

DOI: 10.13376/j.cblls/2016078

文章编号: 1004-0374(2016)05-0602-13



宋晓元, 中国科学技术大学生命科学学院教授、博士生导师, 中组部第三批“青年千人计划”入选者。2006年在美国罗彻斯特大学生物系取得分子、细胞和发育生物学博士学位。2006—2012年在加州大学圣地亚哥分校/HHMI进行博士后研究, 期间获癌症研究所博士后奖学金。2012年底加入中国科学技术大学生命科学学院。主要研究方向: (1) 研究 lncRNA 在基因转录调控中的作用 (以嗜热四膜虫、小鼠模型及临床样品为研究对象); (2) 研究脑衰老过程以及精子发生过程中以基因组染色体相互作用为基础的转录调控 (以嗜热四膜虫、小鼠模型及临床样品为研究对象); (3) 研究细胞老化及机体衰老过程的三维空间转录调控, 包括 lncRNA、基因组表观遗传学调控和三维染色质构象等之间的协同作用 (以小鼠模型及临床样品为研究对象)。已在 *Genes Dev*、*Mol Cell Biol*、*Nature*、*PNAS* 等学术期刊上发表了系列文章。担任科技部重大科学研究计划研究骨干, 主持 1 项国家自然科学基金面上项目和 1 项国家自然科学基金委重大研究计划培育项目。

长链非编码RNA与脑衰老的研究进展

王 斐, 宋晓元*

(中国科学技术大学生命科学学院, 合肥 230027, 中国科学院分子细胞科学卓越创新中心, 中国科学院脑功能与脑疾病重点实验室)

摘要: 长链非编码 RNAs (long non-coding RNAs, lncRNAs) 参与调控诸多生命过程, 与大脑发育、脑衰老以及脑衰老相关疾病的发生与发展密切相关。因此, 发现 lncRNAs, 预测并研究 lncRNAs 的特征、结构、功能和作用机制, 有助于我们更加系统、深入地了解 lncRNAs 参与调控生命发育、衰老和疾病发生发展等过程的功能机理, 促进 lncRNAs 的相关成果向临床治疗转化。现就 lncRNAs 和脑衰老相关 lncRNAs 的研究现状和研究思路进行综述。

关键词: 长链非编码 RNA; 脑; 衰老

中图分类号: Q527; R394.32; R743.9 **文献标志码:** A

Recent advances in long non-coding RNAs and brain aging

WANG Fei, SONG Xiao-Yuan*

(CAS Key Laboratory of Brain Function and Disease, CAS Center for Excellence in Molecular Cell Science, School of Life Sciences, University of Science and Technology of China, Hefei 230026, China)

Abstract: Long non-coding RNAs (lncRNAs) involve in the regulation of many biological processes, and are closely related to the development and aging of the brain as well as the occurrence and development of brain aging-associated diseases. Therefore, the work of finding lncRNAs, predicting and analyzing lncRNAs' features,

收稿日期: 2016-03-15

基金项目: 国家自然科学基金重大研究计划培育项目(91540107); 科技部重大科学研究计划(2015CB943000); 中国科学院合肥物质科学中心创新方向项目培育基金(2014FXCX009); 中央高校基本科研业务费专项资金创新团体培育基金(WK2070000034); 中央高校基本科研业务费专项资金青年创新基金(WK2070000023)

*通信作者: E-mail: songxy5@ustc.edu.cn

structures, functions and mechanisms will help us to comprehensively and thoroughly understand the functional mechanisms of lncRNAs in regulating development, aging and the pathogenesis of aging-associated diseases. Furthermore, it will promote the translation of lncRNAs' basic research to clinical therapy. Here we summarize the latest studies of lncRNAs and brain aging-associated lncRNAs.

