

基于流式细胞术和基因组 Survey 的茜草基因组大小测定及特征分析

李文文, 练从龙*, 郭晨宁, 蒙丹丹, 陈随清

(河南中医药大学药学院, 河南省中药资源与中药化学重点实验室, 豫药全产业链研发河南省协同创新中心, 河南 郑州 450046)

摘要: **目的** 测定茜草的基因组大小并分析其特征。**方法** 以番茄为内参, 流式细胞术测定茜草基因组大小; 基因组 Survey 分析获得茜草基因组大小、杂合率、GC 含量等生物学信息。**结果** 流式细胞术估算茜草基因组大小约为 1.72 Gb。通过 Survey 分析获得茜草的高质量数据为 23.7 Gb, 估算茜草基因组大小约为 1.60 Gb。根据 K-mer 分布情况, 估算茜草重复序列比例为 20.24%, 杂合率为 0.89%, 基因组 GC 含量为 37.59%。从基因组结构特征分析来看, 茜草基因组属于低重复、高杂合基因组。**结论** 本实验初步确定了茜草基因组大小约为 1.60~1.72 Gb, 为进一步绘制茜草全基因组图谱奠定了基础, 可为后续开展全基因组测序时的策略选择提供理论参考, 也为茜草品种选育、规模化种植提供科学依据。

关键词: 茜草; 基因组; 流式细胞术; K-mer 分析

中图分类号: R282

文献标志码: A

文章编号: 1001-1528(2026)06-2111-05

doi: 10.3969/j.issn.1001-1528.2026.06.047

茜草 *Rubia cordifolia* L. 为茜草科茜草属植物, 具凉血止血、祛瘀通经等功效^[1-3]。目前, 茜草在分子药理学与分子遗传学方面的研究主要集中于叶绿体基因组、基因克隆等领域, 但关于茜草属基因组大小的研究鲜有报道, 目前针对茜草的全基因组测序相关工作尚未完成, 在一定程度上限制了茜草活性成分生物合成的发展。因此, 对茜草开展全基因组测序, 掌握其全基因组信息尤其必要, 而了解其基因组大小和特征信息是其开展全基因组解析的首要步骤。

基因组大小是基因组学研究基础。多数植物进化中经历多倍化以稳定基因组。评估植物基因组大小可助力基因组组装, 揭示变化机制, 深入研究遗传、进化和次生代谢过程, 为转录和代谢产物研究提供资料, 推动药用植物分子生物学发展。药用植物领域中, 参考基因组参数研究的重要性日益凸显^[4-6]。测定基因组大小的方法有实时荧光定量 PCR 法、流式细胞法、基因组 Survey 分析等^[7], 其中流式细胞仪测定法有高速、准确性好、效率高的优点^[8]。基因测序技术发展推动基因组调查技术广泛应用, 该技术高效、可靠、经济, 可获取基

因组大小、GC 含量、重复序列比例、杂合率及外源污染等信息, 相比传统技术, 能提供更丰富的遗传学背景^[9]。因此, 本实验拟运用流式细胞术与基因组 Survey 技术评估茜草基因组大小并分析基因组特征, 以期为后期茜草基因组的测序、组装策略、图谱绘制、功能基因挖掘解析等研究奠定理论基础。

1 材料

1.1 药材 茜草幼嫩叶片采自河南中医药大学药用植物园, 由河南中医药大学陈随清教授鉴定为正品。

1.2 试剂 RNAase (批号 9001-99-4)、Tris (批号 1115GR500)、SDS (批号 3250GR500)、EDTA (批号 1108GR500)、氯化钠 (批号 1249GR500)、Urea (批号 1152GR500) (德国 BioFROXX GmbH 公司); CATB (批号 BS950-500, 北京兰杰柯科技有限公司); 蛋白酶 K (批号 RP02503)、DNA Marker D15000 (批号 MD110-02)、碘化丙啶 (propidium iodide, PI) (武汉爱博泰克生物科技有限公司)。异丙醇 (批号 80109218)、无水乙醇 (批号 10009218) (国药集团化学试剂有限公司)。

