

基于 SWChen 系统预测靶向林麝 (*Moschus berezovskii*) *FASN*、*HMGCR* 和 *CYP11A1* 基因 miRNAs 的研究*

谌颖莲, 曾德军, 赵贵军, 封孝兰, 张承露, 吴佳勇, 竭航**
(重庆市药物种植研究所, 特色生物资源研究与利用川渝共建重点实验室, 重庆 408435)

摘要:【目的】预测靶向林麝 (*Moschus berezovskii*) *FASN*、*HMGCR* 和 *CYP11A1* 基因的 miRNAs, 为麝香的生物合成机制研究提供参考。【方法】通过林麝麝香的代谢组学测序, 获取麝香中的主要化学组分, 确定相应化学组分合成的关键酶; 通过转录组测序获取林麝基因 mRNA 序列及 miRNAs 序列, 利用自主编写的离线 SWChen 系统对调控关键酶基因表达的 miRNA 进行预测, 并以人 *FASN* 基因对系统准确性进行验证, 同时以牛 (*Bos taurus*)、山羊 (*Capra hircus*) 和绵羊 (*Ovis aries*) 预测结果进行比较和集合论分析, 获取调控林麝基因的候选 miRNAs。【结果】麝香的化学成分主要包括脂肪酸碳链类衍生物、芳香族、固醇类和甾类等衍生物; 预测调控 *FASN* 基因表达的 miRNA 主要为 miR-24-3p 和 miR-30b-3p, 调控 *HMGCR* 基因表达的为 miR-29b、miR-146a、miR-181b-3p 和 miR-429 家族, 调控 *CYP11A1* 基因表达的为 miR-532-3p 和 let-7 家族, 共同调控 3 个基因表达的主要为 miR-205 家族和 miR-143 家族。【结论】上述 miRNAs 为靶向调控林麝 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的候选 miRNAs, 对解析林麝麝香的生物合成机制具有参考价值。

关键词: 林麝; *FASN*; *HMGCR*; *CYP11A1*; miRNA

中图分类号: S 865.411

文献标志码: A

文章编号: 1004-390X (2021) 05-0880-11

Predicting miRNA Targeting *FASN*, *HMGCR* and *CYP11A1* Genes in *Moschus berezovskii* Based on SWChen System

CHEN Yinglian, ZENG Dejun, ZHAO Guijun, FENG Xiaolan,
ZHANG Chenglu, WU Jiayong, JIE Hang

(Key Laboratory of Sichuan and Chongqing Co-construction for Characteristic Biological Resources Research and Utilization, Chongqing Institute of Drug Cultivation, Chongqing 408435, China)

Abstract: [Purpose] To predict the miRNAs targeting *FASN*, *HMGCR* and *CYP11A1* genes in *Moschus berezovskii* for providing the reference about mechanism of musk biosynthesis. [Method] We acquired chemical components of musk to seek out key enzymes by metabolomics sequencing. Then, sequences of *FASN*, *HMGCR* and *CYP11A1* mRNA sequences as well as miRNA sequences data of *M. berezovskii* were obtained by transcriptomic sequencing of musk gland. SW-

收稿日期: 2021-01-15

修回日期: 2021-04-19

网络首发时间: 2021-10-19 12:23:34

*基金项目: 国家自然科学基金项目 (81973428); 重庆市基本科研业务费项目 (cstc2020-jbky001); 重庆市自然科学基金项目 (cstc2019jcyj-msxmX0024)。

作者简介: 谌颖莲 (1992—), 男, 贵州毕节人, 硕士, 研究实习员, 主要从事细胞分子生物学研究。

E-mail: 1361697944@qq.com

**通信作者 Corresponding author: 竭航 (1987—), 男, 四川达县人, 硕士, 副研究员, 主要从事药用动物养殖技术研究。E-mail: jiehangisgood@126.com

网络首发地址: <https://kns.cnki.net/kcms/detail/53.1044.S.20211019.0921.003.html>



Chen system written by ourselves had been used to screen candidate miRNAs of pertinent genes and had been verified the veracity. Meanwhile, the predicted results of *Bos taurus*, *Capra hircus* and *Ovis aries* species were used to compare to get candidate miRNAs regulating genes of *M. berezovskii*.

[**Result**] The chemical substances of *M. berezovskii* mainly included in carbon chain fatty acid derivatives, aromatic, sterols and steroid derivative matters. The predicted results revealed that the miRNAs regulating *FASN* gene expression of *M. berezovskii* mainly constituted as miR-24-3p and miR-30b-3p. The predicted miR-29b, miR-146a, miR-181b-3p and miR-429 family preferentially had influence on *HMGCR* gene expression. Then, miR-532-3p and let-7 family had been displayed as the modulation to *CYP11A1* gene expression. Further, the miR-205 family and miR-143 family could jointly regulate the expression of the three genes. [**Conclusion**] The above predicted miRNAs are candidate regulators targeting the expression of *FASN*, *HMGCR* and *CYP11A1* in *M. berezovskii*, which had the referential value for the analysis of the biosynthesis mechanism of musk in *M. berezovskii*.

Keywords: *Moschus berezovskii*; *FASN*; *HMGCR*; *CYP11A1*; miRNA

麝香是一种珍贵的动物源药材, 具有较高的药用价值和香料价值^[1-4], 但其形成过程复杂且未知, 限制了麝香产业的发展和产香效率的提高。研究表明: 麝香的化学组分包括脂肪酸类的烷烃物质或衍生的醇类、酮类和醛类(如麝香酮、6-甲基庚烷-1,6-二醇和十六烯醛等)以及固醇类、甾类和芳香族类物质(如胆固醇、间甲酚和二氢雌甾酮等)^[2], 这些化学组分对麝香形成至关重要, 同时也是麝香药用和香用价值的关键物质。动物的脂肪酸合酶(fatty acid synthase, *FASN*)、3-羟基-3-甲基-戊二酰辅酶 A 还原酶(3-hydroxy-3-methylglutaryl-CoA reductase, *HMGCR*)和细胞色素 P450 家族 11 亚家族 A1 (*CYP11A1*) 是长链的脂肪酸类衍生物、大环碳链酮类以及芳香族甾类物质合成的关键酶^[5-9]。基于麝香的化学组分和合成相关化学组成的关键酶, 在林麝(*Moschus berezovskii*)中筛选调控 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的调控因子, 对探究麝香可能的生物合成机制具有重要意义。

MicroRNA (miRNA) 是一种长度约为 20 nt 的非编码 RNA, 通过作用靶基因 mRNA 的 3'-UTR 在转录后水平抑制靶基因的表达, 对动物的生长和代谢等过程具有关键的调控作用^[10-11]。筛选调控林麝 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的 miRNAs 能够进一步了解麝香的生物合成过程, 然而现在已知的林麝基因序列信息非常

