SCIENTIFIC DATA

Check for updates

OPEN

ARTICLE

ODFM, an omics data resource from microorganisms associated with fermented foods

Tae Woong Whon[1,4], Seung Woo Ahn[1,3,4], Sungjin Yang[2], Joon Yong Kim[1], Yeon Bee Kim[1], Yujin Kim[1], Ji-Man Hong[2], Hojin Jung[2], Yoon-E Choi[3], Se Hee Lee[1 ✉] & Seong Woon Roh[1 ✉]

ODFM is a data management system that integrates comprehensive omics information for microorganisms associated with various fermented foods, additive ingredients, and seasonings (e.g. kimchi, Korean fermented vegetables, fermented seafood, solar salt, soybean paste, vinegar, beer, cheese, sake, and yogurt). The ODFM archives genome, metagenome, metataxonome, and (meta) transcriptome sequences of fermented food-associated bacteria, archaea, eukaryotic microorganisms, and viruses; 131 bacterial, 38 archaeal, and 28 eukaryotic genomes are now available to users. The ODFM provides both the Basic Local Alignment Search Tool search-based local alignment function as well as average nucleotide identity-based genetic relatedness measurement, enabling gene diversity and taxonomic analyses of an input query against the database. Genome sequences and annotation results of microorganisms are directly downloadable, and the microbial strains registered in the archive library will be available from our culture collection of fermented food-associated microorganisms. The ODFM is a comprehensive database that covers the genomes of an entire microbiome within a specific food ecosystem, providing basic information to evaluate microbial isolates as candidate fermentation starters for fermented food production.

Introduction

Advances in next-generation sequencing technology have led to the rapid expansion of microbial genome sequence data. Easy access, as well as convenient analytical tools, have enabled the exploration of microbial communities in various environmental samples. However, efficient resource usage is becoming increasingly difficult because of the rapid accumulation of sequencing data. Environmental microbiomes in fermented foods, the mammalian gut, and soils comprise not only bacteria, but also archaea, eukaryotic microorganisms, and viruses[1–3]. These microbial entities are all essential in determining the microbial signature and thus, the inherent characteristics of a given ecosystem. In this context, a comprehensive database covering all genomes of a microbiome within a specific ecosystem would aid in improving our understanding of the complex interactions among the microbial populations.

Fermented foods are an integral part of the global human diet. Microbial entities in fermented foods include bacteria, archaea, yeasts, and viruses. Microbial activities, as well as the type of raw materials, ultimately determine the nutritional and organoleptic properties, quality, and safety of the fermentation product[4,5]. Given that consumers and manufacturers alike are interested in tasty, high-quality foods[6] as well as the reliability of geographic origins (i.e. no false indication of the origin of the product)[1], providing standardised microbial profiles and/or genome information for key microorganisms during the fermentation process is important for ensuring the high quality of fermentation products.

Kimchi is a traditional Korean food prepared by fermentation of vegetables, such as kimchi cabbage, along with various added ingredients and seasonings. The global annual consumption of kimchi is 1,500,000 tons[1]. Like other fermented foods, kimchi shows the presence of a distinct microbial community[4,5]. Taxonomic studies using culture-dependent and -independent (e.g. bacterial 16 S rRNA gene sequencing) approaches have revealed that

[1]Microbiology and Functionality Research Group, World Institute of Kimchi, Gwangju, 61755, Korea. [2]Insilicogen Inc., Gyeonggi-do, 16954, Korea. [3]Division of Environmental Science and Ecological Engineering, Korea University, Seoul, 02841, Korea. [4]These authors contributed equally: Tae Woong Whon, Seung Woo Ahn. ✉e-mail: leesehee@wikim.re.kr; swroh@wikim.re.kr

Category	Features
System environment	
Operating system	Centos (v6.5)
Java runtime environment	Java Development Kit (v1.8)
Web application	Apache (v2.2.15) and Tomcat server (v7.0)
Database operation	MySQL (v5.7)
Database content	
Genome	
Bacteria	62 complete and 69 draft genomes covering 96 (sub)species
Archaea	7 complete and 31 draft genomes covering 36 species
Unicellular eukaryotes	14 complete and 14 draft genomes covering 9 species
Metagenome	10 total and 60 viral metagenomes (kimchi)
Metataxonome	113 bacterial metataxonomes (kimchi, fermented seafood, and soybean paste)
(Meta)transcriptome	5 metatranscriptome (kimchi) and 4 archaeal transcriptome
Metabolome	7 metabolomes (kimchi and fermented seafood)
Functionality	
Browsing	JBrowse and GView Java package
Tools	Sequence similarity search (BLAST) and genomic relatedness analysis (ANI)

Table 1. Features of the ODFM.

lactic acid bacteria (LAB), including *Leuconostoc*, *Lactobacillus*, and *Weissella*, are mainly responsible for kimchi fermentation[7–9].

We have developed the Omics Database of Fermentative Microbes (ODFM), a data management system that integrates comprehensive omics information for fermentative microorganisms at the World Institute of Kimchi funded by the Korean government. The ODFM offers not only curated omics sequences of fermented food-associated bacteria, archaea, eukaryotic microorganisms, and viruses, but also several analytical tools that enable gene diversity and taxonomic analyses of an input query on the database at the whole genome level. Our knowledgebase is valuable to researchers who are interested in the functions and spatiotemporal dynamics of microbiomes in fermented foods. In particular, it provides basic information to evaluate microbial strains isolated from fermented foods as candidate starters in terms of food safety and sanitation.

Results

System design and data registration. The ODFM is a web-based application developed in compliance with the HyperText Markup Language (HTML) 5 web standards and, thus, is supported by most web browsers. The program was designed based on the Representational State Transfer (REST) service architecture to support use on various devices, including desktop computers and mobile devices. To support stable web service in a cloud-based service environment, the ODFM is hosted on four servers (web, web application, database, and storage servers). Key specifications for each server are summarised in Table 1.

The software architecture of the ODFM consists of client, server, and database modules. The client module uses Google's AngularJS (version 1.7) as a front-end framework to support cross browsing. The server module operates on JAVA (https://www.oracle.com/technetwork/java/index.html)-implemented Spring framework (https://spring.io) and is additionally equipped with the open-source programs Python (https://www.python.org/) and FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The database module uses MySQL (https://www.mysql.com) to manage the database, while the ODFM user interface architecture consists of Registration, Data Search, Tools, Our Projects, Statistics, and Q&A (Fig. 1a).

To improve data search and the accuracy of search results, submitted genome data and metadata are stored on the storage server and the database server, respectively. The system administrator processes the verification, approval, and release of the registered data. In the data verification step, validation of the file format and conversion of the registered data files are processed in batch using a back-end module. Once the processing is completed, the registered data are presented to the user in the web browser. The ODFM provides an interface to report data, including sequence files, annotated files, and results of sequence quality control according to the data file types. In addition, JBrowse and GView are integrated into the system to provide microbial genome analysis services via the genome browser. A diagram describing the data registration is provided in Supplementary Fig. S1.

Database content. The ODFM database currently contains 131 complete/draft bacterial genome sequences covering 38 genera with 96 (sub)species, of which 24 (sub)species belong to the genus *Lactobacillus*[10], 12 to the genus *Leuconostoc*[11–16], seven to the genus *Acetobacter*[17], six to the genus *Staphylococcus*[18], five to the genus *Enterococcus*, four (sub)species to each of the genera *Pediococcus* and *Weissella*[19,20], three to the genus *Lactococcus*, two (sub)species to each of the genera *Brachybacterium*, *Clostridium*, *Corynebacterium*, and *Pseudomonas*, and one species to each of the genera *Alishewanella*[21], *Bacillus*[22,23], *Brevibacterium*, *Dietzia*, *Escherichia*, *Glutamicibacter*, *Hafnia*, *Halomonas*, *Lentibacillus*, *Listeria*, *Megasphaera*, *Microbacterium*, *Morganella*, *Mycetocola*, *Oceanobacillus*, *Paracoccus*, *Pectinatus*, *Pistricoccus*, *Propionibacterium*, *Salimicrobium*[24], *Streptococcus*, *Tetragenococcus*, and *Vibrio*. As for archaea, 38 complete/draft genome sequences of extremely halophilic archaea (19 genera with

Fig. 1 Search functions in the ODFM knowledgebase. (**a**) Front page of the ODFM knowledgebase. (**b**) Basic search functions of the ODFM. By selecting one of three search functions on the data search tab, users can search omics data categorised by taxonomy, study, and sample.

