1 consumption as compared with TW consumption in the clinical trial. The amounts of reductions of glycoalbumin levels in the HSW1 consumption periods were only 0.93% of initial value as compared with the TW consumption periods in average. In addition, coefficient of variation (CV) of reproducibility of glycoalbumin assay kit used in this study has been reported as 2.8-3.2% [127]. Therefore, reduction of glycoalbumin levels was statistically significant but it is not clear whether there are actually biological and/or medical impacts such as prevention and/or improvement of T2D. Besides, blood glucose levels were not decreased significantly, but tended to be lowered by HSW1 consumption, although glycoalbumin level has been known to reflect blood glucose levels during short-term (at least the last 14 days) [125]. This might be attributed that blood glucose levels are influenced by external factors such as food intake and exercise [128]. Since the current study targeted healthy subjects, alterations of glycoalbumin and glucose levels were expected to be minimal. However, results indicated that glycoalbumin levels of subjects whose initial values exceeded reference intervals by a little were decreased below the upper reference value in week 2 and 4. Therefore, HSW1 consumption may be useful for borderline diabetes patients to control their glycemic status, but sample numbers must be increased to further validate this hypothesis. Additionally, HSW1 consumption did not decrease enough glycoalbumin

level of subject with the highest initial value to a consideration extent, thus HSW1 consumption may not be effective for T2D patients. Moreover, glycoalbumin level of subject with low initial value was decreased below lower reference interval during the test. Therefore, HSW1 consumption may have a risk to induce hypoglycemia in subjects with low glycoalbumin levels from the beginning. Taken together, it is important to perform further studies among borderline diabetes patients to further evaluate the preventive and/or therapeutic effects for T2D and/or glycemic control derived from HSW1 consumption.

For other clinical parameters, insulin concentrations and HOMA-R were not significantly changed by HSW1 consumption. These results suggest that the reduction of glycoalbumin levels observed in this study was not expected to relate changing of insulin secretion and/or insulin resistance. Clinical tests also indicated significant increments in blood calcium levels. This phenomenon was expected to attribute to HSW1 consumption because HSW1 contains calcium ion 29 times greater than TW. According to the previous report, calcium deficiency may apparently lead to insulin resistance [129], therefore calcium supplementation by consumption of calcium-containing mineral water may be important to better manage glycemic control. However, it is also noted that glycoalbumin levels and blood calcium levels were not correlated, thus glycoalbumin reductions and increments of blood calcium levels were not expected to be directly involved.

Focusing on the first TW consumption period (week 1) in the clinical trial, relative glycoalbumin levels were partly decreased, but not significantly, compared with

before the test (week 0). Although it has been previously reported that body weights of the volunteers were decreased by consumption of TW before every meal during 12 weeks [130], the reduction of serum glycoalbumin levels is a novel result. This result suggests that the habit of water consumption before every meal may have the possible potential to decrease serum glycoalbumin levels even if the drinking water is TW.

According to the metabolome analysis of blood samples, ATP and pyruvate were significantly increased whereas 3-phosphoglyceric acid and ADP was significantly decreased. Additionally, other members of glycolysis, fructose 1,6-bisphosphate and phosphoenolpyruvic acid, were tended to decrease, whereas lactate was slightly increased. These results suggest that glycolysis may be upregulated in the HSW1 consumption periods. In the citric acid cycle, citrate, *cis*-aconitate, *iso*-citrate, and malate were tended to increase, but there were no significant differences. Therefore HSW1 consumption may induce enhancement of glycolysis, but not citric acid cycle. Since blood glucose levels were not decreased significantly, but tended to be lowered as observed in our results, it may be attributed to glycolysis enhancement.

On the other hand, concentrations of 3 amino acids (tyrosine, methionine, and glycine) were significantly lowered in the HSW1 consumption periods. A previous study reported that high concentrations of various amino acids especially tyrosine in blood is one of the risk factor of T2D [131]. Additionally, it has been also reported plasma concentrations of several amino acids including tyrosine and methionine were significantly high in hyperinsulinemia (it is often observed in early stage of T2D) patients as compared with healthy subjects [132]. Our results demonstrated that

concentrations of 3 amino acids including tyrosine and methionine in blood were significantly decreased and that of other standard amino acids were tended to be lowered in the HSW1 consumption periods. As such, it can be suggested that the HSW1 consumption may have the possible potential to avoid risk factors of T2D through the alterations of metabolism in the body.

After insulin resistance was increased, it can be expected that proteolysis and ketogenesis would be enhanced [133]. As a result, the increment of blood concentrations of amino acids or ketone bodies is expected. Since metabolome analysis indicated the concentration of 3-hydroxybutyrate, one of the ketone bodies, between TW and HSW1 consumption periods was not changed. Thus the consumption of HSW1 may not affect generation of energy from free fatty acids. As our results demonstrated that concentrations of almost of all amino acids especially tyrosine, methionine, and glycine were decreased, HSW1 consumption may influence energy metabolism through proteolysis but not ketogenesis.

UDP-N-acetylglucosamine is the substrate of O-linked N-acetylglucosamine transferase, and the relationship between O-linked N-acetylglucosamine transferase and insulin resistance has been previously reported [134]. However, it was also reported that concentration of UDP-N-acetylglucosamine in muscle tissue was increased after reaching euglycemia by insulin treatment in obese subjects [135]. These reports suggest that UDP-N-acetylglucosamine is related to glucose control and/or insulin resistance, but the details are still unclear. In the current study, metabolome analysis indicated blood concentrations of UDP-N-acetylglucosamine was significantly increased in the

HSW1 consumption periods, but future studies are required to understand the meaning of this phenomenon.

Since recent studies reported the relationships between gut microbiota and T2D and/or obesity [136-138], we hypothesized that beneficial effect for glycemic control derived from HSW1 consumption might involve gut microbiota. As expected, alterations of gut microbiota compositions derived from HSW1 consumption were observed. In this study, family Christensenellaceae was the most significantly increased taxon. Previous study reported that Christensenellaceae was enriched in lean group (BMI < 25) as compared with obese group (BMI > 30) [139]. Additionally, it was also reported that transplantation of *Christensenella minuta* to germ-free mice reduced weight gain. Moreover, abundance of the family Dehalobacteriaceae that was also increased after HSW1 consumption has been reported that positively correlated with Christensenellaceae. Additionally, the family Porphyromonadaceae was also significantly increased after HSW1 consumption, it has been reported that the taxon was negatively correlated with type 1 diabetes [140], and were decreased in high-fat/high-sucrose diet fed mice [141]. Whereas, family Bifidobacteriaceae that include typical probiotic bacteria such as *Bifidobacterium longum* [142] and *Bifidobacterium animalis* [143] was significantly decreased in the HSW1 consumption periods. Taken together, our results suggest that consumption of HSW1 may have the possible potential to prevent getting obese and/or diabetes via alteration of the gut microbiota compositions such as increasing of lean-associated bacteria, even though potentially beneficial bacteria group, Bifidobacteriaceae was decreased.