少, 缺少林麝基因的 mRNA 和 miRNA 序列信息^[1, 12]。由于牛(*Bos taurus*)、山羊(*Capra hircus*)和绵羊(*Ovis aries*)与林麝编码基因的 3'-UTR 序列并不完全一致, 通过同源物种来推测调控林麝基因的 miRNAs 获取信息并不完整, 缺少适合预测林麝 miRNAs 的工具。本研究通过基于 Smith-Waterman 局部序列比对算法^[13]编写的 SWChen 系统对调控林麝 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的 miRNAs 进行预测筛选, 并以人(*Homo sapiens*)的 *FASN* 基因为例, 与 TargetScan、miRDB 和 miRSystem 预测软件和 miRTarBase 数据库比较, 评价 SWChen 系统对 miRNAs 预测能力; 利用代谢组学测序来确定麝香的化学成分, 对与合成相关化学成分有关的 *FASN*、*HMGCR* 和 *CYP11A1* 基因进行 miRNAs 预测, 以林麝、牛、山羊和绵羊预测的结果进行集合论分析确定候选 miRNAs, 以期对麝香的生物合成机制研究提供参考。

1 材料与方法

1.1 试验动物

118 头健康成年雄性林麝(5 岁)饲养于重庆市药物种植研究所林麝养殖基地(N29°, E107°, 海拔 678 m), 饲养采用精饲料和青绿饲料[红薯藤(*Ipomoea batatas*)和光叶海桐(*Pittosporum glabratum*)嫩叶]相结合的方式, 自由采食。精

饲料包括 65% 玉米、25% 大豆、6% 麦麸、0.4% 食盐、1.5% CaCO_3 、0.8% CaHPO_4 、0.15% 维生素和 0.15% 矿物质。每天 17:30 混合精饲料与红薯藤等补充饲喂。

1.2 样品采集

在 5—6 月泌香盛期，采用微创香腺采集法获取林麝香腺组织样本，立刻放入 RNA 保存液中并转移至 $-80\text{ }^{\circ}\text{C}$ 保存；在 11—12 月麝香成熟阶段，通过“活麝取香”法采取林麝成熟麝香 3 份，置于离心管内保存。

1.3 代谢组检测

将采集到的麝香样品送至上海百趣生物有限公司采用 GC-MS 进行非靶标代谢组学检测，使用甲醇—氯仿 (体积比为 3:1) 溶液进行萃取，通过总离子流图信号获取麝香的主要化学成分及相对含量。

1.4 转录组测定及序列分析

通过总 RNA 提取后，送样至上海生工生物科技有限公司采用 Hi-seq 2500 测序。然后将测序数据进行 Trinity 拼接、序列比对和功能注释等获得 *FASN*、*HMGCR* 和 *CYP11A1* 基因 mRNA 序列并进行结构分析，确定 3'-UTR 序列以及林麝 miRNAs 序列数据；以牛、山羊和绵羊为参考，对 3 个基因的 3'-UTR 序列进行比对，并使用 MEGA 6.0 及最大似然法构建进化树；通过 RNA 二级结构分析 (unafold.rna.albany.edu) 了解各物种基因 3'-UTR 序列的差异情况。

1.5 SWChen 系统对 miRNAs 预测准确性的验证

由于目前常用的 miRNAs 预测软件如 miRSystem (mirsystem.cgm.ntu.edu)、TargetScan (www.targetscan.org) 和 miRDB (mirdb.org) 仅预测部分常见物种的 miRNAs，无法对调控林麝基因的 miRNA 进行准确预测，因此，本研究在 miRBase 数据库 (www.mirbase.org) 下载人的 miRNAs 序列，并在 GenBank (www.ncbi.nlm.nih.gov/genbank/) 查询人 *FASN* 基因 3'-UTR 序列 (NM_0041

04.5)，利用 SWChen 系统 (<https://github.com/YLCHEN1992/SWChen>) 对调控 *FASN* 表达的 miRNAs 进行预测，通过与在线软件 miRSystem、TargetScan 和 miRDB 的预测结果比较，以 miRTarbase 数据库 (mirtarbase.cuhk.edu.cn) 中调控 *FASN* 表达的全部 miRNAs (miRTarget_All) 和具有强力证据可调控人 *FASN* 基因表达的 miRNAs (miRTarget_High) 作为参考，检测 SWChen 系统对 miRNA 预测的准确性。

1.6 调控 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的 miRNA 预测

将得到的林麝 mRNA 3'-UTR 序列及获得的林麝 miRNAs 数据库利用 SWChen 系统进行分析，预测调控林麝 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的 miRNAs；在 GenBank 中获取牛、山羊和绵羊 *FASN*、*HMGCR* 和 *CYP11A1* 基因的 mRNA 序列 (表 1)。在 miRBase 数据库获取牛、山羊和绵羊的 miRNA 序列，由于绵羊 miRNA 数量较少，将山羊和绵羊 miRNA 进行整合得到羊的 miRNA 数据库，共同用于山羊和绵羊 miRNA 的预测分析。利用 SWChen 系统预测调控牛、山羊和绵羊 *FASN*、*HMGCR* 和 *CYP11A1* 基因的候选 miRNAs，并通过双阈值 (种子序列匹配评分及亲和程度评分同时考虑) 预测调控基因表达的 miRNAs，以增加预测精度，最后将预测结果进行集合论分析以期获取麝香合成相关和候选 miRNAs 分子。

2 结果与分析

2.1 SWChen 系统 miRNA 功能预测结果

由图 1 可知：SWChen 通过种子区匹配预测 (single) 调控人 *FASN* 的基因共有 193 条，其中匹配 miRTarget_All 数据 8 条，匹配 miRTarget_High 数据 3 条，而调高 miRNA 与 mRNA 亲和力判定阈值后 (双阈值) 与 miRTarbase 数据库不匹配；TargetScan 在线软件预测有 745 条，其中匹配 miRTarget_All 数据 15 条，匹配 miRTarget_High

表 1 物种基因序列 GenBank 编号
Tab. 1 GenBank No. of species gene sequence

物种 species	<i>FASN</i>	<i>HMGCR</i>	<i>CYP11A1</i>
牛 <i>Bos taurus</i>	NM_001012669.1	NM_001105613.1	NM_176644.2
山羊 <i>Capra hircus</i>	NM_001285629.1	XM_018053703.1	NM_001287574.1
绵羊 <i>Ovis aries</i>	XM_027974304.1	XM_004005125.3	NM_001093789.1