Key words: long non-coding RNA; brain; aging

1 长链非编码RNA

RNA世界假说认为地球上早期的生命分子先以RNA出现,之后才有蛋白质和DNA,且这些早期的RNA分子同时拥有如同DNA的遗传信息储存功能,以及如蛋白质般的催化功能。RNA作为生命初期最关键的分子,支持了早期的细胞或前细胞生命的运作;后来,当DNA和蛋白质的功能远远超过最初RNA的作用时,它才退到了次要地位。1958年,由Francis Crick提出的中心法则认为蛋白质是调控细胞功能的主要参与者,而RNA作为信使将遗传信息从DNA传递到它所编码的蛋白质。然而,近些年RNA的很多其他功能逐渐被人们发现。事实上,哺乳动物的基因组被普遍转录,但只有少于5%的序列编码蛋白质,绝大多数基因组DNA被转录成非编码RNA(non-coding RNA, ncRNA)。科学家发现,其中很多ncRNAs参与细胞周期调控、应激反应、疾病发生等过程,并且与生物体的发育、进化密切相关。

长链非编码RNA(long non-coding RNA, lncRNA)是长度大于200个核苷酸的非编码RNA。lncRNAs相对比较容易被发现,但是,要探讨lncRNAs的作用机制是一个很大的挑战。在20世纪90年代末的人类基因组计划启动前,John Rinn及同事在人22号染色体上寻找蛋白质编码基因时偶然发现了大量的lncRNAs。这些lncRNAs来源于不编码蛋白质的基因组区域,这些转录子缺少开放阅读框和翻译成蛋白质的其他必要特征。当时,大多数科学家都将这些转录子作为一种“噪声”而忽略掉。但John Rinn及同事将它们克隆出来,并最终发现了lncRNA HOTAIR^[1-3]。这个2200 nt的长链剪接RNA转录子与蛋白质复合物相互作用,从而修饰染色质并抑制人类HOX基因的转录,进而调控生物体的生长发育。但到目前为止,它的具体调控机制仍然不清楚。

1.1 ncRNA的分类与特点

目前,关于ncRNA的分类大概有4种方式:(1)根据ncRNA的转录本长度分为小非编码RNAs(包括参与基因转录、转录后调控以及翻译的小RNA,

如small interfering RNAs、Piwi-associated RNAs、small nuclear RNAs、tRNAs等)和lncRNAs(包括参与X染色体失活的Xist、调控丝氨酸/精氨酸剪接因子磷酸化的MALAT1、增加mRNA稳定性的BACE1-AS、调控mRNA前体剪接的hsr ω -n和sat III等);(2)根据ncRNA的生物学功能将其分为持家非编码RNA[包括核糖体RNA(rRNA)、转运RNA(tRNA)、引导RNA(gRNA)、端粒酶RNA、小核RNA(snRNA)、小核仁RNA(snoRNA)等]和调控性非编码RNA[包括一些lncRNA、小干扰RNA(siRNA)、与Piwi蛋白相互作用的piRNAs(Piwi-associated RNAs)等];(3)根据ncRNA的亚细胞定位可分为细胞核非编码RNA和细胞质非编码RNA;(4)根据ncRNA是否具有polyA尾结构可分为具有polyA尾的ncRNAs(polyA-plus ncRNAs)和不具有polyA尾的ncRNAs(polyA-minus ncRNAs)。

由于lncRNAs功能研究仍处于初步阶段,目前通常根据lncRNAs的转录位点来进行分类。根据lncRNA在基因组上转录的位置以及lncRNA与编码蛋白质基因和小RNAs的位置关系可将lncRNAs进一步分为以下几类:基因间长链非编码RNA(long intergenicncRNA, lincRNA)、与注释基因重叠的转录子(反义RNA或内含子序列内的转录子)、小RNA前体、启动子相关lncRNAs(pRNA)、增强子相关lncRNA(eRNA)、转录的假基因(pseudogene)、内含子lncRNA(intronic RNA)、3'UTR相关lncRNA(3'UTR-associated lncRNA)、端粒相关lncRNA(telomere-associated lncRNA),等等(图1)。大部分lncRNAs在细胞核内,主要位于染色质附近以及作为核RNA组分发挥功能,也有一部分lncRNAs分布在细胞质中参与生物学过程。RNA-Seq结果表明,生物体中存在成千上万的lncRNAs,且lncRNAs表达具有组织特异性和时空特异性^[4-5],说明lncRNAs在不同的生物学过程中很可能具有不同的功能^[6]。相对于编码蛋白质的基因,lncRNAs初始序列的保守性较差,但它们的启动子区域和外显子区域比内含子区域及未转录的基因间区域的保守性高很多^[7-8]。而

