frontiers | Frontiers in Sustainable Food Systems

TYPE Original Research
PUBLISHED 29 April 2025
DOI 10.3389/fsufs.2025.1560970

OPEN ACCESS

EDITED BY
Pradeep K. Dubey,
International Rice Research Institute South Asia Regional Centre, India

REVIEWED BY
Viriato Silvio Isac Cossa,
Eduardo Mondlane University, Mozambique
Neha Chakrawarti,
International Rice Research Institute South Asia Regional Centre, India
Marina Padrão Temudo,
University of Lisbon, Portugal

*CORRESPONDENCE
Suoling Zhu
✉ zhusuoling@njau.edu.cn

RECEIVED 15 January 2025
ACCEPTED 26 March 2025
PUBLISHED 29 April 2025

CITATION
Zhao W, Zhu S and Cao Y (2025) Exploration of crop germplasm resources knowledge mining in Chinese ancient books: a route toward sustainable agriculture.
Front. Sustain. Food Syst. 9:1560970.
doi: 10.3389/fsufs.2025.1560970

Exploration of crop germplasm resources knowledge mining in Chinese ancient books: a route toward sustainable agriculture

Wenjuan Zhao[1], Suoling Zhu[1]* and Yuan Cao[2]

[1]Institution of Chinese Agricultural Civilization, College of Humanities and Social Development, Nanjing Agricultural University, Nanjing, China, [2]National Engineering and Technology Center for Information Agriculture, Engineering Research Center of Smart Agriculture, Ministry of Education, Key Laboratory for Crop System Analysis and Decision Making, Ministry of Agriculture, Nanjing Agricultural University, Nanjing, China

Crop germplasm resources, also known as genetic resources, include local varieties, bred varieties, wild resources and special genetic materials. In China, a large number of local varieties and wild resources have been recorded in ancient books. In this paper, Chinese ancient local chronicles, a unique kind of Chinese ancient books, is taken as an example to discover the knowledge of crop germplasm resources and explore their practical applications. Using the local chronicles of Jiangsu Province, China (1251–1947) as the data source, we constructed a spatio-temporal ontology of rice species according to the writing paradigm of descriptive texts on rice, intelligently extracted the features of rice with entity word lists and rule templates, and compared them with existing rice germplasm resources to analyze the similarities and differences. It is found that the descriptive texts on rice in Chinese ancient local chronicles not only contain the knowledge about the morphological characteristics of rice germplasm resources gained by field observation, but also involve the experience about the inherent attributes summarized by hands-on practice, such as the function, quality, phenology, customs, etc. It is a complement to the current collection and utilization of germplasm resources. Furthermore, some of the excellent varieties discovered still possess good trait stability after nearly a century of natural evolution and artificial selection, which provides a solid germplasm base for sustainable agricultural development. This study provides reference for the genetic breeding and insights for the knowledge discovery in the field of the history of agricultural science, and injects new vitality into the sustainable development of agriculture.

KEYWORDS

ancient books, crop germplasm resources, landrace rice, local chronicles, knowledge mining

1 Introduction

Crop germplasm resources play a pivotal role in agricultural production and are essential for ensuring food security, sustainable agricultural development, and environmental protection. The rational utilization and protection of these resources are vital for the genetic improvement of crops, cultivation of resistance, enhancement of quality, and selection and breeding of new varieties. Contemporary breeding techniques have succeeded with the support of genetic engineering and molecular biology. However, the long-term genetic basis of the superior germplasm increases the genetic vulnerability to unpredictable climatic and environmental conditions (McCouch et al., 2013), which seriously restricts the sustainable

development of agricultural production. It is difficult to ensure biogenetic diversity if new germplasm resources are not consistently and effectively introduced (Reif et al., 2005). Moreover, with changes in the growing environment and the functional needs of human beings for germplasm resources, such as health and nutrition, breeding experts are exploring suitable germplasms with both traditional characteristics and new functionalities (Kim et al., 2023; Yang et al., 2023; Langyan et al., 2022; Zsögön et al., 2017). Under this background, re-examining and utilizing the agricultural production experience in the long history has become an important research direction in the field of scientific breeding. For example, N. Fradgley et al. explored the future direction of wheat breeding through the influence of modern and historic wheat local varieties on root architecture (Fradgley et al., 2020); By determining the fatty acid composition of ancient and modern varieties of olive from southern Italy, Angela Cicatelli et al. (2013) found that the oil of the ancient Salella variety of Oria (Cilento NP) contains very high levels of human health beneficial polyunsaturated Omega-6 fatty acids. In conclusion, crop varieties in the historical period are a rich pool of genetic resources, whose potential needs to be further developed and utilized to meet the challenges in modern agriculture and to facilitate the process of crop improvement.

As witnesses and recorders of history, ancient books record the experiences and wisdom of human beings in planting, utilizing, and selecting varieties of crops, thus providing valuable information for us to understand ancient agricultural civilization. With the continuous development of science and technology and the increasing attention paid to agricultural genetic resources, more and more studies have begun to focus on the records of crop germplasm resources in ancient books, trying to find beneficial resources and inspiration for modern agricultural science.

Chinese local chronicles, also referred to as local gazetteers, difangzhi, or local records, are comprehensive encyclopedias that comprehensively record the natural, political, economic, cultural, and social history and current status of a certain administrative division. According to the statistics of Union Catalog of Chinese Local Chronicles, there are 8,264 kinds of local chronicles from the Song Dynasty to the Republic of China, with more than 110,000 volumes, accounting for approximately one-tenth of the ancient Chinese books (Lai, 2005). These historical materials of local chronicles not only record information on varieties, planting, quantity, and distribution of crops, but also describe the main traits of traditional germplasm resources, especially some important features such as high quality, high yield, and resistance to adversity and disease, etc. In recent years, scholars have already conducted researches on the organization of agricultural germplasm resources using the local chronicles. Yu (1993) manually consulted nearly 3,000 kinds of local chronicles to investigate the local variety of watermelon resources in the Ming Dynasty, Qing Dynasty, and the Republic of China. Chen (1992) and Lu (2014) used the historical materials of local chronicles to investigate the geographical distribution and cultivation technology of fungus. Others also studied the characteristics of rice (You, 1981), tea (Jiang, 2017; Wang, 1986), miscellaneous grains (Wang, 1997), fisheries (Ji, 2009), chili peppers (Ding and Hu, 2015), alfalfa (Sun et al., 2017), pears (Li and Bao, 2018) and other products based on the local chronicles. These findings indicate that Chinese local chronicles are helpful in the study of the history of local crops.

