

· 综 述 ·

长非编码 RNA 编码微肽的研究进展

潘剑锋¹, 尚方正¹, 马荣¹, 王敏¹, 戎友俊¹, 梁丽丽¹, 牛舒冉¹, 李彦伯¹, 齐云鹏¹, 张燕军¹, 李金泉^{2,3,4}

- 1 内蒙古农业大学 动物科学学院, 内蒙古 呼和浩特 010018
- 2 农业部肉羊遗传育种重点实验室, 内蒙古 呼和浩特 010018
- 3 内蒙古自治区动物遗传育种与繁殖重点实验室, 内蒙古 呼和浩特 010018
- 4 内蒙古自治区山羊遗传育种工程技术研究中心, 内蒙古 呼和浩特 010018

潘剑锋, 尚方正, 马荣, 王敏, 戎友俊, 梁丽丽, 牛舒冉, 李彦伯, 齐云鹏, 张燕军, 李金泉. 长非编码 RNA 编码微肽的研究进展. 生物工程学报, 2022, 38(9): 3194-3214.

PAN JF, SHANG FZ, MA R, WANG M, RONG YJ, LIANG LL, NIU SR, LI YB, QI YP, ZHANG YJ, LI JQ. Advances of long non-coding RNA encoded micro-peptides. Chin J Biotech, 2022, 38(9): 3194-3214.

摘 要: 长非编码 RNA (long non-coding RNA, lncRNA) 是长度超过 200 nt 的非编码 RNA, 具有一个或多个短开放阅读框, 可编码功能性微肽。这些功能性微肽在各种生物过程中扮演着重要角色, 例如 Ca^{2+} 转运、线粒体代谢、肌细胞融合和细胞衰老等过程。同时, 这些生物过程又在机体稳态调控、疾病和癌症的发生与发展、胚胎发育等重要生理过程中起关键作用。因此, 研究由 lncRNA 编码的微肽在生物体的潜在的调控机制, 将有助于进一步解析生物体潜在调控过程, 并为后续疾病的靶向治疗及动物生长性能的提高提供新的理论依据。本文综述了现阶段 lncRNA 编码微肽领域的最新研究进展, 并对当前微肽在肌肉生理、炎症与免疫、人类常见癌症、胚胎发育等领域的研究进展进行描述与总结, 最后简单阐述了 lncRNA 编码微肽现阶段面临的问题和存在的挑战, 以期后续微肽的深入研究提供科学参考及新思路。

关键词: 长非编码 RNA; 短开放阅读框; 微肽; 肌肉生理; 炎症; 癌症; 胚胎发育

Received: December 9, 2021; **Accepted:** March 29, 2022; **Published online:** July 20, 2022

Supported by: National Natural Science Foundation of China (31860627); Major Science and Technology Projects of Inner Mongolia Autonomous Region, China (2021ZD0012)

Corresponding authors: ZHANG Yanjun. E-mail: imauzyj@163.com

LI Jinquan. E-mail: lijinqun_nd@126.com

基金项目: 国家自然科学基金 (31860627); 内蒙古自治区科技重大专项 (2021ZD0012)

Advances of long non-coding RNA encoded micro-peptides

PAN Jianfeng¹, SHANG Fangzheng¹, MA Rong¹, WANG Min¹, RONG Youjun¹,
LIANG Lili¹, NIU Shuran¹, LI Yanbo¹, QI Yunpeng¹, ZHANG Yanjun¹, LI Jinquan^{2,3,4}

1 College of Animal Science, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia Autonomous Region, China

2 Key Laboratory of Meat Sheep Genetics and Breeding, Ministry of Agriculture, Hohhot 010018, Inner Mongolia Autonomous Region, China

3 Key Laboratory of Animal Genetics, Breeding and Reproduction in Inner Mongolia, Hohhot 010018, Inner Mongolia Autonomous Region, China

4 Goat Genetics and Breeding in Inner Mongolia Autonomous Region Engineering Technology Research Center, Hohhot 010018, Inner Mongolia Autonomous Region, China

Abstract: Long non-coding RNA (lncRNA) refers to non-coding RNA longer than 200 nt, with one or more short open reading frames (sORF), which encode functional micro-peptides. These functional micro-peptides often play key roles in various biological processes, such as Ca²⁺ transport, mitochondrial metabolism, myocyte fusion, cellular senescence and others. At the same time, these biological processes play a key role in the regulation of body homeostasis, diseases and cancers development and progression, embryonic development and other important physiological processes. Therefore, studying the potential regulatory mechanisms of micro-peptides encoded by lncRNA in organisms will help to further elucidate the potential regulatory processes in organisms. Furthermore, it will provide a new theoretical basis for the subsequent targeted treatment of diseases and improvement of animal growth performance. This review summarizes the latest research progress in the field of lncRNA-encoded micro-peptides, as well as the progress in the fields of muscle physiological regulation, inflammation and immunity, common human cancers, and embryonic development. Finally, the challenges of lncRNA-encoded micro-peptides are briefly described, with the aim to facilitate subsequent in-depth research on micro-peptides.

Keywords: long non-coding RNA; short open reading frames; micro-peptide; muscle physiology; inflammation; cancer; embryonic development

人类基因组有 3/4 能够被转录, 但只有约 2% 的基因具有编码蛋白质的能力^[1]。DNA 元件百科全书 (encyclopedia of DNA elements, ENCODE) 的数据表明, 在编码蛋白质基因外的转录物中, 含有大量的非编码 RNA (non-coding RNA, ncRNA)^[2]。近年的研究发现, 这些 ncRNA 中可能含有一个或多个短开放阅读框 (short open reading frames, sORF), 并且可能编码一种小于 100 个氨基酸的微小蛋白质 (肽), 称为微肽 (micro-peptide)^[3-4]。长非编码

RNA (long non-coding RNA, lncRNA) 是 ncRNA 的一个亚类, 其长度超过 200 nt, 由 RNA 聚合酶 II (RNA polymerase II) 转录, 经剪切具有 5' 端帽子结构和 3' 端 poly A 尾巴, 并且已被证实其序列上保守的 sORF 可编码微肽^[5-7]。此外, 由于这类微肽的分子量小、基于保守性分析的筛选机制以及检测技术水平的限制, 导致其常被人们忽略, 而这可能造成许多极其重要的调控机制被“隐藏”, 所以这类微肽的表征就显得尤为重要。

早在 2002 年, 具有编码能力的 lncRNA 就已崭露头角。Rohrig 等^[8]在豆类根瘤器官中发现一种含有两个 sORF 的 lncRNA, 并且其编码的微肽可以与蔗糖合酶结合。2007 年, Galindo 和 Kondo 两个团队分别发现多个与果蝇胚胎发育相关的微肽, 并且这些微肽由 lncRNA 的 sORF 编码^[9-10]。2008 年到 2010 年, 有多个研究团队发现由 lncRNA 编码的微肽还能通过参与机体的免疫过程, 调控癌症的发生与发展^[11-12]。2015 年, Tonkin 等^[13]发现在改善肌肉生理性能方面, lncRNA 编码的微肽同样扮演着重要角色, 并可作为改善肌肉性能的药物靶点。而在近五年, 人们对 lncRNA 编码微肽这一机制的研究更是如雨后春笋般涌现, 并取得许多开创性的成果。而在这些研究中发现, 由 lncRNA 编码的微肽可通过参与 Ca^{2+} 转运^[14-15]、废物降解^[16-17]、线粒体代谢调控^[18]、转录调控^[19]、翻译调控^[20]、mRNA 剪接^[21]、信号传导^[22-24]、肌细胞融合^[25]和细胞衰老^[26]等过程, 调节机体的稳态、癌症和疾病的发生与发展、胚胎发育等重要生理过程 (图 1)。此外, 由 lncRNA 编码的微肽更是与多种人类常见癌症发生与发展密切相关, 其可通过充当抑癌因子或致癌因子影响癌症的发展, 这一发现可为日后靶向治疗提供极具价值的策略。可见, lncRNA 编码的微肽在生物体调控过程中扮演着重要的角色, 对其深入的解析将有助于人们重新认识机体内深层的调控机理, 并对微肽类靶向治疗药物的研发以及相应的临床治疗提供有价值的策略。

本文描述了由 lncRNA 编码的功能性微肽在肌肉生理调控、炎症免疫调节、疾病和癌症的发生与发展、昆虫及脊椎动物的胚胎发育等过程的最新研究进展, 并对微肽领域现阶段面临的问题和存在的挑战进行了简单阐述, 旨在为未来微肽领域的发展提供有价值的科学参考

及研究思路。

1 lncRNA 编码微肽的驱动机制

越来越多的证据表明, 大多数 lncRNA 具有一个或多个 sORF, 并且这些 sORF 是 lncRNA 编码微肽的前提。sORF 的长度小于 300 nt, 是从起始密码子 (ATG 或 AUG) 开始延伸到终止密码子的核苷酸序列, 通常位于非编码序列上, 如 lncRNA、circRNA 和 pri-miRNA 等^[6,27-28]。通过 ORF Finder、CPC2、CNIT 等一系列 ORF 预测及编码潜力鉴定计算工具和 Ribo-seq、RNC-seq 等翻译组学等现代数据分析及测序手段发现, sORF 可在不同物种中被翻译, 并且越长的 sORF 越容易被翻译^[4,29-35]。此外, sORF 的翻译过程还受到 IRES、m6A 修饰的保守位点等开放阅读框上游的调控元件介导^[6]。其中, 内部核糖体进入位点 (internal ribosome entry sites, IRES) 是一种独特的 RNA 调控元件, 它可使具有 sORF 结构的 RNA 招募核糖体并驱动蛋白质合成, 并且 IRES 不同于依赖于 m⁷G 帽子结构的翻译机制的经典真核翻译起始途径的高度复杂性, 其通过与真核起始因子结合发挥作用并促进翻译起始^[36-37]。此外, m6A 修饰的保守位点被发现可以驱动内源 ncRNA 的翻译, 并且在许多可翻译的内源环状 RNA (circular RNA, circRNA) 中也含有 m6A 位点^[38], 因此 m6A 修饰的保守位点可能也是驱动 lncRNA 翻译的关键调控元件。此外, 在研究中发现翻译的 lncRNA 优先定位于细胞质, 而非翻译的 lncRNA 则优先定位于细胞核。并且细胞质 lncRNA 的翻译效率几乎与 mRNA 的翻译效率相当, 而且具有翻译能力的 lncRNA 的 sORF 被核糖体所占据^[39-40]。并且核糖体占据的 lncRNA 与没有核糖体占据的 lncRNA 相比具有显著独特的翻译特性^[41]。表明具有编码潜力的

lncRNA 在细胞质中被翻译, 并且其翻译与其 sORF 被核糖体占据这一机制密切相关。lncRNA 编码微肽的机制示意图如图 1 所示。

2 与肌肉相关的功能性微肽

2.1 与 SERCA 转运 Ca^{2+} 相关的微肽

骨骼肌是人体运动系统中最大也是最重

要的构成组织, 在机体运动和糖脂代谢稳态调控方面发挥着关键作用^[42]。肌浆/内质网 Ca^{2+} ATP 酶 (sarcoplasmic reticulum Ca^{2+} ATPase, SERCA) 是一种可将 Ca^{2+} 从细胞质转运至肌浆网 (sarcoplasmic reticulum, SR) 的膜泵, 其通过调控肌细胞内的 Ca^{2+} 稳态控制肌肉收缩与舒张状态, 从而影响肌肉生理学功能^[43]。2015 年,

