植物遗传资源学报 2023，24（5）：1257-1266
Journal of Plant Genetic Resources DOI：10.13430/j.cnki.jpgr.20230322002

工业大麻中全基因组关联分析研究进展

李洪超，王晓楠，李紫薇，孙宇峰，曹 焜，赵 越
（黑龙江省科学院大庆分院麻类作物研究所，大庆 163316）

摘要：工业大麻作为一种重要的特种经济作物，已经实现全产业链发展，尤其是其韧皮纤维作为重要原料而被广泛应用于纺织、造纸、建筑、家居、绝缘材料、建材以及汽车零部件和复合材料。国内对工业大麻的研发和应用较早，工业大麻纺织产品在全球市场上具有极大地竞争力。在纺织工业中随着市场对纤维产量的需求越来越高，对品质的要求也越来越严格，科研人员需要不断选育优质高产的纤维型工业大麻品种。全基因组关联分析（GWAS，genome-wide association study）能够揭示作物表型与基因型之间的联系，利于挖掘性状相关基因和遗传基础信息，进而分析群体遗传结构，推动分子育种与品种性状改良有机结合培育高产高质、抗病抗逆的优良品种。目前该技术方法已广泛应用在棉花、水稻、玉米、小麦等农作物中，在工业大麻等主要麻类纤维作物中的研究也取得一定进展。因此，本研究对目前工业大麻中GWAS的应用情况及研究结果进行阐述，并对未来应用该技术的研究方向进行了简要说明，以期为工业大麻后续采用GWAS分析提供见解。

关键词：工业大麻；基因；全基因组关联分析

Advances in Genome-Wide Association Study in Industrial Hemp

LI Hongchao，WANG Xiaonan，LI Ziwei，SUN Yufeng，CAO Kun，ZHAO Yue
（*Institute of Bast Fiber Crops，Daqing Branch of Heilongjiang Academy of Sciences，Daqing 163316*）

Abstract： The industrial hemp（*Cannabis sativa* L.）as an important special economic crop，has implemented the development of the whole industrial chain. Its bast fiber is a significant raw material and has been widely used in textiles，papermaking，home furnishings，insulation materials，construction materials，and auto parts and composites. China has stronger academic and applied research experiences in industrial hemp，and the derived textile products are highly competitive in the global market. Due to the increasing yield and quality demand for textile fibers，it is necessary for researchers to breed high-quality，high-yield fiber varieties. The genome-wide association study（GWAS）can reveal the association between phenotype and genotype，thus becoming of interest in gene mining，stacking of multiple elite genes alleles，analyzing the population genetic structure and finally breeding for novel varieties showing high-yield，high-quality，disease resistance，and stress resistance. At present，GWAS is widely used in cotton，rice，corn，wheat and other crops，and increasingly used in industrial hemp and other major hemp fiber crop. Thus，the review summarizes the application of GWAS in industrial hemp and briefly proposes the research focuses that remained yet conducted，in order to provide insights for the application of GWAS technique in industrial hemp.

Key words： industrial hemp；genes；genome-wide association study

收稿日期：2023-03-22 **修回日期：**2023-05-05 **网络出版日期：**2023-05-19
URL：https://doi.org/10.13430/j.cnki.jpgr.20230322002
第一作者研究方向为麻类作物遗传育种与分子生物学，E-mail：L15136244595@163.com
通信作者：王晓楠，研究方向为麻类作物遗传育种及分子生物学，E-mail：wxn_fern@163.com
基金项目：黑龙江省属科研院所科研业务费重点项目（CZKYF2021-2-B011）；黑龙江省省属科研院所科研业务费项目（KYYW2022DQ01）
Foundation projects：Key Projects of Scientific Research Business Cost of Scientific Research Institutes Affiliated to Heilongjiang Province（CZKYF2021-2-B011）；Research Business Fee Project of Provincial Scientific Research Institutes in Heilongjiang Province（KYYW2022DQ01）

工业大麻（*Cannabis sativa* L.）是无毒品利用价值的一种类型，其与毒性相关的四氢大麻酚（THC，delta-9-tetrahydrocannabinol）成分的含量低于0.3%，是人类最早种植的纤维作物之一[1-2]，作为一种可持续和高产的工业作物，具有种植简便、适应环境能力强、生长速度快等种植特点和低碳环保等价值，工业大麻已成为世界公认的重要特种经济作物之一[3]。不仅如此，工业大麻还具有极高的利用价值，如花叶提取物中富含多种化学成分，可用于医药领域和日化行业[4-7]；茎皮由韧皮纤维构成，具有较强的韧度，在纺织、服装、新材料、造纸等不同领域得到广泛利用[8]；种子可开发功能食品；根可作为中药材。随着市场需求量的不断增加，对品种和原料的要求也越来越严格，为满足市场及农业可持续发展的要求，了解作物生长发育过程的遗传基础以提高产量和改良品质，同时利用有限的作物遗传种质资源来挖掘表型性状相关的基因[9]，通过调控相关基因的表达，从而培育出高产、优质和抗逆的作物新品种显得尤为重要[10]。本研究旨在总结全基因组关联分析（GWAS，genome-wide association study）在工业大麻中的最新研究进展并发现研究空白，进而探讨GWAS在挖掘更多性状关联基因中的应用，推动工业大麻优质品种培育。

1 全基因组关联分析技术

在对人类疾病遗传学的研究中，Risch等[11]通过人类基因组计划数据库与关联研究相结合的方法确定遗传疾病基础，将该方法称为GWAS。后续研究者们将GWAS引入作物的遗传育种等研究领域，通过标记特定性状基因和连锁作图来探索与目标性状相关的关键基因、分析群体遗传结构以及推动分子标记辅助育种等[12]。GWAS在水稻、棉花、玉米等大作物中得到广泛应用，在亚麻、工业大麻等小作物中的应用也在逐渐增多。如结合GWAS和转录组学方法鉴定与耐热性相关的候选基因，该结果为通过分子育种提高水稻耐热性提供了目标基因[13]。利用GWAS群体遗传结构分析追溯亚洲棉在中国的起源，与结构变异联合探索棉花的遗传变异以及地理分化，为棉花遗传育种、群体遗传结构分析以及群体驯化提供了遗传基础[14]。Yasir等[15]对棉花中GWAS在生物抗性、非生物胁迫的耐受性、纤维产量和品质性状、现状、前景、瓶颈等方面进行了全面综述。GWAS在亚麻中可用于挖掘干旱及其他相关非生物胁迫响应过程中的关键基因位点[16]，确定与种子耐盐性相关的数量性状位点（QTL，quantitative trait locus）[17]，对白粉病抗性的遗传结构和基因组进行预测[18]，确定进化趋势和农艺性状相关候选基因[19]。