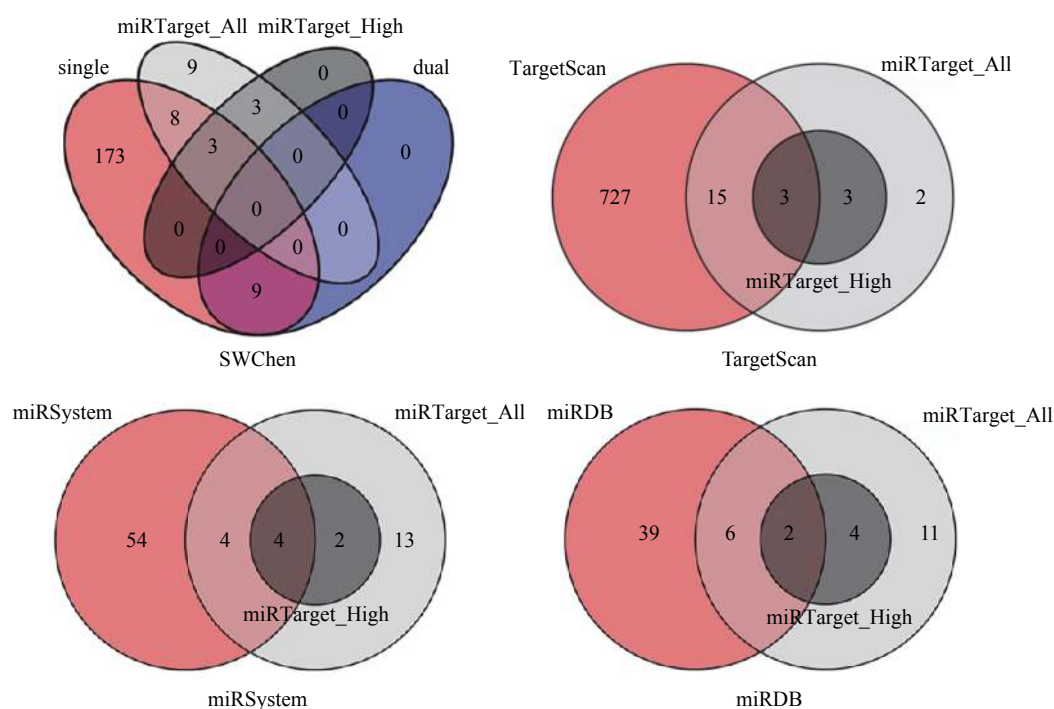


图 1 4 种 miRNA 预测软件与 miRTarget 数据库匹配情况

Fig. 1 Comparing prediction results of four miRNA prediction softwares by using miRTarget database match

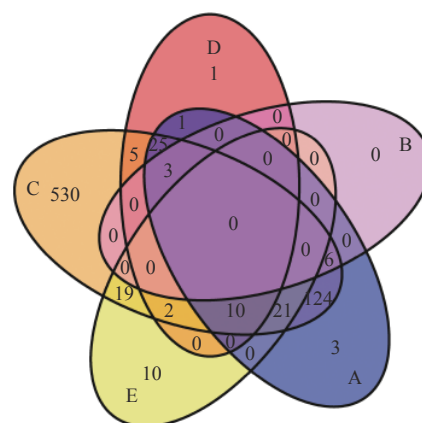
数据 3 条; miRSystem 在线软件预测有 62 条, 其中匹配 miRTarget_All 数据 4 条, 匹配 miRTarget_High 数据 4 条; miRDB 在线软件预测有 47 条, 其中匹配 miRTarget_All 数据 6 条, 匹配 miRTarget_High 数据 2 条。4 个软件预测的 miRNAs 结果集合统计 (图 2) 显示: TargetScan 出现 530 条独立 miRNAs, miRSystem 和 miRDB 分别出现 10 条和 1 条独立的 miRNAs; SWChen 系统总共出现 3 条独立的 miRNAs, 有 186 条与 TargetScan 结果重叠, 31 条与 miRSystem 结果重叠, 39 条与 miRDB 结果重叠, 双阈值 SWChen 结果除 miRSystem 均包含在其他软件预测结果中。

2.2 林麝成熟麝香代谢组测序分析结果

林麝麝香化学组分含量最高的前 10 个物质列于表 2。结果显示: 成熟麝香中的主要成分为长碳链的酮和醛等物质、特殊脂肪酸碳链的环酮类物质以及芳香族和甾类物质。

2.3 物种基因序列进化树构建及 RNA 二级结构分析

由图 3 可知: 同一基因之间 3'-UTR 序列存在差异较小, 其中林麝 *FASN* 和 *HMGCR* 基因 3'-UTR 序列与牛和羊被分为两类, 而林麝 *CYP11A1*



注: A. SWChen 系统种子匹配预测; B. SWChen 系统种子+结合强度匹配预测; C. TargetScan 保守及非保守 miRNA 预测; D. miRDB 预测的 miRNA; E. miRSystem 预测的 miRNA。

Note: A. seed sequence match prediction of SWChen; B. seed sequence and binding intensity match prediction of SWChen; C. conserved and non-conserved miRNA prediction by TargetScan tool; D. miRNA prediction of miRDB tool; E. miRNA prediction of miRSystem tool.

图 2 4 种 miRNA 预测软件预测结果集合统计

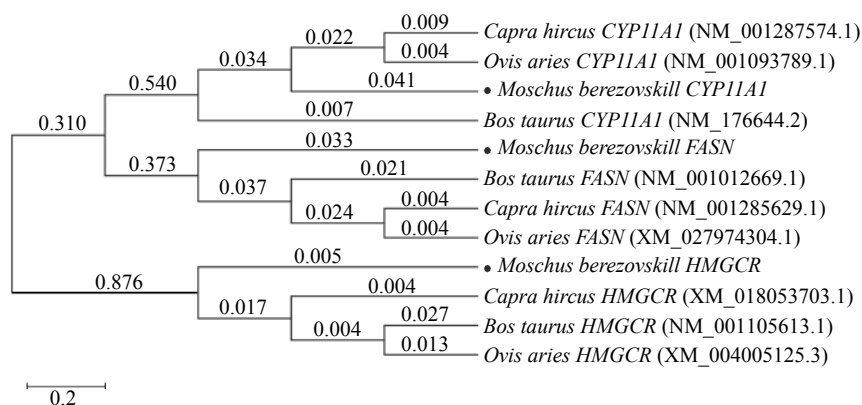
Fig. 2 The gather statistic to results of four miRNA prediction softwares

基因 3'-UTR 序列与山羊和绵羊被归为一类。进一步对各基因的 3'-UTR 序列进行 RNA 二级结构进行预测, 结果 (图 4) 显示: 林麝、牛、山羊和绵羊 *FASN* 基因 3'-UTR 序列 RNA 二级结构的最

表 2 成熟麝香主要化学组分

Tab. 2 The main chemical components of matured musk

化学组分 chemical components	CAS编号 CAS No.	检测峰面积 detected area	样本方差 sample deviation
反式-脱氢雄甾酮 trans-dehydroandrosterone	53-43-0	11 168 282.67	1 121 583.40
6-甲基庚烷-1,6-二醇 6-methylheptane-1,6-diol	5392-57-4	11 110 678.00	811 962.87
麝香酮 muscone	541-91-3	10499 690.67	10790 615.92
对甲酚 p-cresol	106-44-5	4090 610.00	294 912.52
胆固醇 cholesterol	57-88-5	4071 492.00	1409 731.95
13-十四烯醛 13-tetradecenal	85896-31-7	3 715 398.33	2517935.89
8-环十六烯-1-酮 8-cyclohexadecen-1-one	3100-36-5	1 771 978.00	16964.98
顺式-13-十八烯醛 cis-13-octadecenal	58594-45-9	1 749 918.33	2317990.10
苯乙酸 benzeneacetic acid	103-82-2	1 712 839.67	1 217 397.87
4-甲基环十五烷-1-酮 4-methyl-cyclopentadecanone	34894-60-5	1 266 641.67	788 677.07



注：数字代表进化树分支的差异程度。

Note: The numbers represent the degree of divergence in the branches of the evolutionary tree.

