



ADDAGMA: A database for domestic animal gut microbiome atlas

Yueren Xu^{a,1}, Bingbing Lei^{a,1}, Qingfeng Zhang^{b,1}, Yunjiao Lei^a, Cunyuan Li^c, Xiaoyue Li^a, Rui Yao^a, Ruirui Hu^a, Kaiping Liu^a, Yue Wang^a, Yuying Cui^a, Limin Wang^d, Jihong Dai^a, Lei Li^a, Wei Ni^a, Ping Zhou^{d,*}, Ze-Xian Liu^{b,*}, Shengwei Hu^{a,*}



^a College of Life Sciences, Shihezi University, Shihezi, Xinjiang 832003, China

^b State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

^c College of Animal Science and Technology, Shihezi University, Shihezi, Xinjiang 832003, China

^d State Key Laboratory of Sheep Genetic Improvement and Healthy Production, Xinjiang Academy of Agricultural and Reclamation Sciences, Shihezi, Xinjiang 832003, China

ARTICLE INFO

Article history:

Received 14 October 2021

Received in revised form 8 February 2022

Accepted 8 February 2022

Available online 10 February 2022

Keywords:

Gut microbiota
Metagenomics
Domestic animals
Phenotypes
Database

ABSTRACT

Animal gut microbiomes play important roles in the health, diseases, and production of animal hosts. The volume of animal gut metagenomic data, including both 16S amplicon and metagenomic sequencing data, has been increasing exponentially in recent years, making it increasingly difficult for researchers to query, retrieve, and reanalyze experimental data and explore new hypotheses. We designed a database called the domestic animal gut microbiome atlas (ADDAGMA) to house all publicly available, high-throughput sequencing data for the gut microbiome in domestic animals. ADDAGMA enhances the availability and accessibility of the rapidly growing body of metagenomic data. We annotated microbial and metadata from four domestic animals (cattle, horse, pig, and chicken) from 356 published papers to construct a comprehensive database that is equipped with browse and search functions, enabling users to make customized, complicated, biologically relevant queries. Users can quickly and accurately obtain experimental information on sample types, conditions, and sequencing platforms, and experimental results including microbial relative abundances, microbial taxon-associated host phenotype, and P-values for gut microbes of interest. The current version of ADDAGMA includes 290,422 quantification events (changes in abundance) for 3215 microbial taxa associated with 48 phenotypes. ADDAGMA presently covers gut microbiota sequencing data from pig, cattle, horse, and chicken, but will be expanded to include other domestic animals. ADDAGMA is freely available at (<http://addagma.omicsbio.info/>).

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

A great number and variety of microorganisms colonize the mammalian intestine. The gastrointestinal microbiome is a consortium of multiple bacteria, archaea, fungi, and protozoa [1]. Increasing evidence has recently demonstrated that the gut microbiota is crucial to many processes in vertebrate hosts, including development [2,3], disease [4–9], immunization [10–13], and metabolism [14–17]. For example, Fuyong Li et al. [18] found significantly different populations of *Firmicutes* and *Chloroflexi* between high- and low-RFI (residual feed intake) animals. Chong Liu et al. [19] reported that the taxonomic composition of gut bacteria changes with age and is correlated with age-related changes in methane

emission. Within the context of constraining factors such as age and diet, the composition of microbial species early in life stages plays a crucial role in the dynamics of late-succession taxa [20]. A broad range of factors is responsible for differences in the intestinal microflora among hosts, including breed, diet, age, health status, environment, genetic background, and antibiotics [21–27].

Metagenomic sequencing technology, with its advantages of high throughput, low cost, and high speed, makes it possible to study gastrointestinal microbiomes solely by extracting DNA from a gut sample, without having to rely on the cultivation of microorganisms from the sample. The development of high-throughput sequencing and the rapid accumulation of genomic and 16S data from human and animal gut microbiota have given rise to various interrelated databases such as the European Nucleotide Archive (ENA) [28], National Center for Biotechnology Information (NCBI), Sequence Read Archive (SRA) [29], and DNA Data Bank of Japan (DDBJ) [30]. These databases store public sequencing results as

* Corresponding authors.

E-mail addresses: zhpqxqf@163.com (P. Zhou), liuzzx@sysucc.org.cn (Z.-X. Liu), hushengwei@163.com (S. Hu).

¹ These authors equally contributed to this work.

well as some biological information. Other comprehensive, well-annotated repositories for human gut metagenomic data, including gut MEtaGenome Atlas (gutMEGA), gutMDisorder, GMrepo, and HumanMetagenomeDB, allow researchers to access processed data and comparative information on intestinal microbiomes between different human disease conditions. These databases are pivotal research tools for elucidating relationships between human disorders and the gut microbiota [31–34].