在前人的研究中对基因进行标记多采用微卫星标记（SSR，simple sequence repetition）[20]，而GWAS多利用单核苷酸多态性（SNP，single nucleotide polymorphisms）进行分子标记[21]。目前已有利用结构变异（SV，structural variation）与GWAS结合的研究，能够在作物中筛选与表型性状相对应的标记基因，从而加快对性状基因的定位及对可用的分子标记进行筛选。随着分子生物学测序技术的进步及测序成本的降低，工业大麻[22]的基因组陆续公布，基因组数据库不断扩大。高通量基因-表型数据库的构建为GWAS挖掘优异基因资源、解析控制复杂性状的遗传机理、探索种质资源的进化机制以及开展系统的分子育种等系列研究提供了更为高效便捷的技术参考和研究思路[23]。

2 工业大麻中GWAS研究进展

2.1 GWAS在群体结构分析中的应用进展

据记载，大麻作为药物已有近6000年的使用历史，因含有活性THC，其种植和使用在许多国家受到监管和限制。随着相关政策的不断放宽，许多国家开始允许种植药用型和纤维型等不同类型的大麻品种。GWAS除用于探索表型和基因之间的联系外，还被广泛应用于群体结构遗传分析中。Soorni等[24]结合短串联重复序列（STR，short tandem repeat）标记对商业品种进行基因分型测序（GBS，genotyping-by-sequencing），结果表明工业大麻和药用大麻之间存在显著的全基因组分化，但因缺乏全面的基因组数据导致目前对大麻进化史的了解存在巨大差异。Kovalchuk等[25]对此前发表的所有大麻基因组学数据进行详尽分析，发现可用的大麻基因组仍然是不完整的，而Murray[26]的研究弥补了一部分空缺，通过大麻素基因不同的进化程度将纤用大麻与药用大麻分为不同遗传群体，该研究结果为药用型和纤维型等非毒品大麻的分子育种和功能研究提供了前所未有的基因组资源基础。在最近的报道中，Ren等[27]为了探究大麻的驯化历史对大规模种质进行了全基因重测序并进行GWAS分析，所得到的基因组数据与参考基因组CBDRx（https://www. ncbi. nlm. nih. gov/data-hub/genome/GCF_900626175.2/）比对，结果发现大麻种

质显著地聚类成4个分离良好的遗传群体，其中中国的一些地方品种和野生品种与大麻祖先的亲缘关系最近，工业大麻和毒品大麻的地方品种和栽培品种就是由此衍生的。同时发现在驯化过程中两个主要大麻素合酶基因的功能发生丢失，也反映了工业大麻在合法化种植过程中所进行的自然选择可能会成为一种育种手段，该研究为了解工业大麻起源与种群结构提供了新的见解，并为分子遗传领域的研究奠定了基础。为进一步探索可利用的工业大麻资源，采用单分子超长测序（SMUS，single molecule ultra-long sequencing）和高通量染色体构象捕获技术（Hi-C，high-throughput chromosome conformation capture）的组合，Gao等[28]以西藏的野生大麻为材料更新了现有工业大麻基因组序列。

2.2　GWAS在品质性状中的应用进展

工业大麻的纤维品质与其应用价值密切相关，深入探索与品质相关的关键基因是培育高品质品种的前提。为了探究工业大麻的开花时间和性别与纤维品质是否存在密切联系，Petit等[29]结合SNP标记利用GWAS分析了123个大麻种质资源在3种环境下的工业大麻开花时间和性别决定的遗传结构，共确定了8个QTL，其中6个与开花时间相关，可能参与调节光周期和开花温度，2个与性别决定相关，可能参与调节植物激素赤霉酸和生长素平衡，该研究结果为了解工业大麻开花时间和性别决定的遗传学基础提供了研究思路。同年在其另一篇研究中提出对该种质群体的7个纤维品质相关性状进行GWAS筛选相关的QTL，鉴定出16个相关QTL，其中6个候选基因区域可能参与与细胞壁形成相关的单糖、多糖和木质素的生物合成与修饰途径[30]。Woods等[31]对纤用和籽用杂交的F_2群体进行了全基因组测序，共鉴定出69个与农艺（34个）和生化（35个）性状变异相关的位点，在连锁图中鉴定到4个农艺和2个生化性状QTL共定位簇，这表明选用的纤用和籽用之间的表型区别在很大程度上由少数位点控制，但农艺和生化性状之间是否具有关联性和特异性无法确定。在探究影响纤维品质的关键基因的过程中，相比于GWAS技术，利用转录组学相关技术方法的研究更为成熟，报道的研究结果也相对较多。有研究发现工业大麻纤维是由下胚轴二次生长形成的初生和次生韧皮纤维构成，转录因子*NST1*、*MYB46*和*WLIM1*的上调与其二次生长相关[32-34]。Behr等[33]对播种后6、9、15、20 d（H6、H9、H15、H20）的工业大麻进行转录组学分析，发现了次级细胞壁（SCW，secondary cell walls）沉积的主要调控因子*SND2*、*VND1*和*NST1*，在H15中*MYB46*的表达量最高，在H20中*KNAT7*和*MYB85*以及与*AtWLIM1*的直系同源物表达更丰富，而发育后期上调的SCW相关基因表达更为丰富。将特定位点扩增片段测序（SLAF-seq，specific-locus amplified-fragment sequencing）和批量分离分析（BSA，batch separation analysis）结合快速鉴定影响大麻纤维含量的基因，通过实时荧光定量多聚核苷酸链式反应（qRT-PCR，real-time quantitative polynucleotide chain reaction）确定了6个与大麻纤维含量呈高度正相关的基因，并对这些基因进行表达分析，发现*LOC115705530*、*LOC115705530*、*LOC115707511*、*LOC115704794*、*LOC115705371*和*LOC115708688*的表达水平随总纤维含量增加而增加，而*LOC115705875*的表达水平则相反，且*LOC115705875*在苗期的表达水平高于工艺成熟期，研究发现这6个基因参与了转录调控、生长素和水的运输、一碳和糖的代谢[35]。