收稿日期: 2025-01-20

基金项目: 河南省创新生态支持专项项目 (豫财科 [2024] 8 号)

作者简介: 李文文 (2001—), 女, 硕士在读, 研究方向为中药学。E-mail: 17838123682@163.com

* **通信作者:** 练从龙 (1988—), 男, 博士, 副教授, 研究方向为药用植物生理与分子调控机制。E-mail: lian1988@hactcm.edu.cn

网络出版日期: 2025-5-26

网络出版地址: <http://kns.cnki.net/kcms/detail/31.1368.R.20250625.1753.004.html>

1.3 仪器 Tissuelyser 多样品组织研磨仪 (上海净信实业发展有限公司); VRTEX-5 涡旋混匀器 (海门市其林贝尔仪器制造有限公司); H1850 离心机 (湖南湘仪实验室仪器开发有限公司); BD FACScalibur 流式细胞仪 (美国 BD 公司); DYY-6C 电泳仪 (北京六一生物科技有限公司); ROS-UV15 水机纯水仪 (武汉吉百瑞科技有限责任公司); Tanon-1600 紫外拍胶仪 (上海天能科技有限公司)。

2 方法

2.1 细胞核悬浮液制备 分别取番茄、茜草幼嫩叶片,洗净后置于预冷的培养皿中,加入预冷的细胞裂解液,静置 10 min。切碎叶片,移液枪上下吹打混合,42 μm 尼龙筛网过滤,低速冷冻离心,弃去上清液,收集细胞核沉淀,重新悬浮于细胞裂解液中,4 $^{\circ}\text{C}$ 保存备用,即得。

2.2 流式细胞术测定茜草基因组大小 在制备好的细胞核悬浮液中加入 PI 和 RNAase 溶液,混匀,置于冰上遮光染色 0.5~1 h 后备用。流式细胞仪检测细胞核悬浮液在 488 nm 波长处对 FL2 通道发出的荧光强度,每次检测均涵盖 10 000 个颗粒。确保变异系数在 5% 以内。采用 Modifit 3.0 软件对数据进行分析,以番茄的基因组为内参,根据荧光强度峰值测定茜草基因组大小,公式为茜草 DNA 量=番茄 DNA 量 \times (茜草荧光强度/番茄荧光强度)^[10]。

2.3 基因组 DNA 提取与基因组测序 采用传统 CTAB 法提取茜草叶片的基因组 DNA,检测基因组 DNA 浓度、纯度、完整性。检测参数为胶浓度 0.7%;电压 150 V;电泳时间 30 min;以 D15000 DNA 为 Marker;上样量 60 ng。Covaris 超声波破碎仪将 $\text{OD}_{260}/\text{OD}_{280} \geq 1.8$ 的茜草基因组 DNA 样品精

准、随机地打断成所需的片段,经末端修复、A 尾添加、测序接头连接、纯化、PCR 扩增等构建有效片段测序文库,DNBSEQ-T7 平台对文库进行双端测序。将原始数据经过二代过滤,将包括 PCR 扩增造成的重复序列,低质量(质量值 ≤ 5)碱基数目、N 碱基含量超过 5%、含有接头序列污染的短片段序列测序数据剔除。将得到的高质量测序数据将用于进一步的信息分析。

2.4 基因组大小、重复序列和杂合率预测 检测合格的 DNA 样品由武汉贝纳科技有限公司进行建库测序。选取 K 值为 19 的序列进行分析,Jellyfish v2.2.10 软件对序列进行频数与深度作图,根据 K-mer 频数-深度分布曲线,计算基因组大小,公式为基因组大小=总碱基数/平均测序深度=总 K-mer 数量/平均 K-mer 深度。

2.5 GC 含量分布分析 将过滤得到的测序数据采用 SOAPdenovo 2.04-r241 软件进行拼接组装,统计 GC 含量,判断基因组精确组装难易程度。

2.6 样品污染评估 从过滤后的有效数据文库中随机抽取 50 000 条单端 Reads 数据,与 NCBI 核苷酸数据库(NT 库)进行 BLASTN 比对^[11],查看近缘物种,判断是否存在外源污染。