图 3 基因 3'-UTR 序列进化树比对

Fig. 3 Sequence alignment of gene 3'-UTR by evolution tree construction

小自由能分别为-222.80、-234.60、-229.10 和 -234.70 kcal/mol，*HMGCR* 基因 3'-UTR 序列 RNA 二级结构的最小自由能分别为-466.00、-477.60、-267.20 和-1417.30 kcal/mol，*CYP11A1* 基因 3'-UTR 序列 RNA 二级结构的最小自由能分别为-110.10、-112.90、-105.50 和-112.00kcal/mol，其中林麝和牛 *FASN* 基因 3'-UTR 与山羊和绵羊相比出现较多的单链空泡。

2.4 SWChen 系统分析各物种调控 *FASN*、*HMGCR* 和 *CYP11A1* 基因表达的 miRNAs

通过 miRNA 转录组序列保守分析发现：林麝 miRNAs 数据总有 1391 条，miRBase 数据库中查询牛的 miRNAs 数据为 1030 条，山羊和绵羊的数据为 519 条。SWChen 系统的种子匹配预测结果显示：调控林麝、牛、山羊和绵羊 *FASN* 基因表达的 miRNAs 在相应数据库中的占比分别为 5.25%、5.53%、1.70% 和 2.38%；调控 *HMGCR* 基

因表达的 miRNAs 占比分别为 12.55%、13.11%、10.88% 和 6.12%；调控 *CYP11A1* 基因表达的 miRNAs 占比分别为 3.54%、3.69%、4.76% 和 4.22%；在林麝和牛中，调控 *FASN* 和 *HMGCR* 基因表达的 miRNAs 占比高于山羊和绵羊，而调控 *CYP11A1* 基因表达的 miRNAs 占比低于山羊和绵羊。集合论分析结果 (图 5 和表 3) 显示：在 4 个物种中，调控 *FASN* 基因表达的 miRNAs 交集为 miR-24-3p 和 miR-30b-3p；调控 *HMGCR* 基因表达的 miRNAs 交集为 miR-29b、miR-146a 和 miR-2284d；调控 *CYP11A1* 基因表达的 miRNAs 交集为 let-7 家族。

双阈值分析结果显示 (表 4)：调控 *FASN* 基因表达的 miRNAs 仅在山羊和绵羊中存在 miR-874-5p 交集；调控 *HMGCR* 基因表达的 miRNAs 在林麝、牛和山羊中存在 miR-545-3p、miR-429 和 miR-181b-3p 交集，调控 *CYP11A1* 基因表达

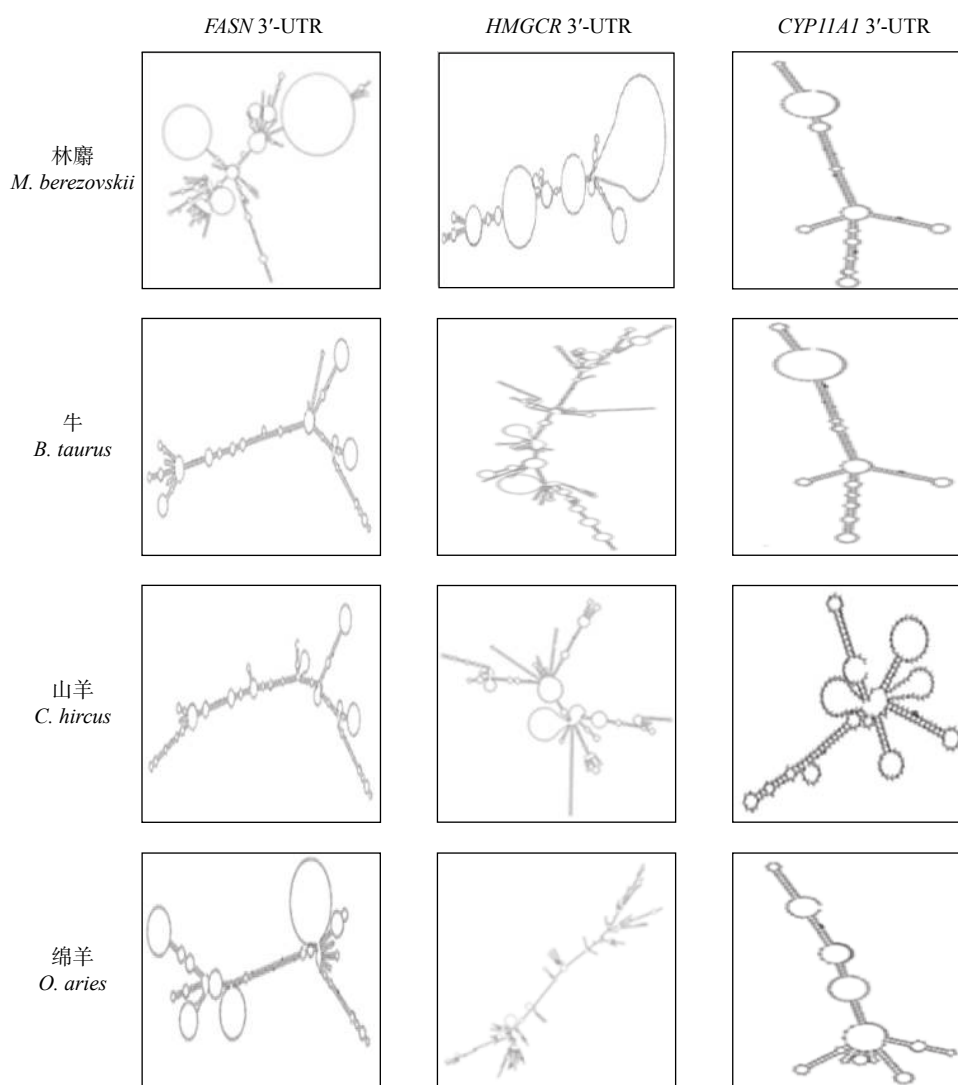


图 4 各物种基因 3'-UTR RNA 二级结构预测

Fig. 4 Predicted RNA secondary structures of gene 3'-UTR sequences in respective species

的 miRNAs 仅在林麝、山羊和绵羊中存在 miR-532-3p 的交集。

3 讨论

SWChen 系统是基于 Smith-Waterman 局部序列比对算法和碱基互补配对原理^[11, 13], 预测 2 条序列之间的相互作用能力和相似性情况的程序, 主要基于 R 语言软件编写, 用于 RNA 之间的作用情况预测^[14-17]。计分规则参数使得序列比对过程中出现连续 3~4 个的相互匹配才能抵消 1 个错配, 符合 miRNA 序列种子匹配规则, 通过矩阵最大值进行重新回数算分即为 SWChen 系统的最终评分, 选择适合的评分参数对于提高相互作用

预测准确性具有重要意义。

Smith-Waterman 算法是评价 2 条序列之间功能域和保守域的比对算法, 相比 BLAST 算法^[18]和马尔可夫模型^[19]推算, Smith-Waterman 算法更加准确, 能够满足预测需求, 但是这种算法的速度最慢, 且会忽略第 2 评分结合位置。miRTarBase 数据库包含了现有已验证调控功能的 miRNAs, 而在所有物种的研究中, 人 *FASN* 研究证据较多, 因此以人种为对象评估 SWChen 系统的准确性具有更高的可信度。程序评估结果显示: 4 种预测工具的结果差异并不大, 其中 SWChen 相比 miRSystem 和 miRDB 工具能预测出较多的候选 miRNAs, 匹配 miRTarBase 数据库的 miRNAs