2.3　GWAS在白粉病等病虫害中的应用进展

工业大麻的生产种植过程存在感染病虫害的风险，容易导致减产和品质降低。目前，对于工业大麻虫害的研究主要集中在发现、预防虫害以及探究虫害对生长发育的影响等。Bakro等[36]系统地描述了波兰工业大麻品种中常见的虫害类型以及虫害对纤维型工业大麻生长发育和产量的影响，并强调可通过选择空气流通、低湿度、适宜种植密度和水分的种植措施来保护植株正常生长及其产量品质。Cranshaw等[37]先后报道了蚜虫、甲虫、蚱蜢、蜘蛛螨、棉铃虫等对工业大麻的危害以及易感部位，确定了几种潜在的关键害虫，并初步提出了虫害防治的建议[38]。Ajayi等[39]认为棉铃虫为美国不同地区工业大麻种植的主要虫害，并阐明了气候变化会影响棉铃虫的发育。白粉病（PM，powdery mildew）是由真菌引起、导致作物减产同时引起人的过敏反应的植物疾病。虽然已有大麻PM抗性的相关报道，但迄今为止尚未有该性状的遗传学描述[40]。揭示PM抗性的遗传基础有助于进行针对性育种，提高产量并减少员工过敏原暴露，还可以用于开发既有抗病性又不伤植株的高效药物。对工业大麻白粉病抗性的研究中还未应用GWAS技术，现有的研究成果大多通过对全基因组序列进行表征以及通过基因注释等方法分析可能的相关基因。后续运用GWAS分析性状关键基因可以与目前已经报道

的基因相互佐证。自Lyngkjær等[41]在大麦中发现霉菌抗性位点O基因(*MLO*，mold resistance site O)与PM抗性相关后，陆续在水稻、玉米、烟草和番茄等物种中鉴定出*MLO*的同源基因。Adams[42]在几个对PM有抗性的大麻品种中观察到重叠基因上存在一个或多个类甜蛋白(TLP，thaumatin-like protein)拷贝，TLP具有β-1，3葡聚糖酶活性，能够产生广泛的病原体抗性，说明该蛋白能够对PM产生抗性；而在转基因番茄中发现几丁质酶和TLP具有协同活性[43]，因此推测编码几丁质酶CH25与PM抗性相关[44]。Pépin等[45]在5个不同的大麻参考基因组Purple Kush（https://www.ncbi.nlm.nih.gov/assembly/GCA_000230575.5）、Finola（https://www.ncbi.nlm.nih.gov/assembly/GCA_003417725.2）、Jamaican Lion Male（https://www.ncbi.nlm.nih.gov/assembly/GCA_013030025.1）、Jamaican Lion Female（https://www.ncbi.nlm.nih.gov/assembly/GCA_016415525.1）和CBDRx（https://www.ncbi.nlm.nih.gov/data-hub/genome/GCF_900626175.2/）中都鉴定并表征了*MLO*基因家族成员，但各染色体所鉴定出的基因数量不同，除5号和9号染色体不携带*MLO*基因外，其余8条染色体上均含有*CsMLO*基因。在5个基因组中均鉴定出*CsMLO14*，在CBDRx基因组未识别到*CsMLO15*。对MLO蛋白序列的系统发育分析表明，分支V的*CsMLO1*和*CsMLO4*在感染PM后会显著上调，初步证据表明它们可能与PM易感性有关。尽管目前报道了部分与白粉病相关的基因，但其遗传基础和作用机制尚不明确，在其他病虫害的研究方面也存在大量的研究空白，未来可以将GWAS应用于挖掘影响病虫害抗性的关键基因，不断扩大工业大麻的功能基因数据库，为培育优良品种提供可能。

2.4 GWAS在抗逆性中的应用进展

目前GWAS在工业大麻抗逆性研究中还未得到广泛应用，多数研究是结合生理生化指标和转录组学技术进行分析。干旱胁迫会导致工业大麻病虫害发病率增加，生长缓慢，延迟纤维和种子成熟，最终影响产量和品质[46]。目前对干旱胁迫下工业大麻生理指标及形态变化的研究较多，但对抗旱调控分子机制的研究较少。烯效唑作为植物生长调节剂之一，能够影响非生物胁迫下植物生长发育的不同生理指标[47]，依据该结论姜颖等[48]提出烯效唑可调节干旱胁迫条件下的植物生长和胁迫耐受性，并在后续的研究中证实烯效唑能够通过参与多种途径对干旱胁迫进行响应[49]。Gao等[50]利用RNA测序技术研究了干旱胁迫和水分充足的工业大麻之间的转录组差异，共鉴定出1292个差异表达基因(DEG，differentially expressed gene)，并通过对植物激素信号转导通路中15个DEG的分析推测该通路所受干旱影响最大。除干旱胁迫外，盐碱胁迫也是威胁工业大麻生长以及产量品质的一种常见非生物胁迫。在最近的报道中，Sun等[51]首次利用GWAS发现了工业大麻种子萌发过程中对耐盐性调节可能起到关键作用的基因，该研究结果将有助于对耐盐性基因的克隆和分子标记辅助育种，同时也为后续利用GWAS探索工业大麻的耐盐性提供了研究方向。Cao等[52]为了确定与耐碱性相关的候选基因，利用RNA测序对受$NaHCO_3$胁迫诱导的工业大麻进行转录组分析，发现DEG富集于与苯丙类生物合成、植物激素信号转导与合成、蔗糖、氮、氨基酸等调控通路，且发现大多数的DEG在受到$NaHCO_3$胁迫时表达量上调，编码GTP结合蛋白的基因*GTPbp*可能在$NaHCO_3$胁迫诱导的内吞作用中起关键作用，推测$NaHCO_3$胁迫通过下调木质素合成途径相关基因来影响工业大麻苯丙烷的生物合成。除干旱胁迫和盐碱胁迫外，因工业大麻可以用于修复镉污染的土地，因此探究工业大麻镉应激反应的关键基因十分重要。Yin等[53]为了揭示MYB基因家族在工业大麻的镉应激反应的作用，将从水稻和大豆数据库中下载的MYB基因序列与工业大麻基因组数据库中的基因信息进行比对鉴定出MYB基因家族的基因序列。结果发现3个MYB亚家族共115个*CsMYB*基因，与镉胁迫下的转录组数据结合，鉴定出7个与镉胁迫响应相关的基因，该结果有助于后续对工业大麻中MYB基因家族的生物学功能研究。