3 结果

3.1 茜草基因组大小 针对番茄(基因组大小为 0.9 Gb)和茜草的荧光发射强度开展测定与分析,结果见图 1。由此可知,番茄和茜草的峰形细而尖锐,两者测定峰位不存在相互干扰的重叠现象,且粒子团聚现象明显,背景碎片化程度较低,表明番茄与茜草区分度良好,确保采用番茄做为内参的可靠性。番茄的荧光强度为 25.45,样品荧光强度为 48.52,比值为 1.91,据此测算出茜草的基因组大小为 1.72 Gb。

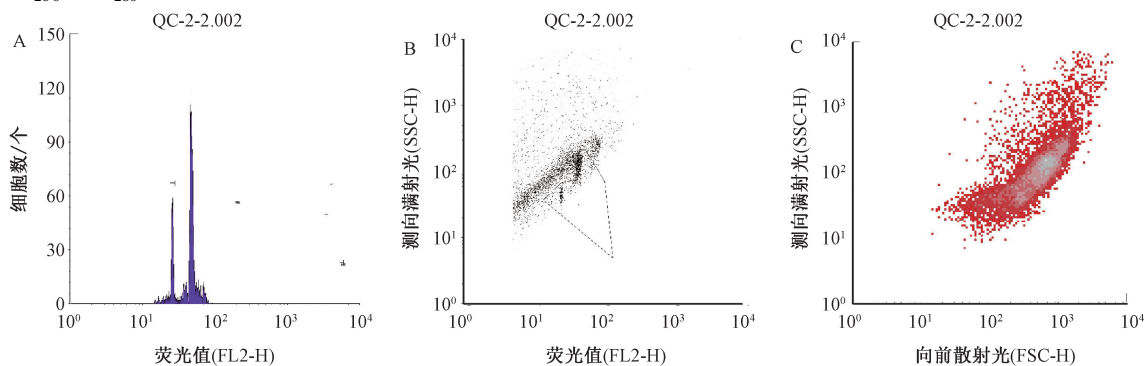
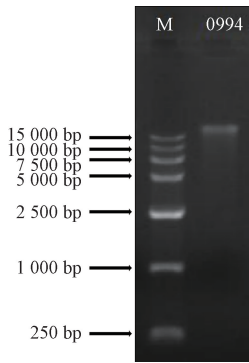


图 1 茜草流式细胞术检测结果

3.2 基因组 DNA 质量分析 茜草基因组 DNA 电泳图(图 2)主带展现清晰、亮度均匀,证明茜草

DNA 具有较高的完整性,但有轻微降解。进一步检测发现,所提取的茜草 DNA 质量浓度为 159.8

ng/ μ L, 总量 2.71 μ g, OD₂₆₀/OD₂₈₀ 为 2.03, OD₂₆₀/OD₂₃₀ 为 1.57。提示所提取的茜草基因组 DNA 整体质量合格, 满足建库要求, 可以用于后续实验。



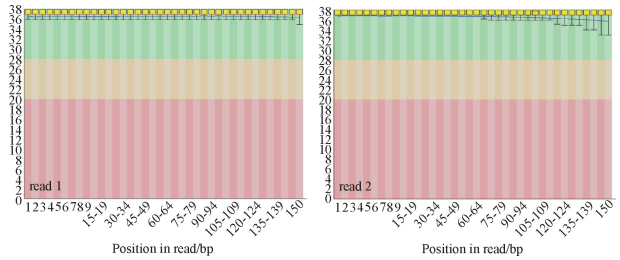
注: 左侧为 DNA 分子量标准, 右侧为茜草基因组 DNA 样品。

图2 茜草基因组 DNA 电泳图

3.3 测序数据统计及 GC 质量分析 本研究共获取 160 622 466 条原始序列, 24 093 369 900 个碱基。G/C 碱基数占总碱基数量的百分比为 37.77%, *Q* 值大于 20、30 的碱基占总体碱基的 98.52%、96.08%。经过高通量测序与高效的过滤、修正技术, 最终得到高质量数据 160 552 364 条序列, 占原始序列的 99.96%; 23 741 484 231 个碱基, G/C 碱基数占总碱基数量的百分比为 37.59%, 显示出极高的数据质量和可靠性。同时, 总测序深度约为 14.84。测序数据 *Q* 值大于 20、30 的碱基占总体碱基的 98%、96% 以上, 超过正常数值。表明测序质量优异, 满足进一步组装和 Survey 分析要求。