36 species) are available, including seven species belonging to the genus *Haloarcula*[25], four belonging to the genus *Halorubrum*[26], three belonging to each of the genera *Haloferax* and *Natronomonas*, two belonging to each of the genera *Halapricum*[27], *Halobacterium*[28], *Halolamina*[29] and *Haloplanus*, and one species belonging to each of the genera *Haladaptatus*[30], *Halalkalicoccus*, *Halarchaeum*, *Halobellus*[31], *Halococcus*, *Halogeometricum*, *Halopenitus*, *Halorhabdus*, *Halostella*, *Haloterrigena*[32], and *Natrinema*[33]. As for eukaryotic microorganisms, 28 genome sequences are currently available, including five undesirable white colony-forming yeasts of the species *Candida*, *Hanseniaspora*, *Kazachstania*, *Pichia*, and *Yarrowia*. These spoilage yeasts can grow on the surface of the kimchi and affect its odour, appearance, and texture[34]. Genome sequences of *Brettanomyces*-, *Penicillium*-, and *Saccharomyces*-belonging species isolated from beer, cheese, and sake, respectively, are available. The database also contains 70 metagenomes, 113 bacterial metataxonomes, nine (meta)transcriptomes, and seven metabolomes for various fermented foods (Table 1). The viral metagenomic sequences have been deposited in the European Bioinformatics Institute (EMBL-EBI) database[35] and are available under accession number PRJEB23957. Details on ODFM database contents are provided in Supplementary Table S1 [1,7,9–14,16,–34,36–108].

Functional omics archive for fermented food-associated microorganisms. The primary purpose of the ODFM is to provide integrative functional omics information on fermented food-associated microorganisms. The ODFM and online resource provide omics information for microbial isolates from food materials (e.g. kimchi, fermented seafood, solar salt, soybean paste, vinegar, beer, cheese, sake, and yogurt). Recent microbial community analyses based on metataxonomics have revealed that hundreds of bacterial operational taxonomic units/amplicon sequence variants can be detected in fermented foods, and that the number of species varies according to the fermentation process[1,7]. To cover the entire microbial populations involved in food fermentation, since 2018, we have been constantly isolating and sequencing fermented food-associated microorganisms, and updating the database with new data to expand the ODFM archive.

Search function. For easy access of omics information, several search tools with simple (i.e. exact-match keyword) and lexical (i.e. partial-match keyword) search options are available on the front page (Fig. 1a). These tools allow users to search for different combinations of search terms. Users can search microbial taxa at the species and strain levels. Once a taxon name is provided by users, the system returns categorised search results
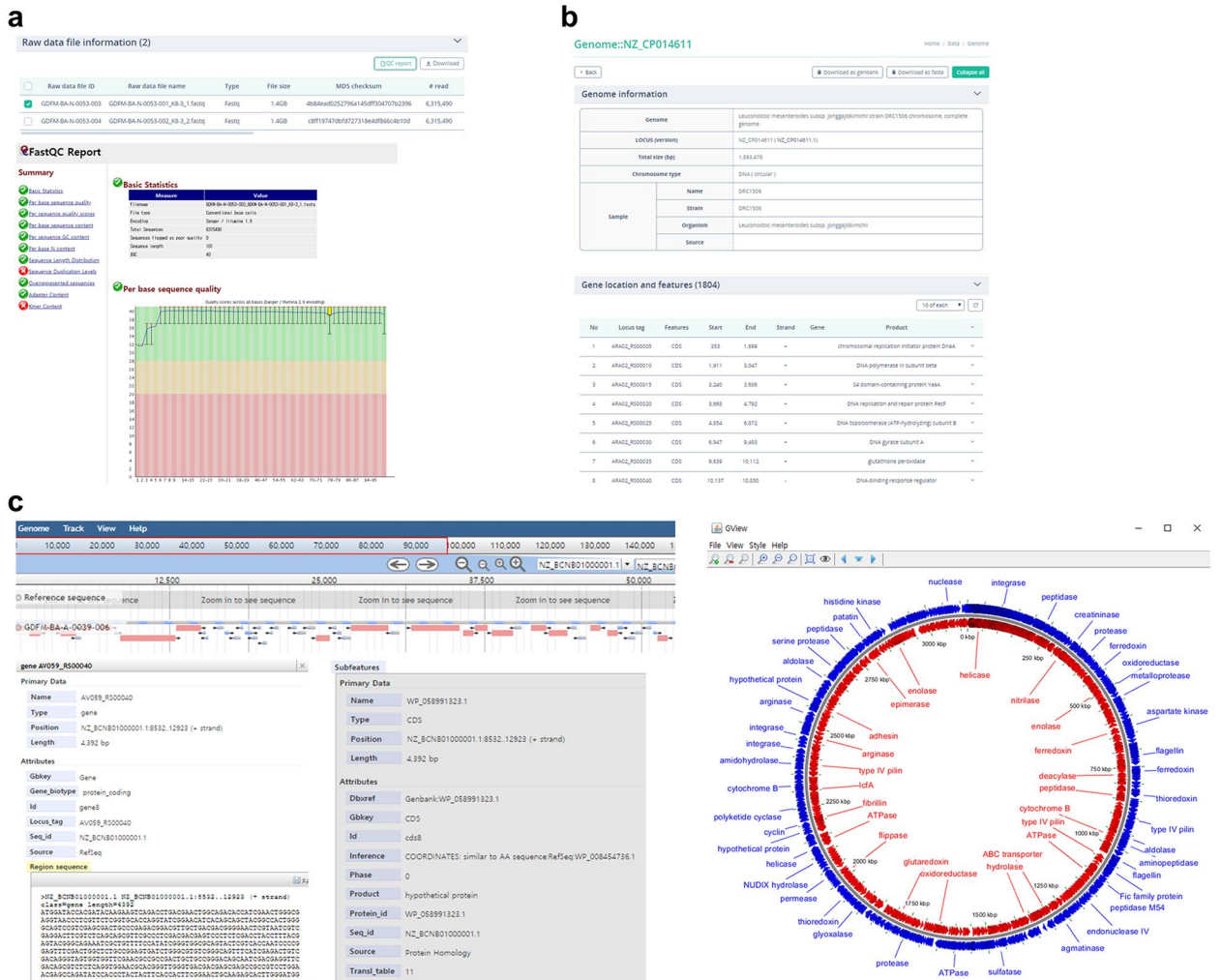
Fig. 2 Genome browsing function in the ODFM. (**a**) Screen image of raw data details comprising raw data information and QC information tabs. The raw data file information tab provides basic information on raw data. The QC report tab utilises FastQC to facilitate simple quality control checks of raw sequence data. (**b**) Details on annotation results, including location and product name of the gene/CDS in table format by using the GenBank (gbk) file format. (**c**) The ODFM converts the GFF file to graphical genome data with linear (left) and circular genome views (right).

(Supplementary Fig. S2a). Results are presented in a tabular format, with each row depicting a microbial taxon that contains the query gene. In the annotation data detail page, a split function in each row shows detailed information on the submitters, isolation sources, sequencing, and annotation results (Supplementary Fig. S2b). The annotation results presented in columns link to both available datasets and additional functions, such as genome viewer. For easy integration of omics data, the ODFM provides three basic search functions. Users can search omics data by taxonomy (i.e. bacteria, archaea, eukaryotic microorganisms, and viruses), study (i.e. publicly available studies highlighting the fermentative microbes), and sample (e.g. kimchi, fermented seafood, solar salt, soybean paste, vinegar, beer, cheese, sake, and yogurt) (Fig. 1b).