Such records about local crops in local chronicles are a repository of knowledge for tracing the evolution of traditional germplasm resources, which can create conditions for finding out the situation of traditional agricultural germplasm resources, discovering excellent varieties, and clarifying evolutionary trends. This paper chooses the local chronicles of the past dynasties in Jiangsu Province, which ranks first among China's major rice-producing provinces, as research materials, and takes the rice variety resources recorded in them as research objects. By summarizing and analyzing the characteristics of rice and its evolution patterns in local chronicles, we attempt to discover the application value of ancient books in the study of crop germplasm resources. Main Focus:

- The linguistic and content features of the description of crop germplasm resources in Chinese ancient books.
- Excellent varieties with strong resistance, good quality, and a wide range of uses mentioned in ancient books, especially the evolution patterns of these varieties and their relationship with existing local varieties.

2 Materials and methods

Jiangsu is located in the transition zone between the subtropical and warm-temperate zones, characterized by a dense water network, numerous lakes, and vast plains. It has been a major and high-yielding region for rice production in China since ancient times. Archeological excavations have revealed that the Dongshan Village site, located within Jiangsu, began primitive rice cultivation as early as 8,000 years ago (Wang and Ding, 1999). Jiangsu has produced a large number of local rice varieties that are well-adapted to local natural conditions and possess excellent agronomic traits, as evidenced by the rich ancient written records.

2.1 Materials

The contents on local products in local chronicles are particularly rich and unique to China. It systematically records the production of a certain product in a certain area, as well as important information related to the product, such as its names, characteristics, effects, customs, and culture. The data for this research mainly comes from the Institution of Chinese Agricultural Heritage. These data were collected from 1955 to 1963 by a dozen researchers in the field of agricultural history, who led more than 100 people to hand-copy and collate the product records in more than 8,000 local chronicles available in China at that time, totaling 36 million Chinese characters (Zeng, 2020). This is the most complete collection of product records in local chronicles that can be seen so far.

2.2 Methods

The method of our research contains four main steps, data acquisition and preprocessing, ontology construction, information extraction, and knowledge discovery. Detailed introductions are provided below.

2.2.1 Data acquisition and preprocessing

Data acquisition is the basic work of this study. The main task of this stage is to retrieve information on rice germplasm resources from numerous local chronicles as comprehensively and accurately as possible. However, collection work often encounters difficulties owing to factors such as regional dialects and the variety of Chinese character fonts. To solve this problem, we prioritized the screening of common rice varieties based on their names, classifications, and writing positions in the text, and extracted alternative names of rice based on rules for multiple rounds of searching. For example, the description of the variety "早红莲" (Zao Hong Lian) mentions that it is also called "救工饥" (Jiu Gong Ji) and "六十日" (Liu Shi Ri), which will also be searched as the variety name. In total, nearly 500 such variety names implicitly mentioned in the text descriptions were identified. The preprocessing stage is to clean the collected data and remove noise (irrelevant) and redundant (repetitive) information to provide a reliable database for the subsequent exploration of rice germplasm resources in local chronicles.

2.2.2 Ontology construction

The content regarding rice recorded in local chronicles covers various knowledge of rice. The Ontology builds a structured knowledge system for these recorded contents to describe the various attributes and relationships of rice germplasm resources. First, we conducted a comprehensive analysis and collation of the knowledge related to rice cultivation, utilization, culture, etc. Second, we refer to the existing specification for rice germplasm resource description,[1] and the Plant and Rice Germplasm Ontology [Rice Trait Ontology (RTO)][2] to ensure the generality and compatibility of constructed the ontology. Third, experts in the field of rice germplasm resources will be invited to further refine the relationship and hierarchy of knowledge. Fourth, the ontology model is used to present the writing paradigm of descriptive texts on rice, which encompasses four categories: crop (e.g., variety, rice type, etc.), geography (e.g., location, soil, etc.), literature (e.g., title, type, etc.), and culture (e.g., quotation, custom). Finally, through expert reviews and practical applications, the ontology structure was continuously revised and improved.

2.2.3 Information extraction

Based on the constructed ontology model, this study used two methods, entity word list-based and rule-based templates, to extract the multi-dimensional rice characteristics involved in local chronicles. The entity word list-based method aims to quickly and efficiently extract the proprietary vocabulary involved in rice description, including rice type (Japonica, Indica, and Glutinous), lowland or upland, season, maturity (early, medium, and late), and abiotic stress, etc. The experiment based on rule template is an extension of the previous method, focusing on the construction rules and lexical combinations of the information to be extracted when describing it, utilizing several trigger word vocabularies and word combinations to extract with regular expressions. For example, abiotic stress ("耐*", "抗*", "不畏*"), sowing period ("*月种", "*月初种", "*可种"), morphology ("*芒*", "*茎*", "*秆*", "*穗*", "*粒*"), and so on ("*" indicates the possible location of the rice features to be extracted in the text). Finally, the extracted information was processed and transformed into structured data, which will be classified, labeled and manually calibrated for subsequent analysis.

2.2.4 Knowledge discovery

Various data analysis techniques and visualization presentation methods are applied in this section to explore the implicit laws and unique features of crop germplasms in ancient books. Through statistical analysis, we identified and uncovered excellent ancient varieties with strong resistance, superior quality, and a wide range of uses, and matched these ancient varieties with modern germplasm databases. The differences and connections between the two were compared to verify the possibility of applying ancient data. The multidimensional presentation of rice characteristics in ancient books using visual charts and graphs can intuitively demonstrate various characteristics and trends. Simultaneously, combining the characteristics of crop germplasm resources recorded in Chinese books with the needs of modern agricultural development, we explored the potential application of ancient rice varieties in breeding improvement and ecological cultivation to further promote the sustainable development of agriculture.

3 Results

A total of 5,274 items of records of Jiangsu rice were collected from ancient local chronicles. Each one consists the names of rice variety, the time of the record, the region of the record, the classification to which the rice belongs, and the description of the variety. In terms of time, the earliest was recorded in Yu Feng Zhi in 1251 and the latest was in the Jin Shan Xian Jian in 1947. In terms of region, there were significantly more records and varieties of rice in southern Jiangsu than in northern, which was probably influenced by natural factors, such as soil, water resources, and climate, as well as the habit of compiling and editing local chronicles.