Recently, many studies reported the importance and positive or negative effects of gut microbiota-derived metabolites such as SCFAs and bile acids [118, 120, 144-146]. Therefore, relative changes of short- and medium-chain fatty acids and bile acids between TW and HSW1 consumption periods were analyzed in this study. 2 SCFAs (butyrate and pentanoate) were significantly increased after HSW1 consumption. Previous studies reported that SCFAs enhance incretin secretion [146] and insulin sensitivity in the muscle and liver [118]. Therefore, beneficial effects for glycemic control derived from HSW1 consumption may be mediated by increasing of intestinal SCFAs concentrations. Additionally, it has been reported that butyrate induces the differentiation of regulatory T cells in the colon and suppresses colonic inflammation [120], therefore HSW1 consumption may have the possible potential to prevent and/or improve not only T2D but also inflammatory bowel disease and allergic diseases. Whereas, deoxycholate is one of the secondary bile acids produced by gut microbiota especially *Clostridium* cluster XI and XIVa [145, 147], and it has been reported that deoxycholate is the key molecule to develop obesity-associated hepatocellular carcinoma. Additionally, it has been reported that butyrate-producing bacteria were also included in the *Clostridium* cluster XIVa [148]. In the present study, butyrate and deoxycholate was increased after HSW1 consumption and it can be attributed to increase relative abundances of 3 *Clostridium* cluster XIVa species (*Clostridium aminophilum*, *Clostridium lavalense*, and *Ruminococcus obeum*). In the HSW1 consumption periods, SCFAs and deoxycholate were increased, and it has been reported that SCFAs have mainly positive effects whereas deoxycholate has mainly negative

effects, therefore HSW1 consumption may have both positive and negative effects. Previous studies have reported that high-fiber diet consumption induced increments of intestinal SCFAs [120], whereas high-fat diet consumption induced high concentrations of deoxycholate in human feces [149]. Therefore, combination of HSW1 and high-fiber/low-fat diet consumption may contribute to gain only positive effects derived from HSW1 consumption. For bile acid analysis, only cholate, taurocholate, and deoxycholate were measured by CE-TOFMS in this study. However, it is required to analyze various types of bile acids by using Ultra-performance liquid chromatography mass-spectrometry (UPLC-MS) as reported previously [150] to further consider the relationships between HSW1 consumption and bile acids.

In the clinical trial, molecular basis of HSW1 consumption in human was investigated and speculated mechanisms of beneficial effects of HSW1 consumption were as summarized (Fig. 3.16). However, the blinded test could not be performed in this study because all types of HSW have specific tastes therefore volunteers would have been to recognize whether the water is tap water or HSW upon consumption. For this reason, placebo effect could not be excluded. However, the placebo effect might be minimal because it would be difficult humans to control blood glucose levels consciously. However, specific tastes of HSW1 such as bitterness and/or saltiness might repress the appetites of volunteers. As such, HSW1 consumption may have the potential to reduce food intake, although the volunteers were strictly instructed to keep to their normal dietary habits during the test. Therefore, it may be required that volunteers have to consume identical meals during the study.

In the animal experiment, it was investigated whether beneficial effects for glycemic control can be observed by consumption of HSW1 and other types of HSW. Results indicated that HSW consumption could not suppress weight gain, therefore there is no potential in preventing obesity in mice. Since the increments of lean-associated bacteria were estimated as one of the triggers to prevent getting obese in the clinical trial, but lean-associated bacteria detected from human volunteers were not present in the mice cecum according to the microbiome analysis of cecal contents. Therefore, weight gain could not be suppressed by HSW consumption even though the potential to prevent getting obese derived from HSW1 consumption was suggested by clinical trial.

Although the body weights were not significantly changed, glucose tolerance was slightly improved in HSW1, 3, 7, and 8 consumption groups but not in other HSW consumption groups. It suggests that HSW1 consumption may have the possible beneficial effect to improve glycemic control in the current clinical trial, therefore murine experiments might be able to use for the screening of effectiveness for glycemic control that were derived from HSW consumption. It has been reported that the consumption of hydrogen carbonate- or sulfur-containing water could prevent or improve T2D [36-38]. Because HSW1 and 7 contain hydrogen carbonate, and HSW8 contains sulfur, their beneficial effects were reasonable. However, HSW2, 4, and 5 also contain hydrogen carbonate but improvement of glucose tolerance was not observed in the mice experiment. Additionally, glucose tolerance was slightly improved in HSW3 consumption group, but the concentrations of both hydrogen carbonate and sulfur of

HSW3 were almost same as TW. Moreover, results of OPLS-DA indicated that it was difficult to discriminate effective and non-effective HSW by their mineral profiles. These results suggest that some types of HSW containing hydrogen carbonate or sulfur actually have the beneficial effects for glycemic control as reported previously, but the concentrations of hydrogen carbonate and sulfur were not sufficient enough to be discriminated of having the potential benefit for glycemic control. For future research, the important factor in HSW consumption leading to the beneficial effects in glycemic control derived from its consumption should be considered by using deionized HSW and/or deionized water supplemented with each mineral included in the effective HSW. In this study, importance of minerals in HSW for glycemic control improvement was evaluated. However, these beneficial effects might be derived from factors yet to be considered such as lipids, sugars, and/or microbes in the HSW. As such, further analyses of these factors is important in the future studies.

According to the results of OGTT, blood glucose levels were significantly decreased at 15 min after glucose administration in the effective HSW consumption groups. As the previous studies reported that GLP-1 induces the food-mediated early insulin secretion [151, 152], these results were expected to occur by increasing of insulin secretion via increasing of GLP-1. As such, GLP-1 and insulin concentrations in blood and insulin sensitivity were analyzed to investigate the reason behind the slight improvement of glucose tolerance in effective HSW consumption groups, but there are no significant differences as compared with TWAF consumption group. Instead, insulin concentrations at 15 and 60 min after glucose administration were significantly

decreased in the HSW8 consumption group, although the glucose levels were significantly decreased at 15 min after glucose administration in this group. Taken together, insulin sensitivity of effective HSW consumption group might not have been improved, thus we hypothesize that the suppression of glucose uptake from gut occurred. However, the reason why glucose tolerance was slightly improved in HSW1, 3, 7, and 8 consumption groups remains unclear in the current study. Therefore, glucose clamp test, 2-deoxyglucose uptake assay [153], and expression analysis of glucose receptors such as GLUT2 [154] should be required in further analyses to evaluate insulin sensitivity and glucose uptake activity to a stricter extent.