Genome browsing function. In the data search page, we provide several browsing functions for annotation data of the database resources. The raw data file information tab provides basic information on raw data (i.e. information regarding the experiment, library preparation, sequencing, and FASTQ file). The QC report tab utilises FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to facilitate simple quality control checks of raw sequence data (Fig. 2a). On the sequence information tab, the browser provides basic (i.e. FASTA) as well as detailed information about annotation data (Fig. 2b). The program uses the GenBank (gbk) file format and returns annotation results, such as location and product name of the gene/coding sequence (CDS). Furthermore, the ODFM provides linear and circular genome views representing graphical genome data based on GFF file information (Fig. 2c). The linear view utilises JBrowse[109], enabling users to browse local annotation results, while the circular view provides microbial genome visualisation in a circular context with an interactive pan and zoom interface using the GView JAVA package[110].
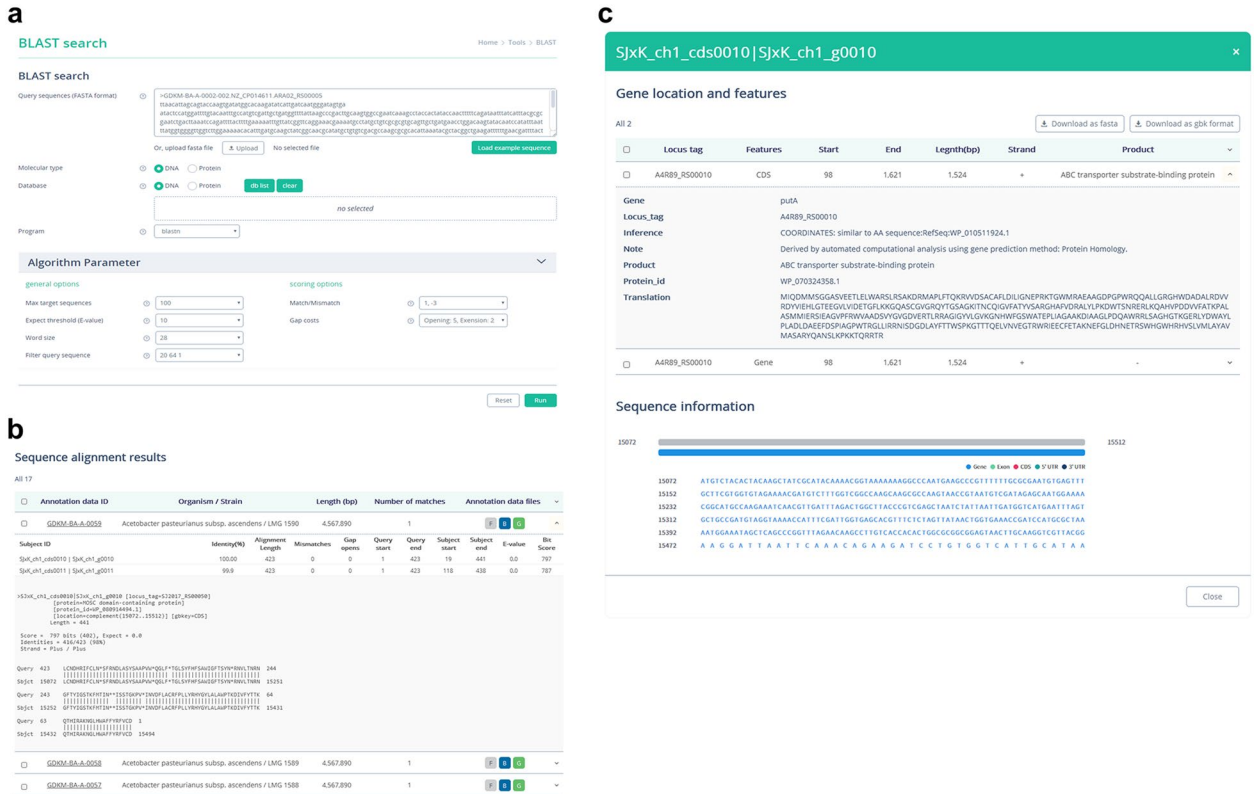
Fig. 3 BLAST search function in the ODFM (**a**) Screen image of the BLAST search tab. Users can provide nucleotide/amino acid sequences, and select a BLAST program (blastn, blastp, blastx, tblastnl, or tblastx), expected threshold, and filter query sequence (true or false). (**b,c**) The system returns alignment results in order of match (**b**), and provides sequence/CDS information for the annotated data as a downloadable text format (**c**).

Archive expansion.   We are currently seeking ways to expand the ODFM archive. Though the system currently does not allow data registration by other researchers, we are willing to accommodate deposition of fermented food-associated omics data generated by others upon request. In addition to the genome resources, the ODFM currently covers various outputs of genome annotation results generated from other databases, such as Kyoto Encyclopaedia of Genes and Genomes (KEGG)[111] and Clusters of Orthologous Groups of proteins (COGs)[112]. In our projects tab on the front page, a list of fermented food-associated microbial studies is provided (Supplementary Fig. S3).

Comparative genomics.   The microbial resources included in the ODFM are all candidate culture starters for fermented foods. Given that metabolic capabilities (e.g. lactate, lactose, and citrate metabolism) and resistance to bacteriophages, but not antibiotic resistance and virulence potential, are desirable functions for a candidate starter[113,114], preliminary screening by means of comparative genomics analysis between query and subject genome resources can be a practical way to select strains for the production of fermented foods. Accordingly, several open-source analytical tools for comparative genomics analysis are integrated in the ODFM.

Sequence similarity search against the ODFM.   A sequence similarity search against public nucleotide databases is generally the first step in the identification of microbial isolates. The ODFM provides the Basic Local Alignment Search Tool (BLAST)[115] search-based local alignment function. Once nucleotide/amino acid sequences are provided by users, the system aligns the query sequences with the local DNA/protein database (Fig. 3a). The system returns alignment results with statistical indicators, including bit score and E-value (Fig. 3b), and provides sequence/CDS information for the annotated data in a downloadable text format (Fig. 3c). Given that the ODFM comprises primarily fermentation-associated microbial genomes, this function is particularly helpful for an initial similarity search of strong candidate fermentative starter strains.

Genetic relatedness analysis.   For average nucleotide identity (ANI)-based genome clustering and genetic relatedness measurement, the dRep tool[116] was integrated into the ODFM. The ANI value is calculated from two genome sequences. Users can create comparative datasets by selecting FASTA files from the registered data in the ODFM or by uploading their own sequences (Fig. 4a). The query genome sequence (complete or draft) is cut into 1,020 bp-long sequences[117], and each fragment is annotated against the whole sequence of the subject genome. Using the MinHash distance to estimate similarity between genomes[118], the dRep tool calculates nucleotide identity between each of the query fragments and the subject genome and returns the ANI value, allowing for simple
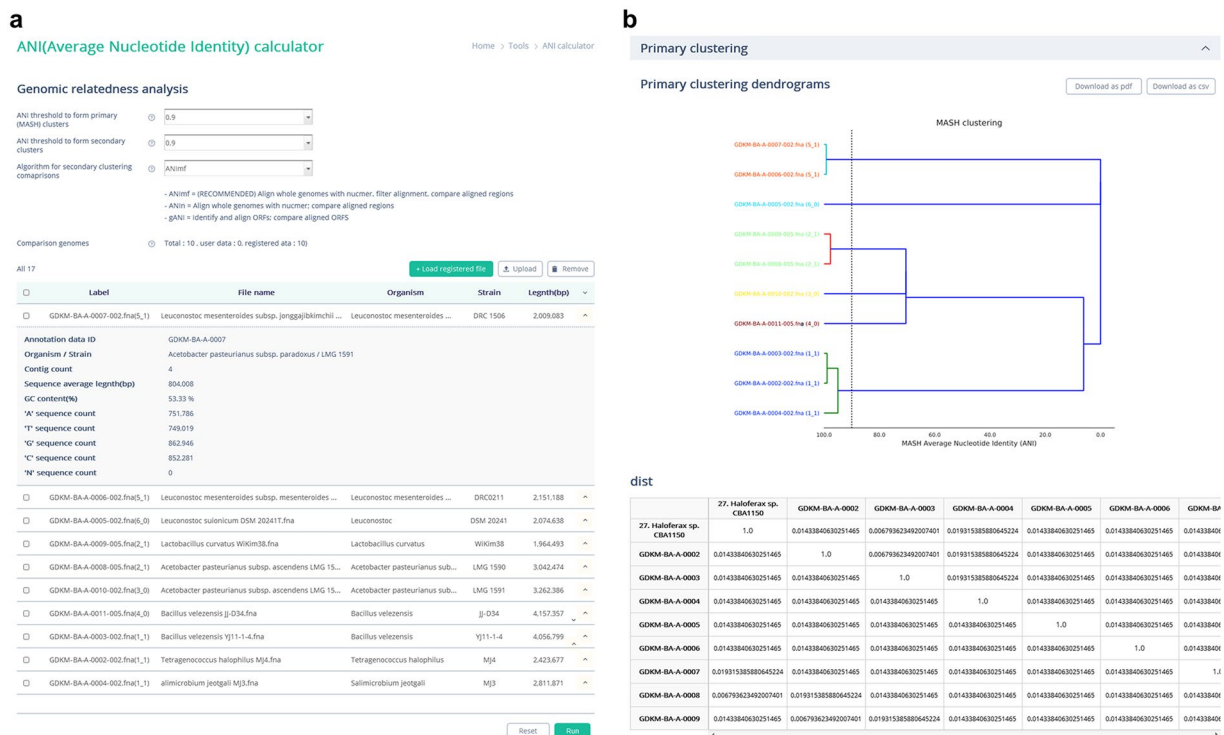
Fig. 4 ANI calculator function in the ODFM. (**a**) Screen image of the ANI calculator tab. Users can form comparative datasets by selecting FASTA files from the registered data in the ODFM or by uploading their own sequences. (**b**) The calculation results are returned as a downloadable table and image.