3.1 Characterization of rice germplasm resource descriptions in local chronicles

Ontology is a formal and detailed description of concepts within a domain using classes, relations, and attributes. In this way, we formalized the rice-related concepts and their relationships in local chronicles, and constructed an ontology model through the ontology editor Protégé. In view of the spatial and temporal nature of the crops recorded in ancient books, the constructed model was defined as the "spatio-temporal ontology model of local rice varieties," as shown in Figure 1. The model consists of four major classes: crop, geography, culture, and literature, which include 24 classes of rice in local chronicles. Among them, the three classes of crop, geography, and culture summarize the characteristics of the content described in rice, while the literature class describes the attributes of the ancient books themselves.

The class of crop contains eight sub-classes: variety, rice type, abiotic stress, phenological, morphological, agronomic, quality, and usage. The crop class and its sub-classes describe the general concepts of rice, such as the phenological period, botanical morphological, and

1 https://www.cgris.net/

2 https://cropontology.org/term/CO_320:ROOT

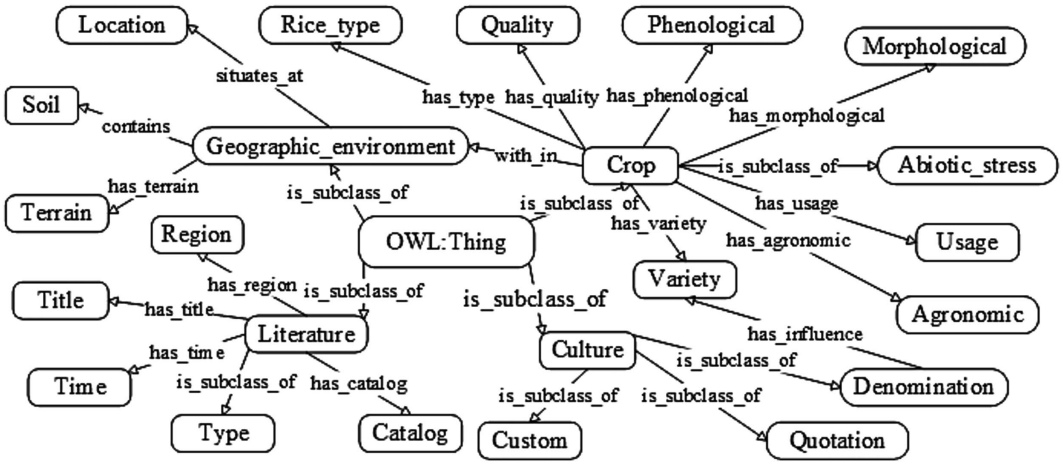

FIGURE 1
Spatio-temporal ontology model of local rice varieties.

yield traits. Among these, the classes of abiotic stress, morphological, and agronomic refer to crop ontology (Shrestha et al., 2010) and are adapted to the characteristics of rice descriptions in local chronicles.

The geography class is an abstract representation of geospatial elements such as geographic location, soil characteristics, and geologic terrain. It can provide clues for the location and discovery of high-quality rice species with three sub-classes: location, soil, and terrain.

The literature class describes the characteristics of the ancient books themselves, including five sub-classes: time, region, type, catalog and other contents of the recording media of rice. For example, the types of literature such as local chronicles, notebooks and novels, nautical diaries, agricultural books, etc., and catalogs such as the genera of cereals, rice, and food and drink.

The culture class covers cultural factors related to rice, including local customs, proverbs, events, and historical documents cited in the description. This class contains the experiences of ancient people in recognizing and utilizing rice varieties for production and life. It is a special historical material that showcases local rice-growing culture, and includes three sub-classes: quotation, custom, and denomination.

Based on the ontology model of local rice varieties constructed above, this study adopts the method of "machine recognition first, manual proofreading later" to extract the knowledge involved in the text of rice varieties, leveraging the efficiency of machine processing along with the interpretive strengths of human oversight. The various entities obtained are shown in Table 1.

As shown in Table 1, local chronicles mainly focus on documenting multi-dimensional crop characteristics of rice, followed by geography and culture. Descriptions of certain rice characteristics in ancient texts, like rice type, lowland or upland, and abiotic stress, show some connections with the way modern germplasm resources are described. For example, ancient Chinese people often distinguished rice types based on crop phenotypic characteristics and eating experience, such as whether it was sticky, whether it had awns, its size, and when it matured. According to local chronicles, rice is divided into three types: japonica rice, indica rice, and glutinous rice, while modern science divides China's cultivated rice into japonica and indica subspecies. It can be seen that there is obviously a certain connection between the ancient and modern classification systems. Moreover, ancient books recorded crop-related information including time of production, origin, denomination, customs, usage, endowing crop germplasm resources with unique spatio-temporal characteristics, cultural and customary backgrounds, as well as traditional naming and historical use. These ancient records have provided invaluable historical data and cultural perspectives for crop research.

3.2 Identification of ancient superior varieties and comparisons with current varieties

Precise identification of key traits in crop varieties provides a solid foundation for discovering superior germplasms. From the extracted characteristics of rice germplasm, several outstanding varieties among Jiangsu's local rice were noticed, characterized by strong stress resistance, high quality, and broad applicability. The following figures display rice varieties with abiotic stress (Figure 2), quality (Figure 3), and usage characteristics (Figure 4).

From Figures 2–4, we have the following findings:

- Many rice varieties have been recorded in local chronicles of Jiangsu Province which are resistant to various abiotic stress factors. According to statistics, rice varieties with drought tolerance are significantly more than those in other types. Furthermore, the same rice variety can have multiple stress resistances, such as "金城稻" (Jin Cheng Dao) and "散稻" (San Dao), which are both salt-tolerant and drought-tolerant. The stress resistance characteristics of these rice varieties mainly consist of two categories: positive characteristics (e.g., drought tolerance, cold tolerance, submergence tolerance) and negative characteristics (e.g., wind intolerance, submergence intolerance, and rain intolerance).
- The naming of varieties is related to their characteristics, for instance, rice varieties with a fragrant aroma often include the character"香" (xiang, meaning fragrant in Chinese) in their names, while the rice variety "小娘糯," (Xiao Niang Nuo, Xiao

Niang means a little weak girl) which is not resistant to wind, rain, and water, is named for its "delicate and fragile" nature.