2.5 GWAS在药用成分中的应用进展

根据工业大麻用途可将其品种分为药用型、纤用型、籽用型、籽纤兼用及其他兼用型。大麻素是大麻植物中存在的酚类化合物的统称，其主要成分有THC、四氢大麻酚酸(THCA，tetrahydrocannabinoid acid)、大麻二酚(CBD，cannabidiol)和大麻二酚酸(CBDA，cannabidiol acid)，THC主要用于精神药物，而CBD也被证实具有很高的医药保健功能[54]。尽管大麻基因组尚不完整，但与大麻素合成相关的关键酶的基因已经被鉴定。通过生物信息学、转录组和蛋白质组数据已鉴定出与CBD含量相关的组蛋白乙酰转移酶基因[55]，编码THCA和CBDA的基

因也通过基因克隆和异源表达被鉴定和表征[56-57]。Laverty 等[58]利用基因分型测序（GBS，genotyping-by-sequencing）开发了大麻的物理和遗传图谱，重点关注了与大麻素合酶相关的基因，鉴定并表征了编码大麻色烯酸合酶（CBCAS，cannabichromenic acid synthase）的四氢大麻酚酸合酶（THCAS，tetrahydrocannabinoid acid synthase）基因，并验证了大麻色烯酸（CBCA，cannabichromenic acid）的含量与大麻组织中编码CBCAS基因的转录水平相关，并基于Weiblen等[59]的研究基础，在与总大麻素含量相关的ANUCS501位点上鉴定出能够催化大麻萜酚酸产生的芳香异戊烯基转移酶，得出大麻萜酚酸是THCAS、CBDAS和CBCAS合成底物的结论。Henry等[60]基于NCBI数据库中的基因功能注释，表征了研究中所用的SNP的功能，在THCAS基因簇标记中发现了2个SNP标记未激活的THCAS/CBCAS，涉及到的2个标记在高CBD品种中被定位在1个等位基因上。基于前人对大麻素生化途径的总结[61-63]，Innes等[64]对关键大麻素生化途径进行了总结描述，确定了4个与大麻素合成相关的关键酶，在9个参考基因组中对编码这些酶的基因进行表征发现了编码橄榄醇合酶的*OLS1*和*OLS2*，橄榄酸环化酶的*OAC*，大麻酸合酶的*CsPT1*、*CsPT4*和*CsPT7*，大麻素氧化环化酶的*CBCAS*、*THCAS*和*CBCAS*。该研究中大麻素氧化环化酶基因的数量较多，在不同参考基因组中的序列相近，该结果与Laverty等[58]研究结果相一致。为了探究CBCAS的作用，通过基因分型将工业大麻品种分为5种基因型，并对大麻素合酶（CBNS，cannabinoid synthase）编码基因在5种分型中进行序列变异性及转录水平分析，发现在所有基因型中*CBCAS*的转录水平低但稳定，说明该基因是活跃有功能的，可能与大麻素的含量相关。该研究首次通过转录组数据分析*CBCAS*基因的表达情况来验证其在大麻素生物合成途径中的作用[65]。Welling等[66]将极端表型与GWAS方法结合富集烷基大麻素（ACB，alkyl cannabinoid）多态性区域，在其中一个区域中识别到编码β-酮酰基载体蛋白还原酶的候选基因*BKR*，该基因被认为与聚酮脂肪酸起始单元合成和烷基侧链长度相关，但很难推测出该基因促进烷基侧链长度的机制。除此之外还揭示了跨越该基因长度的22个变体（包括两个非同义替换）；印证了先前报道过的参与大麻素合成途径的位点以及与化学变异相关的其他位点[58]；确定了LG9.40 QTL中有橄榄醇合酶的候选基因（表1），该基因参与大麻素合成途径中的关键步骤，其数量或效率的变化可能会极大地影响大麻素的生产[67]。Yin等[53]依据Cd胁迫下的*CsMYB*基因在CBD含量不同的品种中的表达情况和特异性表达，推测*CsMYB024*通过Cd胁迫的影响来介导CBD合成途径。Campbell等[68]评估环境对工业大麻的影响及基因型与环境之间的相互作用，分析种质资源的遗传特性，发现THC和CBD的产生几乎完全由遗传因素控制，明确籽粒产量受基因-环境相互作用的影响，进行相关性和显著性分析后发现籽粒产量与THC及CBD含量性状呈负相关，即当籽粒产量升高时，大麻素含量下降。上述系列研究为明晰大麻素生物合成途径调控提供了重要参考，将基因组信息与选择性育种结合能够为今后高效开展定向分子育种奠定基础。

表1　工业大麻中已报道的部分基因/QTL和分子标记

Table 1　Some genes/QTLs and molecular markers reported in industrial hemp

序号 No.	项目 Items		取样部位 Sampling location	基因/QTL/分子标记 Gene/QTL/molecular marker	参考文献 Reference
1	开花时间		嫩叶	6个QTL	[29]
2	性别决定		嫩叶	2个QTL	[29]
3	纤维质量		嫩叶	16个QTL	[30]
4	纤维含量		嫩叶	*LOC115705530*、*LOC115707511*、*LOC115704794*、*LOC115705371*、*LOC115708688*和*LOC115705875*	[35]
5	白粉病		茎叶	*CH25*和*MLO*	[44]
	白粉病		不同时间点接种病毒的叶片	*CsMLO1*和*CsMLO4*	[45]
6	抗旱	吲哚乙酸	植物基部完全展开的第三对真叶	*AUX1*(1)、*IAA*(2)、*SAUR*(4)、*GH3*(5)	[49]
		细胞分裂素		*AHK2_3_4*(4)和*ARR-A*(1)	[49]