In addition, gut microbial function of human and mice were not completely same [155]. Microbiome analysis of human feces indicated lean-associated bacteria was significantly increased in the HSW1 consumption periods, but same bacteria were not detected from the any of the mice cecal contents. Therefore, it was speculated that the mechanisms of the beneficial effects of human and mice were different if the mechanisms involve gut microflora. As such, human flora-associated mice and/or antibiotics treatment may be useful to further evaluate the relationships between the beneficial effects for glycemic control derived from HSW consumption and gut microbiota.

Microbiome analysis in mice experiment showed that relative abundances of microbes were influenced by HSW consumption. Previous study has reported that magnesium and calcium concentrations in culture medium affected the growth of rumen bacteria [156], and as such, mineral consumption is expected to provide various effects

for gut microbiota. However, there are no minerals that were present in high concentrations in HSW7, even though the relative abundances of unclassified Bacteroidales and S24-7 were significantly higher only in HSW7 consumption group as compared with TWAF consumption group. This result suggests that the possible presence of other unknown factors that could have influenced the compositions of cecal microbiota, and/or that the balance of the mineral components is actually important rather than the concentrations of particular single mineral.

In addition, there are many positive or negative correlations between mineral contents and compositions of OTUs. These results suggest that various types of minerals have the potential to increase or decrease relative abundances of gut microbiota in species level. Although recent studies have reported that dietary habits are important in shaping the structure of gut microbiome [157, 158] thereby influencing host health status, there are abundant numbers of studies focusing on the effects of gut microbiota derived from dietary fat and/or fiber [159, 160]. Our present study shows the importance of mineral consumption to control the structure of gut microbiome. Therefore, the impact for gut microbiota compositions and their functions derived from consumption of various types of minerals should be analyzed in the further studies.

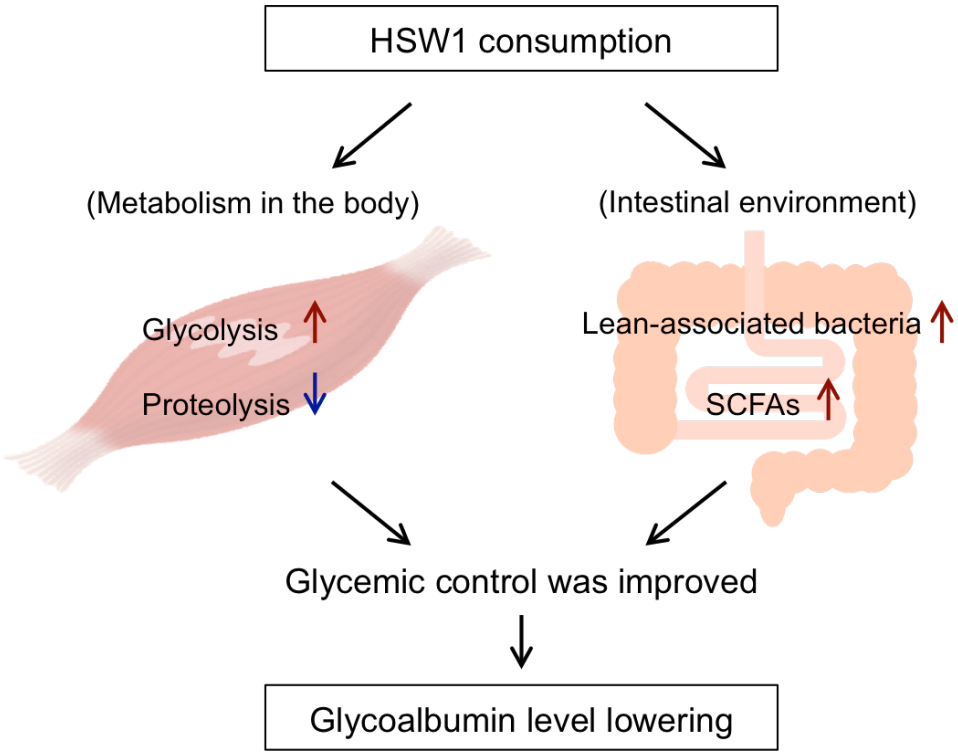


Figure 3.16 Overview of the beneficial effects derived from HSW1 consumption

We speculated the mechanisms how HSW1 consumption lead reduction of serum glycoalbumin levels. According to the results of metabolome analysis, glycolysis may be upregulated. Lowering of blood amino acids may be attributed to depression of proteolysis. Additionally, lean-associated bacteria and SCFAs were increased in the HSW1 consumption periods. Therefore, combination of these effects might improve glycemic control and thereby result in reduced serum glycoalbumin levels.

Chapter 4

Concluding remarks

Hot springs are natural environments that are found globally. They are familiar to people in various countries such as Japan. However, there are a lot of debatable and unresolved questions with regards to hot springs.

In Chapter 2, the microbes living in the Yunohama hot spring and their genes were analyzed [161], as hot spring environments are one of the most important and interesting research targets in environmental microbiology. The study revealed the presence of unique archaea phylogenetically closest to ARMAN from the Yunohama HSW. This finding is the first example of ARMAN living in an alkaline environment. In addition, 12 archaea and 15 bacteria that were detected for the first time in the Yunohama HSW. Collectively, these findings show that more types of microorganisms than previously expected inhabit the hot spring environments. However, we could not confirm whether these microbes actually live in the Yunohama hot spring environment because the results were based on microbiome analysis without cultivation, isolation, and microscopic observation. Therefore, isolation or at least genome reconstruction is needed for further consideration of genome sequence, function, and crosstalk between these microbes.