TABLE 1 Information on the rice extracted.

Class I Crop 12,276		
◆ Variety 5,900		
• Variety Name 5,405	• Alternative Name 495	
◆ Rice Type 1,202		
• Biological 1,043 (Indica 138, Japonica 171, Glutinous 734)	• Lowland or Upland 33 (Lowland 10, Upland 23)	• Season 126 (Early 68, Middle 5, Late 53)
◆ Abiotic Stress 144		
• Submergence Tolerance 37	• Rain Tolerance 14	• Cold Tolerance 15
• Drought Tolerance 37	• Wind Tolerance 30	• Alkali or Salt Tolerance 11
◆ Usage 299		
• Flood/Drought Prevention 19	• Food Processing 219	• Pay Rent and Taxes 27
• Agricultural by-product 26	• Health Care 8	
◆ Agronomic 502		
• Seed set 106	• Plumpness 396	
◆ Quality 637		
• Aroma 222	• Softness 304	• Evaluation 111
◆ Phenological 1,368		
• Maturity 325 (Early 150, Middle 27, Late 138, Other 10)	• Time to Sowing 460	• Time to maturity 583
◆ Morphological 2,224		
• Grain 1,430 (Shape 621, Color 809)	• Awn 493 (Length 273, Color 220)	• Stem 162 (Length 28, Diameter 9, Softness 74, Color 51)
• Husk 69 (Shape 9, Color 60)	• Panicle 55	• Leaf 15
Class I Culture 420		
◆ Quotation 368	◆ Denomination 45	◆ Custom 7
Class I Geography 5,211		
◆ Location 5,145		
• Origin 34	• High Yield and Good Quality 60	• Other 5,051
◆ Soil 32		
• Soil Type 14	• Soil Property 18	
◆ Terrain 34		
• High Terrain 22	• Low Terrain 12	
Class I Literature 23,453		
◆ Title 5,243	◆ Time 5,243	◆ Region 5,243
◆ Type 5,243	◆ Catalog 2,481	

The number represents the number of extractions. ◆ indicates class II, and • indicates the attributes of class.

- As early as 600 years ago, people were already able to grade the quality of crop varieties into superior and inferior levels based on sensory experiences, such as aroma and taste, combined with cultivation knowledge and an understanding of the nutritional value of crops.
- The whole body of rice is a treasure. Not only is it edible, but its by-products such as rice bran and straw can be used to produce animal feed, straw ropes, etc. Additionally, rice has made great contributions to social stability, economic development, and agricultural production in ancient times in terms of alleviating famine, preparing for droughts and floods, and paying rent and taxes. For example, "不道糯" (Bu Dao Nuo), a kind of sticky rice, was often used to pay rent and taxes because of its characteristics of "easy to grow, high yield and low price." "红稻" (Red Dao), a rice which was often used as a replanting variety to prevent flooding because it can be planted late.

Do the local varieties recorded in ancient books still exist today? By examining the Jiangsu Local Germplasm Resources Information System (CGRIS, 1998),[3] we identified 41 rice varieties with the same name as mentioned in local chronicles. Among them, the seed origins of seven rice varieties recorded in CGRIS are very close to those recorded in local chronicles (within the same district or county). The existence of varieties with the same name represents the continuity of rice naming conventions over centuries. These rice varieties with the same name are highly likely to share certain genetic characteristics or cultivation traits. It reflects the role of naming habits in variety identification and classification. Although genetic differences between varieties with the same name cannot be ruled out, these contemporary varieties with the same names as those in ancient books are still worth studying. Because these varieties can still be planted in certain areas during the long history of farming, after natural selection and artificial preference selection. Such inheritance and continuity can provide strong evidence for agricultural conservation, genetic resource protection, and cultural heritage preservation. In order to investigate the possible utilization of crop germplasm resources recorded in ancient books in modern society, the data of variety names and origins matched and analyzed here are all from local chronicles and CGRIS.

Are these inherited varieties consistent with trait characteristics of historical period? By comparing ancient books with modern germplasm resource databases, we found that even after hundreds of years of evolution and cultivation practices, some key traits may remain their similarities. For instance, the most frequently documented rice variety, "六十日" (Liu Shi Ri), was widely cultivated in nine cities in Jiangsu Province, including Suzhou, Songjiang, Yangzhou, and Taicang. The description of "六十日" (Liu Shi Ri) in the local chronicles involves a variety of trait information such as maturity, alias, grain shape, grain color, Season. We counted the four categories with the most concentrated occurrences, alternative name, grain shape, grain color, time to sowing, and time to maturity, as shown in Figure 5. It is evident that the traits of these local varieties have changed relatively little over the centuries, which helps maintain the characteristics and stability of these varieties. We compared the information extracted from the local chronicles