表1（续）

序号 No.	项目 Items		取样部位 Sampling location	基因/QTL/分子标记 Gene/QTL/molecular marker	参考文献 Reference
		赤霉素		*DELLA*（1）	[49]
		脱落酸		*PP2C*（1）和*SnRK2*（2）	[49]
		脱落酸	叶、根、茎皮、茎芽	*PYL4*、*PP2C-1*～*PP2C-6*、*SAPK3*	[50]
		生长素	叶、根、茎皮、茎芽	*X15-1*、*X15-2*、*IAA1*、*IAA-2*	[50]
7	耐盐性		种子	*XM_030641043.1*、*XM_030641906.1*、*XM_030648362.1*、*XM_030648308.1* 和*XM_030646898.1*	[51]
8	镉胁迫		幼苗	*CsMYB045*，*CsMYB016*，*CsMYB067*，*CsMYB098*，*CsMYB010*，*CsMYB061* 和*CsMYB005*	[53]
9	大麻素合成	/	杂交后代	*THCAS*样基因	[58]
		/	雌花	ANUCS501位点	[59]
		/	顶端花序	*OLS*	[31]
		/	顶端总状花序	*BKR*	[66]
		大麻素氧化环化酶	/	SW6和VSSL_BtBD	[60]
			/	*CBCAS*、*THCAS*和*CBCAS*	[64]
		橄榄醇合酶	/	*OLS1*和*OLS2*	[64]
		橄榄酸环化酶	/	*OAC*	[64]
		大麻酸合酶	/	*CsPT1*、*CsPT4*和*CsPT7*	[64]

/：表示文献中未涉及取样部位或具体性状

/：Indicate that the sampling site or specific traits is not mentioned in the literature

3 问题与展望

因测序技术的更新替代以及成本的不断下降，GWAS将应用于更多作物的不同领域研究中。GWAS方法的不断发展为大量性状基因的挖掘和研究提供了极大便利，所发掘的大量优异基因为分子设计育种及培育优良新品种奠定了坚实的基础。但目前GWAS技术也存在不足，其只能粗略预测候选基因的区段，还需结合转录组学等不同方法对候选基因区域内的基因进行功能注释、分析及对关键基因进一步筛选确定。随着数据统计分析的多元化，遗传学、分子生物学以及生物信息学的迅猛发展，更多技术方法不断与GWAS结合，如利用新开发的分子标记、泛基因组技术等与GWAS技术相结合将更准确地锁定目标性状基因和开展群体遗传进化等研究。最新系列研究表明，GWAS数据与泛基因组等技术相结合，可深度挖掘玉米等作物基因组组成及特征，发现核心基因家族及揭示其杂种优势的分子机制[69]。在棉花中，将结构变异（SV，structural variation）与GWAS结合来分析SV对异倍体陆地棉的形态和多样性的影响[70]；在大麦中基于短读长测序与SV结合检测到23个大麦自交系的SV与全基因组和基因特异性表达相关，并对表型进行评估和预测，说明了SV对表型的预测能力要高于研究表型性状的SNP[71]，提示SV与GWAS结合将会在更多作物研究中得到应用。近年来，水稻[72]、番茄[73]、玉米[74]等常见作物中，基因编辑技术（CRISPR/Cas9）已被逐步应用。充分结合GWAS、分子标记、基因工程等系列技术开展的基因编辑创制新种质为农业精准育种提供了重要基础。

目前，综合GWAS在工业大麻中的应用情况，可以发现该技术在纤维品质、产量、抗性以及药用成分等方面的研究较少，这对后续的研究提供了研究方向同时也会给研究带来一定的挑战。本研究中论述的部分基因大多是通过转录组或基因组等分子生物学技术挖掘的，对已知基因的功能、调控机制及遗传机理等还缺少相关研究。随着工业大麻全基因组数据库的不断更新，探索种群遗传变异和具有重要应用价值的基因位点信息，对研究群体进化、优良性状遗传机理和分子遗传育种是必不可少的，是今后GWAS在工业大麻研究中应用的重点。目前的研究只能为改良特定性状的分子遗传育种提供理论基础，还不足以充分支持选育目标性状的优良品种。故而结合不同技术方法，继续挖掘

工业大麻重要性状相关基因、验证基因功能、阐明调控通路和遗传机制及开展基因编辑，将为工业大麻开展定向分子育种选育符合市场需求的新品种提供支撑。

参考文献

[1] Li H L. The origin and use of cannabis in eastern asia linguistic-cultural implications. Economic Botany，1974，28(3)：293-301

[2] 孙宇峰，张晓艳，王晓楠，田媛. 汉麻籽油的特性及利用现状. 粮食与油脂，2019，32(3)：9-11
Sun Y F，Zhang X Y，Wang X N，Tian Y. Characteristics and utilization status of hemp seed oil. Grain and Oil，2019，32(3)：9-11

[3] 苏芳芳，杨光，郑玉光. 工业大麻种植及育种现状研究. 中国中药杂志，2022，47(5)：1190-1195
Su F F，Yang G，Zheng Y G. Study on the current situation of planting and breeding of industrial Cannabis. Chinese Journal of Traditional Chinese Medicine，2022，47(5)：1190-1195

[4] 陈来成，杨占红，何秋星，黄浩骞，宋相志. 工业大麻法规现状及其在化妆品中开发应用概况. 日用化学品科学，2020，43(1)：20-24
Chen L C，Yang Z H，He Q X，Huang H H，Song X Z. Current status of industrial Cannabis regulations and its development and application in cosmetics. Detergent & Cosmetics，2020，43(1)：20-24

[5] 徐锋奎. 大麻药理作用研究与临床应用新进展. 中国医药情报，2004，10(2)：31-32
Xu F K. New progress in pharmacological research and clinical application of cannabis. Chinese Medical Intelligence，2004，10(2)：31-32

[6] Nalli Y，Arora P，Riyaz-Ul-Hassan S，Ali A. Chemical investigation of *Cannabis sativa* leading to the discovery of a prenylspirodinone with anti-microbial potential. Tetrahedron Letters，2018，59(25)：2470-2472

[7] Nuutinen T. Medicinal properties of terpenes found in *Cannabis sativa* and *Humulus lupulus*. European Journal of Medicinal Chemistry，2018，157：198-228

[8] 徐峻. 大麻芯秆硫酸盐法制浆工艺参数的建模与优化. 中国造纸学报，2007，22(4)：28-31
Xu J. Modeling and optimization of process variables for kraft pulping of hemp core. Journal of China Paper Industry，2007，22(4)：28-31

[9] Zhao K Y，Tung C W，Eizenga G，Wright M H，Ali M L，Price A，Norton G J，Islam M R，Reynolds A，Mezey J，McClung A M，Bustamante C D，McCouch S. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nature Communications，2011，2(1)：467