A comprehensive database that can intuitively show the association between domestic animal traits and gut microbiota is currently lacking, and it remains difficult for scientists to query and retrieve certain types of published data. We created a novel database, the domestic animal gut microbiome atlas (ADDAGMA), which provides detailed meta-data such as sample type, sequencing platform, conditions of experimental, and analytical software. We re-analyzed the results of high-throughput sequencing of domestic animal microbiomes and provide such information as the log₂-transformed ratios, P-values, and microbial taxa. In addition, we categorize data in four ways according to the focus of the original studies (health traits, production traits, life-history traits, and microbial diversity). For convenience, we term these categories as “phenotypes”, as defined for the animalQTL [35] database (<https://www.animalgenome.org/cgi-bin/QTLdb/index>). With ADDAGMA, users can quickly obtain data on the gut microbiota associated with particular animals or animal traits by “browse for phenotype” or “browse for taxa”, and can compare their results with those from other, published studies. ADDAGMA thus provide a useful resource for exploring relationships between the phenotypes of domestic animals and their intestinal microbial biotas.

2. Materials and methods

2.1. Collection of raw data

To find potentially relevant papers, we searched the PubMed and Google Scholar databases through 30 September 2020 with the keywords “[animal] AND gut AND metagenome” and “[animal] AND gut AND microbiota”, with pig, cattle, horse, and chicken replacing [animal] in respective searches. Among more than 1300 papers detected, we then checked which papers documented 16S amplicon sequencing raw data or microbial relative-abundance data, which eventually yielded 356 publications. We divided the selected documents into two categories, those with and those without original sequencing data. For those with original data, we downloaded amplicon data for the bacterial 16S rRNA gene from the European Nucleotide Archive (ENA), and by using the command-line tools SRA Explorer and SRA-Tools, from the Sequence Read Archive (SRA), facilitated with Aspera (a high-speed data transfer tool). Other papers were selected which gave effective relative abundances for the intestinal flora, and the microbial information needed for further processing was obtained from the articles or supplementary materials. In the end, we selected 132, 102, 16, and 106 metagenomic, animal-gut studies for pig, cattle, horse, and chicken, respectively (Table 1). We subsequently recovered from these studies’ important fundamental

experimental information such as sequencing method, sample type, processing software, processing platform, and so forth.

2.2. Reprocessing and analysis of original sequence metadata

Different analytical methods may produce completely different analytical results, leading to poor consistency of these results. As a result, an identical standard was established to re-analyze the data, making the data in ADDAMAG more representative and scientific. We either reanalyzed the downloaded data using bioinformatics to detect abundance changes and bacterial p-values, or manually curated metadata from the literature.

We consistently reanalyzed original sequence data downloaded from the ENA database, using USEARCH (v10.0.240) [36] to denoise the data and VSEARCH (v2.15.1) [37] to merge data and assure quality control. Fig. 1 shows the schematic representations of the pipeline for amplicon data. We removed the barcodes and primers by using vsearch after the raw reads from paired-end sequences were merged, while the single-end sequences reads were used directly for subsequent analysis. And then we acquired the clean amplicon sequences following we carried out controlling the quality of sequences (controlling sequencing error rate ≤ 0.01) with the vsearch. The commands of vsearch we used were ‘--fastq_stripleft’ and ‘--fastq_maxee_rate’, respectively. The parameter of command (‘--fastq_stripleft’) was setted according to the collected meta-data values of primer length. We constructed a feature table by quantifying the frequency of features among the sequences in each sample after picking representative sequences and denoising. We simultaneously assigned taxonomic identifiers to the feature sequences, typically at the levels of kingdom, phylum, class, order, family, genus, and species, allowing a dimensionality-reduction perspective on the microbiota (mapping the sequences in the RDP database) [38]. From our analytical results, we obtained the taxonomy and relative abundances of microorganisms in the microbiota, to which we manually annotated biological information on the corresponding hosts taken from the studies examined. These data were included in the ADDAGMA. The studies in which we found data within the articles or in supplementary materials fell into three categories. Some provided the complete relative abundance or reads count, and we could determine fold-changes and calculate P-values. For other studies that provided the average (or median) abundance or fold-changes, and P-values, we calculated the fold-changes and used the P-values directly. Finally, some studies lacked only P-values.

3. Results

3.1. Contents of the database

We collected data from a total of 356 papers and further processed the meta-data as shown in Fig. 2. A total of 3,215 different microbial taxa and 290,422 quantitative events (changes in abundance) at six levels (phylum, class, order, family, genus, and species) were associated with 48 phenotypes. Among these 290,422 quantitative events, 26% were in cattle, 3% in horse, 51% in pig, and 20% in chicken. Table 1 shows the number of events related

Table 1

Number of events related to health traits, production traits, life history traits, and microbial diversity for four domestic animals, and the number of source papers.