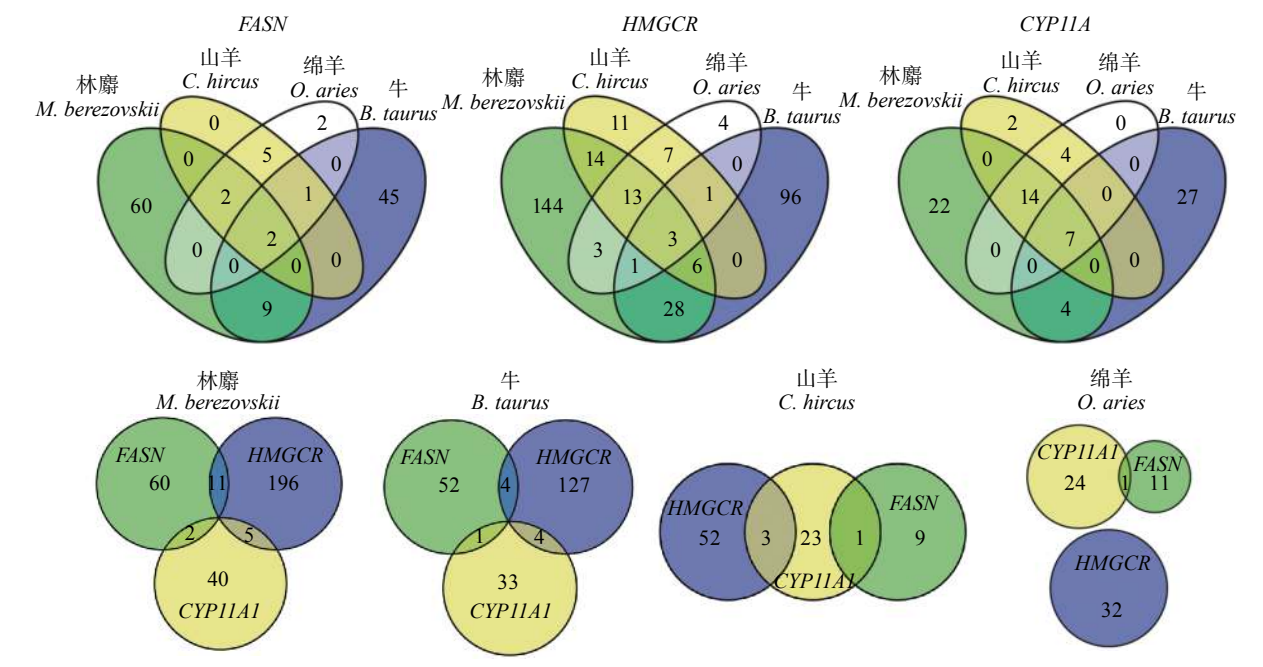


图 5 4 个物种中调控 *FASN*、*HMGR* 和 *CYP11A1* 基因表达的 miRNAs 集合分析维恩图

Fig. 5 Venn diagram of miRNAs regulating *FASN*, *HMGR* and *CYP11A1* genes expression intersection analysis in four species

表 3 调控基因的 miRNA 交集

Tab. 3 miRNA intersection of targeting genes

基因 genes	miRNA交集 miRNA interesection
<i>FASN</i>	miR-24-3p, miR-30b-3p
<i>HMGR</i>	miR-29b, miR-146a, miR-2284d
<i>CYP11A1</i>	let-7g, let-7i, let-7f, let-7b, let-7d, let-7c, let-7a-5p
<i>HMGR</i> ∩ <i>CYP11A1</i> (<i>M. berezovskii</i>)	miR-205a-5p, miR-143-5p, miR-205-5p, miR-205, miR-205b
<i>HMGR</i> ∩ <i>CYP11A1</i> (<i>B. taurus</i>)	miR-125a, miR-125b, miR-205, miR-670
<i>HMGR</i> ∩ <i>CYP11A1</i> (<i>C. hircus</i>)	miR-1224, miR-125a-5p, miR-125b-5p
<i>HMGR</i> ∩ <i>FASN</i> (<i>M. berezovskii</i>)	miR-146b, miR-146a-5p, miR-146c-5p, miR-146d-5p, miR-146a, miR-27b, miR-146b-5p, miR-30b-3p, miR-30e-3p, miR-2424, miR-96-3p
<i>HMGR</i> ∩ <i>FASN</i> (<i>B. taurus</i>)	miR-10165-5p, miR-2358, miR-2394, miR-669
<i>CYP11A1</i> ∩ <i>FASN</i> (<i>M. berezovskii</i>)	miR-143-5p, miR-205a-5p, miR-205-5p, miR-205, miR-205b
<i>CYP11A1</i> ∩ <i>FASN</i> (<i>B. taurus</i>)	miR-2406
<i>CYP11A1</i> ∩ <i>FASN</i> (<i>C. hircus</i> or <i>O. aries</i>)	miR-342-5p

表 4 在各物种中预测调控基因的 miRNAs (双阈值)

Tab. 4 Predicted miRNAs targeting genes in species respectively (dual thresholds)

物种名称 species name	预测调控 <i>FASN</i> 基因的 miRNA spredicted miRNAs for regulating <i>FASN</i> gene	预测调控 <i>HMGR</i> 基因的 miRNA spredicted miRNAs for regulating <i>HMGR</i> gene	预测调控 <i>CYP11A1</i> 基因的 miRNA spredicted miRNAs for regulating <i>CYP11A1</i> gene
林麝 <i>M. berezovskii</i>	—	miR-545-3p, miR-429-3p, miR-200-3p, miR-181b-3p, miR-29d	miR-532-3p
牛 <i>B. taurus</i>	miR-10182-3p, miR-1777a, miR-669	miR-2350, miR-429, miR-545-3p	—
山羊 <i>C. hircus</i>	miR-874-5p	miR-181b-3p, miR-186-5p, miR-429, miR-497-3p	miR-532-3p
绵羊 <i>O. aries</i>	miR-874-5p	—	miR-532-3p

条数更多; 而 TargetScan 预测的 miRNAs 条数约为 SWChen 的 4 倍, 但匹配数据库的 miRNAs 约为 SWChen 的 2 倍, 提示 SWChen 的背景噪音相对 TargetScan 较小。所有的 miRNA 预测工具均基于 miRNA 对靶基因的识别机制^[11], 其中 TargetScan 在计分过程中考虑结合位置的保守性和 miRNA 亲和程度, 再去除物种序列保守性判断之后预测的 miRNAs 会增加, 而 SWChen 并不考虑序列保守性, 同时将评分控制为 80 分, 即种子区域的完全匹配, 占据识别机制中的大部分识别情况; 而 TargetScan 的算法包括了 miRNA 第 2~8 位和第 3~8 位的匹配规律^[11], 因此预测出较多的候选 miRNAs, 但却增加了错误 miRNAs 的噪音, 这是 TargetScan 和 SWChen 之间预测结果差异的原因。miRSystem 系统为集成多种预测工具的 miRNA 预测系统, 在预测结果中获得强有力证据的 miRNAs 匹配条数最多, 背景噪音相对较小, 但匹配 miRTarget_All 数据的数量也减少, 在 miRNA 的靶基因研究中, miRNA 预测主要基于多工具的交集筛选, 因此集成的 miRSystem 会得到更多的 miRTarget_High 匹配结果; 双阈值预测能够得到较多的重叠 miRNAs, 暗示双阈值的准确性可能更高。SWChen 系统能够满足 miRNA 靶基因预测的需求, 能够得到相对其他工具同等价值的结果, 且离线预测、数据库及参数的可变动使其更加灵活, 能够预测林麝 miRNAs 对靶基因的调控情况, 可为特种动物、昆虫、植物和真菌等关键 miRNA 的筛选提供便利。