[10] Tung C W，Zhao K Y，Wright M H，Ali M L，Jung J K，Kimball J，Tyagi W，Thomson M，McNally K L，Leung H，Kim H J，Ahn S N，Reynolds A，Scheffler B，Eizenga G，McClung A，Bustamante C D，McCouch S. Development of a research platform for dissecting phenotype-genotype associations in rice (*Oryza* spp.). Rice，2010，3：205-217

[11] Risch N J，Merikangas K R. The future of genetic studies of complex human diseases. Science，1996，273：1516-1517

[12] Huang X H，Wei X H，Tao S，Zhao Q，Qi F，Zhao Y，Li C Y，Zhu C L，Lu T T，Zhang Z W，Li M，Fan D L，Guo Y L，Wang A H，Wang L，Deng L W，Li W J，Lu Y Q，Weng Q J，Liu K Y，Huang T，Zhou T Y，Jing Y F，Li W，Zhang L，Buckler E，Qian Q，Zhang Q F，Li J Y，Han B. Genome-wide association studies of 14 agronomic traits in rice landraces. Nature Genetics，2010，42(11)：961-967

[13] Li P P，Jiang J，Zhang G G，Miao S Y，Lu J B，Qian Y K，Zhao X Q，Wang W S，Qiu X J，Zhang F，Xu J L. Integrating GWAS and transcriptomics to identify candidate genes conferring heat tolerance in rice. Frontiers in Plant Science，2023，13：1102938

[14] He S P，Sun G F，Geng X L，Gong W F，Pan H D，Jia Y H，Shi W J，Zhao P，Wang J D，Wang L Y，Xiao S H，Chen B J，Cui S F，You C Y，Xie Z M，Wang F，Sun J，Fu G Y，Peng Z，Hu D W，Wang L R，Pang B Y，Du X M. The genomic basis of geographic differentiation and fiber improvement in cultivated cotton. Nature Genetics，2021，53：916-924

[15] Yasir M，Kanwal H H，Hussain Q，Riaz M，Sajjad M，Rong J K，Jiang Y R. Status and prospects of genome-wide association studies in cotton. Frontiers in Plant Science，2022，13：1019347

[16] Sertse D，You F M，Ravichandran S，Braulio，Soto-Cerda B J，Duguid S，Cloutier S. Loci harboring genes with important role in drought and related abiotic stress responses in flax revealed by multiple GWAS models. Theoretical and Applied Genetics，2021，134：191-212

[17] Li X，Guo D，Xue M，Li G，Yan Q，Jiang H，Liu H，Chen J，Gao Y，Duan L，Xie L. Genome-wide association study of salt tolerance at the seed germination stage in flax (*Linum usitatissimum* L.). Genes，2022，13(3)：486

[18] You F M，Rashid K Y，Zheng C，Khan N，Li P，Xiao J，He L，Yao Z，Cloutier S. Insights into the genetic architecture and genomic prediction of powdery mildew resistance in flax (*Linum usitatissimum* L.). International Journal of Molecular Sciences，2022，23(9)：4960

[19] Xie D W，Dai Z G，Yang Z M，Tang Q，Sun J，Xue Y，Song X X，Lu Y，Zhao D B，Zhang L G，Su J G. Genomic variations and association study of agronomic traits in flax. BMC Genomics，2018，19：512

[20] Li G，Wang Y N，Kwon S W，Park Y J. Association analysis of seed longevity in rice under conventional and high-temperature germination conditions. Plant Systematics and Evolution，2014，300(3)：389-402

[21] Huang X H，Kurata N，Wei X H，Wang Z X，Wang A H，Zhao Q，Zhao Y，Liu K Y，Lu H Y，Li W J，Guo Y L，Lu Y Q，Zhou C C，Fan D L，Weng Q J，Zhu C L，Huang T，

Zhang L, Wang Y C, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X P, Xu Q, Dong G J, Zhan Q L, Li C Y, Fujiyama A, Toyoda A, Lu T T, Qi F, Qian Q, Li J Y, Han B. A map of rice genome variation reveals the origin of cultivated rice. Nature, 2012, 490(7421):497-501

[22] Bakel H V, Stout J M, Cote A, Clarke C M, Sharpe A G, Hughes T R, Page J E. The draft genome and transcriptome of *Cannabis sativa*. Genome Biology, 2011, 12: R102

[23] 徐益,张力岚,祁建民,张列梅,张立武.主要麻类作物基因组学与遗传改良:现状与展望.作物学报, 2021, 47(6): 997-1019

Xu Y, Zhang L L, Qi J M, Zhang L M, Zhang L W. Genomics and genetic improvement of major hemp crops: Current status and prospects. Acta Agronomica Sinica, 2021, 47(6): 997-1019

[24] Soorni A, Fatahi R, Haak D, Salami S A, Bombarely A. Assessment of genetic diversity and population structure in Iranian Cannabis germplasm. Scientific Reports, 2017, 7(1): 15668

[25] Kovalchuk I, Pellino M, Rigault P, Velzen R V, Ebersbach J, Ashnest J, Mau M, Schranz E, Alcorn J, Laprairie R B, McKay J K, Burbridge C, Schneider D, Vergara D, Kane N C, Sharbel T F. The genomics of *Cannabis* and its close relatives. Annual Review of Plant Biology, 2020, 71(1): 713-739

[26] Murray A W. Can gene-inactivating mutations lead to evolutionary novelty? Current Biology, 2020, 30(10): R465-R471

[27] Ren G P, Zhang X, Li Y, Ridout K, Serrano-Serrano M L, Yang Y Z, Liu A, Ravikanth G, Nawaz M A, Mumtaz A S, Salamin N, Fumagalli L. Large-scale whole-genome resequencing unravels the domestication history of *Cannabis sativa*. Science Advances, 2021, 7(29): eabg2286

[28] Gao S, Wang B S, Xie S S, Xu X Y, Zhang J, Li P, Yu Y Y, Yang W F, Zhang Y. A high-quality reference genome of wild *Cannabis sativa*. Horticulture Research, 2020, 7: 73

[29] Petit J, Salentijn E M J, Paulo M J, Denneboom C, Trindade L M. Genetic architecture of flowering time and sex determination in hemp (*Cannabis sativa* L.): A genome-wide association study. Frontiers in Plant Science, 2020, 11: 569958

