

引用格式: 郑倩. 基于文献计量学和机器学习的小麦生物育种文献分析[J]. 浙江农林大学学报, 2025, 42(1): 210-217.
ZHENG Qian. Literature analysis of *Triticum aestivum* bio-breeding based on bibliometrics and machine learning[J]. *Journal of Zhejiang A&F University*, 2025, 42(1): 210-217.

基于文献计量学和机器学习的小麦生物育种文献分析

郑倩

(华中农业大学 图书馆, 湖北 武汉 430070)

摘要: 【目的】分析全球小麦 *Triticum aestivum* 生物育种研究进展, 揭示其研究主题、核心知识元素和前沿热点, 为小麦育种的理论研究和学科发展提供参考。【方法】从 Web of Science 核心合集数据库检索 2013—2024 年全球小麦生物育种研究领域发表的文献, 利用文献计量学方法分析该领域的发文量、学术影响力、发文主体等, 利用 VOSviewer 生成国家合作网络和关键词共现网络, 利用机器学习算法潜在狄利克雷分配 (LDA) 分析文献摘要, 建立语言模型, 识别研究主题。【结果】2013 年以来全球小麦生物育种研究的科研产出大幅增加, 共发表文献 16 151 篇。中国是全球在该领域发文最多的国家, 其次是美国。关键词共现图谱显示, 产量、数量性状位点、全基因组关联分析、干旱胁迫、基因表达、单核苷酸多态性是小麦生物育种研究的热点, 而规律间隔成簇短回文重复序列 (CRISPR)、基因组编辑、高通量表型分析、无人机、机器学习等是近年来兴起的研究领域。LDA 分析结果显示: 小麦生物育种可大致分为 5 个研究领域, 包括遗传定位、基因组和育种、生物逆境、非生物逆境和产量形成。其中, 小麦锈病、数量性状位点定位、面粉品质、干旱、基因组等是高度关注的研究主题。【结论】未来小麦生物育种需要充分利用包括组学、自动表型、人工智能、基因编辑、基因组育种等现代生物技术和信息技术, 发掘和利用重要基因, 开展智慧育种。图 4 表 2 参 25

关键词: 小麦; 生物育种; 文献计量分析; VOSviewer; 潜在狄利克雷分配 (LDA); 机器学习

中图分类号: S337 **文献标志码:** A **文章编号:** 2095-0756(2025)01-0210-08

Literature analysis of *Triticum aestivum* bio-breeding based on bibliometrics and machine learning

ZHENG Qian

(Library of Huazhong Agricultural University, Wuhan 430070, Hubei, China)

Abstract: [Objective] This study aims to analyze the research progress of global wheat (*Triticum aestivum*) bio-breeding, and reveal its research themes, core knowledge elements and cutting-edge hotspots, so as to provide reference for theoretical research and discipline development in wheat breeding. [Method] Literature published in the field of global wheat breeding research from the Web of Science Core Collection Database from 2013 to 2024 was retrieved. Bibliometric analysis was applied to analyze the number of publications, academic influence, and main contributors in this field. VOSviewer was used to generate national collaboration networks and keyword co-occurrence networks. The Latent Dirichlet Allocation (LDA) machine learning algorithm was applied to analyze abstracts, establish language models, and identify research topics. [Result] Global research output in wheat biological breeding significantly increased since 2013, with a total of 16 151 publications. China was the country with the most publications, followed by the United States. Keyword co-occurrence analysis showed that yield, quantitative trait locus (QTL), genome-wide association studies (GWAS), drought stress, gene expression, and single nucleotide polymorphism (SNP) were hot topics in

收稿日期: 2024-08-10; 修回日期: 2024-10-29

基金项目: CALIS 全国农学文献信息中心研究项目 (2023028)

作者简介: 郑倩 (ORCID: 0000-0001-7846-1307), 馆员, 从事学科情报服务与研究。E-mail: qianzheng@mail.

hzau.edu.cn

research, while in recent years research fields such as clustered regularly interspaced short palindromic repeats (CRISPR), genome editing, high-throughput phenotyping, drones, and machine learning emerged. LDA analysis revealed that wheat biological breeding could be roughly divided into five research areas, including genetic mapping, genome and breeding, biotic stress, abiotic stress, and yield formation. Among them, wheat rust, QTL mapping, flour quality, drought, and genome were highly concerned. [Conclusion] In the future, wheat biological breeding needs to fully employ modern biotechnology and information technology, including omics, automated phenotyping, artificial intelligence, gene editing, and genome breeding, to identify and utilize key genes and facilitate intelligent breeding. [Ch, 4 fig. 2 tab. 25 ref.]