Metatranscriptome analysis revealed sRNAs that could form stable secondary structures (SURFYs) in the environmental samples. Previous studies have reported the variations and functions of sRNAs derived from model organisms, but the current study

indicated further diversity of sRNAs derived from previously uncultured microorganisms. Moreover, detailed analysis of fragmented tRNA molecules revealed the probable existence of type-specific tRNA degradation. Thus, many types of tRNAs tended to truncate around the anticodon region, while several others were truncated at the D-loop or the variable region. Additionally, this analysis also revealed the nucleotide preference at the first base of the anticodon (34th nucleotide of tRNAs) for site-specific tRNA degradation events. The degradation patterns of bacterial tRNAs and tRNA degradation enzymes have also been evaluated previously using model organisms. The current study reports various novel findings related to tRNA degradation in uncultured naturally occurring bacteria isolated from the environmental samples. Although the present study implied the existence of several uncharacterized RNases related to type- and/or site-specific tRNA degradation derived from the microbes living in the Yunohama hot spring environments, further biochemical assays are required to identify these enzymes. Taken together, the current study yielded many unique findings from the oceanfront deep subsurface hot spring environment. This indicates that hot spring environments are important and interesting research targets in environmental microbiology, and I believe that further novel microbes and genes will be elucidated from hot spring environments in the future.

In Chapter 3, the physiological effects of HSW consumption were evaluated based on the results of metabolome and microbiome analyses [162], since the molecular underlying the effects of balneotherapy are not well understood. In this study, we have shown that serum glycoalbumin levels were significantly decreased after consumption

of HSW1, which was collected from the Nagayu hot spring area. In addition, it was also suggested that glycolysis may be enhanced and proteolysis may be suppressed in humans in the HSW1 consumption periods. Lean-associated bacteria and concentrations of several organic acids in the intestine were also significantly increased. Therefore, HSW1 consumption may exert beneficial effects for glycemic control in humans by alteration of metabolic dynamics and intestinal environments. As previous studies of balneotherapy were often limited to epidemiological and clinical evaluations, findings of the current research may become very important as the first step for towards understanding the molecular basis of balneotherapy. However, it should be noted that many points still require to be considered to confirm and clarify the beneficial effects of HSW consumption on glycemic control, and subsequently, clinical trials among borderline diabetes patients and/or large-scale clinical trials would also be required in the future.

In the experiment on mice, the effects of HSW consumption on glycemic control were compared using 8 types of HSW. HSW1 consumption also showed some beneficial effects for glycemic control in mice, and it was expected that 4 out of 8 types of HSW including HSW1 have the potential to improve glycemic control. Therefore, murine model could be useful for screening to evaluate the effectiveness of HSW consumption. However, obesity could not be prevented in mice by consumption of any type of HSW. Since consumption of HFD32 might have excessively induced obesity to evaluate the effects of HSW consumption, it is important for future studies to perform similar experiments using mildly obese mice. Although the reasons pertaining to the

slight improvement in glucose tolerance in the HSW1, 3, 7, and 8 consumption groups remain unclear, we suspect that inhibition of glucose uptake may be one of the mechanisms, especially in the HSW8 consumption group. To confirm this, glucose clamp test, 2-deoxyglucose uptake assay, and gene expression analyses of glucose receptors in the intestinal tissue should be conducted. In the current study, importance of minerals in HSW for glycemic control improvement was evaluated, but it is still obscure according to our results. Therefore, further analyses focusing on the lipids, sugars, and/or microbes in the HSW may be important to reveal the important factor(s) for improvement of glycemic control.

In addition, we also propose the potential importance of minerals in the HSW, which may be able to control the balance of gut microbiota. In this study, several OTUs, mainly corresponding to Bacteroidales and Clostridiales, were highly correlated with mineral concentrations in drinking water. Recently, dietary habits, especially abundance of fat and/or fiber, have been reported to be important in determining the composition of gut microbiome, but dietary minerals are also expected to impact the compositions of gut microbiota. Finally, we believe that elucidation of the molecular basis of balneotherapy in humans and animals is vital for understanding of balneotherapy, and our study is the first step towards this goal.

In conclusion, the findings in this dissertation contribute to advancing the understanding of ecosystems in hot spring environments and the molecular basis of balneotheraphy. I hope these findings impact not only the scientific world, but also find applications in the real world.

Acknowledgments

First, I would like to thank Professor Masaru Tomita, who guided me to the field of bioscience and provided a great research environment at the Institute for Advanced Biosciences (IAB), Keio University. His passion and wisdom were always valuable to me. I would also like to thank Professor Akio Kanai, who taught me the important things in science. He was kind at times and strict at others, but his teachings will always be treasured. I am deeply grateful to Project Associate Professor Shinji Fukuda, who always gave me insightful comments and suggestions. My sincere gratitude also goes to Professors Mitsuhiro Watanabe, Tomoyoshi Soga, Mitsuhiro Itaya, and Toru Takebayashi; Associate Professors Yasuhiro Naito, Hiroki Kuroda, and Miki Akiyama; and Assistant Professor Hitomi Sano.

Dr. Wanping Aw always kindly taught me about scientific writing. I really appreciate everything that she did for me. I would also like to thank Project Associate Professor Joe Inoue, Dr. Yumiko Nakanishi, Ms. Yoko Takashina, Ms. Chiharu Ishii, Ms. Noriko Fukuda, Ms. Yuka Ohara, and Mr. Yuki Yoshida for their experimental and/or analytical support during my research. I would like to show my greatest appreciation to Mr. Yasuaki Goto, Dr. Shinya Hayasaka, and President Shigeo Kurihara, members of the Japan Health and Research Institute, for their substantial help and advice during this study. For their support during the clinical trial, I wish to thank Dr. Kyo Ito, Ms. Mitsuyo Matsuo, and members of the Ito Medical Office. I would like to offer my gratitude to the officials of Taketa-city, Ms. Ryoko Washitsukasa, Ms. Kyoko

Nakashita, Ms. Nozomi Mai, Mr. Toru Miyazaki, Mr. Yasuyuki Morita, Mr. Takahiro Kudo, Mr. Teiji Shimizu, Mr. Toshinori Hayashi, Mr. Hironori Shin, and Mayor Katsuji Shuto, for the governmental support during the clinical trial.

As my successive research advisers, I extend my sincere thanks to Dr. Rintaro Saito, Mr. Shinya Murata, Dr. Kosuke Fujishima, and Dr. Yoshiki Ikeda. I really appreciate my past and present colleagues, Mr. Fujitaka Baba, Ms. Chiharu Ishii, Ms. Yuri Yamamoto, and Mr. Keita Kamezaki, for their friendship. They always made me happy and the days we spent together were among my most favorite.

Finally, I would like to express my deepest gratitude to my dear parents, Yutaka Murakami and Noriko Murakami, and to my brother, Issei Murakami, for their continued and invaluable support to my student life. I thank them again from the bottom of my heart.