[30] Pedró J P, Salentijn E M J, Paulo M J, Denneboom C, Loo E N V, Trindade L M. Elucidating the genetic architecture of fiber quality in hemp (*Cannabis sativa* L.) using a genome-wide association study. Frontiers in Genetics, 2020, 11: 566314

[31] Woods P, Campbell B J, Nicodemus T J, Cahoon E, Mullen J L, McKay J K. Quantitative trait loci controlling agronomic and biochemical traits in *Cannabis sativa*. Genetics, 2021, 219(2):iyab099

[32] Zhong R Q, Ye Z H. Secondary cell walls: Biosynthesis, patterned deposition and transcriptional regulation. Plant and Cell Physiology, 2014, 56(2):195-214

[33] Behr M, Legay S, Žižková E, Motyka V, Dobrev P I, Hausman J F, Lutts S, Guerriero G. Studying secondary growth and bast fiber development: The hemp hypocotyl peeks behind the wall. Frontiers in Plant Science, 2016, 7:1733

[34] Han L B, Li Y B, Wang H Y, Wu X, Li C L, Luo M, Wu S J, Kong Z, Pei Y, Jiao G L, Xia G X. The dual functions of WLIM1a in cell elongation and secondary wall formation in developing cotton fibers. The Plant Cell, 2013, 25(11):4421-4438

[35] Zhao Y, Sun Y F, Cao K, Zhang X Y, Bian J, Han C W, Jiang Y, Xu L, Wang X N. Combined use of specific length amplified fragment sequencing (SLAF-seq) and bulked segregant analysis (BSA) for rapid identification of genes influencing fiber content of hemp (*Cannabis sativa* L.). BMC Plant Biology, 2022, 22(1):250

[36] Bakro F, Wielgusz K, Bunalski M, Jedryczka M. An overview of pathogen and insect threats to fibre and oilseed hemp (*Cannabis sativa* L.) and methods for their biocontrol. Integrated Control in Oilseed Crops, 2018, 136:9-20

[37] Cranshaw W, Halbert S, Favret C, Britt K, Miller G. *Phorodon cannabis* Passerini (Hemiptera: Aphididae), a newly recognized pest in North America found on industrial hemp. Insecta Mundi, 2018, 662:1-12

[38] Cranshaw W, Schreiner M, Britt K, Kuhar T P, McPartland J, Grant J. Developing insect pest management systems for hemp in the United States: A work in progress. Journal of Integrated Pest Management, 2019, 10(1):26

[39] Ajayi O S, Samuel-Foo M. Hemp pest spectrum and potential relationship between helicoverpa zea infestation and hemp production in the United States in the face of climate change. Insects, 2021, 12(10):940

[40] Stack G M, Toth J A, Carlson C H, Cala A R, Marrero-González M I, Wilk R L, Gentner D R, Crawford J L, Philippe G, Rose J K C, Viands D R, Smart C D, Smart L B. Season-long characterization of high-cannabinoid hemp (*Cannabis sativa* L.) reveals variation in cannabinoid accumulation, flowering time, and disease resistance. Global Change Biology Bioenergy, 2021, 13:546-561

[41] Lyngkjær M F, Newton A C, Atzema J L, Baker S J. The barley mlo-gene: An important powdery mildew resistance source. Agronomie, 2000, 20:745-756

[42] Adams D J. Fungal cell wall chitinases and glucanases. Microbiology, 2004, 150:2029-2035

[43] Jongedijk E, Tigelaar H, Roekel J S C V, Bres-Vloemans S A, Dekker I, Elzen P J M, Cornelissen B J C, Melchers L S. Synergistic activity of chitinases and β-1,3-glucanases enhances fungal resistance in transgenic tomato plants. The Methodology of Plant Genetic Manipulation: Criteria for Decision Making, 1995, 3:173-180

[44] McKernan K, Helbert Y, Kane L T, Ebling H, Zhang L, Liu B, McLaughlin S, Eaton Z, Kingan S, Baybayan P,

Concepcion G，Jordan M，Riva A，Barbazuk B，Harkins T. Sequence and annotation of 42 cannabis genomes reveals extensive copy number variation in cannabinoid synthesis and pathogen resistance genes. Bio-archive，2020，DOI：10.1101/2020.01.03.894428

［45］ Pépin N，Hebert F O，Joly D L. Genome-wide characterization of the MLO gene family in *Cannabis sativa* reveals two genes as strong candidates for powdery mildew susceptibility. Frontiers in Plant Science，2021，DOI：10.3389/fpls. 2021.729261

［46］ Abot A，Bonnafous C，Touchard F，Thibault F，Chocinski-Arnault L，Lemoine R，Dédaldéchamp F. Effects of cultural conditions on the hemp（*Cannabis sativa*）phloem fibres：Biological development and mechanical properties. Journal of Composite Materials，2013，47（8）：1067-1077

［47］ Mackay C E，Christopher Hall J，Hofstra G，Fletcher R A. Uniconazole-induced changes in abscisic acid，total amino acids，and proline in *Phaseolus vulgaris*. Pesticide Biochemistry Physiology，1990，37（1），74-82

［48］ 姜颖，赵越，孙全军，李振伟.植物生长调节剂在植物生长发育中的应用.黑龙江科学，2018，9(24)：4-7，11
Jiang Y，Zhao Y，Sun Q J，Li Z W. Application of plant growth regulators in plant growth and development. Heilongjiang Science，2018，9(24)：4-7，11

［49］ Jiang Y，Sun Y F，Zheng D F，Han C W，Cao K，Xu L，Liu S X，Cao Y Y，Feng N J. Physiological and transcriptome analyses for assessing the effects of exogenous uniconazole on drought tolerance in hemp（*Cannabis sativa* L.）. Scientific Reports，2021，11（1）：14476

［50］ Gao C C，Cheng Z H，Zhao L，Yu Y T，Tang Q，Xin P F，Liu T M，Yan Z，Guo Y，Zang G G. Genome-wide expression profiles of hemp（*Cannabis sativa* L.）in response to drought stress. International Journal of Genomics，2018，2018：3057272

［51］ Sun J，Chen J Q，Zhang X Y，Xu G C，Yue Y，Dai Z G，Su J G. Genome-wide association study of salt tolerance at the germination stage in hemp. Euphytica，2023，219：5