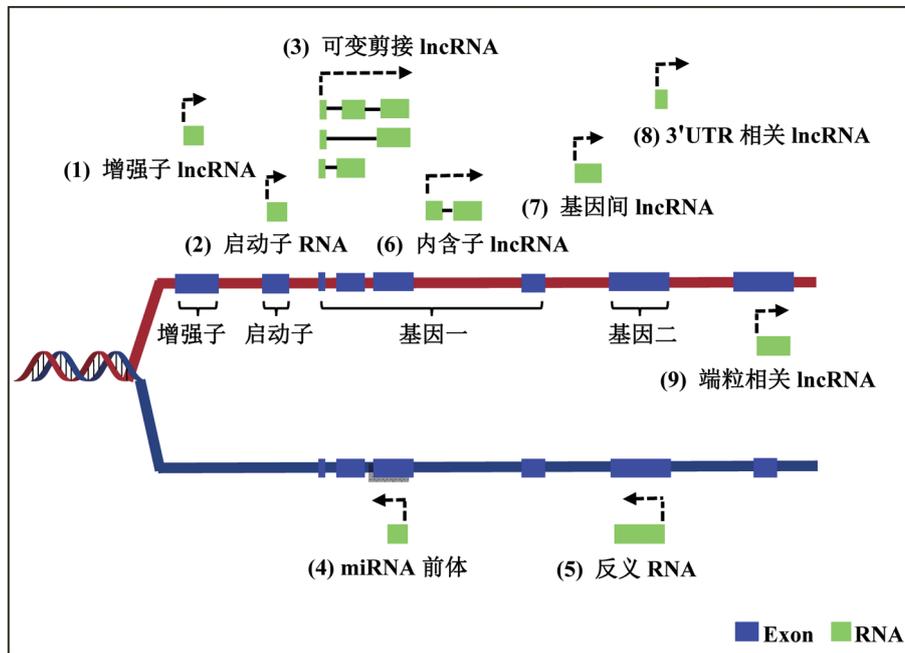


图1 lncRNA的在基因组上的位置分类图

且, lncRNAs 的基因组位置在各个物种间是相对比较保守的^[9]。lncRNAs 结构具有多样性,可塑性较强,它们能通过碱基配对与 DNA 和 RNA 相互作用,从而参与基因调控^[10-11],也能通过特定的结构基序(motif)与蛋白质结合发挥功能。lncRNAs 这种独特的结构特性使其作为染色质修饰复合物和基因组之间的重要连接物^[12],以引导分子^[13]、支架分子^[3]、诱导分子^[14]等形式发挥作用^[15],或直接调控染色质的高级结构形式和基因表达^[18]。很多 lncRNAs 被证明可以直接募集染色质复合物,抑制或激活染色质修饰,从而改变染色质状态及基因表达^[16-18](图2)。