3 https://www.cgris.net/cgris_english.html

10.3389/fsufs.2025.1560970

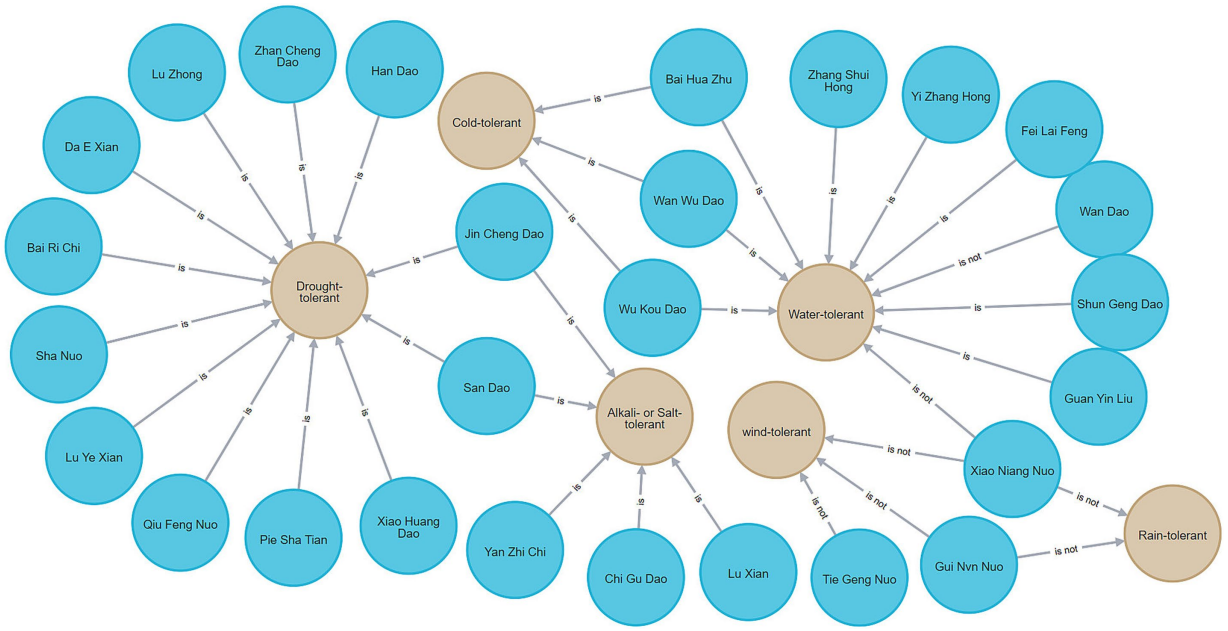

FIGURE 2
Rice varieties with resistance. The blue circle indicates the name of the rice variety. Yellow circles indicate the stress-resistant quality of the rice.

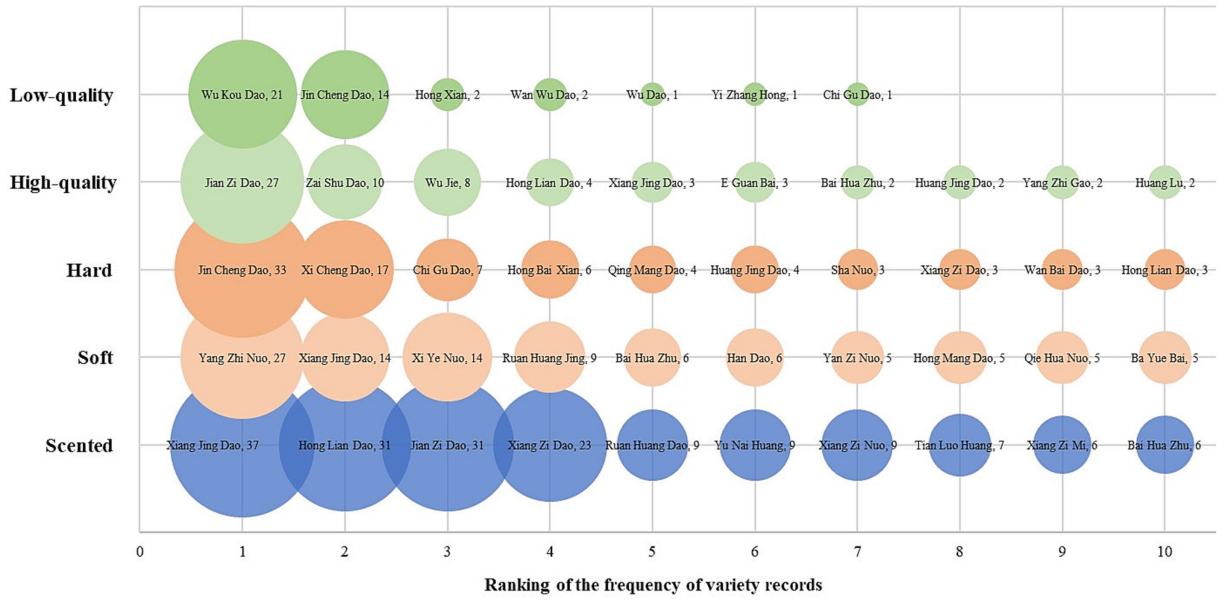

FIGURE 3
Quality characteristics and variety frequency statistics. The top 10 varieties of each quality characteristic. Among them, the softness feature is divided into hard and soft, and the evaluation feature is counted separately for high and low.

with two varieties of the same name retrieved from CGRIS and found that:

- The planting regions are similar. The two varieties in CGRIS come from Changshu County and Hai'an County, respectively. By comparing with the local chronicles, we found that the local chronicles of Changshu County in 1254, 1,499, and 1,539 had already recorded of "六十日" (Liu Shi Ri). As for Hai'an County, we have not been able to find its planting history records in ancient books. Since Hai'an County was established in 1948, which is later than the scope recorded in ancient books. Therefore, it is reasonable that the ancient books did not mention the planting in Hai'an County. However, according to the current location of the Hai'an area, the adjacent north, south and west (the east is the ocean) all have a history of planting "六十日" (Liu Shi Ri). We can speculate that it was likely cultivated in Hai'an.