［52］ Cao K，Sun Y F，Han C W，Zhang X Y，Zhao Y，Jiang Y，Jiang Y Z，Sun X L，Guo Y X，Wang X N. The transcriptome of saline-alkaline resistant industrial hemp（*Cannabis sativa* L.）exposed to $NaHCO_3$ stress. Industrial Crops and Products，2021，170：113766

［53］ Yin M，Pan G，Tao J，Zhao L N，Zheng L，Jiang H，Tang H J，Li C，Li J J，Chen A G，Deng Y，Zhang C P，Li D F，Husng S Q. Genome-wide identification of MYB gene family reveals their potential functions in cadmium stress response and the regulation of cannabinoid biosynthesis in hemp（*Cannabis sativa* L.）. Preprint，2021，http：//dx.doi.org/10.21203/rs.3.rs-156720/v2

［54］ Allen K D，Torres A，Cesare K D，Gaudino R. Evolution，expansion and characterization of cannabinoid synthase gene family in Cannabis sativa.Bio-archive，2022，http：//dx.doi.org/10.1101/2022.11.18.517131

［55］ Cheng Y F，Kang N，Chen Y Z，Hou C，Yu H B，Yu H T，Chen S L，Guo X T，Dong L L. Identification of histone acetyltransferase genes responsible for cannabinoid synthesis in hemp. Chinese Medical Journal，2023，18（1）：16

［56］ Sirikantaramas S，Morimoto S，Shoyama Y，Ishikawa Y，Wada Y，Shoyama Y，Taura F. The gene controlling marijuana psychoactivity. Journal of Biological Chemistry，2004，279：39767-39774

［57］ Taura F，Sirikantaramas S，Shoyama Y，Yoshikai K，Shoyama Y，Morimoto S. Cannabidiolic-acid synthase，the chemotype-determining enzyme in the fiber-type *Cannabis sativa*. FEBS Letters，2007，581：2929-2934

［58］ Laverty K U，Stout J M，Sullivan M J，Shah H，Gill N，Holbrook L，Deikus G，Sebra R，Hughes T R，Page J E，Bakel H V. A physical and genetic map of *Cannabis sativa* identifies extensive rearrangements at the THC/CBD acid synthase loci.Genome Research，2019，29（1）：146-156

［59］ Weiblen G D，Wenger J，Craft K J，ElSohly M A，Mehmedic Z，Treiber E L，Marks M D. Gene duplication and divergence affecting drug content in *Cannabis sativa*. New Phytologist，2015，208(4)：1241-1250

［60］ Henry P，Khatodia S，Kapoor K，Gonzales B，Middleton A，Hong K，Hilyard A，Johnson S，Allen D，Chester Z，Jin D，Rodriguez Jule J C，Wilson I，Gangola M，Broome J，Caplan D，Adhikary D，Deyholos M K，Morgan M，Hall O W，Guppy B J，Orser C. A single nucleotide polymorphism assay sheds light on the extent and distribution of genetic diversity，population structure and functional basis of key traits in cultivated North American cannabis. Journal of Cannabis Research，2020，2(1)：26

［61］ Page J E，Stout J M. US20190382734A1 Cannabichromenic acid synthase from *Cannabis sativa*. United States：National Research Council of Canada，2017

［62］ Vergara D，Huscher E L，Keepers K G，Givens R M，Cizek C G，Torres A，Gaudino R，Kane N C. Gene copy number is associated with phytochemistry in *Cannabis sativa*. AoB Plants，2019，11：plz074

［63］ Gülck T，Møller B L. Phytocannabinoids：Origins and biosynthesis. Trends in Plant Science，2020，25（10）：985-1004

［64］ Innes P，Vergara D. Genomic description of critical upstream cannabinoid biosynthesis genes. Bio-archive，2022，DOI：10.1101/2022.12.15.520586

［65］ Fulvio F，Paris R，Montanari M，Citti C，Cilento V，Bassolino L，Moschella A，Alberti I，Pecchioni N，Cannazza G，Mandolino G. Analysis of sequence variability and transcriptional profile of cannabinoid synthase genes in *Cannabis sativa* L. chemotypes with a focus on cannabichromenic acid synthase. Plants，2021，10：1857

［66］ Welling M T，Liu L，Kretzschmar T，Mauleon R，Ansari O，King G J. An extreme-phenotype genome - wide association

study identifies candidate cannabinoid pathway genes in *Cannabis*. Scientific Reports，2020，10(1)：18643

[67] Gagne S J，Stout J M，Liu E，Boubakir Z，Clark S，Page J E. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. Proceedings of the National Academy of Sciences of the United States of America，2012，109(31)：12811-12816

[68] Campbell B J，Berrada A F，Hudalla C，Amaducci S，McKay J K. Genotype × environment interactions of industrial hemp cultivars highlight diverse responses to environmental factors. Agrosystems，Geosciences & Environment，2019，2：1-11

[69] Cagirici H B，Andorf C M，Sen T Z. Co-expression pan-network reveals genes involved in complex traits within maize pan-genome. BMC Plant Biology，2022，22：595

[70] Jin S K，Han Z G，Hu Y，Si Z F，Dai F，Lu H，Cheng Y，Li Y Q，Zhao T，Fang L，Zhang T Z. Structural variation（SV）-based pan-genome and GWAS reveal the impacts of SVs on the speciation and diversification of allotetraploid cottons. Molecular Plant，2023，16(4)：678-693

[71] Weisweiler M，Arlt C，Wu B Y，Inghelandt D V，Hartwig T，Stich B. Structural variants in the barley gene pool：Precision and sensitivity to detect them using short-read sequencing and their association with gene expression and phenotypic variation. Theoretical and Applied Genetics，2022，135(10)：3511-3529

[72] Sukegawa S，Nureki O，Toki S，Saika H. Genome editing in rice mediated by miniature size Cas nuclease SpCas12f. Frontiers in Genome Editing，2023，DOI：http://dx.doi.org/10.3389/fgeed.2023.1138843

[73] Tiwari J，Singh A，Behera T. CRISPR/Cas genome editing in tomato improvement：Advances and applications. Frontiers in Plant Science，2023，14：1121209

[74] Farooq S，Shahzadi A，Razzaq A，Saleem F，Wani S H，Sandhu K. Advances in genome editing for maize improvement. Springer，2023，DOI：10.1007/978-3-031-21640-4_9