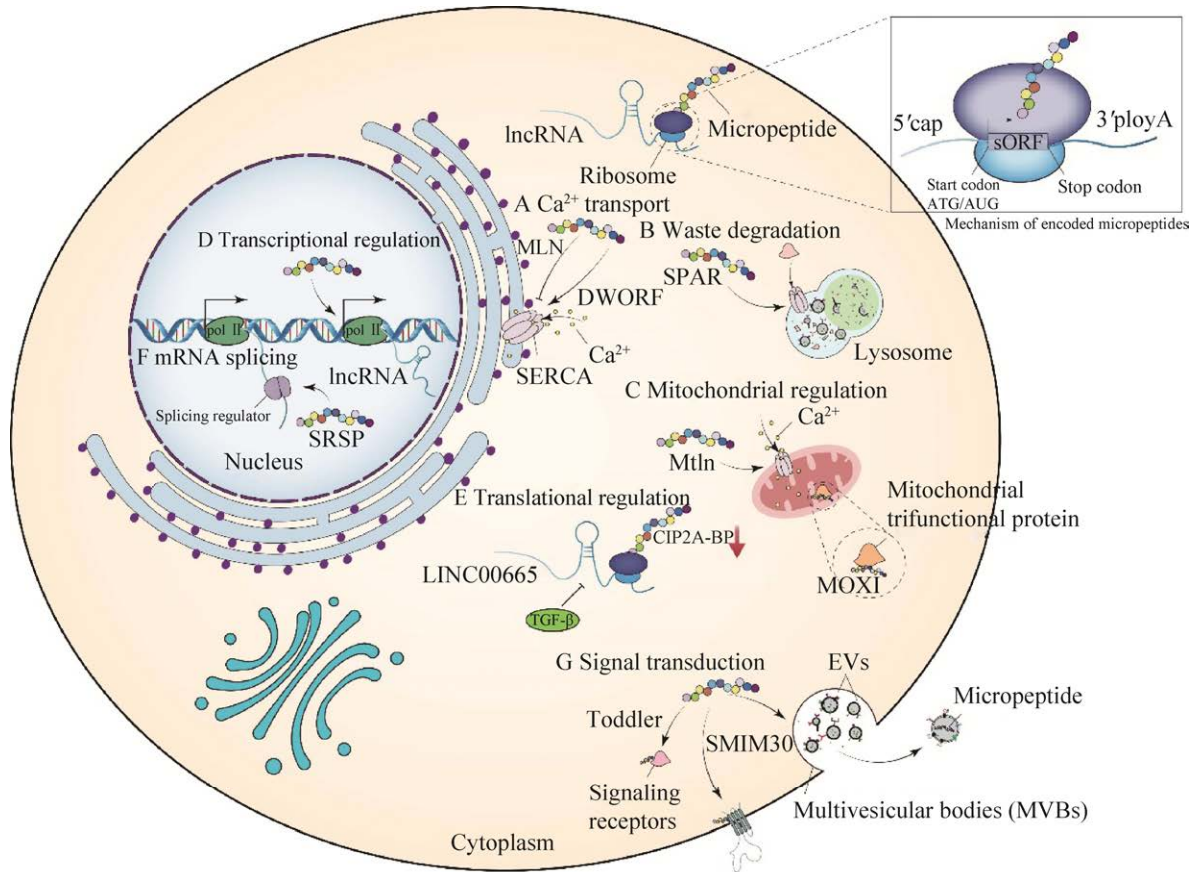


图 1 微肽的各种生物学功能及机制示意图

Figure 1 Schematic diagram of the various biological functions of micro-peptides. A: Ca^{2+} transport: MLN and DWORF respectively mediate Ca^{2+} transport by inhibiting or activating SERCA; B: waste degradation: SPAR regulates waste degradation by lysosomes; C: mitochondrial regulation: Mtn promotes mitochondrial Ca^{2+} transport, MOXI binds to mitochondrial trifunctional protein in mitochondrial inner membrane to regulate mitochondrial metabolism; D: transcriptional regulation. lncRNA Six1-encoded micro-peptide activates *Six1* gene transcription; E: translation regulation: TGF- β inhibits LINC00665-encoded CIP2A-BP; F: mRNA splicing: SRSP interacts with multiple splicing regulators to regulate mRNA splicing; G: signaling: SMIM30 activates signaling by driving membrane anchoring of the non-receptor tyrosine kinase SRC/YES1, micro-peptides conduct intercellular signaling through extracellular vesicles (EVs), toddler activates signaling by driving internalization of signaling receptors.

Anderson 等^[14]在研究骨骼肌的生理功能中发现一种由骨骼肌特异性 lncRNA 编码的保守肽 MLN (myoregulin), 该肽与 PLN (phospholamban) 和 SLN (sarcolipin) 具有相似的结构和功能。MLN 可通过直接与 SERCA 结合阻碍 Ca^{2+} 吸收至 SR, 从而控制肌肉的舒张机制^[14]。并且缺失 MLN 可增强骨骼肌中 Ca^{2+} 处理能力并提高运动表现^[13]。表明 MLN 是骨骼肌生理学功能中重要的调节器。随后, 该团队在 MLN 研究的基础上又发现了另外两种跨膜微肽 ELN (endoregulin) 和 ALN (another-regulin), 它们与 MLN 一样都是通过抑制 SERCA 阻碍 Ca^{2+} 吸收至 SR, 从而控制肌肉的状态^[44]。Anderson 团队的这两个研究揭示了 SERCA 抑制性微肽家族在肌肉和非肌肉细胞类型中控制细胞内 Ca^{2+} 动力学的保守机制。Nelson 等^[15]发现一种定位于 SR 膜的微肽 DWORF, 其由肌肉特异性 lncRNA 编码。该微肽可通过置换 SERCA 抑制剂、PLN、SLN 和 MLN 增强 SERCA 活性^[15]。此外, DWORF 对 SERCA 具有比 PLN 更高的表观结合亲和力, 并且过表达的 DWORF 可减轻与过表达 PLN 相关的收缩功能障碍^[45]。通过扩张型心肌病小鼠模型发现, DWORF 可增强扩张型心肌病小鼠的心收缩力并预防心力衰竭, 减轻小鼠的心肌病^[46]。这些数据表明 DWORF 有望成为一种极具吸引力的心衰治疗候选药物。综上所述, lncRNA 编码的功能性微肽可通过充当 SERCA 的激活剂或抑制剂调控 SERCA 转运 Ca^{2+} 这一过程, 调节肌肉生理学状态, 并且这一规律可为日后相关治疗药物的研发提供理论支持 (图 2A)。

2.2 与 mTORC1 相关的微肽

雷帕霉素靶蛋白复合物 1 (mechanistic target of rapamycin complex 1, mTORC1) 是真核细胞中高度保守的丝氨酸/苏氨酸蛋白激酶复合物, 可以在生长因子、营养因素等调控下,

促进细胞生长, 抑制细胞自噬等过程^[47-48]。2017 年, Matsumoto 等^[16]在急性损伤后的骨骼肌中发现 LINC00961 可编码一种在人和小鼠之间保守的微肽, 称为氨基酸反应小调节肽 (small regulatory polypeptide of amino acid response, SPAR)。该微肽可在氨基酸匮乏或刺激下, 使溶酶体处的 v-ATPase 和 Regulator 复合物保持紧密结合的状态, 从而导致 mTORC1 无法募集至溶酶体表面, 抑制 mTORC1 的激活 (图 2B)^[16]。而 mTORC1 的激活能有效地促进肌肉再生, 表明在急性损伤后骨骼肌的肌肉再生中 SPAR 充当着抑制剂的作用^[49-50]。

2.3 与线粒体代谢调控相关的微肽

线粒体是机体运动的重要供能细胞器, 在肌肉中大量存在, 并在肌肉运动能力的表现上起关键作用。2018 年, Makarewich 等^[18]在肌肉组织中发现一种可增强脂肪酸 β -氧化的线粒体微肽 MOXI (micro-peptide regulator of beta-oxidation), 其由一种分别在小鼠和人类基因组中注释为 1500011K16Rik 和 LINC00116 的 lncRNA 编码。该微肽可在线粒体内膜与线粒体三功能蛋白 (mitochondrial trifunctional protein) 结合, 形成一种在脂肪酸 β -氧化中起关键作用的酶复合物^[18]。在敲除 MOXI 的小鼠模型中发现心脏和骨骼肌线粒体表现出代谢脂肪酸能力减弱, 并伴随着运动能力显著降低的现象, 突出了 MOXI 在代谢调控中的作用^[18]。同年, Stein 等^[17]在骨骼肌和心脏中发现一种可编码微肽的 LINC00116, 其编码一种线粒体调节肽 (mitoregulin, MtlN)。该微肽在线粒体内膜通过与心磷脂 (cardiolipin) 结合, 维持膜和嵌入包括呼吸超复合物在内的蛋白质复合物的结构完整性, 以及支持蛋白质复合物的组装, 提高线粒体的呼吸效率^[17]。过表达 MtlN 可使线粒体膜电位、呼吸速率和 Ca^{2+} 保留能力增加, 线粒体

活性氧 (reactive oxygen species, ROS) 和无基质 Ca^{2+} 减少, 而这些结果与呼吸链复合、超级复合和耦合的改善相一致^[51]。这表明 Mtl α 可通过支持蛋白质复合物组装和/或稳定性提高呼吸效率。肌源性分化是肌肉发育和再生过程中的一个重要事件。2019 年, Lin 等^[52]在肌肉中发现一种定位在线粒体的微肽 MPM, 其可通过促进 PGC-1 α 基因 (线粒体呼吸相关基因) 表达增强线粒体生物发生和线粒体呼吸, 并促使线粒体的耗氧量和 ATP 产生增加, 从而刺激肌源性分化和肌肉纤维生长, 提高肌肉的运动能力。表明 MPM 可能是肌原分化、骨骼肌发育和再生的新促进剂, 并可能成为抗肌肉萎缩症和抗衰老的潜在治疗靶点。综上所述, lncRNA 编码的微肽可通过调控线粒体代谢、线粒体呼吸活动及线粒体内蛋白质复合物的组装等过程, 影响心脏和肌肉的生理状态。