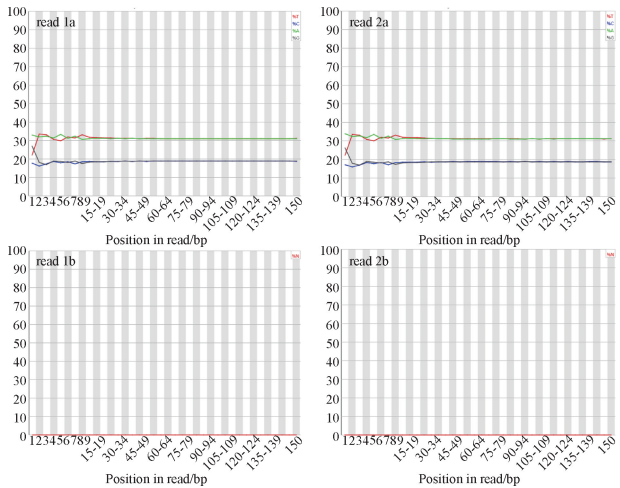
碱基质量值越高, 被测错的概率就越小。由图 3 可知, 茜草基因组中大部分测序数据质量值均大于 34, 表明茜草基因组测序序列的质量较高, 测序结果可信度较高。由图 4 可知, 前部分位置的分布线出现较大波动, 这可能是 PCR 扩增过程导致的差异。后续碱基 G 和 C 的占比趋于一致, A 和 T 的占比同样呈现出相似的规律, 并且碱基 N 占比几乎为 0, 说明测序的碱基质量正常。由图 5 可知, 平均 GC 含量实际分布呈单峰, 曲线形态接近正态, 与理论结果相似, 表明测序数据未受到外源污染, 质量良好。根据 GC 含量测定结果可以初步判断二代基因组测序过程中影响偏向性的随机性影响程度。经计算分析, 得到修正后茜草基因组 GC 含量为 37.59%, 较为适中, 没有明显的 GC 偏向性, 结果可靠。

3.4 基因组大小、杂合度和重复序列比例评



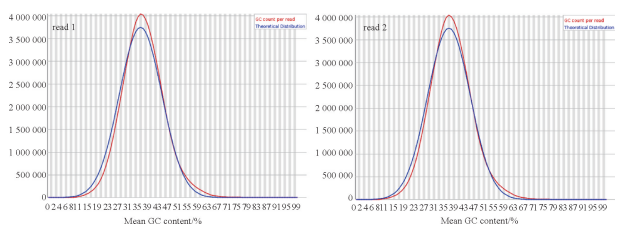
注: 正、反向测序数据分别为 read 1、read 2, 红色线条为质量值中位数, 蓝色线条为质量值平均数, 箱图线的上下线分别是质量值的最大值和最小值。