and standardised procedures for genome-related analysis of microbial isolates with closely related strains. The results are provided as a downloadable table and image (Fig. 4b).

Discussion

We developed the ODFM, a web-based knowledgebase featuring archival and analytical functions for genome data for bacteria, archaea, eukaryotic microorganisms, and viruses associated with fermented foods. The ODFM is freely available on the website https://odfm.wikim.re.kr. This easily accessible online-browsable resource facilitates rapid and functional explorations of genomes of fermentation-associated microorganisms. Genome sequence and annotation results for reference microorganisms, as well as analytical results, are directly download-able. All microbial strains registered in the ODFM will be made available. Our team operates a culture collection of fermented food-associated microorganisms, the Microorganism and Gene Bank (https://mgb.wikim.re.kr), at the World Institute of Kimchi.

By using the omics resources combined with the search tools, users are able to evaluate microbial strains isolated from fermented foods as candidate starters, and/or select microbial strain(s) among the deposited resources for use as starters. However, the process of fermentation is difficult to control because the fermentation phenotypes of different isolates are influenced by environmental conditions (e.g., temperature, humidity, and type of ingredients) and interactions with other, pre-colonised microbial communities. Previous studies have reported findings that allow linking of certain fermentative microbes with expected key features/metabolites in kimchi fermentation. *Leuconostoc* and *Lactobacillus* species are the major mannitol- and gamma-aminobutyric acid-producing LAB, respectively[119]. *Lactococcus* and some *Lactobacillus* species are homo-fermentative LAB responsible for the production of lactate from pyruvate by lactate dehydrogenase[120]. *Leuconostoc mesenteroides*, *Lactobacillus sakei*, and *Weissella koreensis* convert pyruvate to diacetyl/acetoin by using acetolactate synthase, acetolactate decarboxylase, and diacetyl reductase and thus contribute to the flavour of kimchi[15,20,121].

We expect the ODFM to provide a framework for the analysis of genome characteristics of microorganisms isolated from various fermented foods. To increase the usage of the data and information contained in the ODFM knowledgebase, we will continuously improve the features and performance of each function. By adding categories of fermented foods based on global consumer preferences and encompassing microbial resources, our long-term goal with the ODFM is to facilitate the genomic characterisation of food microorganisms and their application as fermentation starters, as well as further functional probiotics and biological agents.

Data availability

The ODFM is licensed under a Creative Commons Attribution 4.0 International License. The genome sequences in the ODFM are freely available on the website https://odfm.wikim.re.kr. All genome sequences have been deposited in NCBI GenBank and are available under the accession numbers listed in Supplementary Table S1. The viral metagenomic sequences have been deposited in the European Bioinformatics Institute (EMBL-EBI)

database[35] and are available under accession number PRJEB23957. We plan to deposit additional genome sequences for fermentative microbes to a member of the INSDC (http://www.insdc.org/) to promote sharing activities in the genomics community.

Code availability

The code used to build the systemic architecture of the GDKM is available on GitHub: https://github.com/yang4851/gdkm.