2.4 与肌细胞融合相关的微肽

成肌细胞融合是一个关键过程, 有助于肌肉在发育过程中的生长和损伤后肌纤维的再生^[53]。2017 年, Bi 等^[54]发现一种含 84 个氨基酸的肌肉特异性微肽 myomixer, 该微肽的表达与成肌细胞分化相关, 对胚胎发生过程中的融合和骨骼肌形成至关重要。myomixer 可促进成肌细胞融合, 并与融合膜蛋白 myomaker 结合诱导成纤维细胞-成纤维细胞融合和成纤维细胞-成肌细胞融合^[55]。表明 myomaker 和 myomixer 是哺乳动物成肌细胞融合重要的驱动因素。并且单独的 myomaker 不仅不足以驱动成肌细胞融合, 还会消除干细胞群融合并阻止肌肉再生, 从而导致损伤后严重的肌肉退化^[56]。这点表明在肌肉再生和细胞融合过程中, myomaker 和 myomixer 是重要的结合伴侣, 两者通过相互协同调控肌细胞融合过程。Zhang 等^[25]在研究中发现另一种可与 myomaker 共表达并诱导伴随

快速细胞骨架重排的细胞融合的微肽 minion (microprotein inducer of fusion), 该微肽由 LOC101929726 编码。并且缺乏 minion 的肌源性祖细胞 (myogenic progenitors) 正常分化但无法形成合胞体肌管 (syncytial myotubes)^[25], 表明 Minion 在肌肉发育过程中的重要性, 并定义了诱导哺乳动物细胞融合的双组分程序。

2.5 其他与肌肉调控相关的微肽

2013 年, Magny 等^[57]在果蝇心脏中发现两种可通过调节钙转运影响肌肉规律收缩的微肽, 这些微肽似乎从苍蝇到人类的一系列物种中保存了超过 5.5 亿年, 并且与心脏病有关。2017 年, Cai 等^[19]在肌肉组织中发现一种双功能 lncRNA Six1, 该 lncRNA 一方面可对编码蛋白质的 *Six1* 基因进行顺式调控, 另一方面编码的微肽可激活 *Six1* 基因促进 *Six1* 表达, 同时两方面皆可促进细胞增殖并参与肌肉生长过程。2020 年, Wang 等^[58]发现, 一种在脊椎动物中高度保守的微肽 LEMP (lncRNA encoded micropeptide), 其由 lncRNA MyolncR4 (1500011K16RIK) 编码, 并且缺乏该微肽可使小鼠的肌肉形成和再生过程受阻, 以及斑马鱼体节的肌肉发育受损 (图 2)。综上所述, 在果蝇、小鼠和斑马鱼的肌肉分化中进化保守的微肽发挥着重要作用, 并且这一不断增长的微肽家族的功能具有极其重要的作用。

3 与炎症免疫相关的微肽

炎症是涉及多种心血管疾病 (cardiovascular diseases, CVDs) 的重要过程, 而含有 nod-like 受体家族 Pyrin 域 3 (nod-like receptor family pyrin domain containing 3, NLRP3) 炎性小体是先天免疫和炎症的重要参与者^[59]。2020 年, Bhatta 等^[60]在小鼠巨噬细胞中发现一种 lncRNA 1810058I24Rik, 它在暴露于脂多糖

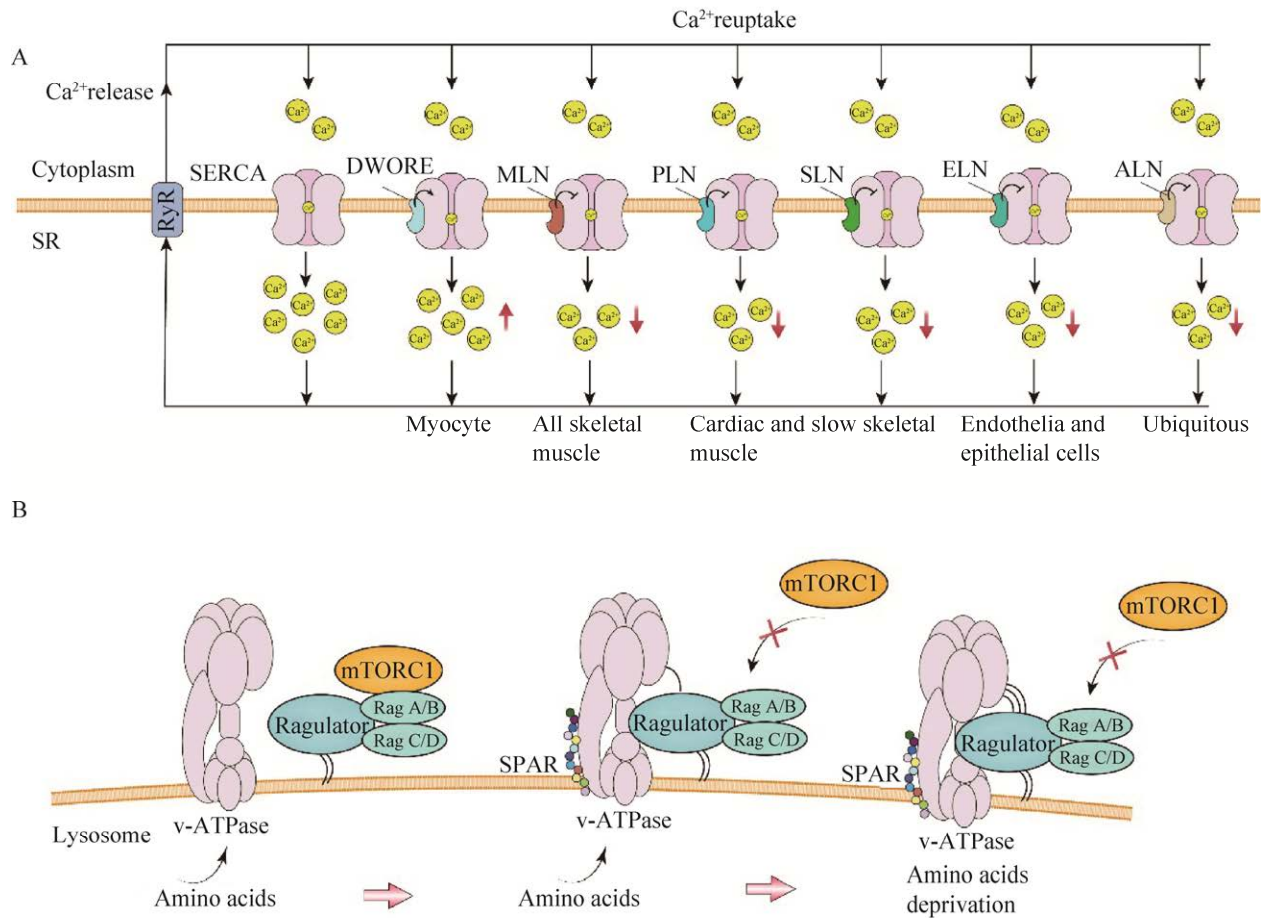


图 2 肌肉相关微肽调控机制示意图

Figure 2 Schematic diagram of the regulatory mechanism of muscle-related micro-peptides. (A) DWORF activates SERCA activity to promote Ca^{2+} uptake to SR, and MLN, PLN, SLN, ELN and ALN inhibit SERCA activity to prevent Ca^{2+} uptake to SR. (B) SPAR binding to v-ATPase prevents mTORC1 from being recruited to the lysosomal surface during amino acid stimulation or deprivation.

(lipopolysaccharide, LPS) 以及其他 Toll 样受体 (Toll-like receptors, TLR) 和炎症细胞因子的人和鼠骨髓细胞中被下调。并且其编码的线粒体微肽-47 (mitochondrial micropeptide-47, Mm47) 被证实可激活 NLRP3 炎性小体, 参与机体的先天免疫反应和炎症进程。此外, 由 lncRNA 编码的微肽还与细胞的抗原呈递和 T 细胞启动有关。Niu 等^[61]在人类上发现一种由 lncRNA MIR155HG 编码的微肽 miPEP155 (P155), 该微肽在炎症抗原呈递细胞中高度表达, 并通过与树突状细胞 (dendritic cells, DC) 中抗原运输

和呈递所需的伴侣, 热休克同源蛋白 70 (heat shock cognate protein 70, HSC70) 的腺苷-5'-三磷酸酯结合域相互作用破坏 HSC70-HSP90 机制, 从而调节细胞的抗原呈递和 T 细胞启动。而外源性注射 P155 被发现可以改善两种经典 DC 驱动的自身炎症小鼠模型, 这证明一种被注释为“非蛋白编码”的转录本编码的微肽的内源性存在, 并且这种微肽可作为抗原呈递的调节剂和炎症性疾病的抑制剂, 用于炎症性疾病的靶向治疗^[61]。此外, Jackson 等^[62]在研究炎症的先天免疫中发现一种微肽, 该微肽可在炎症性

肠病小鼠模型中产生针对感染和炎症的先天免疫反应。有趣的是该微肽是由一种包含非规范 ORF 的 lncRNA Aw112010 编码, 因此将包含非规范 ORF 的基因错误注释为非编码 RNA 可能会掩盖大量以前未发现的蛋白质编码基因在免疫和疾病中的重要作用^[62]。

4 与癌症相关的微肽

真核转录物的很大一部分被认为是 lncRNA, 可调节各种癌症特征。越来越多的证据表明 lncRNA 包含 sORF, 这些 sORF 可被翻译成具有功能的微肽。研究表明, 这些功能性微肽而非 lncRNA 本身对人类疾病及癌症起着关键的调控作用, 并且可为疾病及癌症治疗提供宝贵的方法及理论依据。但由于现阶段技术水平等因素的限制, 这些功能性微肽与人类疾病及癌症之间的关联在很大程度上仍然未知。此外, 现阶段 lncRNA 编码的微肽主要在这些癌症中被研究: 黑色素瘤、结肠癌、乳腺癌、胶质瘤、肝细胞癌、食管鳞状细胞癌、肺癌、头颈鳞癌和急性髓性白血病等。

4.1 黑色素瘤

黑色素瘤 (melanoma) 是一种恶性程度最高的皮肤癌类型^[63]。在 2008 年到 2010 年, 研究者相继在黑色素瘤细胞中发现 MELOE-1 和 MELOE-2 两种微肽, 该两种微肽都由 lncRNA meloe (melanoma-overexpressed antigen) 的 IRES 依赖性翻译^[64] (图 3A)。并且在研究中发现 MELOE-1 和 MELOE-2 与黑色素瘤患者的复发预防相关。Godet 等^[11]将输注含有 MELOE-1 特异性 T 细胞的肿瘤浸润性淋巴细胞 (tumor-infiltrating lymphocyte, TILs) 与人类组织相容性白细胞抗原 A2 (human histocompatibility leukocyte antigen A2, HLA-A2) 患者的复发预防之间进行关联分析发现, 9 名没有复发的患

者中, 有 5 名输注了含有 MELOE-1 特异性 T 细胞的 TIL, 而 21 名复发的患者中, 有 0 名输注了这种含有 TIL 的淋巴细胞。Godet 等^[12]在输注了 TIL 治疗后没有复发的病人的 TIL 群体中也检测到 MELOE-2 反应性 T 细胞的存在。这些结果表明, meloe 编码的这类黑色素瘤抗原可能通过参与 T 细胞免疫监视过程, 减少患者的复发概率, 这可为黑色素瘤免疫治疗提供一种极具吸引力的方案。