Species	NO. of health traits	NO. of production traits	NO. of life history traits	NO. of microbial diversity	NO. of papers
Cattle	1,766	30,067	0	43,727	102
Pig	61,287	47,125	5,576	34,242	1,320
Horse	5,094	126	0	2,262	16
Chicken	8,289	40,175	0	9,686	106

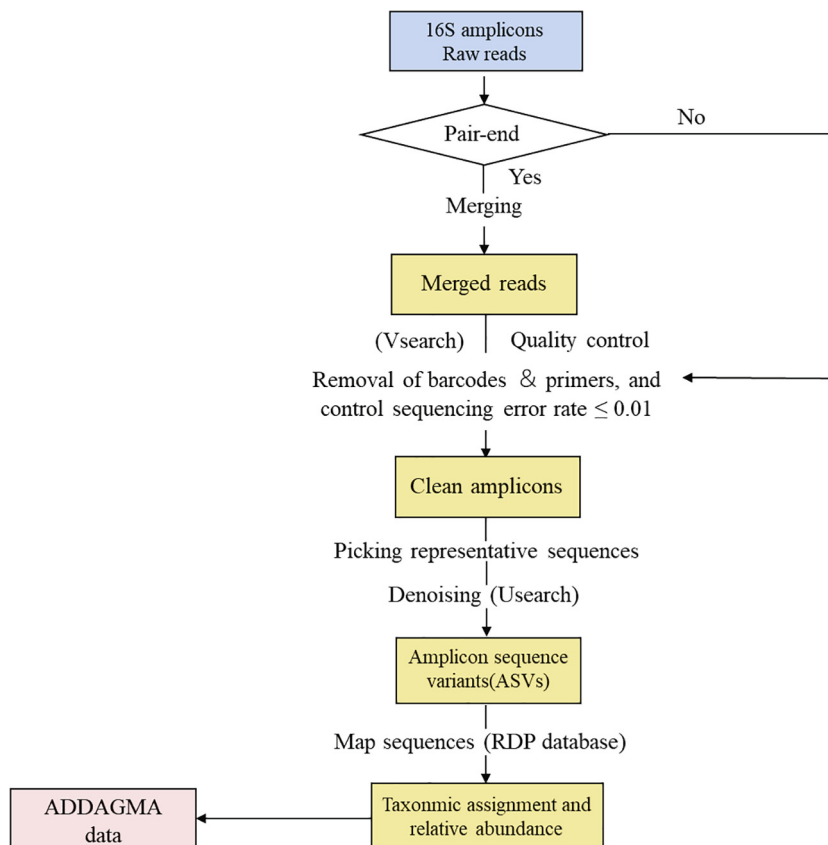


Fig. 1. Schematic representations of the pipeline for amplicon data.

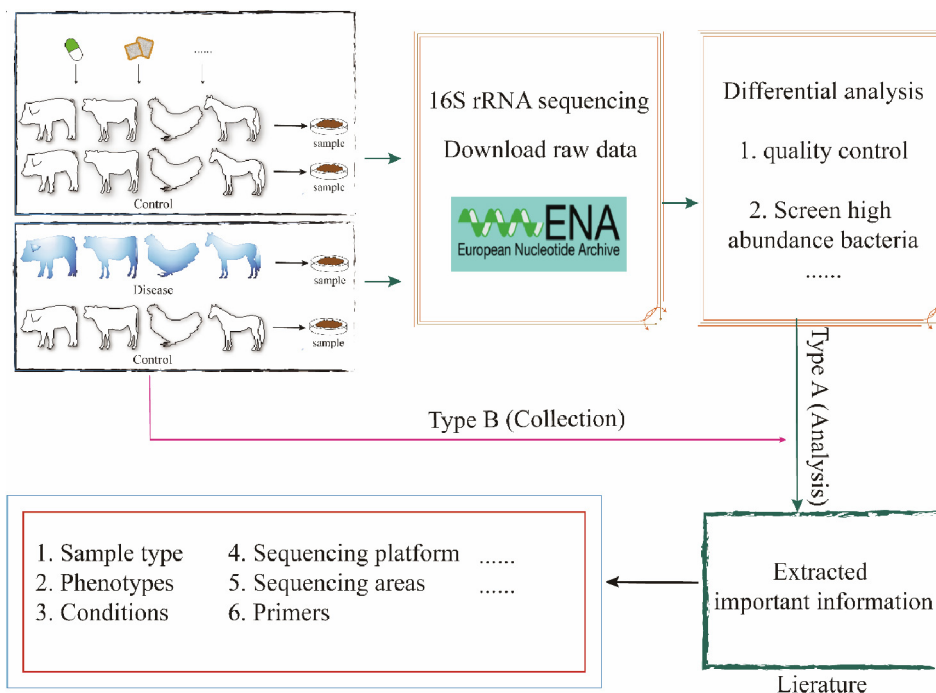


Fig. 2. Flow diagram for the process of data collection and construction of the database.

to health traits, production traits, life-history traits, and microbial diversity in the four species. The majority of studies on production traits were related to animal metabolism, whereas most of those

on health traits were related to animal diseases. Fig. 3A and B shows the number of gut microbial species involved in the four animals and the number of microbial species shared by them,