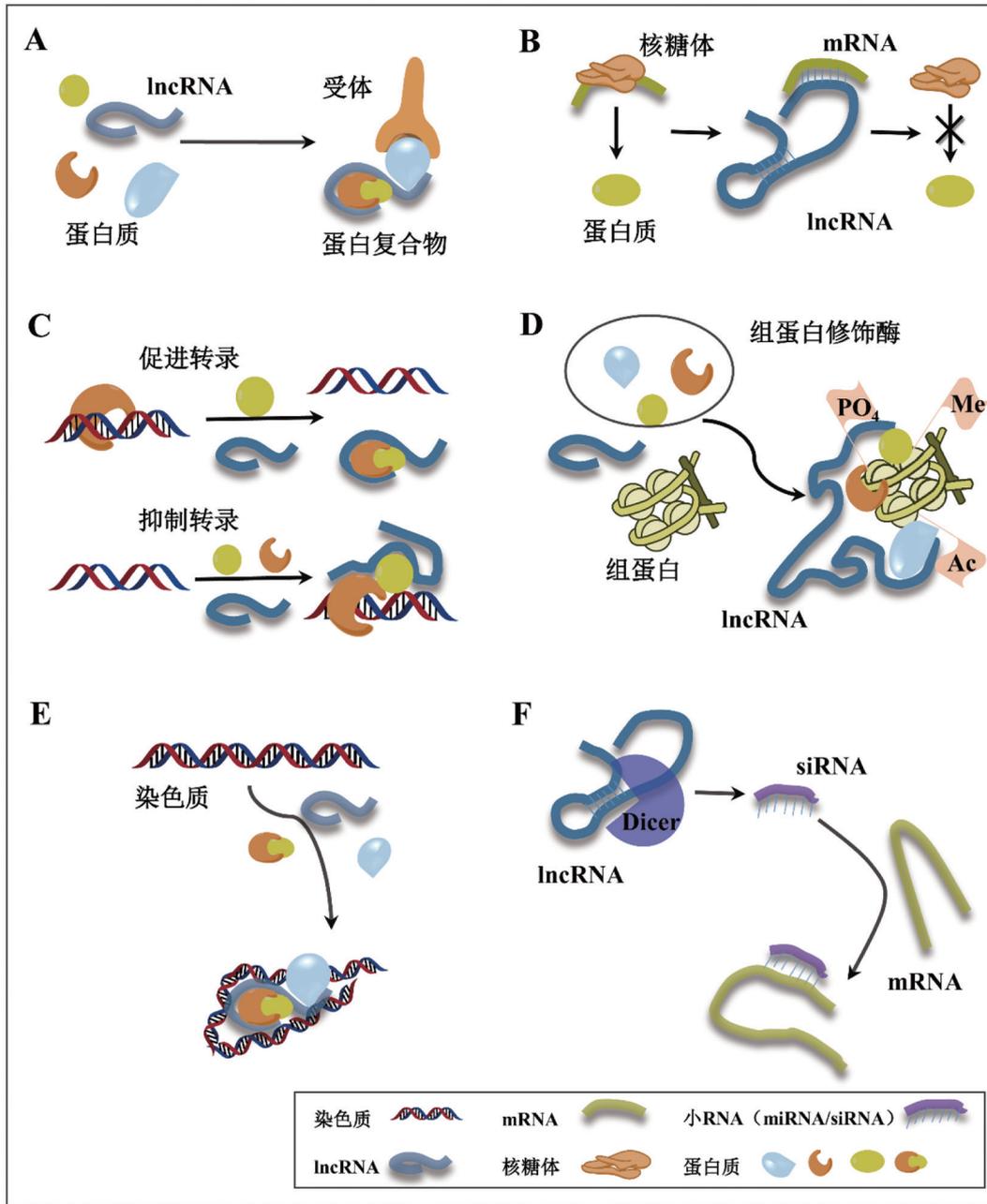
1.2 lncRNA的研究方法

利用 RNA-seq 对细胞转录物进行高通量深度测序,可以预测新的 lncRNA 转录本。通过测序等方法发现,哺乳动物中存在数量众多的非编码转录本,而且被发现的非编码转录本的数量还在不断增加。我们需要思考的是,发现的这些新的 lncRNAs 中有多少是具有生物学功能的。在进化过程中,大量的非编码转录本具有极低的进化保守性,大多数基因间 lncRNAs 几乎不具有进化保守性。然而,没有进化保守性并不能说明 lncRNAs 就没有功能。每种哺乳动物都有可能进化出一系列独特的非编码转录本。目前所发现的这些非编码转录本可能包含了大量的转录噪声,而少数具有功能的 lncRNAs 隐藏

在这种噪声背景下。因此,要展开对功能性 lncRNAs 的理解,主要面临两大挑战:(1)识别出最有可能具有功能的 lncRNAs;(2)推测这些 lncRNAs 的假定功能,并基于假设进行实验验证。目前,在实验验证 lncRNAs 生物学功能的研究过程中主要存在以下几方面的问题:研究 lncRNAs 的方法、研究 lncRNAs 的体系背景、lncRNAs 的结构研究(具体见 1.3)以及 lncRNAs 功能的临床应用前景(具体见 1.4)。图3展示了研究 lncRNAs 的简要思路。

1.2.1 如何识别出最有可能具有功能的lncRNAs

当前,识别可能具有功能的 lncRNAs 主要依赖于 RNA-Seq、Microarray 分析以及 CAGE (cap analysis of gene expression) 和 SAGE (serial analysis of gene expression) 等手段,也可以根据 ChIP (chromatin immunoprecipitation) 得到的特定组蛋白修饰位点在基因组上的分布来间接推测功能性 lncRNAs^[19]。目前被发现的 lncRNAs 已经很多,但是深入研究 lncRNAs 的功能仍然比较困难。传统的方法是通过 Microarray 分析和 RNA-Seq 分析发现各个组织之间显著差异表达的 lncRNAs,通过敲减(过)表达 lncRNAs 观察细胞、组织或生物体表型等发生的变化。从生物学过程产生的众多转录产物中区别出有功能的转录子是非常困难的,但是,来自基因间区域的许多转录子可能更有意义:在不同实验中,它们的剪接点以及数量多等方面都表现出相对的一致