4.2 结直肠癌

结直肠癌 (colorectal cancer, CRC) 是胃肠道常见的恶性肿瘤^[65]。2017 年, Huang 等^[66]在 CRC 组织中发现一种可抑制 CRC 恶性发展的微肽, 该微肽由 lncRNA HOXB-AS3 编码。并且该微肽通过与异型核糖核酸蛋白 A1 (heterogeneous nuclear ribonucleoprotein A1, hnRNP A1) RGG 基序中的甘氨酸残基竞争性结合, 拮抗由 hnRNP A1 介导的丙酮酸激酶 M (pyruvate kinase M, PKM) 剪接调节, 从而调控癌症代谢重编程过程影响癌症的恶性发展^[66] (图 3B)。而另一种由肠道特异性 LINC00675 编码的抑制癌症恶性发展的微肽 FORCP (FOXA1-regulated conserved small protein), 其主要通过响应内质网应激调节 CRC 细胞凋亡这一过程发挥抑癌作用^[67] (图 3C)。此外, 由 lncRNA 编码的微肽在 CRC 中还可以充当致癌因子促进癌症的进展。2020 年, Meng 等^[21]在晚期结直肠癌中发现一种可通过与多种剪接调节因子相互作用促进癌症恶性发展的微肽 SRSP (splicing regulatory small protein), 其由 LOC90024 编码 (图 3D)。并且在同一年 Zhu 等也在 CRC 中发现一种由 LINC00266-1 编码的微肽 RBRP (RNA-binding regulatory peptide), 该微肽可通过与胰岛素样生长因子-2 mRNA 结合蛋白 1 (insulin-like growth factor-2 mRNA-

binding protein 1, IGF2BP1) 相互作用促进肿瘤恶性发展^[68](图 3E)。2021 年, Ge 等^[69]在 CRC 中鉴定一种由 LINC00467 编码的微肽 ASAP (ATP synthase-associated peptide), 其通过与 α 和 γ 亚基 (ATP5A 和 ATP5C) 相互作用增强 ATP 合酶的构建, 从而增加 ATP 合酶活性和线粒体耗氧率, 促进 CRC 细胞增殖 (图 3F)。综上所述, lncRNA 编码的微肽可在 CRC 发生与发展过程中充当抑癌因子或致癌因子起调控作用。

4.3 乳腺癌

乳腺癌 (breast cancer, BC) 是恶性组织在乳腺组织内形成的一种疾病^[70]。2018 年, Polycarpou-Schwarz 等^[71]在 BC 中发现一种由 lncRNA 编码的微肽 CASIMO1 (cancer-associated small integral membrane open reading frame 1), 该微肽在激素受体阳性的乳腺肿瘤中高表达 (图 3G)。并且该微肽可通过与调节脂滴形成的 BC 致癌基因角鲨烯环氧化酶 (squalene epoxidase, SQLE) 相互作用, 调节细胞脂质稳态并介导 BC 的恶性发展^[71]。2022 年, Zhang 等^[72]在 BC 中发现一种由 lncRNA CTD-2256P15.2 编码的微肽 PACMP (PAR-amplifying and CtIP-maintaining micropeptide), 该微肽通过调控 DNA 损伤应答过程影响肿瘤的发展和耐药 (图 3H)。并且靶向 PACMP 能导致 DNA 同源重组修复缺陷, 促进肿瘤细胞对放疗、化疗、靶向治疗等多种药物的敏感性, 并明显改善肿瘤疗效^[72]。三阴性乳腺癌 (triple negative breast cancer, TNBC) 是 BC 的一种亚型, 其具有最具侵袭性的表型和较差的总体生存率^[73]。Wang 等^[74]在 TNBC 组织中鉴定出一种由 LINC00908 编码的微肽 ASRPS (a small regulatory peptide of STAT3), 该微肽在 TNBC 中表达下调并与较差的总体生存率相关。此外, 该微肽可通过抑制 TNBC 的血管生成, 抑制 TNBC 的恶性发

展^[74] (图 3I)。表明 LINC00908 编码的微肽 ASRPS 可作为 TNBC 特异性治疗靶点。2022 年, Wu 等^[75]在 TNBC 组织中发现另外一种与 TNBC 的血管生成和转移相关的微肽 XBP1SBM (XBP1s binding micropeptide), 其由 lncRNA MLLT4-AS1 编码。主要通过改善 TNBC 中的谷氨酰胺 (glutamine, Gln) 水平促进血管生成和转移, 从而促进 TNBC 的恶性发展。有研究表明, 转化生长因子- β (transforming growth factor- β , TGF- β) 信号通路在 BC 转移中起关键作用, 并且发现 TGF- β 主要通过调控 lncRNA 转录与翻译过程调控肿瘤的进展。由 LINC00665 编码的微肽 CIP2A-BP (CIP2A binding peptide), 其可通过与癌性蛋白磷酸酶 2A 抑制剂 (Cancerous inhibitor of protein phosphatase 2A, CIP2A) 结合抑制 TNBC 的恶性发展^[20] (图 3J)。但在 TNBC 细胞系中 CIP2A-BP 的翻译被 TGF- β 下调, 这使得癌症向恶性发展, 而这一过程的发生主要是由于 TGF- β 激活的 Smad 信号通路诱导的翻译抑制蛋白 4E 结合蛋白 1 (4E-binding protein 1, 4E-BP1) 抑制翻译起始因子真核生物启动因子 4E (eukaryotic initiation factor 4E, eIF4E) 导致的^[20]。这提示着在 TNBC 的发展中 TGF- β 可能扮演重要的致癌作用, 并且可为日后 CIP2A-BP 在 TNBC 的靶向治疗提供新的方法。

4.4 胶质瘤

胶质瘤 (glioma) 是一种致命的恶性脑癌。许多报道表明水和离子通道的行为异常在调节肿瘤增殖、迁移、凋亡和分化中起关键作用^[76-77]。研究发现一些含有 sORF 的 lncRNA 可编码形成用于水或离子调节的寡聚体微肽。然而, 由于微肽肽段难以识别, 导致其功能机制还远未被清楚了解。2021 年, Cao 等^[77]在胶质瘤中发现一种可编码形成五聚体通道的微肽的 lncRNA DLEU1, 其包含两个 sORF (ORF1 和 ORF8)。

在模拟水和离子 (Na^+ 和 Cl^-) 通过该五聚体通道的传输机制的预测模型中, 发现 ORF1 编码的肽五聚体可作为一个自组装的水通道而不是离子通道, 而 ORF8 则既不渗透离子也不渗透水^[77]。表明 lncRNA 编码的微肽在水和离子通道的调节中起着重要作用, 这可为后续胶质瘤的治疗及药物开发提供方法和理论的支持。值得注意的是, 许多 lncRNA 被选择性地分类到细胞外囊泡 (extracellular vesicles, EVs) 中, 并以多种方式调节癌症的发生和发展^[78]。因此, 微肽是否也可以与 EVs 一起发挥作用 (例如传播分子和信号), 从而改变或调节局部或远距离的受体细胞。在 2021 年, Cai 等^[24]从神经胶质瘤 EVs 中鉴定出一类长度介于 20–100 个氨基酸的微肽, 其由 lncRNA 编码。并且研究人员分别从构建的胶质瘤患者和健康受试者的 EVs 微肽表达谱中, 鉴定出了 64 种和 48 种微肽。对两者进行对比, 发现癌症患者中缺少 17 种微肽, 但检测到 19 个新的微肽, 表明两者表达谱存在显著差异, 暗示 EVs 中的微肽可能在识别胶质瘤患者方面具有潜在的诊断应用^[24]。此外, 研究者在检查人血浆时, 分别从纯化的 EVs、全血浆和不含 EVs 的血浆中鉴定出 48、11 和 3 种微肽, 表明微肽主要富含于 EVs 中, 并且由 EVs 介导的微生物蛋白转移可能代表了一种新的细胞间通讯机制^[24]。2021 年, Huang 等^[79]在胶质母细胞瘤 (glioblastoma, GBM) 中发现一种由磷酸酶和张力蛋白同源物 (phosphatase and tensin homolog, PTEN) 上游开放阅读框 (upstream open reading frame, uORF) 编码的微肽 MP31, 该微肽通过与线粒体乳酸脱氢酶 (mitochondrial lactate dehydrogenase, mLDH) 竞争, 限制线粒体中乳酸-丙酮酸转化烟酰胺腺嘌呤二核苷酸 (nicotinamide adenine dinucleotide, NAD^+)。此外, 研究者在敲除 MP31 的小鼠体

内发现乳酸代谢整体被增强, 而敲除 MP31 的星形胶质细胞中, 显示出胶质瘤生成被激活和动物的总体存活时间被缩短的情况^[79]。这些结果表明, GBM 中 MP31 可通过协调乳酸代谢重编程过程, 调控胶质母细胞瘤的恶性进展。

4.5 肝细胞癌

肝细胞癌 (hepatocellular carcinoma, HCC) 是一种始发于肝脏的恶性肿瘤^[80]。2020 年, Pang 等^[23]发现癌细胞中有许多 lncRNA 可与核糖体蛋白 S6 (ribosomal protein S6, RPS6) 结合, 并且这些 lncRNA 具有编码微肽的潜能。其中 LINC00998 被广泛关注, 并在研究中发现 LINC00998 可编码一种含 59 个氨基酸的保守肽 SMIM30 (small integral membrane protein 30)^[23] (图 3K)。该微肽可通过调节细胞增殖和迁移促进 HCC 的肿瘤发生, 而且这一调控过程与 lncRNA 本身无关, 表明 SMIM30 具有独立的功能^[23]。另一种具有独立功能的微肽 KRASIM (KRAS-interacting microprotein), 其由 lncRNA NCBP2-AS2 编码, 并通过与 Kirsten 大鼠肉瘤病毒癌基因同源物 (Kirsten ratsarcoma viral oncogene homolog, KRAS) 相互作用降低 KRAS 的蛋白水平, 从而抑制 HCC 细胞中致癌信号的传导^[81] (图 3L)。表明 KRASIM 是一种 KRAS 通路的新型微生物蛋白抑制剂, 可为致癌信号调节和 HCC 治疗提供有价值的策略。另外, 诱导癌细胞凋亡也是微肽抑制癌症向恶性发展的重要方式。由 lncRNA HBVPTPAP 编码的微肽通过调节 Janus 酪氨酸激酶/信号转导与转录激活子 (Janus tyrosine kinase/signal transducer and activator of tran-ions, JAK/STAT) 信号通路诱导 HCC 细胞凋亡, 从而抑制 HCC 恶性发展^[82] (图 3M)。虽然已有许多已知参与人类癌症进展的 lncRNA 被证明可以编码具有生物学功能的肽, 但是 lncRNA 编码的微肽在 HCC 细胞衰老