已有研究表明: 麝香的形成过程主要有 2 种假设^[1-4, 12]。第一, 林麝香腺分泌初香后, 初香物质储存于香囊中经微生物的发酵熟化后形成成熟麝香; 第二, 林麝香腺直接分泌麝香的主要化学成分储存于香囊中积累, 而微生物不是麝香形成的决定因素。无论林麝通过哪种方式形成成熟麝香, 林麝合成和分泌的化学组分对麝香的形成至关重要。通过代谢组学的手段对麝香进行检测, 使用不同的萃取试剂得到的化学组分和含量有一定差异, 在麝香的甲醇和氯仿提取物中主要成分为脂肪酸类的烷烃类、醇类、酮类、醛类、固醇类、甾类及芳香族类物质, 与其他文献报道^[2]一致。脂肪酸衍生类、环酮类、甾类和固醇类为麝

香的主要成分, 而 *FASN*、*HMGCR* 和 *CYP11A1* 是合成上述物质的关键基因。

FASN 基因对于长链脂肪酸的合成和活化过程具有重要作用^[5, 9, 20-22], 如碳链的延伸、末端酰基的活化和双键的还原等, 特别是与麝香酮 (3-甲基环十五酮) 碳链骨架的形成具有较强的关联性; 同时 *FASN* 合成的亚基酶对碳链的转运、去饱和、脱羟基和碳链活化等具有重要作用, 如酰基载体蛋白 β -酮酰 ACP 还原酶和 β -酮酰 ACP 水解酶等。对于固醇类和甾类物质而言, *HMGCR* 是其合成的限速酶^[8-9], 它参与甲羟戊酸的合成, 而甲羟戊酸作为基础物质进一步形成异戊烯基焦磷酸、鲨烯、羊毛固醇和胆固醇等产物, 对于类固醇物质的合成和分泌具有重要作用。*CYP11A1* 是胆固醇的侧链裂解的关键酶^[4, 8, 23], 对于孕烯酮、雄烯二酮和脱氢雄甾酮等物质合成至关重要, 而在麝香检测到大量的雄性激素前体物质, 如反式-脱氢雄甾酮, 其合成同时需要 *HMGCR* 和 *CYP11A1* 的参与。因此, *FASN*、*HMGCR* 和 *CYP11A1* 对于麝香的生物合成有较强的关联性, 具有较大的研究价值和潜力。miRNAs 是基因转录后调控的关键因子^[10-11], 对生物代谢组分的调控具有关键作用。1 个基因可受多条 miRNAs 调控, 同时 1 条 miRNAs 可调控多个基因的表达, 预测共同靶向林麝多个麝香合成相关基因表达的 miRNAs 对探究麝香生物合成机制具有重要意义, 同时也为牛和羊的肉质和繁殖性状相关研究提供参考。

林麝、牛、山羊和绵羊的 *FASN*、*HMGCR* 和 *CYP11A1* 基因 3'-UTR 序列差异较小, 同源物种之间基因的 3'-UTR 有保守性, 保守区域可能为共同的 miRNAs 作用位点, 同时也验证了序列拼接的正确性^[12]; 较高的自由能及更多的单链空泡结构将增加 miRNA 与 mRNA 的结合概率^[14-15]。通过 SWChen 系统分别对 4 个物种中调控 *FASN*、*HMGCR* 和 *CYP11A1* 基因的 miRNAs 进行预测, 在林麝和牛中, *FASN* 基因受控的 miRNAs 在各自 miRNA 数据库中的占比相对山羊和绵羊较高, 暗示林麝和牛 *FASN* 基因受 miRNA 调控的潜力强于山羊和绵羊, 与 RNA 结构分析的结果一致; 林麝 *HMGCR* 基因受 miRNA 调控的数量占比较高, 可能与其 3'-UTR 长度和序列复杂程

度有关；而 *CYP11A1* 基因在 4 个物种之间预测的候选 miRNA 数量差异不大，且 RNA 二级结构空泡结构及最小自由能差距不明显，暗示 4 个物种 *CYP11A1* 基因受 miRNA 调控的潜力一致。候选 miRNAs 数量在相应物种总体数据库中的占比和基因 3'-UTR 的 RNA 二级结构和最小自由能预测，可为分析 miRNA 对基因的调控潜能提供参考。

集合论分析显示：4 个物种 *FASN* 基因共同受到 miR-24-3p 和 miR-30b-3p 调控，miR-24-3p 可通过调控锌指同源基因 *ZHX2* (zinc fingers and homeoboxes 2) 影响 *FASN* 基因的表达^[24]，miR-30c-5p 可调控下调 *FASN* 的表达改善肝性脂肪沉积^[22]，预测 miR-24-3p 和 miR-30b-3p 有很大的潜力调控 4 个物种的 *FASN* 基因表达。预测调控 *HMGCR* 基因表达的 miRNAs 结果显示：miR-29b、miR-146a 和 miR-2284d 为 4 个物种 miRNA 的交集部分，miR-29 能够抑制小鼠 *HMGCR* 的表达而影响在肝脏的胆固醇沉积^[25]，miR-29b 作为 miR-29 的同源 miRNAs，具有较大的潜力调控 4 个物种 *HMGCR* 的表达，miR-146a 与免疫炎症、心血管疾病、糖尿病及总胆固醇含量等具有一定的相关性^[26-29]，这些因素均与固醇类和甾类物质合成相关，miR-146a 存在很大的潜能调控 *HMGCR* 基因的表达影响固醇类物质的合成导致相应疾病发生；通过双阈值分析可知：调控 *HMGCR* 基因表达的 miRNA 有较大可能为 miR-181b-3p 和 miR-429 家族，miR-181d-5p 能够显著下调猪的 *HMGCR* 基因^[30]，而鸡肝脏中的 miR-429 与 *HMGCR* 的表达具有较强的相关性^[31]，因此 miR-29b、miR-146a、miR-181b-3p 和 miR-429 都具有调控 *HMGCR* 表达的潜力。调控 *CYP11A1* 基因表达的 miRNAs 集合分析显示均属于 let-7 家族成员，let-7 家族是细胞中含量最为丰富的 miRNA 家族，参与动物免疫、繁殖和生长发育等过程^[32-34]；*CYP11A1* 属于雄性激素的合成限速酶^[8-9]，在各物种中均有较高的保守性。结合 UTR 进化树、二级结构和 miRNA 占比可知：4 个物种 *CYP11A1* 基因的相关参数基本接近一致，暗示 let-7 对保守性较强的 *CYP11A1* 有较强的调控潜能；进一步的双阈值预测结果显示：miR-532-3p 也能够调控 *CYP11A1* 基因的表达，而 miR-532-3p 被报道涉及女性不孕不育^[35]，miR-532-3p 可能通过调控 *CYP11A1* 基因导致相关生