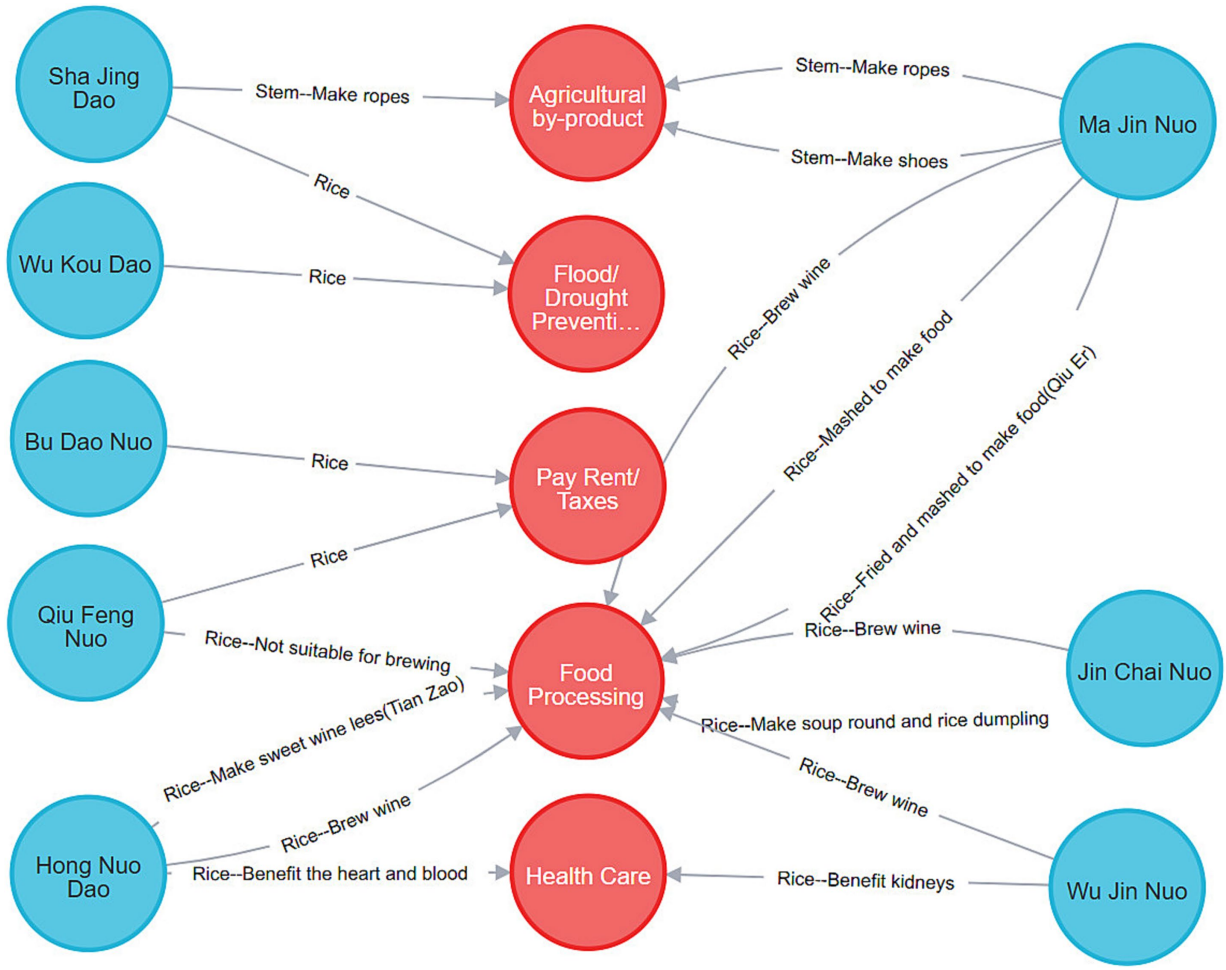

FIGURE 4
Widely used rice. Blue circles indicate rice variety names. The red circle indicates the usage of the rice.

- The traits are similar. "White color and small shape" was recorded many times in the local chronicles, as shown in Figure 5. "六十日" (Liu Shi Ri) in CGRIS has the characteristic of "white color." In addition, the CGRIS has two fields to describe the shape. One is the aspect ratio of the grain, which is divided into four grades (short round; broad ovoid; oval; thin). The other is the length of the grain, which is divided into four grades (short; medium; long; extra long). The grain shapes of the two varieties in CGRIS are "oval and medium" and "thin and long "respectively. Combining these two forms, we can reasonably infer that the grains of these two varieties are not full and medium in length, and are relatively small. It is consistent with the relevant descriptions in the local chronicles.

"六十日" (Liu Shi Ri) in CGRIS and ancient books show similar planting regions and comparable traits. We also compared other varieties in CGRIS. "旱稻" (Han Dao) and "芦叶籼" (Lu Ye Xian) are planted in Yangzhong County and Rugao County in central Jiangsu, respectively. The drought tolerance of these two varieties at the seedling stage is rated as the highest level 9. In the local chronicles, there are eight records between 1750 and 1934 documenting their drought tolerance, with cultivation areas concentrated in central and southern Jiangsu. In addition, "黄粳稻" (Huang Jing Dao) with full grains was collected in Wujin County. We counted two local rice varieties with the same name in the local chronicles between 1,547 and 1840. It has the characteristics of large grains and high quality, and is planted in Jiangyin County adjacent to Wujin County. The above examples all reflect the strong connection between CGRIS and local chronicles. The reliability of the content recorded in the local chronicles is confirmed. It also reflects that the germplasm characteristics of crops are relatively stable over a long period of time.

In the process of comparing ancient and modern records, we also discovered some rice varieties that were popular in ancient times, such as the "红莲稻" (Hong Lian Dao) with large, fragrant grains, the "箭子稻" (Jian Zi Dao) with high quality and fragrance grains, the "赶陈糯" (Gan Chen Nuo) suitable for brewing, and the "秋风糯" (Qiu Feng Nuo), which was commonly used for paying rent and taxes. No corresponding rice varieties were found in CGRIS. If we exclude potential omissions during the germplasm collection process, these varieties may only meet the needs of local social production and people's lives at that time. As societal demands evolved, such as the disappearance of the need for rent and tax payment in the 21st century, these varieties may no longer be cultivated. Of course, we cannot rule out the influence of factors like geographical