图3 质控后的茜草碱基质量分布图



注: 红色、蓝色、绿色、黑色线条分别表示 T、C、A、G 碱基的含量。

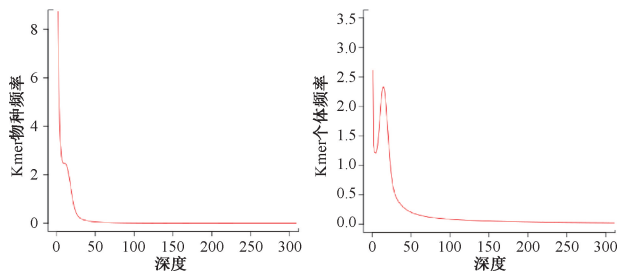
图4 质控后的茜草碱基含量分布图



注: 红色、蓝色线条分别为 GC 含量分布图、正态理论分布图。

图5 质控后的茜草 GC 含量分布图

估 K-mer分布见图 6。由此可知, K-mer 曲线在深度为 12 附近出现一个主峰, 依据分布情况推测可能为单倍体。在主峰后无明显拖尾现象, 表明其有较低的重序列。利用 GCE 软件统计得到 K-mer 长度为 19, K-mer 总数为 19 557 006 962, K-mer 平均深度为 12.22。经过基因组大小公式计算后得到茜草基因组大小为 1 600.11 Mb, 有效数据量为 23.74, 测序的深度为 14.84。根据 K-mer 深度分布, 估算重复序列占比为 20.24%, 通过计算杂合峰值与纯合峰值之比, 得到茜草基因组杂合率为 0.89%。



注：横坐标为 K-mer=19 时的深度，纵坐标为不同深度的频率。

图 6 K-mer 分布图

综上所述，K-mer 分析得到的茜草基因组大小与流式细胞术测定结果接近，测定结果相对准确，具有一定的可靠性。通过基因组基本结构特征的数据分析，茜草基因组具有低重复、高杂合的基因组特性，适合做单倍体基因组分析。

3.5 样品污染评估 50 000 条序列中成功与数据库比对的数量为 5 668 条，物种分类结果显示，比对次数最多的物种是 *R. cordifolia*，共 2 172 条，见表 1。且茜草属植物的 DNA 序列在比对结果中占据多数，并未发现异常比对结果，表明本次茜草 DNA 样品测序过程中未受到外源性污染的影响。

表 1 NT 数据库对比结果物种分类

种名	科名	序列数/条	占比/%
茜草 <i>Rubia cordifolia</i> L.	茜草科	2 172	4.34
田茜 <i>Sherardia arvensis</i> L.	茜草科	1 653	3.31
糙叶野丁香 <i>Leptodermis scabrida</i> Hook. f.	茜草科	239	0.48
茜草属植物 <i>Rubia horrida</i>	茜草科	140	0.28
藜 <i>Chenopodium album</i> L.	藜科	54	0.11
鸡屎藤 <i>Paederia foetida</i> L.	茜草科	51	0.10
篇蓄 <i>Polygonum aviculare</i> L.	蓼科	42	0.08
欧亚路边青 <i>Geum urbanum</i> L.	蔷薇科	39	0.08
苹果 <i>Malus pumila</i> Mill.	蔷薇科	37	0.07
番茄 <i>Solanum lycopersicum</i> L.	茄科	32	0.06
盔状黄芩 <i>Scutellaria galericulata</i> L.	唇形科	31	0.06
洋常春藤 <i>Hedera helix</i> L.	五加科	30	0.06
黑夏至草 <i>Ballota nigra</i>	唇形科	27	0.05
欧洲水青冈 <i>Fagus sylvatica</i> L.	壳斗科	23	0.05
欧洲假繁缕 <i>Theligonum cynocrambe</i> L.	茜草科	22	0.04
北方拉拉藤 <i>Galium boreale</i> L.	茜草科	22	0.04
锦熟黄杨 <i>Buxus sempervirens</i> L.	黄杨科	21	0.04
小花金鱼草 <i>Misopates orontium</i> L.	车前科	20	0.04
脬果沙拐枣 <i>Calligonum junceum</i> Litv.	蓼科	19	0.04

4 讨论

本实验结合流式细胞术和基因组 Survey 分析 2 种方法，以精确测定茜草基因组大小及其特征，从而提升结果的准确性。通过测序结果和 K-mer 分析计算，得到单倍体茜草基因组大小为 1.60 Gb，此结果与流式细胞术估算的 1.72 Gb 相比略微偏小，这可能是由于流式细胞术估测的结果是以已知基因

组大小的对照样品为参考获得的，所得结果存在细微偏差。

本实验预测的茜草科植物茜草的基因组大小与茜草科植物短小蛇根草 (439.90 Mb)^[12]、大红戟 (446.30 Mb)^[13]、巴戟天 (484.85 Mb)^[14] 的基因组大小结果有较大差异，可能是因为茜草科植物基因组在演化中经历了增减的动态平衡以适应环境环境变化，尤其是转座子序列在基因组中的数量和活跃程度，对基因组大小具有显著影响^[15]。