中的作用很大程度上仍是未知。Xiang 等^[26]报道了一种在 HCC 细胞衰老中起重要作用的 microRNA PINT87aa (p53-induced transcript 87aa), 该 microRNA 由 LINC-PINT 编码。PINT87aa 在过氧化氢诱导的 HCC 细胞衰老模型中显著上调, 并且过表达的 PINT87aa 可诱导生长抑制、细胞衰老和减少线粒体自噬 (图 3N)。这一结果表明 PINT87aa 可能成为日后 HCC 治疗的潜在靶点和预后的新型生物标志物。此外, LINC-PINT 的第二个外显子可通过自环化形成环状分子 circPINT, 并编码功能性 microRNA PINT87aa。而该 microRNA 可直接与聚合酶相关因子复合物 (polymerase associated factor complex, PAF1c) 相互作用抑制多个癌基因的转录延长, 从而抑制胶质母细胞瘤肿瘤细胞增殖^[83]。表明 LINC-PINT 在癌症的发生和发展中起着重要作用, 可作为日后癌症治疗及预后的生物标志物。

4.6 食管鳞状细胞癌

食管鳞状细胞癌 (esophageal squamous cell carcinoma, ESCC) 是食管癌中最常见的病理类型, 约占食管癌的 90%, 是一种临床比较严重的疾病^[84]。然而, ESCC 与 lncRNA 编码的 microRNA 之间的关联在很大程度上仍然未知。2020 年, Wu 等^[85]在男性 ESCC 中表征了一种 Y 连锁的 LINC00278, 该 lncRNA 第一个外显子的 sORF 可编码一种阴阳 1 结合 microRNA (Yin Yang 1 binding micropeptide, YY1BM), 并且该 microRNA 是一种潜在的抗癌因子。由吸烟诱导的 ALKBH5 蛋白可通过介导 lncRNA m6A 去甲基化, 降低 LINC00278 的 sORF 翻译效率^[85]。此外, YY1BM 可参与 ESCC 进展并抑制 YY1 和雄激素受体 (androgen receptor, AR) 之间的相互作用, 从而通过 AR 信号通路降低真核延伸因子-2 激酶 (eukaryotic elongation factor-2 kinase, eEF2K) 的表达^[86]。因此 YY1BM 的下调可显著上调 eEF2K 表达并抑

制细胞凋亡, 从而使 ESCC 细胞更适应营养缺乏状态。这些结果提示了吸烟与男性 ESCC 进展中 AR 信号传导之间的机制联系, 当由吸烟诱导的 Y 染色体相关的 lncRNA m6A 去甲基化时, 男性患癌症的风险增加。

4.7 其他癌症

肺癌 (lung cancer) 是肺内细胞出现基因突变, 出现恶性增长的过程。Lu 等^[87]在肺癌中发现一种定位于核仁的 microRNA UBAP1-AST6, 其由 lncRNA 编码并在肺癌细胞系中表达, 且与肺癌细胞的恶性增殖相关。Li 等^[88]在头颈部鳞状细胞癌 (head and neck squamous cell carcinoma, HNSCC) 中发现由 lncRNA RP11-469H8.6 编码的内源性 microRNA MIAC (micropeptide inhibiting actin cytoskeleton), 该 microRNA 通过与水通道蛋白 2 (aquaporin 2) 相互作用抑制 HNSCC 向恶性发展 (图 3O)。此外, 研究者通过结合 500 例 TCGA 数据库中 HNSCC RNA-seq 的数据和 154 例临床样本检测分析证实 MIAC 在 HNSCC 组织中表达显著低于癌旁组织, 其表达水平与患者总生存率呈正相关, 与临床病理分期、区域淋巴结转移呈负相关, 提示 MIAC 可作为 HNSCC 临床诊断和预后的潜在标志物^[88]。急性髓性白血病 (acute myeloid leukemia, AML) 是一种造血系统恶性肿瘤, 源于骨髓内恶性造血前体细胞的克隆性扩增^[89-90]。2021 年, Sun 等^[91]在 AML 中发现一种由 lncRNA AHS1L-AS1 编码的 microRNA APPLE (a peptide located in ER)。该 microRNA 在 AML 中高表达, 并通过促进聚腺苷酸结合蛋白胞质 1 (polyadenylate-binding protein cytoplasmic 1, PABPC1) 与真核生物启动因子 4G (eukaryotic initiation factor 4G, eIF4G) 相互作用、mRNA 环化以及 eIF4F 起始复合物组装, 以支持特定的促癌翻译程序, 从而促进造血系统恶性肿瘤的发展^[91](图 3P)。该

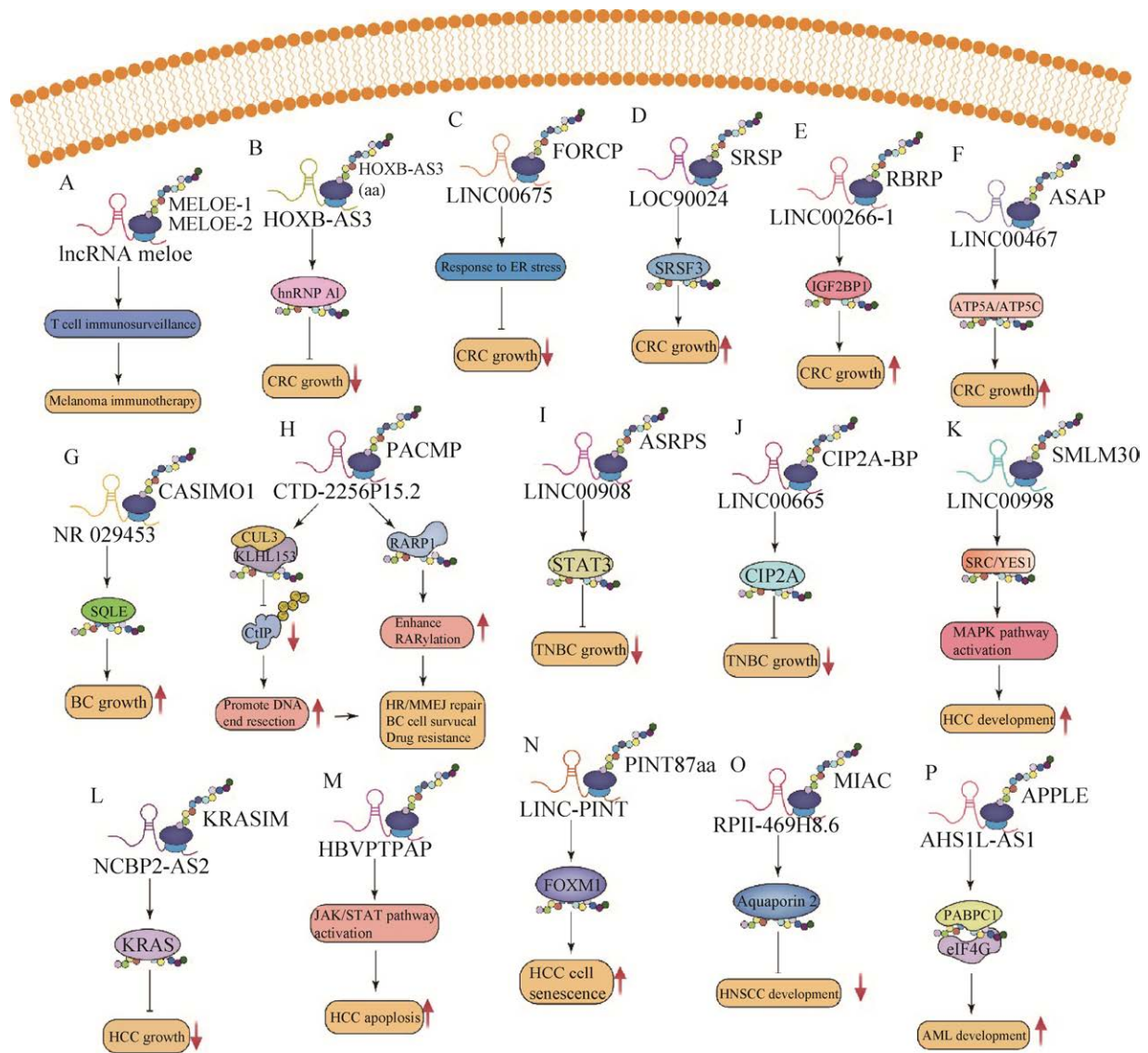


图 3 lncRNA 编码的微肽在癌症中的调控机制示意图

Figure 3 Schematic diagram of the regulatory mechanism of lncRNA-encoded micro-peptides in cancer. (A) The micro-peptides MELOE-1 and MELOE-2 encoded by lncRNA meloe are involved in melanoma immunotherapy by participating in the T cell immune surveillance process. (B) The micro-peptide HOXB-AS3 (aa) encoded by lncRNA HOXB-AS3 inhibits the malignant growth of CRC by interacting with hnRNP A1. (C) The micro-peptide FORCP encoded by LINC00675 inhibits the malignant growth of CRC by responding to endoplasmic reticulum stress. (D) The micro-peptide SRSP encoded by LOC90024 promotes the malignant growth of CRC by interacting with serine- and arginine-rich splicing factor 3 (SRSF3). (E) The micro-peptide RBRP encoded by LINC00266-1 promotes the malignant growth of CRC by interacting with IGF2BP1. (F) The micro-peptide ASAP encoded by LINC00467 regulates ATP synthase activity by interacting with α and γ subunits (ATP5A and ATP5C), and promotes the malignant development of colorectal cancer. (G) The micro-peptide CASIMO1 encoded by lncRNA NR_029453 promotes the malignant growth of BC by interacting with SQLE. (H) lncRNA CTD-2256P15.2 encodes a micro-peptide PACMP that inhibits ubiquitinated degradation of CtIP (CTBP-interacting protein) by binding to Cullin3-KLHL15 to

promote DNA end resection on the one hand, and enhances RARylation by binding to RARP1 on the other hand, thereby regulating DNA repair, breast cancer cell survival and drug resistance. (I) The micro-peptide ASRPS encoded by LINC00908 inhibits the malignant growth of TNBC by interacting with STAT3. (J) The micro-peptide CIP2A-BP encoded by LINC00665 inhibits the malignant growth of TNBC by combining with the oncogene CIP2A. (K) The micro-peptide SMIM30 encoded by LINC00998 promotes the malignant development of HCC by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. (L) The micro-peptide KRASIM encoded by lncRNA NCBP2-AS2 inhibits the malignant growth of HCC by interacting with KRAS. (M) The micro-peptide encoded by lncRNA HBVPTPAP induces apoptosis of HCC cells by activating the JAK/STAT signaling pathway. (N) The micro-peptide PINT87aa encoded by LINC-PINT induces senescence of HCC cells by binding to FOXM1. (O) The micro-peptide MIAC encoded by lncRNA RP11-469H8.6 inhibits the malignant development of HNSCC by interacting with aquaporin 2. (P) The micro-peptide APPLE encoded by lncRNA AHS1L-AS1 enhances translation initiation by interacting with PABPC1-eIF4G and promotes the malignant development of AML.