Key words: *Triticum aestivum*; bio-breeding; bibliometrics; VOSviewer; LDA; machine learning

小麦 *Triticum aestivum* 是世界范围内种植面积最广、产量最多的粮食作物之一。随着自然进化与人类科技进步,世界农业育种已经历了驯化育种、传统育种和分子育种 3 个时代^[1]。目前,仅依靠常规育种技术,包括小麦在内的主要作物的产量已经达到了相当高的水平,但仍然不能满足未来需求^[2]。小麦传统育种以杂交育种为中心,选育出了许多高产、稳产、高抗的优异品种。但传统育种周期长、效率低和遗传背景狭窄等问题日益突出^[3]。基因组学、系统生物学、合成生物学和人工智能等前沿学科的交叉融合,催生了融合遗传学、细胞生物学、现代生物工程技术 and 信息技术等现代生物育种技术^[4-5]。

文献计量学是以文献或文献相关媒介为研究对象,采用数学、统计学等计量方法,研究文献和文献工作系统的数量关系和规律,以及探讨科学技术动态特征的一门学科^[6]。在小麦研究领域,王瑞^[7]、GIRALDO 等^[8]和孟静等^[9]分别对 2018 年以前发表的小麦相关学术文献或专利开展了文献计量学分析,并探索了小麦研究的不同研究主题和研究前沿。2014 年国际小麦研究联盟发表了首个小麦基因组草图,2018 年发布了第 2 版小麦基因组序列^[10]。此后,人工智能、自动表型、基因组育种等现代生物技术得到了广泛应用,小麦生物育种进入了快速发展的时代^[11]。因此,了解这个阶段小麦生物育种的发展过程和特点,有助于预测小麦生物育种的未来研究方向,对保障粮食安全至关重要。近年来,随着人工智能的发展,利用机器学习方法处理大量文字信息,建立文本模型的自然语言处理(NLP)被应用到医学研究和社会科学等领域^[12-13],然而该方法在农业生物领域应用较少。

为全面了解小麦生物育种的发展态势和研究趋势,本研究基于 Web of Science 核心合集数据库检索小麦生物育种相关文献,结合文献计量学和机器学习方法,分析了国内外小麦生物育种相关的学术文献,包括发文的年代分布特点、主要发文国家、发文机构、研究热点与前沿、研究主题等,旨在揭示小麦生物育种领域的学术研究态势,为小麦育种的理论研究和学科发展提供参考。

1 材料与方法

1.1 数据来源

本研究文献数据来源于 Web of Science 核心合集数据库,引文索引数据库包括科学引文索引(SCIE)、社会科学引文索引(SSCI)、会议论文引文索引(CPCI-S、CPCI-SSH)和化学信息数据库(IC、CCR)。通过文献调研、专家咨询和多次试检索,本研究以转基因、基因组育种、基因编辑、合成生物学等生物育种关键技术在小麦中的应用为检索策略,检索年限为 2013—2024,检索得到的 16 151 篇文献作为本研究的数据来源。检索日期为 2024 年 8 月 31 日。

1.2 数据分析

分析国家、研究机构、作者的发文量和影响力时,使用 WoS 数据库平台中自带的分析功能,并结合 Excel 软件进行分析。由于本研究数据来源截止到 2024 年 8 月 31 日,数据库中 2024 年的相关文献仍在更新中,所有年度趋势仅显示 2013—2023 的分析结果。使用 VOSviewer 1.6.20 构建知识图谱。在进行国家合作网络分析时,国家最少发文数量设为 35 篇。在对作者关键词进行共现网络分析时,关键词最小频次设为 30 次,分析前对关键词进行清洗,合并同义关键词。

潜在狄利克雷分配(LDA)用于确定每篇文章中的研究主题^[14]。LDA 模型是一种非监督机器学习的

文本挖掘方法,能够从大规模的文档集中识别出潜藏的主题信息。LDA算法的主题网络分析突出了相互关联的主题集群共同出现的区域,并将相似的主题进行聚类,突出主题之间的关系。通过分析所有索引文章的摘要来对主题进行建模,并将确定的主题数量设置为50。根据算法计算出每个主题的概率,最终确定每篇文章所属的主题,并进行手动检查来命名每个主题。通过计算每个文档中不同主题同时出现的次数,建立1个主题互作网络,明确主题之间的联系。所有R和Python代码参考文献[14]。

2 结果与分析

2.1 年度发文分布

2013—2024年小麦生物育种领域发表的16 151篇文献的年度发文量显示(表1):从2013年开始每年发文量均在900篇以上,且呈逐年上升趋势,在2022年达到了1 931篇,2023年发文量略有下降。从每年的发文量增量来看,2014—2018年以每年33~107篇的速度稳步增加,2019年发文量出现跳跃式增长,比2018年增加了223篇。2019年后小麦生物育种发文量的快速增长可能与2018年发布第2版小麦基因组序列有关。虽然2023年后小麦生物育种发文量出现下降趋势,但小麦生物育种仍处于快速发展的重要阶段,是近年来的热门研究领域。