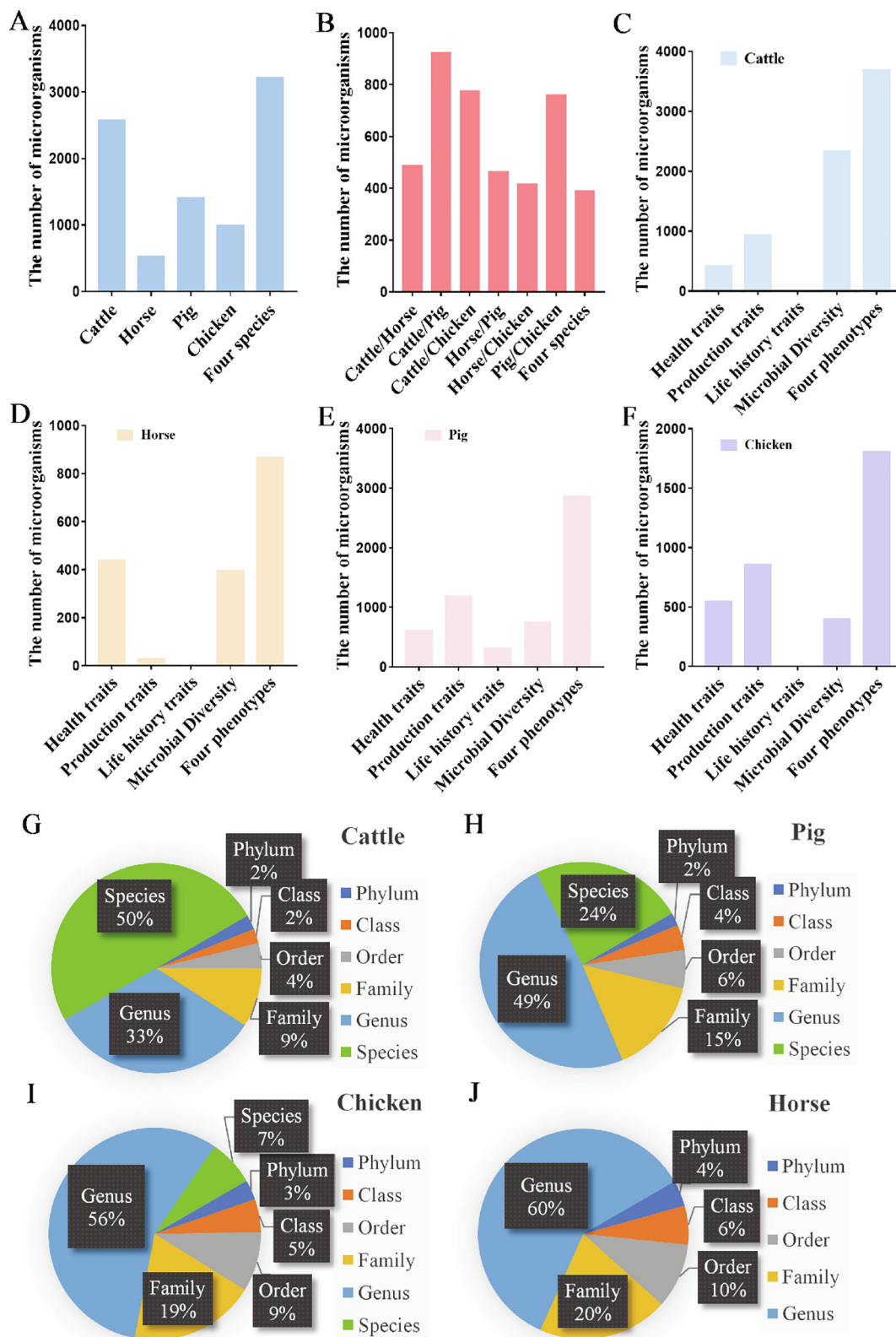


Fig. 3. Summary of the composition of the ADDAGMA database for domestic cattle, horse, pig, and chicken. (A) Number of guts microbial taxa detected in the four species. (B) Number of microbial taxa shared between pairs of the four species, and among all species. (C–F) Number of gut microbial taxa detected in studies related to animal phenotypes or microbial diversity in cattle, horse, pig, and chicken, respectively. (G–J) Proportions of gut microbial taxa identified to different taxonomic levels in cattle, horse, pig, and chicken, respectively.



Fig. 4. Query and search capabilities in ADDAGMA. (A) Quick search function on the home page. (B) Advanced search function on the search page. (C) Browser for phenotype on the browse page. (D) Browser for taxa on the browse page.

respectively. Fig. 3C–F shows the number of gut microbial species involved in the health traits, production traits, life history traits, and microbial diversity phenotypes of the four species, respectively. Fig. 3G–J shows the distribution of gut microbial in cattle, pig, horse, and chicken among different taxa.