A: lncRNA的募集与支架作用。lncRNA通过直接与相关蛋白质特异性结合, 作为结构组分与蛋白质形成核酸蛋白质复合物或调节蛋白的亚细胞定位或靶向定位。B: lncRNA与mRNA互补形成双链, 沉默或抑制mRNA的翻译。C: lncRNA强化或抑制蛋白基因的转录起始。lncRNA与染色质上相关蛋白结合, 蛋白与染色质分离, 有助于RNA聚合酶等起始复合物促进转录过程, 或lncRNA介导抑制性蛋白与染色质结合从而抑制转录发生。D: lncRNA调控表观遗传修饰。lncRNA介导组蛋白修饰酶对组蛋白进行甲基化、乙酰化、磷酸化、泛素化等, 从而影响基因的转录活性; E: lncRNA介导的染色质重塑。lncRNA与蛋白结合共同调控染色质空间构象。F: lncRNA作为miRNA、endo-siRNA前体。lncRNA在Dicer酶的作用下产生miRNA和内源性siRNA。

图2 lncRNA的作用方式

性。转录子的丰度越高表明它具有功能的可能性越大; 但针对不同的科研问题, 识别并筛选出最有可能具有功能的 lncRNAs 的方法也有所不同, 如在进行衰老研究时, 可以根据现有的高通量测序数据与

数据库预测在差异表达的 lncRNAs 中, 筛选端粒酶相关 lncRNAs、线粒体相关 lncRNAs、衰老基因的增强子相关 lncRNAs、衰老基因的启动子相关 lncRNAs、绝缘子相关 lncRNAs, 等等, 并对它们



图3 lncRNA研究思路简图

进行家族归类预测和功能预测。预测靶向 lncRNAs 的方法越来越多样化, 但这些基于理论数据的猜想与预测最终需要通过实验来证明这些显著差异表达的 lncRNAs 是否具有功能。

1.2.2 如何验证预测lncRNAs的生物学功能

研究编码蛋白质的基因可以通过对该基因在基因组中进行突变、替换、截短或删除, 之后观察、检测这些改变带来的后果或者通常所说的表型, 从而发现它们与生长发育以及疾病发生的相关性。虽然绝大多数基因组都被转录成 ncRNAs, 但是上述研究编码蛋白质的基因的方法只适合很少一部分 lncRNAs 的功能研究。这主要是因为以下几个原因: (1) 大多数维持功能的 lncRNAs 对序列保守性要求很低, 不受框架改变或终止密码子的影响, 对碱基的小插入/缺失或碱基替换的遗传干扰不太敏感; (2) 一些 lncRNAs 精密协调基因调控网络并且增加生物体复杂性, lncRNAs 突变所导致的细微表型改变不容易通过遗传筛选被检测到; (3) 很多可能具有功能的 lncRNAs 很可能在其他通路是冗余的或没有功能的; (4) 很多报道的 lncRNAs 是在基因间区域转录的, 而利用常规方法干扰与其他蛋白基因重叠的 lncRNAs 自身而不影响重叠的蛋白基因是很有挑

战性的; (5) 很多功能性 lncRNAs 结构域并不清楚, 无法像对待编码蛋白质基因那样对其功能结构域进行突变或删除, 因此只能通过对 lncRNAs 基因整个片段进行敲除、替换、或启动子敲除、或在 5' 端连续加入 3 个 polyA 提前终止转录来研究整个转录子的转录。另外, lncRNAs 存在的转录亚型、lncRNAs 的编码潜能、表达特异性等方面在筛选靶向的候选 lncRNAs 时都是需要考虑的。