表1 小麦生物育种领域2013—2023年度发文量

Table 1 Annual publications of wheat bio-breeding during 2013—2023

出版年份	发文量/篇	出版年份	发文量/篇	出版年份	发文量/篇	出版年份	发文量/篇
2013	906	2016	1 113	2019	1 520	2022	1 931
2014	1 013	2017	1 217	2020	1 671	2023	1 681
2015	1 080	2018	1 297	2021	1 760		

2.2 主要发文国家

小麦是世界三大粮食作物之一。共有123个国家参与了小麦生物育种研究(图1),说明小麦生物育种研究是全球广泛关注的学科领域。在所有国家中,中国发文量最多(5 389篇),随后是美国(2 695篇)、澳大利亚(1 489篇)和印度(1 398篇)。中国发文量明显高于其他国家,在全球小麦生物育种研究领域具有竞争优势。在文献影响力方面,中国的总被引频次最高,美国次之,随后是澳大利亚、英国和德国。在篇均被引频次方面,挪威、瑞士、英国、德国、法国、荷兰等发达国家普遍高于发展中国家。中国发表论文数量远远超过其他国家,但篇均被引仅为20.5次,说明中国小麦生物育种研究在提高科研产出数量的同时,还需要进一步提高科研影响力。

本研究使用VOSviewer软件对小麦生物育种的主要发文国家进行合作网络分析(图2)。由图2可见:发文量较多的58个国家形成了5个不同合作集群,分别以中国和美国(紫色)、英国和德国(绿色)、土耳其和伊朗(红色)、巴基斯坦和埃及(蓝色)、澳大利亚和印度(黄色)为代表。其中,中国是发文最多的国家,但美国是国际合作最多的国家(链接数为58个,总链接强度为2 855)。与中国合作最多的5个国家是美国(链接强度为536)、澳大利亚(链接强度为375)、巴基斯坦(链接强度为271)、加拿大(链接强度为135)和墨西哥(链接强度为129)。国际合作不仅能促进了中国小麦生物育种科研产出的数量,也能提高论文的国际影响力。

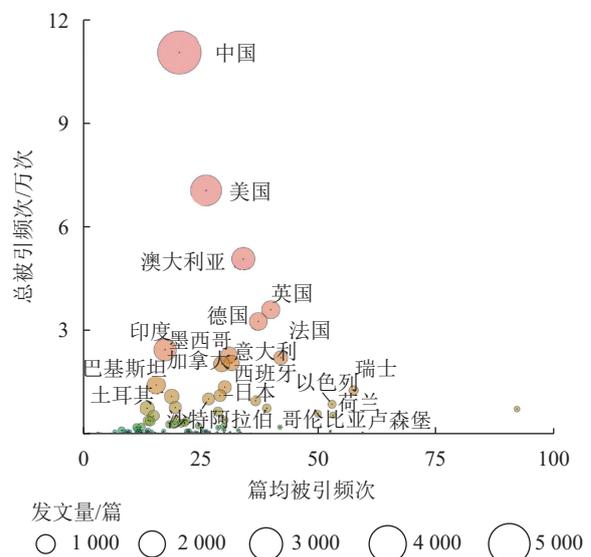
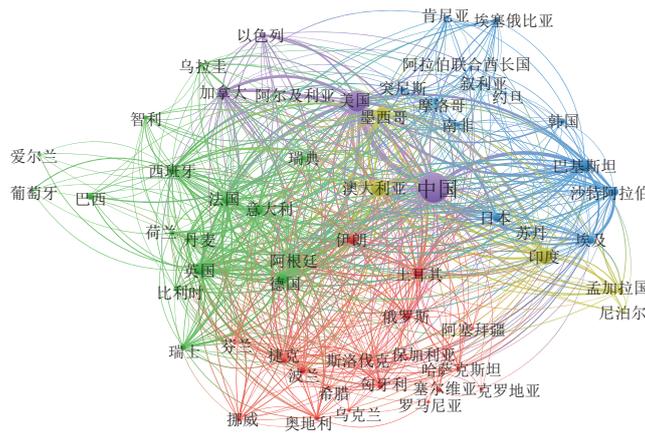


图1 小麦生物育种全球发文情况和被引频次

Figure 1 Worldwide scientific production and citations in the research field of wheat bio-breeding



每个节点代表 1 个国家，节点的大小与发文数量成正比。节点之间的连线表示 2 个国家具有共同合作发文关系，连线越粗表示合作越紧密，合作发文越多。不同颜色代表不同合作集群。

图 2 小麦生物育种国家合作网络

Figure 2 Country collaboration network visualization using VOSviewer

2.3 主要发文机构

从机构发文量来看，在小麦生物育种领域发文较多的机构有美国农业部、中国农业科学院、国际农业研究磋商组织、西北农林科技大学、国际玉米小麦改良中心等 (表 2)。西北农林科技大学是发文量进入前 5 位的高等学校，在此领域的科研实力较强。在学术影响力方面，英国生物技术与生物科学研究理事会 (BBSRC)、英国研究与创新署 (UKRI)、法国国家农业食品与环境研究院 (INRAE)、英联邦科学与工业研究组织 (CSIRO)、堪萨斯州立大学具有较高的篇均被引频次。中国在小麦生物育种领域的主要研究机构是中国农业科学院、西北农林科技大学、中国科学院等，在世界上处于第一梯队。