3.2. Query function and result presentation for ADDAGMA

ADDAGMA is a search engine by which the database can be queried using restricting terms such as animal species, phenotypes, conditions, or microbial names, to obtain detailed information data on the relative abundance of gut microbes. It has three query methods, including quick search, advanced search, and browse (Fig. 4A, B). Users can enter keywords such as microorganisms, conditions, or phenotypes on the home page (Fig. 4A) for a quick search based on specific interests. Additionally, they can input multiple keywords (Fig. 4B) on the search page to obtain a more accurate result using an advanced search function. Users can easily obtain microbes species information, phenotypes, conditions, sample types, log2 conversion ratios, P-values, and so forth. The browse page (Fig. 4C) organizes the data under the root categories 'cattle', 'horse', 'pig', and 'chicken', with numerous sub-categories under the phenotypes 'health traits', 'production traits', 'microbial diver-

sity', and 'life history traits' for each animal species. By clicking 'health traits', 'production traits' or 'life history traits', the user can view specific traits for a specific animal species. 'Microbial diversity' displays all experiments that were conducted specifically to study microbial diversity. Finally, the browse menu for 'Taxa' provides a classification of microorganisms (Fig. 4D), with all microbial taxa arranged by their initials and labeled as to the taxonomic level to which they were identified, and allows the user to browse the microbial taxa in the ADDAGMA database.

Each query will display the following information: organism, resource, phenotype, taxon, condition, sample, log2 ratio, p-value, and details. Two contents, 'about dataset' and 'about quantification', can be accessed under 'Detail' (Fig. 5A–C). 'About dataset' mainly collates and displays experimental information such as sample type, PMID, phenotype, software, sequencing platform, and methods. 'About quantification' shows the raw data as a histogram.

3.3. Application of the database

To better understand how ADDAGMA can be used, we here provide an example. One report [39] indicated that the gut microbiota can contribute to bovine mastitis, and suppose we are interested in