殖类疾病。而在林麝中未发现共同调控 *FASN*、*HMGCR* 和 *CYP11A1* 的 miRNAs 分子，但 miR-205 和 miR-143 家族存在交集，有参与麝香合成调控通路的可能。研究表明：miR-205 主要参与 AKT 和 Wnt 信号通路^[36-40]，miR-143 主要参与 NF- κ B 信号通路^[41-43]，暗示麝香的生物合成机制可能涉及 miR-205 和 miR-143 参与的 AKT、Wnt 和 NF- κ B 信号通路。

4 结论

本研究通过自主开发的 SWChen 系统预测发现：调控林麝 *FASN* 基因表达的 miRNA 包括 miR-24-3p 和 miR-30b-3p，调控林麝 *HMGCR* 基因表达的 miRNA 有 miR-29b、miR-146a、miR-181b-3p 和 miR-429 家族，调控林麝 *CYP11A1* 基因表达的 miRNA 有 miR-532-3p 和 let-7 家族，共同调控林麝 3 个基因表达的 miRNAs 主要为 miR-205 家族和 miR-143 家族。研究结果可为研究林麝、牛、山羊和绵羊 *FASN*、*HMGCR* 和 *CYP11A1* 的转录后调控提供参考，同时对解析林麝的麝香合成机制具有重要意义。

[参考文献]

- [1] JIE H, ZHANG P, XU Z X, et al. MicroRNA and other small RNA sequence profiling across six tissues of Chinese forest musk deer (*Moschus berezovskii*)[J]. Bio-Med Research International, 2019, 2019: 4370704. DOI: 10.1155/2019/4370704.
- [2] LI D Y, CHEN B L, ZHANG L, et al. The musk chemical composition and microbiota of Chinese forest musk deer males[J]. Scientific Reports, 2016, 6: 18975. DOI: 10.1038/srep18975.
- [3] LI Y M, ZHANG T X, QI L, et al. Microbiota changes in the musk gland of male forest musk deer during musk maturation[J]. Frontiers in Microbiology, 2018, 9: 3048. DOI: 10.3389/fmicb.2018.03048.
- [4] ZHANG T X, PENG D, QI L, et al. Musk gland seasonal development and musk secretion are regulated by the testis in muskrat (*Ondatra zibethicus*)[J]. Biological Research, 2017, 50: 10. DOI: 10.1186/s40659-017-0116-9.
- [5] CARMAN G M, HAN G S. Fat-regulating phosphatidic acid phosphatase: a review of its roles and regulation in lipid homeostasis[J]. Journal of Lipid Research, 2019, 60(1): 2. DOI: 10.1194/jlr.S087452.
- [6] SMIT L A, MOZAFFARIAN D, WILLETT W. Review of fat and fatty acid requirements and criteria for developing dietary guidelines[J]. Annals Nutrition and Metabolism, 2009, 55: 44. DOI: 10.1159/000228995.
- [7] DE SIQUEIRA CARVALHO A A, DA SILVA V G,

- ZANOTELI E, et al. Myopathy due to HMGR antibodies in adult mimicking muscular dystrophy associated with cancer and statin exposure-narrative review of the literature-case report[J]. Therapeutics and Clinical Risk Management, 2018, 14: 903. DOI: [10.2147/TCRM.S162931](#).
- [8] KUMAZAWA T, TSUCHIYA N, WANG L Z, et al. Microsatellite polymorphism of steroid hormone synthesis gene *CYP11A1* is associated with advanced prostate cancer[J]. International Journal of Cancer, 2004, 110(1): 140. DOI: [10.1002/ijc.20070](#).
- [9] RAHALI D, DALLAGI Y, HUPKENS E, et al. Spermatogenesis and steroidogenesis disruption in a model of metabolic syndrome rats[J]. Archives of Physiology and Biochemistry, 2020(11): 1. DOI: [10.1080/13813455.2020.1812665](#).
- [10] FILIPOWICZ W, BHATTACHARYYA S N, SONENBERG N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?[J]. Nature Reviews Genetics, 2008, 9(2): 102. DOI: [10.1038/nrg2290](#).
- [11] BARTEL D P. MicroRNAs: target recognition and regulatory functions[J]. Cell, 2009, 136(2): 215. DOI: [10.1016/j.cell.2009.01.002](#).
- [12] ZHANG L, JIE H, XIAO Y P, et al. Genomic identification and expression analysis of the cathelicidin gene family of the forest musk deer[J]. Animals, 2019, 9(8): 481. DOI: [10.3390/ani9080481](#).
- [13] SMITH T F, WATERMAN M S. Identification of common molecular subsequences[J]. Journal of Molecular Biology, 1981, 147(1): 195. DOI: [10.1016/0022-2836\(81\)90087-5](#).
- [14] 崔小芝, 吴巧群, 吴小敏, 等. 从江香猪 *lncGHR1* 基因的克隆、组织表达及超表达研究[J]. 农业生物技术学报, 2020, 28(12): 2165. DOI: [10.3969/j.issn.1674-7968.2020.12.007](#).
- [15] 杨丽红, 吴小敏, 崔小芝, 等. *NONHSAT184603.1* 基因对解旋酶基因 *BLM* 表达的影响[J]. 农业生物技术学报, 2020, 28(9): 1642. DOI: [10.3969/j.issn.1674-7968](#).
- [16] 吴小敏, 倪锴, 朱晓峰, 等. 猪 *PRKAA1* 基因相关 *lncRNAs* 的筛选及其在组织中的表达分析[J]. 农业生物技术学报, 2020, 28(2): 232. DOI: [10.3969/j.issn.1674-7968.2020.02.005](#).
- [17] 吴巧群, 许厚强, 崔小芝, 等. 从江香猪 miR-143-3P 克隆及组织表达[J]. 中国畜牧杂志, 2021, 57(1): 1. DOI: [10.19556/j.0258-7033.20200624-04](#).
- [18] ALTSCHUL S F, MADDEN T L, SCHÄFFER A A, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs[J]. Nucleic Acids Research, 1997, 25(17): 3389. DOI: [10.1093/nar/25.17.3389](#).
- [19] PEDERSEN J S, HEIN J. Gene finding with a hidden markov model of genome structure and evolution[J]. Bioinformatics, 2003, 19(2): 219. DOI: [10.1093/bioinformatics/19.2.219](#).
- [20] KAVIANI S, COOPER J A. Appetite responses to high fat meals or diets of varying fatty acid composition: a comprehensive review[J]. European Journal of Clinical Nutrition, 2017, 71(10): 1154. DOI: [10.1038/ejcn.2016.250](#).
- [21] WOOD J D, ENSER M, FISHER A V, et al. Fat deposition, fatty acid composition and meat quality: a review[J]. Meat Science, 2008, 78(4): 343. DOI: [10.1016/j.meatsci.2007.07.019](#).
- [22] FAN J H, LI H P, NIE X, et al. MiR-30c-5p ameliorates hepatic steatosis in leptin receptor-deficient (db/db) mice via down-regulating FASN[J]. Oncotarget, 2017, 8(8): 14561. DOI: [10.18632/oncotarget.14561](#).
- [23] O'HARA L, YORK J P, ZHANG P M, et al. Targeting of GFP-Cre to the mouse *Cyp11a1* locus both drives cre recombinase expression in steroidogenic cells and permits generation of *Cyp11a1* knock out mice[J]. PLoS One, 2014, 9(1): e84541. DOI: [10.1371/journal.pone.0084541](#).
- [24] YU X G, LIN Q H, WU Z C, et al. ZHX2 inhibits SREBP1c-mediated *de novo* lipogenesis in hepatocellular carcinoma via miR-24-3p[J]. Journal of Pathology, 2020, 252(4): 358. DOI: [10.1002/path.5530](#).
- [25] LIU M X, GAO M, LI C Z, et al. Dicer1/miR-29/HMGR axis contributes to hepatic free cholesterol accumulation in mouse non-alcoholic steatohepatitis[J]. Acta Pharmacologica Sinica, 2017, 38(5): 660. DOI: [10.1038/aps.2016.158](#).
- [26] LUO H C, LUO Q S, WANG C F, et al. Association of miR-146a, miR-149, miR-196a2, miR-499 gene polymorphisms with ischemic stroke in a Chinese people[J]. Oncotarget, 2017, 8(46): 81295. DOI: [10.18632/oncotarget.18333](#).
- [27] MENSÀ E, GIULIANI A, MATAACCHIONE G, et al. Circulating miR-146a in healthy aging and type 2 diabetes: age- and gender-specific trajectories[J]. Mechanisms of Ageing and Development, 2019, 180: 1. DOI: [10.1016/j.mad.2019.03.001](#).
- [28] HUANG S, LV Z Q, DENG Q F, et al. A genetic variant in Pre-miR-146a (rs2910164 C>G) is associated with the decreased risk of acute coronary syndrome in a Chinese population[J]. The Tohoku Journal of Experimental Medicine, 2015, 237(3): 227. DOI: [10.1620/tjem.237.227](#).
- [29] DEL MONTE A, ARROYO A B, ANDRES-MANZANO M J, et al. miR-146a deficiency in hematopoietic cells is not involved in the development of atherosclerosis[J]. PLoS One, 2018, 13(6): e0198932. DOI: [10.1371/journal.pone.0198932](#).
- [30] JIA Y M, LING M F, ZHANG L C, et al. Downregulation of *miR-150* expression by DNA hypermethylation is associated with high 2-hydroxy-(4-methylthio) butanoic acid-induced hepatic cholesterol accumulation in nursery piglets[J]. Journal of Agricultural and Food Chemistry, 2016, 64(40): 7530. DOI: [10.1021/acs.jafc.6b03615](#).
- [31] HU Y, ZHANG R, ZHANG Y H, et al. *In ovo* leptin ad-