References

1. Jung, M. J. *et al*. Viral community predicts the geographical origin of fermented vegetable foods more precisely than bacterial community. *Food Microbiol* **76**, 319–327, https://doi.org/10.1016/j.fm.2018.06.010 (2018).
2. Kim, M. S. & Bae, J. W. Lysogeny is prevalent and widely distributed in the murine gut microbiota. *Isme J* **12**, 1127–1141, https://doi.org/10.1038/s41396-018-0061-9 (2018).
3. Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *Proc Natl Acad Sci USA* **115**, 6506–6511, https://doi.org/10.1073/pnas.1711842115 (2018).
4. Tamang, J. P., Watanabe, K. & Holzapfel, W. H. Review: Diversity of Microorganisms in Global Fermented Foods and Beverages. *Front Microbiol* **7**, 377, https://doi.org/10.3389/fmicb.2016.00377 (2016).
5. Patra, J. K., Das, G., Paramithiotis, S. & Shin, H. S. Kimchi and Other Widely Consumed Traditional Fermented Foods of Korea: A Review. *Front Microbiol* **7**, 1493, https://doi.org/10.3389/fmicb.2016.01493 (2016).
6. Tuomola, E., Crittenden, R., Playne, M., Isolauri, E. & Salminen, S. Quality assurance criteria for probiotic bacteria. *Am J Clin Nutr* **73**, 393S–398S, https://doi.org/10.1093/ajcn/73.2.393s (2001).
7. Jung, J. Y. *et al*. Metagenomic analysis of kimchi, a traditional Korean fermented food. *Appl Environ Microbiol* **77**, 2264–2274, https://doi.org/10.1128/AEM.02157-10 (2011).
8. Park, E. J. *et al*. Bacterial community analysis during fermentation of ten representative kinds of kimchi with barcoded pyrosequencing. *Food Microbiol* **30**, 197–204, https://doi.org/10.1016/j.fm.2011.10.011 (2012).
9. Lee, S. H., Jung, J. Y. & Jeon, C. O. Bacterial community dynamics and metabolite changes in myeolchi-aekjeot, a Korean traditional fermented fish sauce, during fermentation. *Int J Food Microbiol* **203**, 15–22, https://doi.org/10.1016/j.ijfoodmicro.2015.02.031 (2015).
10. Lee, S. H., Jung, M. Y., Song, J. H., Lee, M. & Chang, J. Y. Complete Genome Sequence of *Lactobacillus curvatus* Strain WiKim38 Isolated from Kimchi. *Genome Announc* **5**, https://doi.org/10.1128/genomeA.00273-17 (2017).
11. Jung, J. Y., Lee, S. H. & Jeon, C. O. Complete genome sequence of *Leuconostoc carnosum* strain JB16, isolated from kimchi. *J Bacteriol* **194**, 6672–6673, https://doi.org/10.1128/JB.01805-12 (2012).
12. Jung, J. Y., Lee, S. H. & Jeon, C. O. Complete genome sequence of *Leuconostoc gelidum* strain JB7, isolated from kimchi. *J Bacteriol* **194**, 6665, https://doi.org/10.1128/JB.01806-12 (2012).
13. Lee, S. H., Jung, J. Y., Lee, S. H. & Jeon, C. O. Complete genome sequence of *Leuconostoc kimchii* strain C2, isolated from Kimchi. *J Bacteriol* **193**, 5548, https://doi.org/10.1128/JB.05707-11 (2011).
14. Moon, J. S. *et al*. Genome sequence analysis of potential probiotic strain *Leuconostoc lactis* EFEL005 isolated from kimchi. *J Microbiol* **53**, 337–342, https://doi.org/10.1007/s12275-015-5090-8 (2015).
15. Chun, B. H., Kim, K. H., Jeon, H. H., Lee, S. H. & Jeon, C. O. Pan-genomic and transcriptomic analyses of *Leuconostoc mesenteroides* provide insights into its genomic and metabolic features and roles in kimchi fermentation. *Sci Rep* **7**, https://doi.org/10.1038/s41598-017-12016-z (2017).
16. Chun, B. H., Lee, S. H., Jeon, H. H., Kim, D. W. & Jeon, C. O. Complete genome sequence of *Leuconostoc suionicum* DSM 20241[T] provides insights into its functional and metabolic features. *Stand Genomic Sci* **12**, 38, https://doi.org/10.1186/s40793-017-0256-0 (2017).
17. Jia, B. *et al*. Complete Genome Sequences of Two Acetic Acid-Producing *Acetobacter pasteurianus* Strains (Subsp. *ascendens* LMG 1590[T] and Subsp. *paradoxus* LMG 1591[T]). *Front Bioeng Biotechnol* **5**, 33, https://doi.org/10.3389/fbioe.2017.00033 (2017).
18. Sung, J. S., Chun, J., Choi, S. & Park, W. Genome sequence of the halotolerant *Staphylococcus* sp. strain OJ82, isolated from Korean traditional salt-fermented seafood. *J Bacteriol* **194**, 6353–6354, https://doi.org/10.1128/JB.01653-12 (2012).
19. Lee, S. H., Jung, J. Y., Lee, S. H. & Jeon, C. O. Complete genome sequence of *Weissella koreensis* KACC 15510, isolated from kimchi. *J Bacteriol* **193**, 5534, https://doi.org/10.1128/JB.05704-11 (2011).
20. Jeong, S. E. *et al*. Genomic and metatranscriptomic analyses of *Weissella koreensis* reveal its metabolic and fermentative features during kimchi fermentation. *Food Microbiol* **76**, 1–10, https://doi.org/10.1016/j.fm.2018.04.003 (2018).
21. Jung, J., Chun, J. & Park, W. Genome sequence of extracellular-protease-producing *Alishewanella jeotgali* isolated from traditional Korean fermented seafood. *J Bacteriol* **194**, 2097, https://doi.org/10.1128/JB.00153-12 (2012).
22. Jung, W. Y., Jung, J. Y., Lee, H. J. & Jeon, C. O. Functional Characterization of Bacterial Communities Responsible for Fermentation of Doenjang: A Traditional Korean Fermented Soybean Paste. *Front Microbiol* **7**, 827, https://doi.org/10.3389/fmicb.2016.00827 (2016).
23. Lee, H. J., Chun, B. H., Jeon, H. H., Kim, Y. B. & Lee, S. H. Complete Genome Sequence of *Bacillus velezensis* YJ11-1-4, a Strain with Broad-Spectrum Antimicrobial Activity, Isolated from Traditional Korean Fermented Soybean Paste. *Genome Announc* **5**, https://doi.org/10.1128/genomeA.01352-17 (2017).
24. Lee, S. H., Jung, J. Y. & Jeon, C. O. Draft genome sequence of *Salimicrobium* sp. strain MJ3, isolated from Myulchi-Jeot, Korean fermented seafood. *J Bacteriol* **194**, 6695, https://doi.org/10.1128/JB.01808-12 (2012).
25. Yun, J. H. *et al*. Complete genome sequence of *Haloarcula* sp. CBA1115 isolated from non-purified solar salts. *Mar Genomics* **23**, 19–21, https://doi.org/10.1016/j.margen.2015.03.012 (2015).
26. Lee, H. W. *et al*. Draft genome sequence of *Halorubrum halophilum* B8[T], an extremely halophilic archaeon isolated from salt-fermented seafood. *Mar. Genomics* **18 Pt B**, 117–118, https://doi.org/10.1016/j.margen.2014.08.005 (2014).
27. Song, H. S. *et al*. Draft genome sequence of *Halapricum salinum* CBA1105[T], an extremely halophilic archaeon isolated from solar salt. *Mar Genomics* **18 Pt B**, 133–134, https://doi.org/10.1016/j.margen.2014.09.006 (2014).
28. Lim, S. K. *et al*. Genomic Analysis of the Extremely Halophilic Archaeon *Halobacterium noricense* CBA1132 Isolated from Solar Salt That Is an Essential Material for Fermented Foods. *J Microbiol Biotechnol* **26**, 1375–1382, https://doi.org/10.4014/jmb.1603.03010 (2016).
29. Lee, M. H. *et al*. Draft genome sequence of *Halolamina rubra* CSA1107[T], an agarolytic haloarchaeon isolated from solar salt. *Mar. Genomics* **18**, 127–128, https://doi.org/10.1016/j.margen.2014.09.003 (2014).
30. Lee, H. W. *et al*. Draft genome sequence of the extremely halophilic archaeon *Haladaptatus cibarius* type strain D43[T] isolated from fermented seafood. *Stand Genomic Sci* **10**, 53, https://doi.org/10.1186/s40793-015-0051-8 (2015).

31. Lee, M. H. *et al*. Draft genome sequence of the agarolytic haloarchaeon *Halobellus rufus* type strain CBA1103. *FEMS Microbiol Lett* **362**, 1–3, https://doi.org/10.1093/femsle/fnu005 (2015).

32. Cha, I. T. *et al*. Genome sequence of the haloarchaeon *Haloterrigena jeotgali* type strain A29^T isolated from salt-fermented food. *Stand Genomic Sci* **10**, 49, https://doi.org/10.1186/s40793-015-0047-4 (2015).

33. Kim, Y. B. *et al*. Novel haloarchaeon *Natrinema thermophila* having the highest growth temperature among haloarchaea with a large genome size. *Sci Rep* **8**, 7777, https://doi.org/10.1038/s41598-018-25887-7 (2018).

34. Kim, J. Y. *et al*. Community structures and genomic features of undesirable white colony-forming yeasts on fermented vegetables. *J Microbiol* **57**, 30–37, https://doi.org/10.1007/s12275-019-8487-y (2019).

35. *European Nucleotide Archive* https://identifiers.org/ebi/bioproject:PRJEB23957 (2020).

36. Jung, M. Y., Lee, S. H., Lee, M., Song, J. H. & Chang, J. Y. *Lactobacillus allii* sp. nov. isolated from scallion kimchi. *Int J Syst Evol Microbiol* **67**, 4936–4942, https://doi.org/10.1099/ijsem.0.002327 (2017).

37. Lim, H. I. *et al*. Draft Genome Sequence of *Lactobacillus sakei* Strain wikim 22, Isolated from Kimchi in Chungcheong Province, South Korea. *Genome Announc* **2**, https://doi.org/10.1128/genomeA.01296-14 (2014).

38. Jung, M. Y., Lee, C., Seo, M. J., Roh, S. W. & Lee, S. H. Characterization of a potential probiotic bacterium *Lactococcus raffinolactis* WiKim0068 isolated from fermented vegetable using genomic and *in vitro* analyses. *BMC Microbiol* **20**, 136, https://doi.org/10.1186/s12866-020-01820-9 (2020).

39. Chun, B. H., Kim, K. H., Jeon, H. H., Lee, S. H. & Jeon, C. O. Pan-genomic and transcriptomic analyses of *Leuconostoc mesenteroides* provide insights into its genomic and metabolic features and roles in kimchi fermentation. *Sci Rep* **7**, 11504, https://doi.org/10.1038/s41598-017-12016-z (2017).

40. Whon, T. W. *et al*. *Oceanobacillus kimchii* sp. nov. isolated from a traditional Korean fermented food. *J Microbiol* **48**, 862–866, https://doi.org/10.1007/s12275-010-0214-7 (2010).