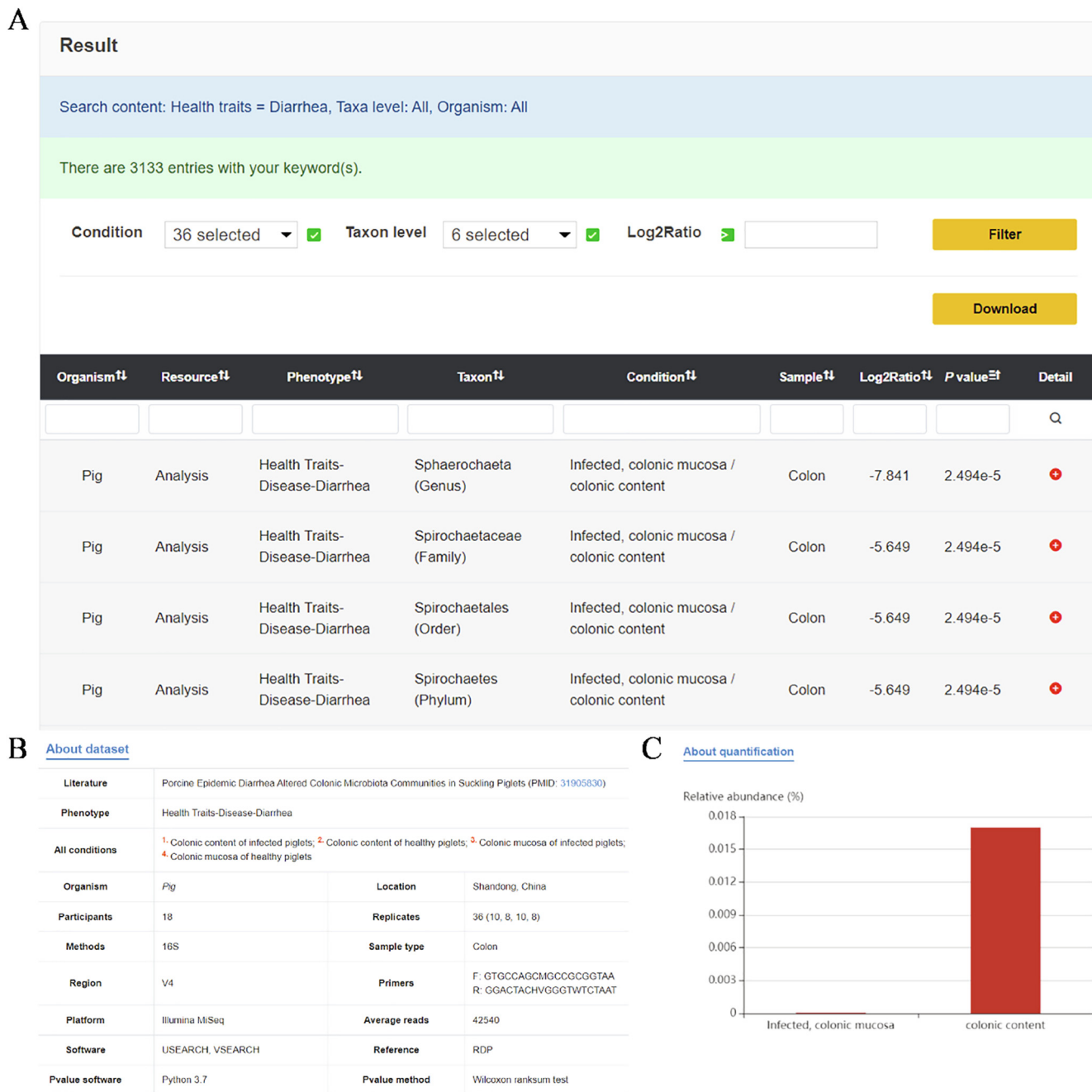


Fig. 5. Example of detailed results from a query in ADDAGMA. (A) Table on the Result page. (B) Summary of experimental information displayed via the “About dataset” option. (C) Histogram showing raw quantitative data, displayed via the “About quantification” option.

whether *Lactobacillus* plays critical roles in bovine mastitis. Using the advanced search function, with the keyword ‘cattle’ for organism, ‘*Lactobacillus*’ for taxon, and ‘mastitis’ for condition. The output result shows that the level of *Lactobacillus* was significantly higher in healthy controls than in animals with bovine mastitis in five datasets, suggesting that *Lactobacillus* plays a major role in bovine mastitis. In fact, *Lactobacillus gasseri* (LA806) has barrier and immunomodulatory effects and plays an active role in the resistance to breast infections caused by *Staphylococcus aureus* [40]. If we subsequently perform an experiment and find the abundance of *Lactobacillus* to be lower in animals with bovine mastitis, we can then use ADDAGMA to help validate the reliability of this result. For each study we incorporated, our database also provides all the information about the phenotypes associated with different

processing conditions and microbial changes, providing researchers with resources for exploration and reference.

4. Discussion

Intestinal microbes play a vital role in animal disease, health, and growth. Although the original sequencing data are stored in multiple databases, information on dynamic changes in the composition and abundance of animal gut microbiota is difficult to access. In this study, we constructed a database that involves the quantifications in the intestinal microbiota in domestic pig, cattle, horse, and chicken and provides associations between intestinal microbiota and various phenotypes. To access these data, we have

constructed a user-friendly website to display detailed results accurately and conveniently.

ADDAGMA is the first database to comprehensively collect intestinal microbial information related to phenotypes such as health and production performance in domestic animals. In general, it provides a tool for researchers interested in exploring the relationship between animal phenotypes and gut microbiota changes. The ADDAGMA database contains two search pages, two browse tables, and several filtering options for users.

There is still much work to be done to improve ADDAGMA as a practical and convenient metagenomic database for domestic animal intestinal microbiota. We plan to expand it to include other domestic animals such as sheep, camel, and rabbit, and to integrate metatranscriptomic, metaproteomic, and metametabolomic data sets into the database. Due to an increasing realization of the close interactions between animals and their gut microbes, the integration of microbial and animal meta-data is highly important to domestic animal research. However, ADDAGMA currently only includes data on bacteria, which is a slight limitation of ADDAGMA. We will update the literature and database regularly and add more data on microorganisms such as fungi, virus and archaea in later versions. As it improves as a comprehensive resource, ADDAGMA will provide intuitive data links for research on livestock intestinal microbes.

CRediT authorship contribution statement

Yueren Xu: Methodology, Investigation, Formal analysis, Data curation, Validation, Writing – original draft. **Bingbing Lei:** Methodology, Investigation, Formal analysis, Data curation, Validation, Writing – original draft. **Qingfeng Zhang:** Methodology, Software, Visualization, Data curation. **Yunjiao Lei:** Data curation. **Cunyuan Li:** Data curation. **Xiaoyue Li:** Data curation. **Rui Yao:** Data curation. **Ruirui Hu:** Data curation. **Kaiping Liu:** Data curation. **Yue Wang:** Data curation. **Yuying Cui:** Data curation. **Limin Wang:** Data curation. **Jihong Dai:** Data curation. **Lei Li:** Data curation. **Wei Ni:** Resources, Supervision. **Ping Zhou:** Funding acquisition, Supervision. **Ze-Xian Liu:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Shengwei Hu:** Conceptualization, Resources, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) (31660718 and U1803111), the Bingtuan Science and Technology Project (2018BC011, 2019AB034, and 2021CB033), Young Innovative Talents funding (2017CB003, CXRC201603 and CXRC201806), the foundation of the State Key Laboratory for Sheep Genetic Improvement and Healthy Production (2021ZD08), and Program for Guangdong Introducing Innovative and Entrepreneurial Teams (2017ZT07S096), Tip-Top Scientific and Technical Innovative Youth Talents of Guangdong Special Support Program (2019TQ05Y351).