References

1. Pentecost A, *et al.* (2003) What is a hot spring? *Canadian Journal of Earth Sciences* 40(11):1443-1446.
2. Jackson R (1990) Waters and spas in the classical world. *Medical History. Supplement* (10):1-13.
3. van Tubergen A & van der Linden S (2002) A brief history of spa therapy. *Annals of the Rheumatic Diseases* 61(3):273-275.
4. Serbulea M & Payyappallimana U (2012) Onsen (hot springs) in Japan--transforming terrain into healing landscapes. *Health & Place* 18(6):1366-1373.
5. Amann R & Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nature Reviews Microbiology* 6(5):339-348.
6. Warnecke F & Hess M (2009) A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. *Journal of Biotechnology* 142(1):91-95.
7. Handelsman J, *et al.* (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* 5(10):R245-249.
8. Baker BJ, *et al.* (2006) Lineages of acidophilic archaea revealed by community genomic analysis. *Science* 314(5807):1933-1935.
9. Nunoura T, *et al.* (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Research* 39(8):3204-3223.
10. DeLong EF & Pace NR (2001) Environmental diversity of bacteria and archaea. *Systematic Biology* 50(4):470-478.
11. Smith MI, *et al.* (2014) Microbiome analysis - from technical advances to biological relevance. *F1000prime Reports* 6:51.
12. Barns SM, *et al.* (1994) Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Sciences of the United States of America* 91(5):1609-1613.
13. Kvist T, *et al.* (2007) Archaeal diversity in Icelandic hot springs. *FEMS Microbiology Ecology* 59(1):71-80.
14. Hall JR, *et al.* (2008) Molecular characterization of the diversity and distribution of a thermal spring microbial community by using rRNA and metabolic genes. *Applied and Environmental Microbiology* 74(15):4910-4922.
15. Inskeep WP, *et al.* (2010) Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS One* 5(3):e9773.
16. Venter JC, *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304(5667):66-74.
17. Xie W, *et al.* (2011) Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. *The ISME Journal* 5(3):414-426.
18. Nunoura T, *et al.* (2005) Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environmental Microbiology* 7(12):1967-1984.
19. Grice EA, *et al.* (2009) Topographical and temporal diversity of the human skin microbiome. *Science* 324(5931):1190-1192.

20. Nakatsuji T, *et al.* (2013) The microbiome extends to subepidermal compartments of normal skin. *Nature Communications* 4:1431.
21. Murakami S, *et al.* (2015) Comprehensive analysis of microbes and metabolites in human tear fluids. *Keio SFC Journal* 15(1):382-400.
22. Arumugam M, *et al.* (2011) Enterotypes of the human gut microbiome. *Nature* 473(7346):174-180.
23. Oshima T & Imahori K (1974) Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. *International Journal of Systematic Bacteriology* 24(1):102-112.
24. Schleper C, *et al.* (1995) A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. *Journal of Bacteriology* 177(15):4417-4426.
25. Itoh T, *et al.* (2002) *Vulcanisaeta distributa* gen. nov., sp. nov., and *Vulcanisaeta souniana* sp. nov., novel hyperthermophilic, rod-shaped crenarchaeotes isolated from hot springs in Japan. *International Journal of Systematic and Evolutionary Microbiology* 52(Pt 4):1097-1104.
26. Suzuki T, *et al.* (2002) *Sulfolobus tokodaii* sp. nov. (f. *Sulfolobus* sp. strain 7), a new member of the genus *Sulfolobus* isolated from Beppu Hot Springs, Japan. *Extremophiles* 6(1):39-44.
27. Lee DJ, *et al.* (2014) Azoarcus taiwanensis sp. nov., a denitrifying species isolated from a hot spring. *Applied Microbiology and Biotechnology* 98(3):1301-1307.
28. Brock TD & Freeze H (1969) *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *Journal of Bacteriology* 98(1):289-297.
29. Henry EA, *et al.* (1994) Characterization of a new thermophilic sulfate-reducing bacterium *Thermodesulfovibrio yellowstonii*, gen. nov. and sp. nov.: its phylogenetic relationship to *Thermodesulfobacterium commune* and their origins deep within the bacterial domain. *Archives of Microbiology* 161(1):62-69.
30. Baross JA & Hoffman SE (1985) Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Origins of Life and Evolution of the Biosphere* 15(4):327-345.
31. Schumacher MA, *et al.* (2015) Structures of archaeal DNA segregation machinery reveal bacterial and eukaryotic linkages. *Science* 349(6252):1120-1124.
32. Saiki RK, *et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839):487-491.
33. Nasermoaddeli A & Kagamimori S (2005) Balneotherapy in medicine: A review. *Environmental Health and Preventive Medicine* 10(4):171-179.
34. Karagulle M & Karagulle MZ (2015) Effectiveness of balneotherapy and spa therapy for the treatment of chronic low back pain: a review on latest evidence. *Clinical Rheumatology* 34(2):207-214.
35. Pagourelias ED, *et al.* (2011) Carbon dioxide balneotherapy and cardiovascular disease. *International Journal of Biometeorology* 55(5):657-663.
36. Gutenbrunner C (1993) Kontrollierte studie über die wirkung einer haustrinkkur mit einem natrium-hydrogencarbonat-säuerling auf die blutzuckerregulation bei gesunden versuchspersonen. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin* 3(04):108-110.
37. Ohtsuka Y, *et al.* (2003) Effect of hot spring water drinking on glucose metabolism. *The Journal of the Japanese Society of Balneology, Climatology and Physical Medicine* 66(4):227-230.