研究所提供的证据表明,通过靶向 APPLE 可为 AML 的潜在靶向治疗提供新的治疗策略。

5 与其他疾病相关微肽

肺动脉高压 (pulmonary hypertension, PH) 是一种罕见且致命的疾病,由缺氧引起的肺动脉平滑肌细胞 (pulmonary artery smooth muscle cells, PASMCs) 增殖是 PH 的重要病理过程,并且防止 PASMCs 增殖可有效降低 PH 死亡率^[92]。Liu 等^[93]在缺氧小鼠模型中发现一种具有高编码能力的 lnc-Rps41,其在低氧条件下可介导 PASMCs 增殖。该团队随后又发现其编码的微肽 RPS4XL (ribosomal protein S4 X isoform-like) 可通过与核糖体蛋白 S6 (ribosomal protein S6, RPS6) 结合调控 RPS6 磷酸化,抑制缺氧引起的 PASMCs 增殖^[94]。这一发现表明具有高编码能力的 lnc-Rps41 可作为缺氧 PH 早期诊断和治疗的潜在靶点。急性心肌梗塞 (acute myocardial infarction, AML) 是冠状动脉发生急性堵塞,引起部分心肌缺血性坏死的一种严重冠心病^[95]。Spencer 等^[96]发现一种在血管生成中起重要作用的微肽 SPAAR (small regulatory polypeptide of amino acid response),其由 LINC00961 编码。在小鼠模型中,发现 LINC00961/SPAAR 基因座丢失可影响小鼠发育、心肌动力学和心肌梗

死心脏反应^[97]。数据表明, LINC00961/SPAAR 有助于生长和发育以及成年期的基础心血管功能、减轻心肌梗塞风险。

6 胚胎发育相关微肽

昆虫中的分节基因 (segmentation genes) 是产生早期胚胎体节划分的必要条件。2006 年, Savard 等^[98]发现一种面粉甲虫 (*Tribolium*) 中分节基因间隙家族成员 mlpt (mille-pattes)。敲低 mlpt 可导致腹节转化为胸节,为胚胎提供多达十对腿。此外, mlpt 和 *Tribolium* 中已知的间隙基因之间存在交叉调节作用,并且其可转录一种含 4 种 sORF 的多顺反子 mRNA,这一发现提示着由其编码的微肽可能在细胞组织发育成胚胎过程中起着关键作用,但其具体机制还有待探索^[98]。Galindo 等^[9]在果蝇中发现一种之前被推定为非编码 RNA 的基因 tal (tarsal-less),其可被翻译成一种只有 11 个氨基酸的微肽,并且该微肽可通过控制果蝇相关基因表达和组织折叠控制胚胎发育。而另一种之前在果蝇中被鉴定为非编码 RNA 的基因 pri (polished rice),其进化上保守的 sORF 可编码 11 或 32 个氨基酸的微肽,并且 pri 在胚胎发生过程中在所有上皮组织中表达^[10]。此外, pri 功能的丧失完全消除了顶端角质层结构,并且还导致气管导

管扩张缺陷^[10]。而这个研究提供的证据表明, pri 非细胞自主地起作用, 并且 pri 上 4 个保守的 sORF 在功能上可能是冗余的。但是在随后的研究中证明, 在果蝇胚胎发生过程中, 由 pri 的 sORF 编码的微肽可通过修饰转录因子 Svb (shavenbaby) 控制果蝇的表皮分化^[99]。并且在果蝇胚胎发生过程中, pri 编码的微肽可为表皮形态发生的转录程序提供严格的控制, 表明 pri 编码的微肽在功能上并不是冗余的。此外, 在果蝇表皮分化中鉴定的分子复合物, 包括 mlpt、泛素连接酶 Ubr3 和 Svb, 代表了早期昆虫胚胎模式化所需的古老发育模块^[100]。而该模块的分节功能的丧失与早期果蝇胚胎中 Svb 的表达受限同时发生^[100]。当人为恢复早期 Svb 的表达会导致依赖于 mlpt 功能的分节缺陷, 这表明尽管胚胎模式在不断演变, 但古老的发育“开关”仍具有持久的效力。这些结果突出保守的分子复合物在基本遗传网络的约束下的进化可塑性。此外, Pauli 等^[22]在斑马鱼胚胎发生过程中发现一种由 LOC100506013 编码的保守性微肽 toddler, 并且该微肽的缺失和过表达都会减少斑马鱼原肠胚形成过程中中内胚层细胞的运动。综上所述, 微肽对果蝇和斑马鱼的胚胎发育起着关键作用, 但是微肽在哺乳动物胚胎发育方面的研究还相对较少, 其具体作用机制还不清晰, 因此探究微肽在哺乳动物胚胎发育方面的研究将更具有挑战性。

7 问题与展望

随着研究不断的深入, 许多具有编码能力的 lncRNA 被发现, 使得人们开始重新审视 lncRNA 在生物体中的作用。而由 lncRNA 编码的微肽大都独立于 lncRNA 在各种生物过程中发挥作用, 因此在研究中应注意将微肽的功能和 lncRNA 本身的功能区分开。另外, 在早前

的研究中发现微肽可通过参与生物体细胞程序性死亡、抗原呈递、钙离子转运、线粒体代谢调控、废物降解和肌细胞融合等过程, 调控机体的稳态、肌肉的生理功能、疾病和癌症的发生与发展、胚胎发育等过程(微肽功能总结见表 1)。因此, 关于微肽的研究将有助于研究者对生物体生理功能的解析和对疾病及癌症靶向治疗药物的研发。

在近 10 年的研究中发现大部分关于微肽的研究主要集中在人类疾病和癌症。并且微肽主要通过充当致癌因子或抑癌因子影响癌症的进展。当前人们致力于开发新型癌症靶向治疗药物, 并从中发现微肽类靶向治疗药物, 相比于当前各类化疗药物或靶向治疗药物, 具有特异性和活性高、细胞毒性小、免疫原性低等优点^[102-103]。所以微肽类靶向治疗药物将有可能成为一类极具吸引力的靶向治疗药物。lncRNA 编码的微肽具有肽段较短、分子量小以及表达丰度低等特点, 而这可能造成微肽类药物提取和合成困难以及相关检测技术无法准确识别等问题。因此相应技术的进一步优化和巧妙的实验设计, 将是微肽领域取得持续进展的关键。此外, 越来越多的研究发现翻译组学的发展可为微肽领域研究提供坚实的助力, 其中 Ribo-seq 更是当前研究 lncRNA 编码能力的重要手段。研究者可将 Ribo-seq 与 RNA-seq 或 lncRNA-seq 进行联合分析, 从中寻找具有编码潜力的关键 lncRNA。同时在此基础上结合蛋白质组学, 可对新微肽进行更深入地挖掘及佐证。并且获得的关键 lncRNA 也可通过相关编码潜力评估及预测工具、软件进行进一步评估与预测, 从而获得更为准确可信的数据。

尽管当前不断有微肽的功能被表征和发现, 但功能性微肽的筛选机制仍饱受争议。主要是由于基于保守性分析和规范化的微肽筛选

表 1 lncRNA 编码的微肽及其功能

Table 1 Micro-peptides encoded by lncRNA and their functions

Species	lncRNA	Micro-peptide	Length (aa)	Function	Physiological processes	References
Soy bean	ENOD40	ENOD40 (aa)	12/24	Interaction with sucrose synthase	Plant physiology	[8]
Human/ mouse	LINC00948/ AK009351	MLN	46	Important regulator of skeletal muscle physiology	Muscle physiology	[14]
Human/ mouse	LOC100507537/ NONMMUG026737	DWORF	34	Enhances muscle contractility, prevents heart failure and attenuates cardiomyopathy in mouse; candidate drug for heart failure treatment	Muscle physiology	[15,46]
Mouse	LINC00961	SPAR	75	Regulate muscle regeneration	Muscle physiology	[16]
Mouse	LOC101929726	Myomixer	84	Promote myoblast fusion, regulate muscle development and muscle fiber regeneration after injury	Muscle physiology	[54-56]
Mouse	LOC101929726	Minion	84	Control of cell fusion and muscle formation	Muscle physiology	[25]
Mouse	LINC00116/ 1500011K16Rik	MOXI	56	Regulation of skeletal muscle mitochondrial metabolism	Muscle physiology	[18]
Mouse	LINC00116	Mtln	56	Regulation of skeletal muscle mitochondrial metabolism	Muscle physiology	[17]
Mouse/ zebrafish	MyolncR4/ 1500011K16RIK	LEMP	56	Promotes muscle formation and regeneration in mouse	Muscle physiology	[58]
Chicken	Six1	Six1 (aa)	74	Promotes cell proliferation and participates in muscle growth	Muscle physiology	[19]
Mouse	1810058I24Rik	Mm47	47	Controls innate immunity and influencing inflammation	Inflammation and immunity	[60]
Human	MIR155HG	miPEP155 (P155)	17	Regulation of antigen presentation suppresses autoimmune inflammation	Inflammation and immunity	[61]
Human	Meloe	MELOE-1/ MELOE-2	46/39	Involved in T cell immune surveillance; optimal T cell target for melanoma immunotherapy	Melanoma	[11-12,64]
Human	HOXB-AS3	HOXB-AS3 (aa)	53	Modulating the process of cancer metabolic reprogramming and inhibiting colorectal cancer growth	Colorectal cancer	[66]
Human	LINC00675	FORCP	79	Regulates colorectal cancer cell apoptosis, inhibits proliferation and tumorigenesis	Colorectal cancer	[67]
Human	LINC00266-1	RBRP	71	Promote colorectal cancer tumorigenesis	Colorectal cancer	[68]
Human	LINC00467	ASAP	94	Promote colorectal cancer cell proliferation	Colorectal cancer	[69]
Human	NR_029453	CASIMO1	83	Promote the malignant development of breast cancer	Breast cancer	[71]
Human	CTD-2256P15.2	PACMP	44	Promote the malignant development of breast cancer	Breast cancer	[72]

待续

续表 1

Species	lncRNA	Micro-peptide	Length (aa)	Function	Physiological processes	References
Human	LINC00908	ASRPS	60	Inhibit the malignant development of triple negative breast cancer	Breast cancer	[74]
Human	MLLT4-AS1	XBP1SBM	21	Promote the malignant development of triple negative breast cancer	Breast cancer	[75]
Human	LINC00665	CIP2A-BP	52	Inhibit the malignant development of triple negative breast cancer	Breast cancer	[20]
Human	LINC00998	SMIM30	59	Regulation of cell proliferation and migration, promotes tumorigenesis in hepatocellular carcinoma	Hepatocellular carcinoma	[23]
Human	NCBP2-AS2	KRASIM	99	Inhibition of oncogenic signaling in hepatocellular carcinoma cells	Hepatocellular carcinoma	[81]
Human	LINC-PINT	PINT87aa	87	Induction of senescence in hepatocellular carcinoma cells; novel biomarkers and potential therapeutic targets	Hepatocellular carcinoma	[26]
Human	LINC00278	YY1BM	21	Inhibit the malignant development of esophageal squamous cell carcinoma	Esophageal squamous cell carcinoma	[85]
Human	UBAP1-AST6	BAP1-AST6 (aa)	–	Involved in the malignant proliferation of lung cancer cells	Lung cancer	[87]
	RP11-469H8.6	MIAC	51	Inhibiting the malignant progression of Head and neck squamous cell carcinoma	Head and neck squamous cell carcinoma	[88]
Human	AHS1L-AS1	APPLE	90	Promote the development of acute myeloid leukemia	Acute myeloid leukemia	[91]
Human	Rps4l	RPS4XL	–	Inhibition of hypoxia-induced proliferation of pulmonary artery smooth muscle cells	Pulmonary hypertension	[94]
Human	LINC00961	SPAAR	75	Reduce the risk of myocardial infarction	Acute myocardial infarction	[96-97]
Fruit fly	Pri (polished rice)	Pri (aa)	11/32	Control of <i>Drosophila</i> epidermal differentiation during <i>Drosophila</i> embryogenesis	Embryonic development	[10]
Fruit fly	Tal (tarsal-less)	Tal (aa)	11	Control of <i>Drosophila</i> -related gene expression and tissue folding and control of embryonic development	Embryonic development	[9]
Zebrafish	LOC100506013	Toddler	58	Promotes zebrafish embryogenesis	Embryonic development	[22]
Mouse	Gm9999	Kastor/Polluks	53/40	Regulation of voltage-dependent anion channels (VDAC) and spermatogenesis	Reproductive physiology	[101]