References

- [1] Barko PC, McMichael MA, Swanson KS, Williams DA. The gastrointestinal microbiome: a review. *J Vet Intern Med* 2018;32(1):9–25.
- [2] Sampson T, Mazmanian S. Control of brain development, function, and behavior by the microbiome. *Cell Host Microbe* 2015;17(5):565–76.
- [3] Luczynski P, McVey Neufeld K-A, Oriach CS, Clarke G, Dinan TG, Cryan JF. Growing up in a bubble: using germ-free animals to assess the influence of the gut microbiota on brain and behavior. *Int J Neuropsychopharmacol* 2016;19(8):pyw020. <https://doi.org/10.1093/ijnp/pyw020>.
- [4] Dinan TG, Cryan JF. Gut instincts: microbiota as a key regulator of brain development, ageing and neurodegeneration. *J Physiol* 2017;595(2):489–503.
- [5] Argüello H, Estellé J, Leonard FC, Crispie F, Cotter PD, O'Sullivan O, et al. Influence of the Intestinal Microbiota on Colonization Resistance to *Salmonella* and the Shedding Pattern of Naturally Exposed Pigs. *mSystems* 2019;4:e000021–19.
- [6] Xu S, Lin Y, Zeng D, Zhou M, Zeng Y, Wang H, et al. *Bacillus licheniformis* normalize the ileum microbiota of chickens infected with necrotic enteritis. *Sci Rep* 2018;8(1). <https://doi.org/10.1038/s41598-018-20059-z>.
- [7] Morrison PK, Newbold CJ, Jones E, Worgan HJ, Grove-White DH, Dugdale AH, et al. The Equine Gastrointestinal Microbiome: Impacts of Age and Obesity. *Front Microbiol* 2018;9:3017.
- [8] Azad E, Derakhshani H, Forster RJ, Gruninger RJ, Acharya S, McAllister TA, et al. Characterization of the rumen and fecal microbiome in bloated and non-bloated cattle grazing alfalfa pastures and subjected to bloat prevention strategies. *Sci Rep* 2019;9(1). <https://doi.org/10.1038/s41598-019-41017-3>.
- [9] Bu D, Zhang X, Ma Lu, Park T, Wang L, Wang M, et al. Repeated Inoculation of Young Calves With Rumen Microbiota Does Not Significantly Modulate the Rumen Prokaryotic Microbiota Consistently but Decreases Diarrhea. *Front Microbiol* 2020;11:1403.
- [10] Rudin A, Lundell A-C. Infant B cell memory and gut bacterial colonization. *Gut Microbes* 2012;3(5):474–5.
- [11] Huang T, Gao B, Chen W-L, Xiang R, Yuan M-G, Xu Z-H, et al. Temporal effects of high fishmeal diet on gut microbiota and immune response in *Clostridium perfringens*-challenged chickens. *Front Microbiol* 2018;9:2754.
- [12] Lu J, Zhang X, Liu Y, Cao H, Han Q, Xie B, et al. Effect of fermented corn-soybean meal on serum immunity, the expression of genes related to gut immunity, gut microbiota, and bacterial metabolites in grower-finisher pigs. *Front Microbiol* 2019;10. <https://doi.org/10.3389/fmicb.2019.02620.10.3389/fmicb.2019.02620.10.3389/fmicb.2019.02620.s001.10.3389/fmicb.2019.02620.s002>.
- [13] Liu H, Zeng X, Zhang G, Hou C, Li N, Yu H, et al. Maternal milk and fecal microbes guide the spatiotemporal development of mucosa-associated microbiota and barrier function in the porcine neonatal gut. *BMC Biol* 2019;17(1). <https://doi.org/10.1186/s12915-019-0729-2>.
- [14] Zmora N, Suez J, Elinav E. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol* 2019;16(1):35–56.
- [15] Yang H, Huang X, Fang S, Xin W, Huang L, Chen C. Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct fitness. *Sci Rep* 2016;6:27427.
- [16] Che L, Hu Qi, Wang Ru, Zhang Du, Liu C, Zhang Y, et al. Inter-correlated gut microbiota and SCFAs changes upon antibiotics exposure links with rapid body-mass gain in weaned piglet model. *J Nutr Biochem* 2019;74:108246. <https://doi.org/10.1016/j.jnutbio.2019.108246>.
- [17] Sun Z, Yu Z, Wang B. *Perilla frutescens* leaf alters the rumen microbial community of lactating dairy cows. *Microorganisms* 2019;7(11):562. <https://doi.org/10.3390/microorganisms7110562>.
- [18] Li F, Hitch TCA, Chen Y, Creevey CJ, Guan LL. Comparative metagenomic and metatranscriptomic analyses reveal the breed effect on the rumen microbiome and its associations with feed efficiency in beef cattle. *Microbiome* 2019;7:6.
- [19] Liu C, Meng Q, Chen Y, Xu M, Shen M, Gao R, et al. Role of age-related shifts in rumen bacteria and methanogens in methane production in cattle. *Front Microbiol* 2017;8:1563.
- [20] Furman O, Shenhav L, Sasson G, Kokou F, Honig H, Jacoby S, et al. Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome assembly dynamics. *Nat Commun* 2020;11(1). <https://doi.org/10.1038/s41467-020-15652-8>.
- [21] Noel SJ, Olijhoek DW, Mclean F, Løvendahl P, Lund P, Højberg O. Rumen and fecal microbial community structure of holstein and jersey dairy cows as affected by breed, diet, and residual feed intake. *Animals* 2019;9(8):498. <https://doi.org/10.3390/ani9080498>.
- [22] Liu S, Zhao L, Zhai Z, Zhao W, Ding J, Dai R, et al. Porcine epidemic diarrhea virus infection induced the unbalance of gut microbiota in piglets. *Curr Microbiol* 2015;71(6):643–9.
- [23] Costa MC, Arroyo LG, Allen-Vercoe E, Stämpfli HR, Kim PT, Sturgeon A, et al. Comparison of the fecal microbiota of healthy horses and horses with colitis by high throughput sequencing of the V3–V5 region of the 16S rRNA gene. *PLoS ONE* 2012;7(7):e41484.
- [24] Pineda-Quiroga C, Borda-Molina D, Chaves-Moreno D, Ruiz R, Atxaerandio R, Camarinha-Silva A, et al. Microbial and Functional Profile of the Cecum from Laying Hens Affected by Feeding Prebiotics, Probiotics, and Synbiotics. *Microorganisms* 2019;7(5):123. <https://doi.org/10.3390/microorganisms7050123>.
- [25] Angelucci F, Cechova K, Amlerova J, Hort J. Antibiotics, gut microbiota, and Alzheimer's disease. *J Neuroinflammation* 2019;16:108.
- [26] Lu L, Luo T, Zhao Y, Cai C, Fu Z, Jin Y. Interaction between microplastics and microorganism as well as gut microbiota: A consideration on environmental animal and human health. *Sci Total Environ* 2019;667:94–100.
- [27] Ussar S, Griffin N, Bezy O, Fujisaka S, Vienberg S, Softic S, et al. Interactions between gut microbiota, host genetics and diet modulate the predisposition to obesity and metabolic syndrome. *Cell Metab* 2015;22(3):516–30.