38. Schoppen S, *et al.* (2007) Does bicarbonated mineral water rich in sodium change insulin sensitivity of postmenopausal women? *Nutricion Hospitalaria* 22(5):538-544.
39. Annegret F & Thomas F (2013) Long-term benefits of radon spa therapy in rheumatic diseases: results of the randomised, multi-centre IMuRa trial. *Rheumatology International* 33(11):2839-2850.
40. Ablin JN, *et al.* (2013) Spa treatment (balneotherapy) for fibromyalgia-a qualitative-narrative review and a historical perspective. *Evidence-Based Complementary and Alternative Medicine* 2013:638050.
41. Choi YJ, *et al.* (2013) Therapeutic effects and immunomodulation of suanbo mineral water therapy in a murine model of atopic dermatitis. *Annals of Dermatology* 25(4):462-470.
42. Perez-Granados AM, *et al.* (2010) Reduction in cardiovascular risk by sodium-bicarbonated mineral water in moderately hypercholesterolemic young adults. *The Journal of Nutritional Biochemistry* 21(10):948-953.
43. Tabish SA (2008) Complementary and Alternative Healthcare: Is it Evidence-based? *International Journal of Health Sciences* 2(1):V-IX.
44. Esteghamati A, *et al.* (2015) Complementary and alternative medicine for the treatment of obesity: a critical review. *International Journal of Endocrinology and Metabolism* 13(2):e19678.
45. Wieland LS, *et al.* (2011) Development and classification of an operational definition of complementary and alternative medicine for the Cochrane collaboration. *Alternative Therapies in Health and Medicine* 17(2):50-59.
46. Lourens-Hattingh A & Viljoen BC (2001) Yogurt as probiotic carrier food. *International Dairy Journal* 11(1):1-17.
47. Tang TH, *et al.* (2002) RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. *Nucleic Acids Research* 30(4):921-930.
48. Gottesman S (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends in Genetics* 21(7):399-404.
49. Jost D, *et al.* (2011) Small RNA biology is systems biology. *BMB Reports* 44(1):11-21.
50. Aiba H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Current Opinion in Microbiology* 10(2):134-139.
51. Gama-Castro S, *et al.* (2008) RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Research* 36(Database issue):D120-124.
52. Argaman L, *et al.* (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Current Biology* 11(12):941-950.
53. Rivas E, *et al.* (2001) Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Current Biology* 11(17):1369-1373.
54. Chen S, *et al.* (2002) A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. *Bio Systems* 65(2-3):157-177.
55. Yachie N, *et al.* (2006) Prediction of non-coding and antisense RNA genes in *Escherichia coli* with Gapped Markov Model. *Gene* 372:171-181.
56. Vogel J, *et al.* (2003) RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Research* 31(22):6435-6443.
57. Kawano M, *et al.* (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Research* 33(3):1040-1050.
58. Tjaden B, *et al.* (2002) Transcriptome analysis of *Escherichia coli* using high-density

oligonucleotide probe arrays. *Nucleic Acids Research* 30(17):3732-3738.

59. Zhang A, *et al.* (2003) Global analysis of small RNA and mRNA targets of Hfq. *Molecular Microbiology* 50(4):1111-1124.
60. Raghavan R, *et al.* (2011) Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Research* 21(9):1487-1497.
61. Shinhara A, *et al.* (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genomics* 12:428.
62. Masse E, *et al.* (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli. Genes & Development* 17(19):2374-2383.
63. Haiser HJ, *et al.* (2008) Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*. *Nucleic Acids Research* 36(3):732-741.
64. Masaki H & Ogawa T (2002) The modes of action of colicins E5 and D, and related cytotoxic tRNases. *Biochimie* 84(5-6):433-438.
65. Ogawa T, *et al.* (2006) Sequence-specific recognition of colicin E5, a tRNA-targeting ribonuclease. *Nucleic Acids Research* 34(21):6065-6073.
66. Thompson DM, *et al.* (2008) tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* 14(10):2095-2103.
67. Emara MM, *et al.* (2010) Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *The Journal of Biological Chemistry* 285(14):10959-10968.
68. Lee YS, *et al.* (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes & Development* 23(22):2639-2649.
69. Kato M, *et al.* (2011) Age-associated changes in expression of small, noncoding RNAs, including microRNAs, in *C. elegans*. *RNA* 17(10):1804-1820.
70. Evguenieva-Hackenberg E (2005) Bacterial ribosomal RNA in pieces. *Molecular Microbiology* 57(2):318-325.
71. Milbury CA, *et al.* (2010) Fragmentation of the large subunit ribosomal RNA gene in oyster mitochondrial genomes. *BMC Genomics* 11:485.
72. Serganov A (2009) The long and the short of riboswitches. *Current Opinion in Structural Biology* 19(3):251-259.
73. Shi Y, *et al.* (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459(7244):266-269.
74. Bomar L, *et al.* (2011) Directed culturing of microorganisms using metatranscriptomics. *mBio* 2(2):e00012-11.
75. Weinberg Z, *et al.* (2009) Exceptional structured noncoding RNAs revealed by bacterial metagenome analysis. *Nature* 462(7273):656-659.
76. Gantner S, *et al.* (2011) Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. *Journal of Microbiological Methods* 84(1):12-18.
77. Wilms R, *et al.* (2006) Specific bacterial, archaeal, and eukaryotic communities in tidal-flat sediments along a vertical profile of several meters. *Applied and Environmental Microbiology* 72(4):2756-2764.
78. Lowe TM & Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25(5):955-964.
79. Chan PP & Lowe TM (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research* 37(Database issue):D93-97.

80. Szymanski M, *et al.* (2002) 5S Ribosomal RNA Database. *Nucleic Acids Research* 30(1):176-178.
81. Cole JR, *et al.* (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* 37(Database issue):D141-145.
82. Pruesse E, *et al.* (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* 35(21):7188-7196.
83. Gardner PP, *et al.* (2009) Rfam: updates to the RNA families database. *Nucleic Acids Research* 37(Database issue):D136-140.
84. Ding Y & Lawrence CE (2003) A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Research* 31(24):7280-7301.
85. Ding Y, *et al.* (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA* 11(8):1157-1166.
86. Larkin MA, *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.
87. Letunic I & Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23(1):127-128.
88. Takai K, *et al.* (2001) Archaeal diversity in waters from deep South African gold mines. *Applied and Environmental Microbiology* 67(12):5750-5760.
89. Baker BJ, *et al.* (2010) Enigmatic, ultrasmall, uncultivated Archaea. *Proceedings of the National Academy of Sciences of the United States of America* 107(19):8806-8811.
90. Massana R, *et al.* (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Applied and Environmental Microbiology* 63(1):50-56.
91. Takai K & Horikoshi K (1999) Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* 152(4):1285-1297.
92. Jurgens G & Saano A (1999) Diversity of soil Archaea in boreal forest before, and after clear-cutting and prescribed burning. *FEMS Microbiology Ecology* 29(2):205-213.
93. Miyoshi T, *et al.* (2005) Phylogenetic characterization of 16S rRNA gene clones from deep-groundwater microorganisms that pass through 0.2-micrometer-pore-size filters. *Applied and Environmental Microbiology* 71(2):1084-1088.
94. Regalia M, *et al.* (2002) Prediction of signal recognition particle RNA genes. *Nucleic Acids Research* 30(15):3368-3377.
95. Wassarman KM & Storz G (2000) 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101(6):613-623.
96. Winkler W, *et al.* (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419(6910):952-956.
97. Schlax PJ, *et al.* (2001) Translational repression of the *Escherichia coli* alpha operon mRNA: importance of an mRNA conformational switch and a ternary entrapment complex. *The Journal of Biological Chemistry* 276(42):38494-38501.
98. Washietl S, *et al.* (2005) Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. *Nature Biotechnology* 23(11):1383-1390.
99. Marck C & Grosjean H (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8(10):1189-1232.
100. Weinberg Z, *et al.* (2010) Comparative genomics reveals 104 candidate structured RNAs from