表 2 小麦生物育种领域的主要发文机构

Table 2 Most publishing institutions in the research field of wheat bio-breeding

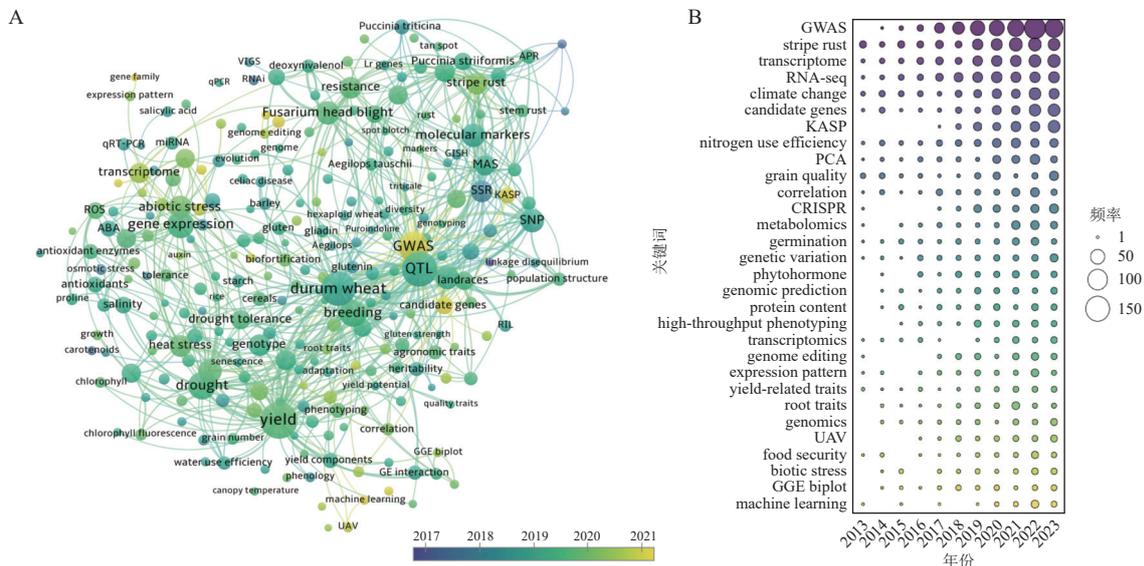
机构	发文量/篇	总被引频次/次	篇均被引频次/次	机构	发文量/篇	总被引频次/次	篇均被引频次/次
美国农业部	1 145	29 407	25.7	堪萨斯州立大学	498	20 780	41.7
中国农业科学院	1 115	25 663	23.0	四川农业大学	432	6 742	15.6
国际农业研究磋商组织	970	29 080	30.0	英联邦科学与工业研究组织	428	17 861	41.7
西北农林科技大学	861	17 803	20.7	法国国家农业食品与环境研究院	388	18 639	48.0
国际玉米小麦改良中心	818	26 104	31.9	埃及知识库	383	7 608	19.9
印度农业研究理事会	732	12 926	17.7	华盛顿州立大学	381	9 639	25.3
中国科学院	702	21 794	31.0	中国农业大学	377	9 910	26.3
中国农业与农村部	605	11 806	19.5	河南农业大学	371	7 505	20.2
英国研究与创新署	504	24 995	49.6	加拿大农业与农业食品部	348	10 435	30.0
生物技术与生物科学研究理事会	503	24 980	49.7	南京农业大学	328	6 900	21.0

2.4 研究热点和趋势

为了探索小麦生物育种的研究热点和趋势，利用 VOSviewer 对小麦生物育种领域发表文献的关键词进行共现和演化分析，生成关键词共现标签视图 (图 3A)。可以看出：在频次大于 30 的关键词中，产量 (yield) 这一关键词的节点最大，说明产量是小麦生物育种的重点，其次是数量性状位点 (QTL)、硬粒小麦 (durum wheat)、全基因组关联分析 (GWAS)、干旱 (drought)、基因表达 (gene expression)、育种 (breeding)、单核苷酸多态性 (SNP) 和分子标记 (molecular markers)。

由图 3A 可见小麦生物育种领域研究热点的演变。代表性的蓝色节点有简单序列重复 (SSR)、小麦叶锈菌 *Puccinia triticina*、RNA 干扰 (RNAi) 等。这些技术或领域在较早的时期就被大量研究，而近年来的研究进展缓慢或有下降趋势，说明小麦生物育种中的相关研究已经相当成熟或有新的技术方法得到应用。代表性的绿色节点，如 yield、breeding、QTL、durum wheat、drought 等关键词增长平缓。黄色节点