- [28] Harrison PW, Alako B, Amid C, Cerdeño-Tárraga A, Cleland I, Holt S, et al. The European Nucleotide Archive in 2018. *Nucl Acids Res* 2019;47(D1):D84–8.
- [29] Kodama Y, Shumway M, Leinonen R. The Sequence Read Archive: explosive growth of sequencing data. *Nucl Acids Res* 2012;40(D1):D54–6.
- [30] Kodama Y, Mashima J, Kosuge T, Kaminuma E, Ogasawara O, Okubo K, et al. DNA Data Bank of Japan: 30th anniversary. *Nucl Acids Res* 2018;46(D1):D30–5.
- [31] Zhang Q, Yu K, Li S, Zhang X, Zhao Qi, Zhao X, et al. gutMEGA: a database of the human gut MEtaGenome Atlas. *Brief Bioinform* 2020;20:bbaa082.
- [32] Cheng L, Qi C, Zhuang He, Fu T, Zhang X. gutMDisorder: a comprehensive database for dysbiosis of the gut microbiota in disorders and interventions. *Nucl Acids Res* 2020;48(D1):D554–60.
- [33] Kasmanas JC, Bartholomäus A, Corrêa FB, Tal T, Jehmlich N, Herberth G, et al. HumanMetagenomeDB: a public repository of curated and standardized metadata for human metagenomes. *Nucleic Acids Res* 2021;49(D1):D743–50.
- [34] Wu S, Sun C, Li Y, Wang T, Jia L, Lai S, et al. GMrepo: a database of curated and consistently annotated human gut metagenomes. *Nucl Acids Res* 2020;48(D1):D545–53.
- [35] Hu Z-L, Fritz ER, Reecy JM. AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. *Nucl Acids Res* 2007;35(Database):D604–9.
- [36] Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26(19):2460–1.
- [37] Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 2016;4:e2584.
- [38] Liu Y-X, Qin Y, Chen T, Lu M, Qian X, Guo X, et al. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* 2021;12(5):315–30.
- [39] Ma T, Shen L, Wen Q, Lv R, Hou Q, Kwok LY, et al. PacBio sequencing revealed variation in the microbiota diversity, species richness and composition between milk collected from healthy and mastitis cows. *Microbiology* 2021;167(7). <https://doi.org/10.1099/mic.0.000968>.
- [40] Blanchet F, Rault L, Peton V, Le Loir Y, Blondeau C, Lenoir L, et al. Heat inactivation partially preserved barrier and immunomodulatory effects of *Lactobacillus gasseri* LA806 in an *in vitro* model of bovine mastitis. *Benef Microbes* 2021;12(1):95–106.