41. Shin, N. R. *et al*. Genome sequence of *Corynebacterium nuruki* S6-4^T, isolated from alcohol fermentation starter. *J Bacteriol* **193**, 4257, https://doi.org/10.1128/JB.05354-11 (2011).

42. Kim, K. H. *et al*. *Acetobacter oryzifermentans* sp. nov., isolated from Korean traditional vinegar and reclassification of the type strains of *Acetobacter pasteurianus* subsp. *ascendens* (Henneberg 1898) and *Acetobacter pasteurianus* subsp. *paradoxus* (Frateur 1950) as *Acetobacter ascendens* sp. nov., comb. nov. *Syst Appl Microbiol* **41**, 324–332, https://doi.org/10.1016/j.syapm.2018.03.003 (2018).

43. Park, S. K., Roh, S. W., Whon, T. W. & Bae, J. W. Genome sequence of *Brachybacterium squillarum* M-6-3^T, isolated from salt-fermented seafood. *J Bacteriol* **193**, 6416–6417, https://doi.org/10.1128/JB.06183-11 (2011).

44. Kim, J., Roh, S. W. & Bae, J. W. Draft genome sequence of *Dietzia alimentaria* 72^T, belonging to the family Dietziaceae, isolated from a traditional Korean food. *J Bacteriol* **193**, 6791, https://doi.org/10.1128/JB.06229-11 (2011).

45. Jung, M. J., Roh, S. W., Kim, M. S., Whon, T. W. & Bae, J. W. Genome sequence of *Lentibacillus jeotgali* Grbi^T, isolated from traditional Korean salt-fermented seafood. *J Bacteriol* **193**, 6414–6415, https://doi.org/10.1128/JB.06139-11 (2011).

46. Kim, J. *et al*. *Paracoccus jeotgali* sp. nov., isolated from Korean salted and fermented shrimp. *J Microbiol* **57**, 444–449, https://doi.org/10.1007/s12275-019-8704-8 (2019).

47. Kim, K. H., Lee, S. H., Chun, B. H., Jeong, S. E. & Jeon, C. O. *Tetragenococcus halophilus* MJ4 as a starter culture for repressing biogenic amine (cadaverine) formation during saeu-jeot (salted shrimp) fermentation. *Food Microbiol* **82**, 465–473, https://doi.org/10.1016/j.fm.2019.02.017 (2019).

48. Cleenwerck, I., Vandemeulebroecke, K., Janssens, D. & Swings, J. Re-examination of the genus *Acetobacter*, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov. *Int J Syst Evol Microbiol* **52**, 1551–1558, https://doi.org/10.1099/00207713-52-5-1551 (2002).

49. Lisdiyanti, P. *et al*. Systematic study of the genus *Acetobacter* with descriptions of *Acetobacter indonesiensis* sp. nov., *Acetobacter tropicalis* sp. nov., *Acetobacter orleanensis* (Henneberg 1906) comb. nov., *Acetobacter lovaniensis* (Frateur 1950) comb. nov., and *Acetobacter estunensis* (Carr 1958) comb. nov. *J Gen Appl Microbiol* **46**, 147–165, https://doi.org/10.2323/jgam.46.147 (2000).

50. Huang, C. H., Chang, M. T., Huang, L. & Chua, W. S. Molecular discrimination and identification of *Acetobacter* genus based on the partial heat shock protein 60 gene (hsp60) sequences. *J Sci Food Agric* **94**, 213–218, https://doi.org/10.1002/jsfa.6231 (2014).

51. Tohno, M. *et al*. Description of *Lactobacillus iwatensis* sp. nov., isolated from orchardgrass (Dactylis glomerata L.) silage, and *Lactobacillus backii* sp. nov. *Int J Syst Evol Microbiol* **63**, 3854–3860, https://doi.org/10.1099/ijs.0.051920-0 (2013).

52. Fraunhofer, M. E. *et al*. Characterization of beta-glucan formation by *Lactobacillus brevis* TMW 1.2112 isolated from slimy spoiled beer. *Int J Biol Macromol* **107**, 874–881, https://doi.org/10.1016/j.ijbiomac.2017.09.063 (2018).

53. Long, G. Y. & Gu, C. T. *Lactobacillus jixianensis* sp. nov., *Lactobacillus baoqingensis* sp. nov., *Lactobacillus jiayinensis* sp. nov., *Lactobacillus zhaoyuanensis* sp. nov., *Lactobacillus lindianensis* sp. nov., *Lactobacillus huananensis* sp. nov., *Lactobacillus tangyuanensis* sp. nov., *Lactobacillus fuyuanensis* sp. nov., *Lactobacillus tongjiangensis* sp. nov., *Lactobacillus fujinensis* sp. nov. and *Lactobacillus mulengensis* sp. nov., isolated from Chinese traditional pickle. *Int J Syst Evol Microbiol* **69**, 2340–2353, https://doi.org/10.1099/ijsem.0.003474 (2019).

54. Geissler, A. J., Behr, J., von Kamp, K. & Vogel, R. F. Metabolic strategies of beer spoilage lactic acid bacteria in beer. *Int J Food Microbiol* **216**, 60–68, https://doi.org/10.1016/j.ijfoodmicro.2015.08.016 (2016).

55. Bergsveinson, J., Thomson, E., Jacoby, D., Coady, Y. & Ziola, B. Genome Sequence of *Megasphaera cerevisiae* NSB1, a Bacterium Isolated from a Canning Line and Able To Grow in Beer with High Alcohol Content. *Genome Announc* **5**, https://doi.org/10.1128/genomeA.01686-16 (2017).

56. Kramer, T. *et al*. Comparative genetic and physiological characterisation of *Pectinatus* species reveals shared tolerance to beer-associated stressors but halotolerance specific to pickle-associated strains. *Food Microbiol* **90**, 103462, https://doi.org/10.1016/j.fm.2020.103462 (2020).

57. Kern, C. C., Usbeck, J. C., Vogel, R. F. & Behr, J. Optimization of Matrix-Assisted-Laser-Desorption-Ionization-Time-Of-Flight Mass Spectrometry for the identification of bacterial contaminants in beverages. *J Microbiol Methods* **93**, 185–191, https://doi.org/10.1016/j.mimet.2013.03.012 (2013).

58. Kajala, I. *et al*. *Lactobacillus backii* and *Pediococcus damnosus* isolated from 170-year-old beer recovered from a shipwreck lack the metabolic activities required to grow in modern lager beer. *FEMS Microbiol Ecol* **94**, https://doi.org/10.1093/femsec/fix152 (2018).

59. Storari, M., Wuthrich, D., Bruggmann, R., Berthoud, H. & Arias-Roth, E. Draft Genome Sequences of *Clostridium tyrobutyricum* Strains FAM22552 and FAM22553, Isolated from Swiss Semihard Red-Smear Cheese. *Genome Announc* **3**, https://doi.org/10.1128/genomeA.00078-15 (2015).

60. Collins, M. D., Falsen, E., Akervall, E., Sjoden, B. & Alvarez, A. *Corynebacterium kroppenstedtii* sp. nov., a novel corynebacterium that does not contain mycolic acids. *Int J Syst Bacteriol* **48**, 1449–1454, https://doi.org/10.1099/00207713-48-4-1449 (1998).

61. Martino, G. P., Quintana, I. M., Espariz, M., Blancato, V. S. & Magni, C. Aroma compounds generation in citrate metabolism of *Enterococcus faecium*: Genetic characterization of type I citrate gene cluster. *Int J Food Microbiol* **218**, 27–37, https://doi.org/10.1016/j.ijfoodmicro.2015.11.004 (2016).

62. Bhakdi, S., Klonisch, T., Nuber, P. & Fischer, W. Stimulation of monokine production by lipoteichoic acids. *Infect Immun* **59**, 4614–4620, https://doi.org/10.1128/IAI.59.12.4614-4620.1991 (1991).

63. Bonham, K. S., Wolfe, B. E. & Dutton, R. J. Extensive horizontal gene transfer in cheese-associated bacteria. *Elife* **6**, https://doi.org/10.7554/eLife.22144 (2017).