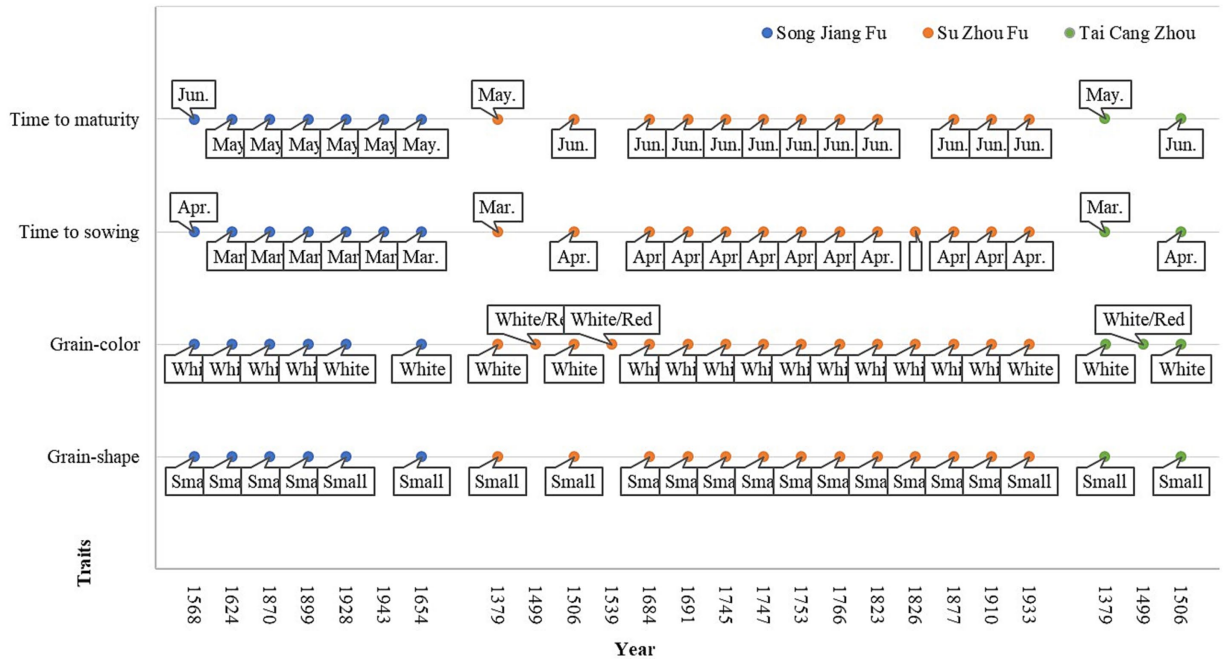

FIGURE 5
Description of the shape, color, time to sowing, and time to maturity of "六十日" (Liu Shi Ri) in different local chronicles. According to the regions recorded in the historical texts, the blue dots represent Songjiang Fu, the yellow dots represent Suzhou Fu, and the green dots represent Taicang Zhou.

environment, climate conditions, and consumer preferences on the survival of these rice varieties.

4 Discussion

The crop germplasm resources in ancient books are abundant and diverse. If properly utilized, they can provide valuable data for germplasm collection and utilization. Taking ancient local chronicles which exhibit remarkable temporal continuity and spatial clarity as the research literature, this study systematically reviewed the writing features of rice varieties recorded in ancient local chronicles of Jiangsu Province, China, during the period 1,251–1947, and constructed a spatio-temporal ontology of local rice varieties that integrates crop, geography, literature, and culture. Based on this, we extracted and analyzed the characteristics of the rice varieties and made several discoveries.

In addition to the characteristics of traditional crop germplasm resources (e.g., season, abiotic stress, morphological), ancient books also document a wealth of field observation experiences regarding crop usage, quality, phenological, custom, culture, etc. (see Figures 2–4). During the long historical period, the ancients distinguished rice varieties through their naked eyes or with the aid of simple tools during spring plowing, summer cultivation, autumn harvesting, and winter storage, providing a valuable supplement to the current collection of germplasm resources and crop breeding.

Many outstanding varieties have been recorded in ancient books. By comparing them with today's varieties, we found that some key traits of these varieties have been preserved, even after hundreds of years of evolution and planting practices. As shown in Figure 5, the variety of "六十日" (Liu Shi Ri) has just undergone minimal changes in grain shape and color, as well as sowing and harvesting times. The stability of traits in local varieties offers potential for the utilization of germplasm. Of course, when comparing ancient and modern times, we also found that many ancient varieties with high quality evaluations did not find corresponding rice seeds in CGRIS. It inspires us to preserve the local varieties with excellent quality in a timely manner, and record the names and characteristics of the local varieties at that time for subsequent use.

According to statistics, over 7 million crop germplasm resources have been preserved globally and are distributed across approximately 1750 germplasm repositories (FAO, 2010). Faced with such a vast germplasm resource bank, it is nearly impossible to comprehensively analyze the phenotypic traits, variety characteristics, and quality of each crop using traditional assessment methods within a short period (Zhao et al., 2019). Many wild relatives and local varieties documented in ancient texts have provided valuable clues. After long-term natural selection and artificial screening, these varieties still exhibit good ecological adaptability and genetic stability, and are the natural gene pools for crop breeding (Thanopoulos et al., 2024). In summary, through extraction and analysis of germplasm resource information in ancient books, combined with the application of modern technology, we can better protect and utilize these precious genetic resources, provide strong support for the sustainable development of agriculture, and help agriculture achieve high quality and sustainable development in the context of the new era.

The advent of modern technologies, such as artificial intelligence, have enabled efficient mining and utilization of ancient books. Through systematic organization and analysis of historical data, the experimental varieties required for breeding can be promptly

identified and screened. In addition, the literature records of crop germplasm resources provide a cultural background and historical context (Bar-Oz and Schmidt, 2025). In-depth studies of these documents are helpful to understand the track of ancient agricultural development and provide historical references for modern agricultural development, so as to develop and utilize germplasm resources more effectively.

The exploration we are engaged in at present only targets rice varieties recorded in ancient local chronicles from Jiangsu. We plan to expand our research to include other crops, regions, and types of literature such as agricultural books and poems. We will also collaborate closely with experts in the field of biology to develop a core seed bank of excellent varieties recorded in historical literature. Our future research will delve deeper into ancient books to uncover the rules governing crop germplasm resources and provide data support for global germplasm collection efforts, as well as clues for the discovery of distinctive rice varieties and enhance the cultural and historical significance of crop germplasm resources.

5 Conclusion

Crop germplasm resources are crucial to the sustainable development of agriculture. Chinese ancient books contain rich detailed information on the germplasm resources of local crops, and their potential is yet to be exploited. The continuity of landraces with the same name across historical and contemporary records, in terms of genetic traits and cultivation practices, reflects a manifestation of sustainable agriculture, as these varieties have been able to adapt to local environments and meet specific farmer needs over a considerable period of time. The genetic stability and ecological adaptability of crops provide valuable clues for the collection and utilization of germplasm resources. The similarity between current local varieties and the information recorded in ancient books in terms of planting regions and crop traits provides important reference values for geographical positioning and trait orientation for the crop germplasm resource census. Moreover, the content reliability of ancient books and the stability of crop germplasm traits play a guiding role in the practical application of germplasm resources, such as establishing in-situ gene banks and reintroducing uncultivated but preferred landraces.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

WZ: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Writing – review & editing. YC: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Research Project of the Jiangsu Provincial Social Science Application Research Elite Project "Research on mining and organizing information of local varieties of soybean resources in the era of AIGC in Local Chronicles" (24SYB-074).