茜草基因组 GC 含量约为 37.59%，此结果与茜草叶绿体基因组 GC 含量结果相似^[16]，与已有报道茜草科植物的 GC 含量接近^[17-18]。重复序列比例为 20.24% 属于低重复基因组，与蒲昕颖等^[17] 所得的结果相似。杂合率为 0.89% (低杂合率为小于 0.5%) 属于高杂合基因组，高杂合率基因组在研究遗传学和进化生物学等领域中具有深远意义，也是生物多样性的重要组成部分之一。通过 DNBSEQ-T7 高通量测序技术完成了茜草基因组 Survey 分析，为下一步开展基因组图谱绘制、挖掘重要功能基因提供了重要的有效数据指导。

5 结论

本实验以番茄作为参考植物，流式细胞术预估茜草基因组大小为 1.72 Gb；全基因组 Survey 分析得到茜草基因组大小为 1.60 Gb，由此估测茜草的基因组大小范围在 1.60~1.72 Gb 之间。茜草重复序列比例为 20.24%，杂合率为 0.89%，基因组具有低重复、高杂合的特点。

参考文献：

[1] 李文文, 郭凯华, 练从龙, 等. 茜草种子质量检验方法与分级标准研究[J]. 种子, 2024, 43(5): 146-156.

[2] 李海峰, 肖凌云, 张菊, 等. 茜草化学成分及其药理作用研究进展[J]. 中药材, 2016, 39(6): 1433-1436.

[3] 刘路, 王佳佳, 崔雪靖, 等. 茜草属植物化学成分研究进展[J]. 中国实验方剂学杂志, 2015, 21(17): 225-229.

[4] 张云燕, 安宇, 林峰, 等. 基于流式细胞术和 K-mer 分析的银缕梅属 (*Parrotia* C. A. Mey.) 植物基因组大小测定[J]. 植物遗传资源学报, 2021, 22(2): 561-570.

[5] 伍文丹. 老芒麦 (*Elymus sibiricus* L.) 基因组 Survey 及 gSSR 标记开发与应用[D]. 雅安: 四川农业大学, 2019.

[6] 毛常清, 沙秀芬, 黄静, 等. 川芎基因组 survey 测序及其特征分析[J]. 中草药, 2023, 54(3): 907-914.

[7] 马秋月, 王玉斌, 李倩中, 等. 基于流式细胞术和 K-mer 方法测定 6 种蕨属植物基因组大小[J]. 南京林业大学学报 (自然科学版), 2023, 47(1): 163-170.

[8] 王利虎, 张琼, 陈凯, 等. 流式细胞术在植物倍性鉴定及基因组大小估测中的应用策略[J]. 分子植物育种,

- 2021, 19(17): 5833-5841.
- [9] 王雪, 周佳熠, 孙会改, 等. 新疆沙冬青基因组调查测序与基因组大小预测[J]. 植物遗传资源学报, 2018, 19(1): 143-149.
- [10] 张苏芳, 张真, 王鸿斌, 等. 我国几种重要松毛虫基因组大小测定[J]. 林业科学研究, 2014, 27(5): 583-589.
- [11] 刘奇, 孟珍, 刘勇, 等. 基于BLAST的数据清洗与质量控制方案[J]. 计算机工程, 2011, 37(4): 73-75.
- [12] Rai A, Hirakawa H, Nakabayashi R, et al. Chromosome-level genome assembly of *Ophiorrhiza pumila* reveals the evolution of camptothecin biosynthesis[J]. *Nat Commun*, 2021, 12(1): 405.
- [13] Zhang Y M, Zhang F, Jin L, et al. A chromosome-level genome assembly of the *Knoxia roxburghii* (Rubiaceae) [J]. *Sci Data*, 2023, 10(1): 803.
- [14] Wang J H, Xu S Q, Mei Y, et al. A high-quality genome assembly of *Morinda officinalis*, a famous native southern herb in the Lingnan region of southern China[J]. *Hortic Res*, 2021, 8(1): 135.
- [15] 陈建军, 王瑛. 植物基因组大小进化的研究进展[J]. 遗传, 2009, 31(5): 464-470.
- [16] 陈晓颖, 胡本祥, 史嘉周, 等. 茜草叶绿体全基因组序列及其系统发育分析[J]. 西北植物学报, 2023, 43(11): 1855-1865.
- [17] 蒲昕颖, 陈晓娇, 王欣格, 等. 基于流式细胞术和基因组Survey的红大戟基因组大小测定和分析[J/OL]. 分子植物育种: 1-13 (2022-10-28) [2025-05-20]. <http://kns.cnki.net/kcms/detail/46.1068.S.20221028.1008.002.html>.
- [18] Wang W C, Shao F, Deng X Q, et al. Genome surveying reveals the complete chloroplast genome and nuclear genomic features of the crocin-producing plant *Gardenia jasminoides* Ellis[J]. *Gene Res Crop Evol*, 2020, 68(3): 1-16.