- ministration affects hepatic lipid metabolism and microRNA expression in newly hatched broiler chickens[J]. Journal of Animal Science and Biotechnology, 2012, 3: 16. DOI: [10.1186/2049-1891-3-16](https://doi.org/10.1186/2049-1891-3-16).
- [32] ETEBARI K, AFRAD M H, TANG B, et al. Involvement of microRNA miR-2b-3p in regulation of metabolic resistance to insecticides in *Plutella xylostella*[J]. Insect Molecular Biology, 2018, 27(4): 478. DOI: [10.1111/imb.12387](https://doi.org/10.1111/imb.12387).
- [33] PENG T F, PAN Y O, GAO X W, et al. Reduced abundance of the *CYP6CY3*-targeting *let-7* and *miR-100* miRNAs accounts for host adaptation of *Myzus persicae nicotianae*[J]. Insect Biochemistry and Molecular Biology, 2016, 75: 89. DOI: [10.1016/j.ibmb.2016.06.002](https://doi.org/10.1016/j.ibmb.2016.06.002).
- [34] LI J B, XUE H, LI T T, et al. Exosomes derived from mesenchymal stem cells attenuate the progression of atherosclerosis in ApoE^{-/-} mice via miR-let7 mediated infiltration and polarization of M2 macrophage[J]. Biochemical and Biophysical Research Communications, 2019, 510(4): 565. DOI: [10.1016/j.bbrc.2019.02.005](https://doi.org/10.1016/j.bbrc.2019.02.005).
- [35] DA SILVA L F I, DA BROI M G, DA LUZ C M, et al. miR-532-3p: a possible altered miRNA in cumulus cells of infertile women with advanced endometriosis[J]. Reproductive BioMedicine Online, 2020, 42(3): 579. DOI: [10.1016/j.rbmo.2020.10.010](https://doi.org/10.1016/j.rbmo.2020.10.010).
- [36] XUE Y D, LIU Y Y, BIAN X H, et al. miR-205-5p inhibits psoriasis-associated proliferation and angiogenesis: Wnt/ β -catenin and mitogen-activated protein kinase signaling pathway are involved[J]. The Journal of Dermatology, 2020, 47(8): 882. DOI: [10.1111/1346-8138.15370](https://doi.org/10.1111/1346-8138.15370).
- [37] ZHANG P, LU X Y, SHI Z F, et al. miR-205-5p regulates epithelial-mesenchymal transition by targeting PTEN via PI3K/AKT signaling pathway in cisplatin-resistant nasopharyngeal carcinoma cells[J]. Gene, 2019, 710: 103. DOI: [10.1016/j.gene.2019.05.058](https://doi.org/10.1016/j.gene.2019.05.058).
- [38] CUI Y, CHEN R, MA L, et al. miR-205 expression elevated with EDS treatment and induced leydig cell apoptosis by targeting RAP2B via the PI3K/AKT signaling pathway[J]. Frontiers in Cell and Developmental Biology, 2020, 8: 448. DOI: [10.3389/fcell.2020.00448](https://doi.org/10.3389/fcell.2020.00448).
- [39] SUN L L, XIAO L, DU X L, et al. MiR-205 promotes endothelial progenitor cell angiogenesis and deep vein thrombosis recanalization and resolution by targeting PTEN to regulate Akt/autophagy pathway and MMP2 expression[J]. Journal of Cellular and Molecular Medicine, 2019, 23(12): 8493. DOI: [10.1111/jcmm.14739](https://doi.org/10.1111/jcmm.14739).
- [40] WANG D J, WANG S M, LIU Q, et al. SZ-685C exhibits potent anticancer activity in both radiosensitive and radioresistant NPC cells through the miR-205-PTEN-Akt pathway[J]. Oncology Reports, 2013, 29(6): 2341. DOI: [10.3892/or.2013.2376](https://doi.org/10.3892/or.2013.2376).
- [41] CHEN Y Z, PENG H, ZHOU S Q, et al. ADAR1 is targeted by miR-143 to regulate IL-1 β -induced endothelial activation through the NF κ B pathway[J]. International Journal Biochemistry and Cell Biology, 2017, 89(2): 25. DOI: [10.1016/j.biocel.2017.05.021](https://doi.org/10.1016/j.biocel.2017.05.021).
- [42] ZHENG Y, YANG F, FU L, et al. The mechanism of miR-143 inducing apoptosis of liver carcinoma cells through regulation of the NF- κ B pathway[J]. Oncology Letters, 2018, 15(6): 9567. DOI: [10.3892/ol.2018.8486](https://doi.org/10.3892/ol.2018.8486).
- [43] WANG Y J, LI H, SHI Y S, et al. miR-143-3p impacts on pulmonary inflammatory factors and cell apoptosis in mice with mycoplasmal pneumonia by regulating TLR4/MyD88/NF- κ B pathway[J]. Bioscience Reports, 2020, 40(7): BSR20193419. DOI: [10.1042/BSR20193419](https://doi.org/10.1042/BSR20193419).

责任编辑: 何馨成