64. Oguntoyinbo, F. A. *et al*. *Halomonas nigrificans* sp. nov., isolated from cheese. *Int J Syst Evol Microbiol* **68**, 371–376, https://doi.org/10.1099/ijsem.0.002515 (2018).

65. Zhuravleva, D. E. *et al*. Complete Genome Sequence of *Lactobacillus hilgardii* LMG 7934, Carrying the Gene Encoding for the Novel PII-Like Protein PotN. *Curr Microbiol* **77**, 3538–3545, https://doi.org/10.1007/s00284-020-02161-6 (2020).

66. Hynonen, U. *et al*. Functional characterization of probiotic surface layer protein-carrying *Lactobacillus amylovorus* strains. *BMC Microbiol* **14**, 199, https://doi.org/10.1186/1471-2180-14-199 (2014).

67. Pepper, S. J. & Britz, M. L. An Acid Up-Regulated Surface Protein of *Lactobacillus paracasei* Strain GCRL 46 is Phylogenetically Related to the Secreted Glucan- (GpbB) and Immunoglobulin-Binding (SibA) Protein of Pathogenic Streptococci. *Int J Mol Sci* **20**, https://doi.org/10.3390/ijms20071610 (2019).

68. Somerville, V. *et al*. Long-read based de novo assembly of low-complexity metagenome samples results in finished genomes and reveals insights into strain diversity and an active phage system. *BMC Microbiol* **19**, 143, https://doi.org/10.1186/s12866-019-1500-0 (2019).

69. Ennahar, S. *et al*. Production of pediocin AcH by *Lactobacillus plantarum* WHE 92 isolated from cheese. *Appl Environ Microbiol* **62**, 4381–4387, https://doi.org/10.1128/AEM.62.12.4381-4387.1996 (1996).

70. Tschoeke, D. A. *et al*. Exploring the Genome of Cheese Starter Lactic Acid Bacterium *Lactococcus lactis* subsp. *lactis* CECT 4433. *Genome Announc* **2**, https://doi.org/10.1128/genomeA.01142-14 (2014).

71. Frantzen, C. A. *et al*. Genomic Characterization of Dairy Associated *Leuconostoc* Species and Diversity of Leuconostocs in Undefined Mixed Mesophilic Starter Cultures. *Front Microbiol* **8**, 132, https://doi.org/10.3389/fmicb.2017.00132 (2017).

72. Makela, P., Schillinger, U., Korkeala, H. & Holzapfel, W. H. Classification of ropy slime-producing lactic acid bacteria based on DNA-DNA homology, and identification of *Lactobacillus sake* and *Leuconostoc amelibiosum* as dominant spoilage organisms in meat products. *Int J Food Microbiol* **16**, 167–172, https://doi.org/10.1016/0168-1605(92)90011-q (1992).

73. Guo, X. *et al*. Detection and Genomic Characterization of a *Morganella morganii* Isolate From China That Produces NDM-5. *Front Microbiol* **10**, 1156, https://doi.org/10.3389/fmicb.2019.01156 (2019).

74. Li, J. *et al*. *Mycetocola zhujimingii* sp. nov., isolated from faeces of Tibetan antelopes (Pantholops hodgsonii). *Int J Syst Evol Microbiol* **69**, 1117–1122, https://doi.org/10.1099/ijsem.0.003280 (2019).

75. Quintieri, L. *et al*. Biofilm and Pathogenesis-Related Proteins in the Foodborne *P. fluorescens* ITEM 17298 With Distinctive Phenotypes During Cold Storage. *Front Microbiol* **11**, 991, https://doi.org/10.3389/fmicb.2020.00991 (2020).

76. Quintieri, L., Caputo, L., De Angelis, M. & Fanelli, F. Genomic Analysis of Three Cheese-Borne *Pseudomonas lactis* with Biofilm and Spoilage-Associated Behavior. *Microorganisms* **8**, https://doi.org/10.3390/microorganisms8081208 (2020).

77. Bertuzzi, A. S. *et al*. Genome Sequence of *Staphylococcus saprophyticus* DPC5671, a Strain Isolated from Cheddar Cheese. *Genome Announc* **5**, https://doi.org/10.1128/genomeA.00193-17 (2017).

78. Wels, M. *et al*. Draft Genome Sequence of *Streptococcus thermophilus* C106, a Dairy Isolate from an Artisanal Cheese Produced in the Countryside of Ireland. *Genome Announc* **3**, https://doi.org/10.1128/genomeA.01377-15 (2015).

79. Toh, H. *et al*. Complete genome sequence of *Lactobacillus acetotolerans* RIB 9124 (NBRC 13120) isolated from putrefied (hiochi) Japanese sake. *J Biotechnol* **214**, 214–215, https://doi.org/10.1016/j.jbiotec.2015.09.006 (2015).

80. Gong, L. *et al*. A New Isolate of *Pediococcus pentosaceus* (SL001) With Antibacterial Activity Against Fish Pathogens and Potency in Facilitating the Immunity and Growth Performance of Grass Carps. *Front Microbiol* **10**, 1384, https://doi.org/10.3389/fmicb.2019.01384 (2019).

81. He, Q. *et al*. Comparative genomic analysis of *Enterococcus faecalis*: insights into their environmental adaptations. *BMC Genomics* **19**, 527, https://doi.org/10.1186/s12864-018-4887-3 (2018).

82. Iartchouk, O., Kozyavkin, S., Karamychev, V. & Slesarev, A. Complete Genome Sequence of *Lactobacillus acidophilus* FSI4, Isolated from Yogurt. *Genome Announc* **3**, https://doi.org/10.1128/genomeA.00166-15 (2015).

83. Laino, J. E., Hebert, E. M., Savoy de Giori, G. & LeBlanc, J. G. Draft Genome Sequence of *Lactobacillus delbrueckii* subsp. *bulgaricus* CRL871, a Folate-Producing Strain Isolated from a Northwestern Argentinian Yogurt. *Genome Announc* **3**, https://doi.org/10.1128/genomeA.00693-15 (2015).

84. Bai, Y. *et al*. Complete genome sequence of *Streptococcus thermophilus* MN-BM-A01, a strain with high exopolysaccharides production. *J Biotechnol* **224**, 45–46, https://doi.org/10.1016/j.jbiotec.2016.03.003 (2016).

85. Roh, S. W. *et al*. Complete genome sequence of *Halalkalicoccus jeotgali* B3^T, an extremely halophilic archaeon. *J Bacteriol* **192**, 4528–4529, https://doi.org/10.1128/JB.00663-10 (2010).

86. Yim, K. J. *et al*. Draft genome sequence of the extremely halophilic archaeon *Halococcus sediminicola* CBA1101^T isolated from a marine sediment sample. *Mar Genomics* **18**, 145–146, https://doi.org/10.1016/j.margen.2014.10.003 (2014).

87. Gallone, B. *et al*. Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts. *Cell* **166**, 1397–1410 e1316, https://doi.org/10.1016/j.cell.2016.08.020 (2016).

88. Salazar, A. N. *et al*. Chromosome level assembly and comparative genome analysis confirm lager-brewing yeasts originated from a single hybridization. *BMC Genomics* **20**, 916, https://doi.org/10.1186/s12864-019-6263-3 (2019).

89. Nguyen, H. V. & Gaillardin, C. Evolutionary relationships between the former species *Saccharomyces uvarum* and the hybrids *Saccharomyces bayanus* and *Saccharomyces pastorianus*; reinstatement of *Saccharomyces uvarum* (Beijerinck) as a distinct species. *FEMS Yeast Res* **5**, 471–483, https://doi.org/10.1016/j.femsyr.2004.12.004. (2005).

90. Rainieri, S., Kodama, Y., Nakao, Y., Pulvirenti, A. & Giudici, P. The inheritance of mtDNA in lager brewing strains. *FEMS Yeast Res* **8**, 586–596, https://doi.org/10.1111/j.1567-1364.2008.00363.x (2008).