1.2.2.1 利用特定组蛋白修饰在染色质上的结合位点预测功能性lncRNAs

Guttman 等^[7]开发了一种有效的方法, 利用 ChIP-Seq 创建全基因组染色质状态图谱 (genome-wide chromatin-state maps), 发现由 RNA 聚合酶 II 活化转录的基因启动子区被 H3K4me3 标记, 并且在转录区域上有 H3K36me3 标记, 并将这一独特的结构称为“K4-K36 domain”。他们通过识别位于已知编码蛋白质基因位点外的 K4-K36 domain 来系统地发现 lincRNAs。此外, 他们通过计算 lincRNAs 与 lincRNAs 之间以及 lincRNAs 与蛋白质编码基因之间的皮尔森相关系数 (Pearson correlation coefficient), 建立 lincRNAs 与 lincRNAs 之间以及 lincRNAs 与蛋白质编码基因之间的相关矩阵, 进行功能相关性分析。Shen 等^[19]利用 ChIP-Seq 结果预测基因组上的顺式调控元件 (cis-regulatory elements), 将 H3K4me3 富集或 polII 结合的信号作为活性启动子标志物, 而启动子区域外的 H3K4me1 或 H3K27ac 被作为增强子标志物, CTCF 结合位点被作为潜在的绝缘子元件 (insulator elements) 标志物。基于此, 可以预测出启动子相关 lncRNAs、增强子相关 lncRNAs、绝缘子相关 lncRNAs 等。

1.2.2.2 以与lncRNAs相互作用的蛋白质作为切入点, 寻找lncRNAs并研究其功能

lncRNAs 并不总是通过碱基配对发挥功能, 因此, lncRNAs 序列信息并不一定能够解释 lncRNAs 如何与基因组相互作用。目前, 可以通过 RIP-Seq 技术 (RNA immunoprecipitation followed by sequencing) 发现与 lncRNAs 相互作用的蛋白质。而另外一种研究 lncRNAs 功能的方法是, 当人们发现一个转录因子结合到一个基因间区域时, 首先研究最近的蛋白基因, 通过最近的蛋白质编码基因和结合的转录因子寻找 lincRNA 的功能。

1.2.2.3 以与lncRNAs相互作用的其他RNA为线索, 研究lncRNAs的分子机制

David Tollervey 实验室开发了一种高通量鉴

定 RNA 与 RNA 相互作用的方法——CLASH 技术 (crosslinking, ligation, and sequencing of hybrids)^[20-21]。该技术通过对细胞进行交联处理, 将 RNA 和 RNA 的结合物与 RNA 诱导沉默复合物 (RISC) 固定连接, 通过免疫共沉淀富集纯化 RISC 复合物, 再通过分子内连接将 RNA 的 3'-OH 末端与靶 RNA 的 5'-PO₄ 末端连接, 形成一条长的单链 RNA, 最后再进行高通量分析相互作用的 RNA。ChIRP 技术 (chromatin isolation by RNA purification) 是近几年新开发的用于鉴定 lncRNA 在染色质上结合位点的一种手段^[11]。它是对细胞进行交联处理将 lncRNA、染色质及相互结合蛋白固定连接, 超声将染色质打成小片段, 通过设计的生物素标记探针与 lncRNAs 杂交, 并利用亲和素磁珠纯化出 lncRNAs 及其结合的染色质片段和蛋白质, 最后经核酸酶消化得到与 lncRNA 结合的 DNA 片段 (及 RNA 结合蛋白), 对这些 DNA 进行建库测序、分析, 从而可以分析出 lncRNAs 结合在基因组上的位置。此外, 紫外交联免疫共沉淀技术 (CLIP)^[22-23] 与高通量测序技术相结合的研究系统^[24] 能在全基因组水平上鉴定与特定 RNA 结合蛋白相互作用的 RNA 分子。而且, 近几年衍生出的 iCLIP 技术可以精确鉴定 RNA 分子中蛋白质结合位点。与此同时, 活细胞内的 RNA 特异性标记手段和成像技术的快速发展也促进了 lncRNAs 的亚细胞定位^[25-29], 促进了 RNA 的功能研究, 但目前的内源 RNA 标记和成像技术都存在不同程度的局限性, 还需要进一步发展与提高。

1.2.2.4 在染色体三维空间结构背景下探索 lncRNAs 的结构与功能

近些年, 人们逐渐认识到染色体的三维空间结构在基因调控中具有重要作用, 细胞核及染色体空间组织结构也已经成为表观遗传学研究的一个新的方面。特定的信号可以诱导细胞核内染色体不同区域间的相互作用, 从而发挥重要的基因调控功能。2002 年, Job Dekker 博士基于染色体间彼此接近区域的物理交联开创了染色体构象捕获技术——3C (capturing chromosome conformation), 用于鉴定细胞在自然状态下的染色质相互作用。随后, 4C (circular chromosome conformation capture)^[30]、5C (chromosome conformation capture carbon copy)^[31]、CHIA-PET (Chromatin Interaction Analysis using Paired End Tag sequencing)^[32-33]、Hi-C^[34]、3D-DSL^[35] 等技术的应用使得染色体构象捕获技术逐渐完善并向多位点, 甚至全基因组方向发展, 从根本上改变