基于文献计量学的白降丹本草考证

梁晓曦¹, 杨宇¹, 王瑾^{2*}, 黄勤挽^{3*}

(1. 成都中医药大学药学院, 四川 成都 611137; 2. 成都中医药大学民族医药学院, 四川 成都 611137; 3. 成都中医药大学现代中药产业学院, 四川 成都 611137)

摘要: 目的 系统梳理白降丹的名称、处方组成、炼制工艺及功能主治的历史沿革, 为其现代临床应用提供文献依据。**方法** 查阅历代本草、医籍、方书及近现代相关文献, 结合文献计量学手段, 对白降丹的相关内容进行了文献整理与考证分析。**结果** 白降丹以“白降丹”为正名, 别称包括“白灵药”“水火丹”等; 基本处方由水银、白矾、皂矾、火硝、盐组成, 常辅以朱砂、雄黄、硼砂增强疗效。炼制工艺分为备料、结胎、封固、冷却、降丹、去火毒及收丹。古代用于外科病症, 现代应用扩展至肿瘤、感染性疾病、骨科、皮肤病等诸多领域。**结论** 白降丹组方与炼制工艺历史沿革清晰, 临床应用广泛。现代使用应在继承传统基础上注意用药安全, 成人连续使用每次不宜超过 78 mg, 以保障疗效与安全性。

关键词: 白降丹; 历史沿革; 文献计量学; 处方组成; 炼制工艺; 可视化

中图分类号: R283

文献标志码: A

文章编号: 1001-1528(2026)06-2115-07

doi: 10.3969/j.issn.1001-1528.2026.06.048

作为中医外科的重要药物形式, 丹药源远流长, 在历代本草有大量记载。其中, 白降丹因其显著的腐蚀与解毒功效, 被誉为“外科三宝”之一, 常用于疮疡、痈疽等病证, 在明清达到应用鼎盛时期, 其主要成分包括水银、火硝、白矾, 精炼后形成了一种主要由氯化汞(HgCl_2)组成的化学制剂, 被视为早期外用抗菌剂的代表。

尽管白降丹在传统医学中地位显著, 但因药物毒性较强且难以精准掌控剂量, 随着现代药物的发展与替代, 白降丹逐渐淡出临床, 迄今仍无同类药物在治疗疮疡肿毒方面能完全取代白降丹的功效^[1]。然而现有研究多集中于丹剂类药物的源流考证, 对白降丹的名称演变、处方组成、炼制工艺及临床应用经验尚缺乏系统梳理与深入探讨。本研

收稿日期: 2025-11-05

基金项目: 国家中医药管理局 2022 年中药炮制技术传承创新项目 (GZY-KJS-2022-052)

作者简介: 梁晓曦 (2002—), 女, 硕士在读, 从事中药炮制与制剂研究。E-mail: liangxiaoxi@stu.cdutcm.edu.cn

* 通信作者: 王瑾 (1983—), 女, 博士, 教授, 从事中药炮制与制剂研究。E-mail: Wangjin0816@126.com

黄勤挽 (1979—), 男, 博士, 教授, 从事中药炮制与制剂研究。E-mail: huangqinwan@cdutcm.edu.cn