91. Hewitt, S. K., Donaldson, I. J., Lovell, S. C. & Delneri, D. Sequencing and characterisation of rearrangements in three *S. pastorianus* strains reveals the presence of chimeric genes and gives evidence of breakpoint reuse. *PLoS One* **9**, e92203, https://doi.org/10.1371/journal.pone.0092203 (2014).

92. Tafer, H., Sterflinger, K. & Lopandic, K. Draft Genome Sequence of the Interspecies Hybrid *Saccharomyces pastorianus* Strain HA2560, Isolated from a Municipal Wastewater Treatment Plant. *Genome Announc* **6**, https://doi.org/10.1128/genomeA.00341-18 (2018).

93. Gibson, B. & Liti, G. *Saccharomyces pastorianus*: genomic insights inspiring innovation for industry. *Yeast* **32**, 17–27, https://doi.org/10.1002/yea.3033 (2015).

94. Bing, J., Han, P. J., Liu, W. Q., Wang, Q. M. & Bai, F. Y. Evidence for a Far East Asian origin of lager beer yeast. *Curr Biol* **24**, R380–381, https://doi.org/10.1016/j.cub.2014.04.031 (2014).

95. Jung, J. Y. *et al*. Effects of *Leuconostoc mesenteroides* starter cultures on microbial communities and metabolites during kimchi fermentation. *Int J Food Microbiol* **153**, 378–387, https://doi.org/10.1016/j.ijfoodmicro.2011.11.030 (2012).

96. Jeong, S. H., Jung, J. Y., Lee, S. H., Jin, H. M. & Jeon, C. O. Microbial succession and metabolite changes during fermentation of dongchimi, traditional Korean watery kimchi. *Int J Food Microbiol* **164**, 46–53, https://doi.org/10.1016/j.ijfoodmicro.2013.03.016 (2013).

97. Jeong, S. H. *et al*. Effects of red pepper powder on microbial communities and metabolites during kimchi fermentation. *Int J Food Microbiol* **160**, 252–259, https://doi.org/10.1016/j.ijfoodmicro.2012.10.015 (2013).

98. Jeong, S. H., Lee, S. H., Jung, J. Y., Choi, E. J. & Jeon, C. O. Microbial succession and metabolite changes during long-term storage of Kimchi. *J Food Sci* **78**, M763–769, https://doi.org/10.1111/1750-3841.12095 (2013).

99. Lee, S. H. & Jung, J. Y. & Jeon, C. O. Source Tracking and Succession of Kimchi Lactic Acid Bacteria during Fermentation. *J Food Sci* **80**, M1871–1877, https://doi.org/10.1111/1750-3841.12948 (2015).

100. Jung, M. Y. *et al*. Role of jeotgal, a Korean traditional fermented fish sauce, in microbial dynamics and metabolite profiles during kimchi fermentation. *Food Chem* **265**, 135–143, https://doi.org/10.1016/j.foodchem.2018.05.093 (2018).

101. Jung, J. Y. *et al.* Metatranscriptomic analysis of lactic acid bacterial gene expression during kimchi fermentation. *Int J Food Microbiol* **163**, 171–179, https://doi.org/10.1016/j.ijfoodmicro.2013.02.022 (2013).
102. Song, H. S. *et al.* Microbial niches in raw ingredients determine microbial community assembly during kimchi fermentation. *Food Chem* **318**, 126481, https://doi.org/10.1016/j.foodchem.2020.126481 (2020).
103. Jung, J. Y., Lee, S. H. & Jeon, C. O. Microbial community dynamics during fermentation of doenjang-meju, traditional Korean fermented soybean. *Int J Food Microbiol* **185**, 112–120, https://doi.org/10.1016/j.ijfoodmicro.2014.06.003 (2014).
104. Chun, B. H., Kim, K. H., Jeong, S. E. & Jeon, C. O. The effect of salt concentrations on the fermentation of doenjang, a traditional Korean fermented soybean paste. *Food Microbiol* **86**, 103329, https://doi.org/10.1016/j.fm.2019.103329 (2020).
105. Jung, J. Y., Lee, S. H., Lee, H. J. & Jeon, C. O. Microbial succession and metabolite changes during fermentation of saeu-jeot: traditional Korean salted seafood. *Food Microbiol* **34**, 360–368, https://doi.org/10.1016/j.fm.2013.01.009 (2013).
106. Lee, S. H., Jung, J. Y. & Jeon, C. O. Effects of temperature on microbial succession and metabolite change during saeu-jeot fermentation. *Food Microbiol* **38**, 16–25, https://doi.org/10.1016/j.fm.2013.08.004 (2014).
107. Lee, S. H., Jung, J. Y. & Jeon, C. O. Microbial successions and metabolite changes during fermentation of salted shrimp (saeu-jeot) with different salt concentrations. *PLoS One* **9**, e90115, https://doi.org/10.1371/journal.pone.0090115 (2014).
108. Jung, J. Y., Lee, H. J., Chun, B. H. & Jeon, C. O. Effects of Temperature on Bacterial Communities and Metabolites during Fermentation of Myeolchi-Aekjeot, a Traditional Korean Fermented Anchovy Sauce. *PLoS One* **11**, e0151351, https://doi.org/10.1371/journal.pone.0151351 (2016).
109. Buels, R. *et al.* JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol* **17**, 66, https://doi.org/10.1186/s13059-016-0924-1 (2016).
110. Petkau, A., Stuart-Edwards, M., Stothard, P. & Van Domselaar, G. Interactive microbial genome visualization with GView. *Bioinformatics* **26**, 3125–3126, https://doi.org/10.1093/bioinformatics/btq588 (2010).
111. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457–D462, https://doi.org/10.1093/nar/gkv1070 (2016).
112. Tatusov, R. L. *et al.* The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**, 41, https://doi.org/10.1186/1471-2105-4-41 (2003).
113. Kelleher, P., Murphy, J., Mahony, J. & van Sinderen, D. Next-generation sequencing as an approach to dairy starter selection. *Dairy Sci Technol* **95**, 545–568, https://doi.org/10.1007/s13594-015-0227-4 (2015).
114. Bennedsen, M., Stuer-Lauridsen, B., Danielsen, M. & Johansen, E. Screening for antimicrobial resistance genes and virulence factors via genome sequencing. *Appl Environ Microbiol* **77**, 2785–2787, https://doi.org/10.1128/AEM.02493-10 (2011).
115. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421, https://doi.org/10.1186/1471-2105-10-421 (2009).
116. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *Isme J* **11**, 2864–2868, https://doi.org/10.1038/ismej.2017.126 (2017).
117. Goris, J. *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**, 81–91, https://doi.org/10.1099/ijs.0.64483-0 (2007).
118. Ondov, B. D. *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* **17**, 132, https://doi.org/10.1186/s13059-016-0997-x (2016).
119. Jung, J. Y., Lee, S. H. & Jeon, C. O. Kimchi microflora: history, current status, and perspectives for industrial kimchi production. *Appl Microbiol Biotechnol* **98**, 2385–2393, https://doi.org/10.1007/s00253-014-5513-1 (2014).
120. Jia, B. *et al.* Catalytic, Computational, and Evolutionary Analysis of the d-Lactate Dehydrogenases Responsible for d-Lactic Acid Production in Lactic Acid Bacteria. *J Agric Food Chem* **66**, 8371–8381, https://doi.org/10.1021/acs.jafc.8b02454 (2018).
121. Kim, K. H. *et al.* Genomic and metabolic features of *Lactobacillus sakei* as revealed by its pan-genome and the metatranscriptome of kimchi fermentation. *Food Microbiol* **86**, 103341, https://doi.org/10.1016/j.fm.2019.103341 (2020).

Acknowledgements

Author contributions

S.W.R. and S.H.L. designed the study. T.W.W., S.W.A., S.Y., J.Y.K., Y.B.K., Y.K., J.-M.H., H.J., Y.-E.C., S.H.L., and S.W.R. conducted the system design and data processing. T.W.W., S.W.A., S.H.L., and S.W.R. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-021-00895-x.

Correspondence and requests for materials should be addressed to S.H.L. or S.W.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.