的关键词在近几年的研究中较为活跃,可能代表新兴的研究方向,代表性的关键词包括 GWAS、竞争性等位基因特异性 PCR (KASP)、候选基因 (candidate genes)、基因组编辑 (genome editing)、转录组 (transcriptome)、机器学习 (machine learning) 等,说明这些技术或领域是近年来的研究热点,科研产出上上升较快。



在关键词共现图谱中,每个节点代表1个关键词,节点越大表示关键词出现的频次越高;节点间的连线越粗表示2个关键词共同出现在同一篇文章中的频率越高,即2个关键词的联系越紧密。节点的颜色表示关键词所在文献的平均出版年,颜色越蓝的关键词表示较早的研究热点或在早期就被大量研究,颜色越黄的关键词代表最新出现的研究热点。ABA. 脱落酸; APR. 成株抗性; CRISPR. 规律间隔成簇短回文重复序列; GE. 基因型环境; GGE. 基因型主效加基因型环境互作效应; GISH. 基因组原位杂交; GWAS. 全基因组关联分析; KASP. 竞争性等位基因特异性聚合酶链式反应; Lr. 抗叶锈病基因; MAS. 分子标记辅助选择; miRNA. 微小核糖核酸; PCA. 主成分分析; qRT-PCR. 实时荧光定量聚合酶链式反应; QTL. 数量性状位点; RIL. 重组自交系; RNAi. RNA 干扰; ROS. 活性氧; SNP. 单核苷酸多态性; SSR. 简单重复序列; UAV. 无人机; VIGS. 病毒诱导的基因沉默。

图3 小麦生物育种研究关键词共现图谱 (A) 和新兴关键词年度分布 (B)

Figure 3 Co-occurrence overlay visualization of author keyword (A) and yearly trends of newly emerging keywords (B) in the research field of wheat bio-breeding

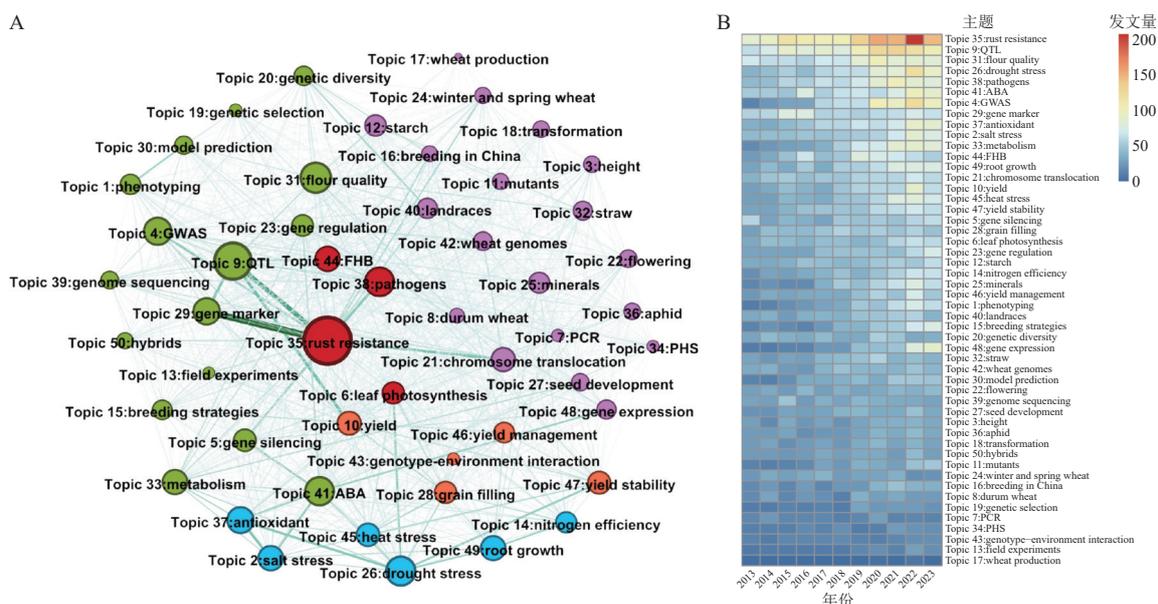
为了清楚地了解这些词的变化趋势,选取了平均出版年数值大于2020(即关键词所在文献的平均出版年在2020年之后),并且出现频次较高的前30个关键词,对其年度分布进行了分析(图3B)。可以看出:GWAS相关的年度产出要明显高于其他关键词。GWAS在2017前发文还较少(出现频次低于10次),2017年后迅速增长,2022年已成为小麦生物育种研究中最大的热点(出现频次高于同年其他关键词)。有些关键词在2017年以前没有或只有少量相关研究,2017年后逐渐兴起,如KASP、CRISPR、高通量表型分析 (high-throughput phenotyping)、基因组编辑、无人机 (UAV)、机器学习等。这与近年来组学、生物技术、高通量表型和人工智能等现代生物育种技术的发展和广泛应用相关。