Acknowledgments

We thank Caiyuan Wang, college of agriculture for the expert interpretation of the rice knowledge. We are especially grateful to Ping Bao for the insightful comments and materials used for experiments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Bar-Oz, G., and Schmidt, J. (2025). Why study the archeo-histories of dryland landraces now? *Trends Plant Sci.* doi: 10.1016/j.tplants.2025.01.004

CGRIS. (1998). Chinese crop germplasm resources information network. Available online at: https://www.cgris.net/cgris_english.html (Accessed May 20, 2024)

Chen, S. Y. (1992). The cultivation of edible mushrooms as seen from ancient China's local records. *Chinese J History Sci Technol* 3, 71–82.

Cicatelli, A., Fortunati, T., De Feis, I., and Castiglione, S. (2013). Oil composition and genetic biodiversity of ancient and new olive (*Olea europea* L.) varieties and accessions of southern Italy. *Plant Sci.* 210, 82–92. doi: 10.1016/j.plantsci.2013.05.011

Ding, X. L., and Hu, Y. Y. (2015). A study on the spread of chili peppers in terms of their Tricial names——focusing on Chinese historical local records. *J Chinese Histor Geography* 30, 104–117.

FAO (2010). The second report on the state of the world's plant genetic resources. Rome: FAO.

Fradgley, N., Evans, G., Biernaskie, J. M., Cockram, J., Marr, E. C., Oliver, A. G., et al. (2020). Effects of breeding history and crop management on the root architecture of wheat. *Plant Soil* 452, 587–600. doi: 10.1007/s11104-020-04585-2

Ji, L. Z. (2009). Study on the marine fishery resources of Shandong during the Ming and Qing dynasties in local chronicles. *J Ocean Univ China* 6, 43–46.

Jiang, X. T. (2017). Tea and life as seen in the Ming and Qing local Records of Wuyi Mountain Area. *Agric Archaeol* 5, 225–236.

Kim, S. H., Subramanian, P., and Hahn, B. S. (2023). Glucosinolate diversity analysis in choy sum (*Brassica rapa* subsp. chinensis var. parachinensis) germplasms for functional food breeding. *Food Secur.* 12:2400. doi: 10.3390/foods12122400

Lai, X. X. (2005). Chinese local histories: historical values and utilization. *J. Natl. Lib. China* 14, 5–8. doi: 10.3969/j.issn.1009-3125.2005.01.002

Langyan, S., Bhardwaj, R., Kumari, J., Jacob, S. R., Bisht, I. S., Pandravada, S. R., et al. (2022). Nutritional diversity in native germplasm of maize collected from three different fragile ecosystems of India. *Front. Nutr.* 9:812599. doi: 10.3389/fnut.2022.812599

Li, K. Q., and Bao, P. (2018). The regional distribution and cultivation types of pear in Ming and Qing dynasties based on "local chronicles products". *Agric. History China* 37, 21–29.

Lu, D. (2014). Production and documentation of edible and medicinal mushrooms in the Ming and Qing dynasties. *A Collection of Essays on the Ming and Qing Dynasties* 2, 262–290.

McCouch, S., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., et al. (2013). Feeding the future. *Nature* 499, 23–24. doi: 10.1038/499023a

Reif, J. C., Zhang, P., Dreisigacker, S., Warburton, M. L., van Ginkel, M., Hoisington, D., et al. (2005). Wheat genetic diversity trends during domestication and breeding. *Theor. Appl. Genet.* 110, 859–864. doi: 10.1007/s00122-004-1881-8

Shrestha, R., Arnaud, E., Mauleon, R., Senger, M., Davenport, G. F., Hancock, D., et al. (2010). Multifunctional crop trait ontology for breeders' data: field book, annotation, data discovery and semantic enrichment of the literature. *AoB plants* 2010:plq008. doi: 10.1093/aobpla/plq008

Sun, Q. Y., Liu, Q., Li, F., and Tao, Y. (2017). Records of alfalfa in local chronicles between the Ming and Qing dynasties. *Acta Pratacul. Sin.* 26, 176–188. doi: 10.11686/cyxb2017097

Thanopoulos, R., Drossinou, I., Koutroumpelas, I., Chatzigeorgiou, T., Stavrakaki, M., and Bebeli, P. J. (2024). Hilly, semi-mountainous and mountainous areas harbor landraces diversity: the case of Messinia (Peloponnese-Greece). *Diversity* 16:151. doi: 10.3390/d16030151

Wang, W. L. (1986). Exploration of ancient tea tree variety resources in China. *J. Tea Commun.* 1, 25–47.

Wang, S. J. (1997). The food grains other than wheat and Rice and their distributions in Jiangsu, Anhui, Zhejiang and Jiangxi provinces in the Ming dynasty. *Agric. History China* 3, 41–47.

Wang, C. L., and Ding, J. L. (1999). Ancient Rice crop research at the Dongshan Village site, Zhangjiagang, China. *Agric. Archaeol.* 3, 88–97.

Yang, X., Niu, Z., Wang, X., Lu, X., Sun, J., Carpena, M., et al. (2023). The nutritional and bioactive components, potential health function and comprehensive utilization of pomegranate: a review. *Food Rev. Intl.* 39, 6420–6446. doi: 10.1080/87559129.2022.2110260

You, X. L. (1981). Historical evidence of Rice variety resources in China. *Agric. Archaeol.* 2, 2–12.

Yu, W. J. (1993). A survey of watermelon quality resources in ancient China. *Agric. Archaeol.* 3, 204–210.

Zeng, J. J. (2020). Chinese agricultural heritage institute in 1950s and 1960s: review of searching-extracting agricultural history data of local chronicles. *Agric. Archaeol.* 4, 213–224. Available at: https://kns.cnki.net/kcms2/article/abstract?v=lwcs1eIaudjs5BO21Frm1EUtbSuUujw5VUVS0PQKet2mgU7DT2vCc6X8rkP3HT7Fr3HRwq1O2iD06ZN2FLH_uYkAi0owNaub19eT7KRjak6b-D6jl4UW4ZStI5zCEUegoZ74A3Buij56iANwiW_Nqd4vjXJ-753NZlh9-NoMSFiLcRl7UvcGNkrtDLd3a2_vgM09Jdybcww=&uniplatform=NZKPT&language=CHS

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop phenomics: current status and perspectives. *Front. Plant Sci.* 10:714. doi: 10.3389/fpls.2019.00714

Zsögön, A., Cermak, T., Voytas, D., and Peres, L. E. P. (2017). Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: case study in tomato. *Plant Sci.* 256, 120–130. doi: 10.1016/j.plantsci.2016.12.012