了对传统常规的二维转录调控的认识, 加深了我们对于基因组空间组织架构以及基因表达精确调控的理解。lncRNAs 的作用与染色体的三维空间结构密切相关。充分利用这些技术, 在染色体三维空间结构背景下, 将更加有利于研究 lncRNA 的结构与功能。

1.2.2.5 利用生物信息学方法实现数据库之间的关联, 预测 lncRNA 的功能与作用机制

Li 等^[36] 通过筛选满足 lncRNA 与 lncRNA 存在共同的共表达基因、在同类功能注释内富集、蛋白质相互作用网络极为相近这三个特征的功能模块, 利用这些共调控功能模块构建 lncRNA-lncRNA 功能协同网络 (lncRNA-lncRNA functional synergistic networks, LFSNs), 在癌症疾病中发现癌症相关的 lncRNAs 富集在构建的癌症特异性 LFSNs 中, 并利用癌症特异性 LFSNs 在癌症 lncRNA 表达数据中集中识别癌症预后生物标志物 (prognostic biomarkers)。定量计算 lncRNA 功能相似性对 lncRNA 功能预测和推断 lncRNA 与疾病的潜在关联性非常重要。Chen 等^[37] 利用计算生物学方法开发了 LNCSIM 模型 (lncRNA functional similarity calculation models), 通过整合已知的 lncRNA- 疾病相关性与疾病相似度, 来大批量地评估 lncRNA 功能相似性, 而且利用 LNCSIM 模型进行的一些癌症相关的预测结果已通过生物学实验被证实。此外, 目前还有很多 lncRNA 功能与结构预测的网络服务器, 如 starBase 提供了非常全面的 CLIP-Seq 实验数据验证的 miRNA-lncRNA 和 lncRNA- 蛋白质的相互作用网络以及 lncRNA 的功能预测的工具^[38]。ncFANs 服务器提供 lncRNA 功能注释与功能富集的分析策略。2016 年, 最新创建的 fRNAdb 服务器可以从碱基配对的可能性与序列同源性等方面, 用不同的方法预测 lncRNA 的二级结构, 在筛选并预测候选 lncRNA 的功能时, 都可以作为一种辅助方法供我们借鉴。

现在, 越来越多的科学家开始研究 lncRNAs, 因此, 更加需要 lncRNAs 的注释以便研究者知道怎样去查找 lncRNAs、了解相关 lncRNAs 的新发现、了解它们的命名与分类等等。通常, Microarrays 分析不能区分一个转录子的不同亚型, 而高通量测序通常可以发现一个转录子的多个亚型, 但无法确定哪一个最具有生物学相关性, 而且新的 lncRNAs 外显子的起始和终止位点等并不是很清楚。近几年, 各个国家的科学家努力建立了很多 lncRNAs 相关数据库: NONCODE、LNCipedia、LncRNADisease

database、ncRNA Expression Database (NRED)、Long Noncoding RNA Database 等等。这些 lncRNAs 数据库为人们提供了 lncRNAs 的位点信息，上下游基因，lncRNAs 在生物体各个组织间特异性差异表达情况，lncRNAs 的保守性分析，lncRNAs 与疾病的相关性，预测 lncRNAs 与 DNA、蛋白质和 miRNA 的相互作用，预测 lncRNAs 的高级结构等。这些数据库的创建与实时更新为研究 lncRNAs 提供了巨大的便利，也打下了坚实的基础。