机制，可能会造成大量非保守和非规范化的微肽被“隐藏”或丢弃，而这可能会造成一些在生物体调控过程中有着重要调控作用的微肽的缺

失。因此如何优化当前的微肽筛选机制将是进一步推动微肽领域研究的有力举措。另外，当前注释具有编码能力的 lncRNA 物种数据库还

相对较少, 只包括人和少数几种模式动物。并且由于物种间差异, 导致许多数据库不足以满足现阶段对微肽研究的要求, 所以基于不同物种间功能注释数据库的建立, 就显得尤为重要。此外, 尽管微肽在人类和模式动物上进行了许多研究, 但是在其他物种如奶牛、山羊、猪等重要经济动物上的研究还相对较少, 具体作用机制还不清晰。然而, 这些经济动物在人们的生活中又是不可或缺的部分, 所以关于这方面研究的推进也将是重中之重。毋庸置疑, lncRNA 编码微肽这一机制将引领新的研究热潮和推动生命科学领域的发展, 并且其提供的新视角将为未来抗癌药物和预后生物标志物的研究提供新的见解。

REFERENCES

- [1] Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature*, 2012, 489(7414): 101-108.
- [2] Consortium ENCOEEP, Moore JE, Purcaro MJ, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, 2020, 583(7818): 699-710.
- [3] Wang JZ, Zhu S, Meng N, et al. ncRNA-encoded peptides or proteins and cancer. *Mol Ther*, 2019, 27(10): 1718-1725.
- [4] Mackowiak SD, Zauber H, Bielow C, et al. Extensive identification and analysis of conserved small ORFs in animals. *Genome Biol*, 2015, 16: 179.
- [5] Statello L, Guo CJ, Chen LL, et al. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*, 2021, 22(2): 96-118.
- [6] Wu P, Mo YZ, Peng M, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol Cancer*, 2020, 19(1): 22.
- [7] Li J, Liu CN. Coding or noncoding, the converging concepts of RNAs. *Front Genet*, 2019, 10: 496.
- [8] Rohrig H, Schmidt J, Miklashevichs E, et al. Soybean ENOD40 encodes two peptides that bind to sucrose synthase. *PNAS*, 2002, 99(4): 1915-1920.
- [9] Galindo MI, Pueyo JI, Fouix S, et al. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol*, 2007, 5(5): e106.
- [10] Kondo T, Hashimoto Y, Kato K, et al. Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nat Cell Biol*, 2007, 9(6): 660-665.
- [11] Godet Y, Moreau-Aubry A, Guilloux Y, et al. MELOE-1 is a new antigen overexpressed in melanomas and involved in adoptive T cell transfer efficiency. *J Exp Med*, 2008, 205(11): 2673-2682.
- [12] Godet Y, Moreau-Aubry A, Mompelat D, et al. An additional ORF on *Meloe* cDNA encodes a new melanoma antigen, MELOE-2, recognized by melanoma-specific T cells in the HLA-A2 context. *Cancer Immunol Immunother*, 2010, 59(3): 431-439.
- [13] Tonkin J, Rosenthal N. One small step for muscle: a new micropeptide regulates performance. *Cell Metab*, 2015, 21(4): 515-516.
- [14] Anderson DM, Anderson KM, Chang CL, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell*, 2015, 160(4): 595-606.
- [15] Nelson BR, Makarewich CA, Anderson DM, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science*, 2016, 351(6270): 271-275.
- [16] Matsumoto A, Pasut A, Matsumoto M, et al. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature*, 2017, 541(7636): 228-232.
- [17] Stein CS, Jadiya P, Zhang XM, et al. Mitoregulin: a lncRNA-encoded microprotein that supports mitochondrial super complexes and respiratory efficiency. *Cell Rep*, 2018, 23(13): 3710-3720.e8.
- [18] Makarewich CA, Baskin KK, Munir AZ, et al. MOXI is a mitochondrial micropeptide that enhances fatty acid β -oxidation. *Cell Rep*, 2018, 23(13): 3701-3709.
- [19] Cai BL, Li ZH, Ma MT, et al. LncRNA-Six1 encodes a micropeptide to activate Six1 in cis and is involved in cell proliferation and muscle growth. *Front Physiol*, 2017, 8: 230.
- [20] Guo BB, Wu SQ, Zhu X, et al. Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. *EMBO J*,

- 2020, 39(1): e102190.
- [21] Meng N, Chen M, Chen D, et al. Small protein hidden in lncRNA LOC90024 promotes “cancerous” RNA splicing and tumorigenesis. *Adv Sci (Weinh)*, 2020, 7(10): 1903233.
- [22] Pauli A, Norris ML, Valen E, et al. Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science*, 2014, 343(6172): 1248636.
- [23] Pang YN, Liu ZY, Han H, et al. Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *J Hepatol*, 2020, 73(5): 1155-1169.
- [24] Cai TX, Zhang Q, Wu BW, et al. LncRNA-encoded microproteins: a new form of cargo in cell culture-derived and circulating extracellular vesicles. *J Extracell Vesicles*, 2021, 10(9): e12123.
- [25] Zhang Q, Vashisht AA, O'Rourke J, et al. The microprotein minion controls cell fusion and muscle formation. *Nat Commun*, 2017, 8: 15664.
- [26] Xiang XH, Fu YN, Zhao K, et al. Cellular senescence in hepatocellular carcinoma induced by a long non-coding RNA-encoded peptide PINT87aa by blocking FOXM1-mediated PHB₂. *Theranostics*, 2021, 11(10): 4929-4944.
- [27] Merry TL, Chan A, Woodhead JST, et al. Mitochondrial-derived peptides in energy metabolism. *Am J Physiol Endocrinol Metab*, 2020, 319(4): E659-E666.
- [28] Sieber P, Platzer M, Schuster S. The definition of open reading frame revisited. *Trends Genet*, 2018, 34(3): 167-170.
- [29] Olexiouk V, Crappé J, Verbruggen S, et al. sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic Acids Res*, 2016, 44(D1): D324-D329.
- [30] Yan YC, Tang R, Li B, et al. The cardiac translational landscape reveals that micropeptides are new players involved in cardiomyocyte hypertrophy. *Mol Ther*, 2021, 29(7): 2253-2267.
- [31] Mishra SK, Wang H. Computational analysis predicts hundreds of coding lncRNAs in zebrafish. *Biology*, 2021, 10(5): 371.
- [32] Ruiz-Orera J, Messeguer X, Subirana JA, et al. Long non-coding RNAs as a source of new peptides. *eLife*, 2014, 3: e03523.
- [33] Zhao J, Qin B, Nikolay R, et al. Translatomics: the global view of translation. *Int J Mol Sci*, 2019, 20(1): 212.
- [34] Kang YJ, Yang DC, Kong L, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res*, 2017, 45(W1): W12-W16.
- [35] Guo JC, Fang SS, Wu Y, et al. CNIT: a fast and accurate web tool for identifying protein-coding and long non-coding transcripts based on intrinsic sequence composition. *Nucleic Acids Res*, 2019, 47(W1): W516-W522.
- [36] Johnson AG, Grosely R, Petrov AN, et al. Dynamics of IRES-mediated translation. *Phil Trans R Soc B*, 2017, 372(1716): 20160177.
- [37] Arhab Y, Bulakhov AG, Pestova TV, et al. Dissemination of internal ribosomal entry sites (IRES) between viruses by horizontal gene transfer. *Viruses*, 2020, 12(6): 612.
- [38] Yang Y, Fan XJ, Mao MW, et al. Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res*, 2017, 27(5): 626-641.
- [39] Mills JD, Ward M, Chen BJ, et al. LINC00507 is specifically expressed in the primate cortex and has age-dependent expression patterns. *J Mol Neurosci*, 2016, 59(4): 431-439.
- [40] Ji Z, Song RS, Regev A, et al. Many lncRNAs, 5'UTRs, and *Pseudogenes* are translated and some are likely to express functional proteins. *eLife*, 2015, 4: e08890.
- [41] Wang HW, Wang Y, Xie SQ, et al. Global and cell-type specific properties of lincRNAs with ribosome occupancy. *Nucleic Acids Res*, 2017, 45(5): 2786-2796.
- [42] Frontera WR, Ochala J. Skeletal muscle: a brief review of structure and function. *Calcif Tissue Int*, 2015, 96(3): 183-195.
- [43] Muthu P, Kumar MS, Kumar SS, et al. Role of SERCA pump in muscle thermogenesis and metabolism. *Compr Physiol*, 2017, 7(3): 879-890.
- [44] Anderson DM, Makarewich CA, Anderson KM, et al. Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. *Sci Signal*, 2016, 9(457): ra119.
- [45] Makarewich CA, Munir AZ, Schiattarella GG, et al. The DWORF micropeptide enhances contractility and