2.5 LDA 研究主题分析

本研究提取了16151篇文献的摘要,进行LDA分析,建立语言模型,识别研究主题,最终将所有文献归为50个研究主题。进一步根据Luovian聚类将50个主题分为了5个主题网络集群,以5种不同的颜色显示(图4A)。

在以遗传定位为代表的主题网络集群(绿色区域)中,研究较多的主题是QTL、GWAS、遗传多样性 (genetic diversity)、基因标记 (gene marker)、表型分析 (phenotyping)、模型预测 (model prediction) 等。其中GWAS、基因标记、基因组测序和QTL之间联系紧密,而表型分析、模型预测和遗传选择 (genetic selection) 之间联系紧密,表明现代小麦遗传育种涉及了表型、基因组、基因定位和育种模型的联合分析以及重要基因的遗传定位。在以基因组和育种研究为代表的主题网络集群(紫色区域)中,包括小麦基因组 (wheat genomes)、染色体易位 (chromosome translocation)、硬粒小麦、地方品种 (landraces)、突变体 (mutants)、淀粉 (starch)、开花 (flowering)、株高 (height) 等研究主题。在以生物逆境为代表的主题集群(红色区域)中,包括锈病抗性 (rust resistance)、赤霉病 (fusarium head blight, FHB)、病原菌 (pathogens)、

叶片光合作用 (leaf photosynthesis) 等研究主题, 其中锈病抗性相关研究发表文章最多。在以非生物逆境为代表的主题集群 (蓝色区域) 中, 研究较多的主题是盐胁迫 (salt stress)、高温胁迫 (heat stress)、干旱胁迫 (drought stress)、氮效率 (nitrogen efficiency)、根系生长 (root growth) 等。在以产量形成为代表的主题网络集群 (橙色区域) 中, 主要包括产量、产量稳定性 (yield stability)、灌浆 (grain filling)、基因型与环境互动 (genotype-environment interaction) 等研究主题。对每个主题的年度发文量进行分析, 发现锈病抗性相关的研究近年来发文量持续增多, 在 2022 年达到了 207 篇, 是小麦生物育种的热点研究主题, 其次是 QTL、面粉品质 (flour quality)、干旱胁迫 (drought stress)、病原菌、脱落酸、GWAS 等相关主题 (图 4B)。2022 年 GWAS 相关文章发表 124 篇, 与 2013 年相比增长了接近 10 倍, 是论文数量增长最快的主题, 该结果与关键词共现和演化分析结果一致。



ABA. 脱落酸; FHB. 赤霉病; GWAS. 全基因组关联分析; PCR. 聚合酶链式反应; PHS. 穗发芽; QTL. 数量性状位点。左图中, 绿色区域表示以遗传定位为代表的主题; 紫色区域表示以基因组和育种研究为代表的主题; 红色区域表示以生物逆境为代表的主题; 蓝色区域表示以非生物逆境为代表的主题; 橙色区域表示以产量形成为代表的主题。节点大小代表相关文献的数量, 连接线的粗细代表主题之间的相关性。右图中横坐标代表年份, 纵坐标代表 LDA 主题, 颜色代表发文量。

图 4 小麦生物育种研究 LDA 主题网络聚类分析 (A) 和年度发文量分析 (B)

Figure 4 Topic cluster network by LDA (A) and yearly trends of different research topics (B) in the research field of wheat bio-breeding

3 讨论

本研究利用文献计量学和机器学习的方法对小麦生物育种的现状、前沿热点和发展趋势进行分析总结。过去 10 a 是小麦生物育种迅速发展时期。然而, 对于国内外小麦生物育种研究论文的文献计量学分析还停留在 2018 年以前^[7-9], 无法反映小麦生物育种研究的现状和趋势。本研究从 Web of Science (WoS) 核心合集数据库中检索 2013—2024 年发表的小麦生物育种相关文献, 共计 16 151 篇。结果显示: 2013 年以来, 全球小麦生物育种研究的科研产出大幅增加, 全球 123 个国家参与小麦生物育种研究, 全球小麦生物育种研究正处于快速发展阶段, 是近年来的热门研究领域。中国是发表小麦生物育种相关论文最多的国家, 还拥有最多的一流大学和研究院所, 对小麦研究的重视程度高。

对 2018 年前发表的小麦研究相关文献分析显示: 小麦遗传育种关键词主要集中在产量、品质、品种、遗传多样性、数量性状位点和分子标记等传统育种和分子育种相关词汇^[7-9]。本研究的关键词共现分析表明: 产量、数量性状位点、全基因组关联分析、干旱、基因表达、单核苷酸多态性等关键词出现的频率最高。随着测序技术的发展, 数量性状位点和全基因组关联分析相关研究迅速增多。目前在小麦中已鉴定超过 27 500 个数量性状位点, 其中大部分的数量性状位点是近 10 a 间鉴定的^[15]。近年来还出现了高通量表型分析、无人机、机器学习、深度学习、基因组编辑、规律间隔成簇短回文重复序列等现代生物技术相关的关键词。多篇文章整合表型组、基因组、代谢组、机器学习和基因编辑等技术对小麦