bacteria, archaea, and their metagenomes. *Genome Biology* 11(3):R31.
101. Kawaji H, *et al.* (2008) Hidden layers of human small RNAs. *BMC Genomics* 9:157.
102. Cole C, *et al.* (2009) Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA* 15(12):2147-2160.
103. Buhler M, *et al.* (2008) TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the *Schizosaccharomyces pombe* siRNA pathway. *Nature Structural & Molecular Biology* 15(10):1015-1023.
104. Thompson DM & Parker R (2009) The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in Saccharomyces cerevisiae. *The Journal of Cell Biology* 185(1):43-50.
105. Walker KZ, *et al.* (2010) Diet and exercise in the prevention of diabetes. *Journal of Human Nutrition and Dietetics* 23(4):344-352.
106. Salas-Salvado J, *et al.* (2011) The role of diet in the prevention of type 2 diabetes. *Nutrition, Metabolism & Cardiovascular Diseases* 21(Supplement 2):B32-48.
107. Samino S, *et al.* (2015) Metabolomics reveals impaired maturation of HDL particles in adolescents with hyperinsulinaemic androgen excess. *Scientific Reports* 5:11496.
108. Sreekumar A, *et al.* (2009) Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 457(7231):910-914.
109. Uetaki M, *et al.* (2015) Metabolomic alterations in human cancer cells by vitamin C-induced oxidative stress. *Scientific Reports* 5:13896.
110. Mishima E, *et al.* (2015) Alteration of the intestinal environment by lubiprostone is associated with amelioration of adenine-induced CKD. *Journal of the American Society of Nephrology* 26(8):1787-1794.
111. Sugimoto M, *et al.* (2010) Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics* 6(1):78-95.
112. Kaur P, *et al.* (2013) Quantitative metabolomic and lipidomic profiling reveals aberrant amino acid metabolism in type 2 diabetes. *Molecular Biosystems* 9(2):307-317.
113. Larsen N, *et al.* (2010) Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One* 5(2):e9085.
114. Qin J, *et al.* (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490(7418):55-60.
115. Norman JM, *et al.* (2015) Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160(3):447-460.
116. Hsiao EY, *et al.* (2013) Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* 155(7):1451-1463.
117. Karlsson FH, *et al.* (2013) Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498(7452):99-103.
118. Kimura I, *et al.* (2013) The gut microbiota suppresses insulin-mediated fat accumulation via the short-chain fatty acid receptor GPR43. *Nature Communications* 4:1829.
119. Xia J & Wishart DS (2010) MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Research* 38(Web Server issue):W71-77.
120. Furusawa Y, *et al.* (2013) Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504(7480):446-450.
121. Kim SW, *et al.* (2013) Robustness of gut microbiota of healthy adults in response to probiotic

intervention revealed by high-throughput pyrosequencing. *DNA Research* 20(3):241-253.

122. Magoc T & Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21):2957-2963.
123. Caporaso JG, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7(5):335-336.
124. Candela M, *et al.* (2010) High taxonomic level fingerprint of the human intestinal microbiota by ligase detection reaction--universal array approach. *BMC Microbiology* 10:116.
125. Koga M & Kasayama S (2010) Clinical impact of glycated albumin as another glycemic control marker. *Endocrine Journal* 57(9):751-762.
126. Kohka K & Maehata E (2004) An evaluation of glucose tolerance, establishment of a glycoalbumin standard range, and a consideration of insulin resistance and diabetes mellitus education. *Journal of Analytical Bio-Science* 27:104-110.
127. Paroni R, *et al.* (2007) Performance characteristics and clinical utility of an enzymatic method for the measurement of glycated albumin in plasma. *Clinical Biochemistry* 40(18):1398-1405.
128. Moebus S, *et al.* (2011) Impact of time since last caloric intake on blood glucose levels. *European Journal of Epidemiology* 26(9):719-728.
129. Fujita T (2000) Calcium paradox: consequences of calcium deficiency manifested by a wide variety of diseases. *Journal of Bone and Mineral Metabolism* 18(4):234-236.
130. Dennis EA, *et al.* (2010) Water consumption increases weight loss during a hypocaloric diet intervention in middle-aged and older adults. *Obesity* 18(2):300-307.
131. Wang TJ, *et al.* (2011) Metabolite profiles and the risk of developing diabetes. *Nature Medicine* 17(4):448-453.
132. Nakamura H, *et al.* (2014) Plasma amino acid profiles are associated with insulin, C-peptide and adiponectin levels in type 2 diabetic patients. *Nutrition & Diabetes* 4:e133.
133. Sonksen P & Sonksen J (2000) Insulin: understanding its action in health and disease. *British Journal of Anaesthesia* 85(1):69-79.
134. Yang X, *et al.* (2008) Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature* 451(7181):964-969.
135. Pouwels MJ, *et al.* (2002) Muscle uridine diphosphate-hexosamines do not decrease despite correction of hyperglycemia-induced insulin resistance in type 2 diabetes. *The Journal of Clinical Endocrinology and Metabolism* 87(11):5179-5184.
136. Turnbaugh PJ, *et al.* (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444(7122):1027-1031.
137. Musso G, *et al.* (2010) Obesity, diabetes, and gut microbiota: the hygiene hypothesis expanded? *Diabetes Care* 33(10):2277-2284.
138. Ridaura VK, *et al.* (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341(6150):1241214.
139. Goodrich JK, *et al.* (2014) Human genetics shape the gut microbiome. *Cell* 159(4):789-799.
140. Wen L, *et al.* (2008) Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* 455(7216):1109-1113.
141. Parks BW, *et al.* (2013) Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metabolism* 17(1):141-152.
142. Xiao JZ, *et al.* (2007) Clinical efficacy of probiotic Bifidobacterium longum for the treatment of symptoms of Japanese cedar pollen allergy in subjects evaluated in an environmental exposure unit. *Allergology International* 56(1):67-75.