- prevents heart failure in a mouse model of dilated cardiomyopathy. *eLife*, 2018, 7: e38319.
- [46] Makarewich CA, Bezprozvannaya S, Gibson AM, et al. Gene therapy with the *DWORF* micropeptide attenuates cardiomyopathy in mice. *Circ Res*, 2020, 127(10): 1340-1342.
- [47] Cho S, Lee G, Pickering BF, et al. mTORC1 promotes cell growth via m6A-dependent mRNA degradation. *Mol Cell*, 2021, 81(10): 2064-2075.e8.
- [48] Villa E, Sahu U, O'Hara BP, et al. mTORC1 stimulates cell growth through SAM synthesis and m6A mRNA-dependent control of protein synthesis. *Mol Cell*, 2021, 81(10): 2076-2093.e9.
- [49] Rion N, Rüegg MA. LncRNA-encoded peptides: more than translational noise? *Cell Res*, 2017, 27(5): 604-605.
- [50] Tajbakhsh S. lncRNA-encoded polypeptide SPAR(s) with mTORC1 to regulate skeletal muscle regeneration. *Cell Stem Cell*, 2017, 20(4): 428-430.
- [51] Chugunova A, Loseva E, Mazin P, et al. LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. *PNAS*, 2019, 116(11): 4940-4945.
- [52] Lin YF, Xiao MH, Chen HX, et al. A novel mitochondrial micropeptide MPM enhances mitochondrial respiratory activity and promotes myogenic differentiation. *Cell Death Dis*, 2019, 10(7): 528.
- [53] Hindi SM, Tajrishi MM, Kumar A. Signaling mechanisms in mammalian myoblast fusion. *Sci Signal*, 2013, 6(272): re2.
- [54] Bi PP, Ramirez-Martinez A, Li H, et al. Control of muscle formation by the fusogenic micropeptide myomixer. *Science*, 2017, 356(6335): 323-327.
- [55] Shi J, Bi PP, Pei JM, et al. Requirement of the fusogenic micropeptide myomixer for muscle formation in zebrafish. *Proc Natl Acad Sci USA*, 2017, 114(45): 11950-11955.
- [56] Bi PP, McAnally JR, Shelton JM, et al. Fusogenic micropeptide myomixer is essential for satellite cell fusion and muscle regeneration. *PNAS*, 2018, 115(15): 3864-3869.
- [57] Magny EG, Pueyo JI, Pearl FMG, et al. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science*, 2013, 341(6150): 1116-1120.
- [58] Wang LT, Fan J, Han LL, et al. The micropeptide LEMP plays an evolutionarily conserved role in myogenesis. *Cell Death Dis*, 2020, 11(5): 357.
- [59] Wang YC, Liu XX, Shi H, et al. NLRP3 inflammasome, an immune-inflammatory target in pathogenesis and treatment of cardiovascular diseases. *Clin Transl Med*, 2020, 10(1): 91-106.
- [60] Bhatta A, Atianand M, Jiang ZZ, et al. A mitochondrial micropeptide is required for activation of the Nlrp3 inflammasome. *J Immunol*, 2020, 204(2): 428-437.
- [61] Niu LM, Lou FZ, Sun Y, et al. A micropeptide encoded by lncRNA MIR155HG suppresses autoimmune inflammation via modulating antigen presentation. *Sci Adv*, 2020, 6(21): eaaz2059.
- [62] Jackson R, Kroehling L, Khitun A, et al. The translation of non-canonical open reading frames controls mucosal immunity. *Nature*, 2018, 564(7736): 434-438.
- [63] Sanchez JA, Robinson WA. Malignant melanoma. *Annu Rev Med*, 1993, 44: 335-342.
- [64] Charpentier M, Dupré E, Fortun A, et al. hnRNP-A1 binds to the IRES of MELOE-1 antigen to promote MELOE-1 translation in stressed melanoma cells. *Mol Oncol*, 2022, 16(3): 594-606.
- [65] Montalban-Arques A, Scharl M. Intestinal microbiota and colorectal carcinoma: implications for pathogenesis, diagnosis, and therapy. *EBioMedicine*, 2019, 48: 648-655.
- [66] Huang JZ, Chen M, Chen D, et al. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Mol Cell*, 2017, 68(1): 171-184.e6.
- [67] Li XL, Pongor L, Tang W, et al. A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. *eLife*, 2020, 9: e53734.
- [68] Zhu S, Wang JZ, Chen D, et al. An oncopeptide regulates m6A recognition by the m6A reader IGF2BP1 and tumorigenesis. *Nat Commun*, 2020, 11(1): 1685.
- [69] Ge QW, Jia D, Cen D, et al. Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. *J Clin Invest*, 2021, 131(22): e152911.
- [70] Liang YR, Zhang HW, Song XJ, et al. Metastatic

- heterogeneity of breast cancer: molecular mechanism and potential therapeutic targets. *Semin Cancer Biol*, 2020, 60: 14-27.
- [71] Polycarpou-Schwarz M, Groß M, Mestdagh P, et al. The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene*, 2018, 37(34): 4750-4768.
- [72] Zhang C, Zhou B, Gu F, et al. Micropeptide PACMP inhibition elicits synthetic lethal effects by decreasing CtIP and poly(ADP-ribosylation). *Mol Cell*, 2022: 2022Feb17;S1097-2022Feb17;S2765(22)00085-5.
- [73] Vagia E, Mahalingam D, Cristofanilli M. The landscape of targeted therapies in TNBC. *Cancers*, 2020, 12(4): 916.
- [74] Wang YR, Wu SQ, Zhu X, et al. LncRNA-encoded polypeptide ASRPS inhibits triple-negative breast cancer angiogenesis. *J Exp Med*, 2020, 217(3): jem.20190950.
- [75] Wu S, Guo B, Zhang L, et al. A micropeptide XBP1SBM encoded by lncRNA promotes angiogenesis and metastasis of TNBC via XBP1s pathway. *Oncogene*, 2022: 2022Feb23.
- [76] Stroka KM, Jiang HY, Chen SH, et al. Water permeation drives tumor cell migration in confined microenvironments. *Cell*, 2014, 157(3): 611-623.
- [77] Cao YP, Yang R, Lee I, et al. Prediction of lncRNA-encoded small peptides in glioma and oligomer channel functional analysis using *in silico* approaches. *PLoS One*, 2021, 16(3): e0248634.
- [78] Dai J, Su YZ, Zhong SY, et al. Exosomes: key players in cancer and potential therapeutic strategy. *Signal Transduct Target Ther*, 2020, 5(1): 145.
- [79] Huang NN, Li FY, Zhang ML, et al. An upstream open reading frame in phosphatase and tensin homolog encodes a circuit breaker of lactate metabolism. *Cell Metab*, 2021, 33(1): 128-144.e9.
- [80] Guo ZW, Meng Y, Zhai XM, et al. Translated long non-coding ribonucleic acid ZFAS1 promotes cancer cell migration by elevating reactive oxygen species production in hepatocellular carcinoma. *Front Genet*, 2019, 10: 1111.
- [81] Xu WL, Deng B, Lin PH, et al. Ribosome profiling analysis identified a KRAS-interacting microprotein that represses oncogenic signaling in hepatocellular carcinoma cells. *Sci China Life Sci*, 2020, 63(4): 529-542.
- [82] Lun YZ, Pan ZP, Liu SN, et al. The peptide encoded by a novel putative lncRNA HBVPTPAP inducing the apoptosis of hepatocellular carcinoma cells by modulating JAK/STAT signaling pathways. *Virus Res*, 2020, 287: 198104.
- [83] Zhang ML, Zhao K, Xu XP, et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat Commun*, 2018, 9(1): 4475.
- [84] Zheng LH, Li LZ, Xie J, et al. Six novel biomarkers for diagnosis and prognosis of esophageal squamous cell carcinoma: validated by scRNA-seq and qPCR. *J Cancer*, 2021, 12(3): 899-911.
- [85] Wu SQ, Zhang LY, Deng JQ, et al. A novel micropeptide encoded by Y-linked LINC00278 links cigarette smoking and AR signaling in male esophageal squamous cell carcinoma. *Cancer Res*, 2020, 80(13): 2790-2803.
- [86] Banday AR, Papenberg BW, Prokunina-Olsson L. When the smoke clears m6A from a Y chromosome-linked lncRNA, men get an increased risk of cancer. *Cancer Res*, 2020, 80(13): 2718-2719.
- [87] Lu SH, Zhang J, Lian XL, et al. A hidden human proteome encoded by 'non-coding' genes. *Nucleic Acids Res*, 2019, 47(15): 8111-8125.
- [88] Li MW, Li X, Zhang YN, et al. Micropeptide MIAC inhibits HNSCC progression by interacting with aquaporin 2. *J Am Chem Soc*, 2020, 142(14): 6708-6716.
- [89] De Kouchkovsky I, Abdul-Hay M. 'Acute myeloid leukemia: a comprehensive review and 2016 update'. *Blood Cancer J*, 2016, 6(7): e441.
- [90] Newell LF, Cook RJ. Advances in acute myeloid leukemia. *BMJ*, 2021: n2026.
- [91] Sun LY, Wang WT, Han C, et al. The oncomicropeptide APPLE promotes hematopoietic malignancy by enhancing translation initiation. *Mol Cell*, 2021, 81(21): 4493-4508.e9.
- [92] Hu L, Wang J, Huang HJ, et al. YTHDF₁ regulates pulmonary hypertension through translational control of MAGED1. *Am J Respir Crit Care Med*, 2021, 203(9): 1158-1172.
- [93] Liu Y, Zhang HY, Li YY, et al. Long noncoding RNA

- Rps4l mediates the proliferation of hypoxic pulmonary artery smooth muscle cells. *Hypertension*, 2020, 76(4): 1124-1133.
- [94] Li YY, Zhang JT, Sun HL, et al. Lnc-Rps4l-encoded peptide RPS4XL regulates RPS₆ phosphorylation and inhibits the proliferation of PASMCs caused by hypoxia. *Mol Ther*, 2021, 29(4): 1411-1424.
- [95] Piamsiri C, Maneechote C, Siri-Angkul N, et al. Targeting necroptosis as therapeutic potential in chronic myocardial infarction. *J Biomed Sci*, 2021, 28(1): 25.
- [96] Spencer HL, Sanders R, Boulberdaa M, et al. The LINC00961 transcript and its encoded micropeptide, small regulatory polypeptide of amino acid response, regulate endothelial cell function. *Cardiovasc Res*, 2020, 116(12): 1981-1994.
- [97] Spiroski AM, Sanders R, Meloni M, et al. The influence of the LINC00961/SPAAR locus loss on murine development, myocardial dynamics, and cardiac response to myocardial infarction. *Int J Mol Sci*, 2021, 22(2): 969.
- [98] Savard J, Marques-Souza H, Aranda M, et al. A segmentation gene in *Tribolium* produces a polycistronic mRNA that codes for multiple conserved peptides. *Cell*, 2006, 126(3): 559-569.
- [99] Kondo T, Plaza S, Zanet J, et al. Small peptides switch the transcriptional activity of shavenbaby during *Drosophila* embryogenesis. *Science*, 2010, 329(5989): 336-339.
- [100] Ray S, Rosenberg MI, Chanut-Delalande H, et al. The mlpt/Ubr3/Svb module comprises an ancient developmental switch for embryonic patterning. *eLife*, 2019, 8: e39748.
- [101] Mise S, Matsumoto A, Shimada K, et al. Kastor and polluks polypeptides encoded by a single gene locus cooperatively regulate VDAC and spermatogenesis. *Nat Commun*, 2022, 13(1): 1071.
- [102] Ye MJ, Zhang JJ, Wei M, et al. Emerging role of long noncoding RNA-encoded micropeptides in cancer. *Cancer Cell Int*, 2020, 20: 506.
- [103] Budamgunta H, Olexiouk V, Luyten W, et al. Comprehensive peptide analysis of mouse brain striatum identifies novel sORF-encoded polypeptides. *Proteomics*, 2018, 18(10): e1700218.

(本文责编 郝丽芳)