重要性状进行遗传解析^[16-20]。该结果说明,现代生物育种和信息技术已经开始应用到小麦育种^[21-22]。也需要注意的是,这些新兴关键词出现的频率还不高,说明小麦生物育种研究未来需要进一步加强现代生物技术的应用。另外,LDA主题词汇分析表明:锈病、数量性状位点、面粉品质、干旱、全基因组关联分析等相关研究的发文量增长较快。这些主题成为近年来的热点研究领域,可能是由于小麦基因组序列的发布和小麦遗传转化技术的突破。比较而言,分子标记、育种选择等主题增长缓慢或停滞。叶锈病是研究最多的主题,有超过80个叶锈病抗性基因应用于小麦抗病育种^[23]。随着全球气候变化加剧,干旱和盐胁迫越来越受到重视,多基因聚合和转基因技术被用来改良综合抗性,实现多性状协同的目标^[24-25]。这些结果说明小麦生物育种研究更加关注抗性与产量的协同提升。

4 结论

在过去10 a间,小麦基因组测序的完成和现代生物技术的应用极大地推动了小麦生物育种的发展。然而还存在几点不足:高影响力和高水平基础研究占比小;关于基因编辑技术开发、基因组选择算法、转基因技术的基本机制的研究很少;对代谢物营养品质、抗病和抗逆机理研究不足;关键基因在育种中的应用研究很少。未来还需要充分利用包括组学、自动表型、人工智能、基因编辑、基因组育种等现代生物技术,发掘和利用重要基因,开展多基因聚合育种,大幅提高小麦产量的同时,改良综合抗性和适应性,实现多性状协同提升。

5 参考文献

- [1] WALLACE J G, RODGERS-MELNICK E, BUCKLER E S. On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics [J]. *Annual Review of Genetics*, 2018, **52**: 421-444.
- [2] 张健. 中国重要农作物生物育种产业化应用的展望[J]. 中国农业科技导报, 2022, **24**(12): 15-24.
ZHANG Jian. Prospects for commercialization of biotech breeding technology of important crops in China [J]. *Journal of Agricultural Science and Technology*, 2022, **24**(12): 15-24.
- [3] 何中虎, 夏先春, 陈新民, 等. 中国小麦育种进展与展望[J]. *作物学报*, 2011, **37**(2): 202-215.
HE Zhonghu, XIA Xianchun, CHEN Xinmin, *et al.* Progress of wheat breeding in China and the future perspective [J]. *Acta Agronomica Sinica*, 2011, **37**(2): 202-215.
- [4] 贺志勇. 生物育种对国民经济与粮食安全的贡献分析[J]. 分子植物育种, 2024, **22**(13): 4505-4510.
HE Zhiyong. The contribution analysis of biobreeding to national economy and food security [J]. *Molecular Plant Breeding*, 2024, **22**(13): 4505-4510.
- [5] 黄耀辉, 焦悦, 吴小智, 等. 生物育种对种业科技创新的影响[J]. 南京农业大学学报, 2022, **45**(3): 413-421.
HUANG Yaohui, JIAO Yue, WU Xiaozhi, *et al.* The influence of biological breeding on the science and technology innovation of seed industry [J]. *Journal of Nanjing Agricultural University*, 2022, **45**(3): 413-421.
- [6] 邱均平, 段宇锋, 陈敬全, 等. 我国文献计量学发展的回顾与展望[J]. *科学学研究*, 2003, **21**(2): 143-148.
QIU Junping, DUAN Yufeng, CHEN Jingquan. The retrospect and prospect on bibliometrics in China [J]. *Studies in Science of Science*, 2003, **21**(2): 143-148.
- [7] 王瑞. 基于文献计量分析的小麦科研实力国际比较研究[D]. 合肥: 安徽农业大学, 2017.
WANG Rui. *International Comparison of Wheat Research Level Based on Bibliometric Analysis* [D]. Hefei: Anhui Agricultural University, 2017.
- [8] GIRALDO P, BENAVENTE E, MANZANO-AGUGLIARO F, *et al.* Worldwide research trends on wheat and barley: a bibliometric comparative analysis [J/OL]. *Agronomy*, 2019, **9**(7): 352 [2024-07-10]. DOI: [10.3390/agronomy9070352](https://doi.org/10.3390/agronomy9070352).
- [9] 孟静, 唐研, 徐淑良, 等. 基于文献计量的国内外小麦遗传育种研究进展[J]. 江苏农业科学, 2020, **48**(3): 64-72.
MENG Jing, TANG Yan, XU Shuliang, *et al.* Progress of genetic breeding research in wheat at home and abroad based on bibliometrics [J]. *Jiangsu Agricultural Sciences*, 2020, **48**(3): 64-72.
- [10] International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome [J/OL]. *Science*, 2018, **361**(6403): eaar7191 [2024-07-10]. DOI: [10.1126/science.eaar7191](https://doi.org/10.1126/science.eaar7191).
- [11] 李洪杰, 陈明, 李少雅, 等. 小麦生物育种: 进展、机遇和挑战[J]. 中国基础科学, 2022, **24**(4): 1-8, 28.