143. Ishizuka A, *et al.* (2012) Effects of administration of Bifidobacterium animalis subsp. lactis GCL2505 on defecation frequency and bifidobacterial microbiota composition in humans. *Journal of Bioscience and Bioengineering* 113(5):587-591.
144. Fukuda S, *et al.* (2011) Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* 469(7331):543-547.
145. Yoshimoto S, *et al.* (2013) Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* 499(7456):97-101.
146. Yadav H, *et al.* (2013) Beneficial metabolic effects of a probiotic via butyrate-induced GLP-1 hormone secretion. *The Journal of Biological Chemistry* 288(35):25088-25097.
147. Ridlon JM, *et al.* (2006) Bile salt biotransformations by human intestinal bacteria. *Journal of Lipid Research* 47(2):241-259.
148. Van den Abbeele P, *et al.* (2013) Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an in vitro gut model. *The ISME Journal* 7(5):949-961.
149. Rafter JJ, *et al.* (1987) Cellular toxicity of fecal water depends on diet. *The American Journal of Clinical Nutrition* 45(3):559-563.
150. Sayin SI, *et al.* (2013) Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metabolism* 17(2):225-235.
151. Drucker DJ (1998) Glucagon-like peptides. *Diabetes* 47(2):159-169.
152. Lugari R, *et al.* (2002) Evidence for early impairment of glucagon-like peptide 1-induced insulin secretion in human type 2 (non insulin-dependent) diabetes. *Hormone and Metabolic Research* 34(3):150-154.
153. Hwang JS, *et al.* (2015) Glucosamine enhances body weight gain and reduces insulin response in mice fed chow diet but mitigates obesity, insulin resistance and impaired glucose tolerance in mice high-fat diet. *Metabolism* 64(3):368-379.
154. Thorens B (2015) GLUT2, glucose sensing and glucose homeostasis. *Diabetologia* 58(2):221-232.
155. Xiao L, *et al.* (2015) A catalog of the mouse gut metagenome. *Nature Biotechnology* 33(10):1103-1108.
156. Morales MS & Dehority BA (2014) Magnesium requirement of some of the principal rumen cellulolytic bacteria. *Animal* 8(9):1427-1432.
157. Claesson MJ, *et al.* (2012) Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488(7410):178-184.
158. Voreades N, *et al.* (2014) Diet and the development of the human intestinal microbiome. *Frontiers in Microbiology* 5:494.
159. Trompette A, *et al.* (2014) Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature Medicine* 20(2):159-166.
160. Caesar R, *et al.* (2015) Crosstalk between Gut Microbiota and Dietary Lipids Aggravates WAT Inflammation through TLR Signaling. *Cell Metabolism* 22(4):658-668.
161. Murakami S, *et al.* (2012) Metatranscriptomic analysis of microbes in an Oceanfront deep-subsurface hot spring reveals novel small RNAs and type-specific tRNA degradation. *Applied and Environmental Microbiology* 78(4):1015-1022.
162. Murakami S, *et al.* (2015) The Consumption of Bicarbonate-Rich Mineral Water Improves Glycemic Control. *Evidence-Based Complementary and Alternative Medicine* 2015:824395.

Abbreviations

Amino acids

A	Ala	alanine
R	Arg	arginine
N	Asn	asparagine
D	Asp	aspartic acid
C	Cys	cysteine
Q	Gln	glutamine
E	Glu	glutamic acid
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
L	Leu	leucine
K	Lys	lysine
M	Met	methionine
F	Phe	phenylalanine
P	Pro	proline
S	Ser	serine
T	Thr	threonine
W	Trp	tryptophan
Y	Tyr	tyrosine
V	Val	valine

Others

AAG	ancient archaeal group
ADP	adenosine diphosphate
Alpha RBS	alpha operon ribosome-binding site
ARMAN	archaeal Richmond Mine acidophilic nanoorganisms
ATP	adenosine triphosphate

BC	before Christ
Beta-proteo	Betaproteobacteria
BLAST	Basic Local Alignment Search Tool
BMI	body mass index
bp	base pair
CAM	complementary and alternative medicine
cDNA	complementary DNA
CE-TOFMS	capillary electrophoresis with electrospray ionization time-of-flight mass spectrometry
CSA	D-camphor-10-sulfonic acid
CV	coefficient of variation
DDBJ	DNA Data Bank of Japan
Delta-proteo	Deltaproteobacteria
DNA	deoxyribonucleic acid
DPP-IV	dipeptidyl peptidase-4
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EMBL	European Molecular Biology Laboratory
FDR	false discovery rate
FLASH	Fast Length Adjustment of Short Reads
FSCG	forest soil crenarchaeotic group
Gamma-proteo	Gammaproteobacteria
GLP-1	glucagon-like peptide-1
GLUT2	glucose transporter 2
GtRNAdb	genomic tRNA database
HDL	high-density lipoprotein
HOMA-R	homeostasis model assessment ratio
HSW	hot spring water
HWCG	hot water crenarchaeotic group
IPGTT	intraperitoneal glucose tolerance test

IPITT	intraperitoneal insulin tolerance test
iTOL	Interactive Tree Of Life
LDL	low-density lipoprotein
min	minute
MCFA	medium-chain fatty acid
MCG	miscellaneous crenarchaeotic group
miRNA	micro RNA
mRNA	messenger RNA
MSEA	Metabolite Set Enrichment Analysis
NCBI	National Center for Biotechnology Information
ND	not detected
NM	not measured
NS	not significant
nt	nucleotide
OGTT	oral glucose tolerance test
OPLS-DA	Orthogonal Partial Least Squares Discriminate Analysis
OTU	operational taxonomic unit
PCR	polymerase chain reaction
QIIME	Quantitative Insights Into Microbial Ecology
RDP	Ribosomal Database Project
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal RNA
RT	room temperature
SAGMCG	south african gold mine crenarchaeotic group
SCFA	short-chain fatty acid
SD	standard deviation
SDS	sodium dodecyl sulfate
sec	second

sRNA	small noncoding RNA
SRP bact	bacterial signal recognition particle RNA
SURFY	small unique RNA from Yunohama
T2D	type 2 diabetes
temp	temperature
TPP	thiamine pyrophosphate
tRNA	transfer RNA
TW	tap water used in the clinical trial (purchased from Nishikawa water plant)
TWAF	tap water used in the animal experiment (obtained from the water tap in the animal facility)
UDP	uridine diphosphate
UPLC-MS	ultra-performance liquid chromatography mass-spectrometry
USA	United States of America