- LI Hongjie, CHEN Ming, LI Shaoya, *et al.* Wheat biotech-breeding: progresses, opportunities and challenges [J]. *China Basic Science*, 2022, **24**(4): 1–8, 28.
- [12] 梁爽, 刘小平. 基于文本挖掘的科技文献主题演化研究进展[J]. 图书情报工作, 2022, **66**(13): 138–149.
LIANG Shuang, LIU Xiaoping. Research progress on topic evolution of scientific and technical literatures based on text mining [J]. *Library and Information Service*, 2022, **66**(13): 138–149.
- [13] ZHANG Yujie, YU Chaoran, ZHAO Feng, *et al.* Landscape of artificial intelligence in breast cancer (2000–2021): a bibliometric analysis [J/OL]. *Frontiers in Bioscience-Landmark*, 2022, **27**(8): 224[2024-07-10]. DOI: [10.31083/j.fbl2708224](https://doi.org/10.31083/j.fbl2708224).
- [14] ZHANG Zeyu, WANG Zhiming, HUANG Yun. A bibliometric analysis of 8 276 publications during the past 25 years on cholangiocarcinoma by machine learning [J/OL]. *Frontiers in Oncology*, 2021, **11**: 687904[2024-07-10]. DOI: [10.3389/fonc.2021.687904](https://doi.org/10.3389/fonc.2021.687904).
- [15] SINGH K, SAINI D K, SARIPALLI G, *et al.* WheatQTLdb V2.0: a supplement to the database for wheat QTL [J/OL]. *Molecular Breeding*, 2022, **42**(10): 56[2024-07-10]. DOI: [10.1007/s11032-022-01329-1](https://doi.org/10.1007/s11032-022-01329-1).
- [16] MA Shengwei, WANG Meng, WU Jianhui, *et al.* WheatOmics: a platform combining multiple omics data to accelerate functional genomics studies in wheat [J]. *Molecular Plant*, 2021, **14**(12): 1965–1968.
- [17] SONG Long, LIU Jie, CAO Beilu, *et al.* Reducing brassinosteroid signalling enhances grain yield in semi-dwarf wheat [J]. *Nature*, 2023, **617**(7959): 118–124.
- [18] GAO Jie, HU Xin, GAO Chunyan, *et al.* Deciphering genetic basis of developmental and agronomic traits by integrating high-throughput optical phenotyping and genome-wide association studies in wheat [J]. *Plant Biotechnology Journal*, 2023, **21**(10): 1966–1977.
- [19] ZHANG Zhen, QU Yunfeng, MA Feifei, *et al.* Integrating high-throughput phenotyping and genome-wide association studies for enhanced drought resistance and yield prediction in wheat [J]. *New Phytologist*, 2024, **243**(5): 1758–1775.
- [20] ZHU Anting, LIU Mengmeng, TIAN Zhitao, *et al.* Chemical-tag-based semi-annotated metabolomics facilitates gene identification and specialized metabolic pathway elucidation in wheat [J]. *The Plant Cell*, 2024, **36**(3): 540–558.
- [21] XIAO Qinlin, BAI Xiulin, ZHANG Chu, *et al.* Advanced high-throughput plant phenotyping techniques for genome-wide association studies: a review [J]. *Journal of Advanced Research*, 2022, **35**: 215–230.
- [22] SINGH B, KUMAR S, ELANGO VAN A, *et al.* Phenomics based prediction of plant biomass and leaf area in wheat using machine learning approaches [J/OL]. *Frontiers in Plant Science*, 2023, **14**: 1214801[2024-07-10]. DOI: [10.3389/fpls.2023.1214801](https://doi.org/10.3389/fpls.2023.1214801).
- [23] PRASAD P, SAVADI S, BHARDWAJ S C, *et al.* The progress of leaf rust research in wheat [J]. *Fungal Biology*, 2020, **124**(6): 537–550.
- [24] LANGRIDGE P, REYNOLDS M. Breeding for drought and heat tolerance in wheat [J]. *Theoretical and Applied Genetics*, 2021, **134**(6): 1753–1769.
- [25] LI Zihan, ZHONG Fan, GUO Jianrong, *et al.* Improving wheat salt tolerance for saline agriculture [J]. *Journal of Agricultural and Food Chemistry*, 2022, **70**(48): 14989–15006.