目前，绝大多数 lncRNAs 的分子生物学功能都是通过细胞培养等体外实验推导出来的，这些 lncRNAs 的生理意义仍无法得知。很多 lncRNAs 在发育期间的转录具有时空特异性，体外难以模拟，而 RNAi 手段由于存在对 lncRNAs 敲减不彻底、脱靶效应、lncRNAs 具有一定的空间结构等问题并不完全适用于 lncRNAs 功能研究。为克服这些困难，建立 lncRNAs 敲除的动物模型进行 lncRNAs 的功能研究是非常必要的，有助于理解它们在各个复杂组织器官中不同的调控作用，对器官发育、生物体生存能力、免疫力产生的影响，以及对多个器官相互关联的影响，还有它们与疾病的发生、发展及治疗的相关性。目前，只有极少数 lncRNAs 进行了 lncRNAs 敲除的动物模型研究，这些研究表明，lncRNAs 对哺乳动物的发育及生理有重要作用，而这一系列工作基础都为未来 lncRNAs 大规模遗传研究提供了一个框架和资源参考^[39]。

近些年，很多科学家一直致力于通过各种实验手段来发现 lncRNAs 的相互作用蛋白质以及研究其功能机制。越来越多的证据表明，lncRNAs 对生物体生长、发育、衰老和疾病的发生等过程具有重要性。2010年，Loewer 等^[40]设计了一种芯片，探测人类基因组中大约 900 个 lincRNA 的表达情况，并且分析了这些 lincRNA 在 iPSCs、ESCs 和成纤维细胞等细胞系中的表达情况。他们发现 lincRNAs 在 iPSCs 中的表达情况与 ESCs 中的表达情况非常相似，但与成纤维细胞中的表达情况截然相反。当成纤维细胞被重编程为多潜能性细胞时，有 133 个 lincRNAs 被诱导，104 个 lincRNA 表达受到抑制。由于其中一些 lncRNAs 在 iPSCs 中的表达水平比在 ESCs 中的上调程度更加显著，推测这些 lncRNAs 对重编程尤为重要。因此，该团队进一步集中于几十个 lncRNAs 的功能研究。此前公布的数据结果表明，多能性相关的必需转录因子 Oct4 结合在基因组上编码 lncRNAs 的位点上。他们通过实验发现，

lincRNA-RoR 影响重编程过程并调控一系列与氧化应激、DNA 损伤和 P53 等相关的基因表达。2011 年，斯坦福大学 Howard Y. Chang 研究组发现，lncRNA HOTTIP 从 *HOXA* 基因 5' 端转录后，能激活体内 *HOXA* 基因的转录^[41]。HOTTIP RNA 结合接头蛋白 WDR5 靶向 WDR5/MLL 复合物，促进组蛋白修饰和基因转录。当敲减 HOTTIP 后，*HOXA* 基因转录起始位点不再存在 MLL1 和 WDR5。当在基因组其他区域再表达 HOTTIP 时，HOTTIP 敲减所造成的影响并没有改变，表明 lncRNA HOTTIP 必须原位地在它转录的染色体上发挥功能。

1.3 lncRNA 高级结构研究的兴起

目前，lncRNAs 领域一个急需解决的问题是相关 lncRNAs 的结构研究。通常，小 RNA 的结构可以利用各种算法工具来评估，但是 lncRNAs 物种间保守性较差，因而难以从一个物种的 lncRNAs 的序列、结构和功能推测另一个物种的相关 lncRNAs 信息。同时，即使序列保守性不好，也有可能结构相似。因此，根据结构信息进行 lncRNAs 分类是一种非常有意义的分类方式。通过结构能有效预测 lncRNAs 功能，可以引导研究者更好地进行后续实验。

现在预测 RNA 二级结构的算法主要是基于热力学、同源比对和统计学习等原理的各种算法。热力学模型是根据 RNA 序列计算出最小自由能的“最优”结构和自由能较低的“次优”结构，预测出 RNA 二级结构。RNA 的功能与其结构密切相关，一些重要的 RNA 结构会在进化中具有一定的保守性。结构保守区域的配对碱基序列会发生共突变，同源比对模型就是通过从同源序列中寻找共突变碱基对，进行二级结构预测。同源比对方法可以直接反映 RNA 在细胞中的结构状态，但需要提供同源序列。机器学习算法中的统计学习模型是利用已知的 RNA 结构对新的 RNA 进行结构预测，统计学习模型准确率比热力学的最小自由能算法预测结果更准确，但统计学习模型在一定程度上存在对训练数据集过学习的缺点。

目前，化学标记也是检测 RNA 二级结构的实验方法之一。2010 年，Kertesz 等^[42]通过高通量测序，利用 PARS 技术 (parallel analysis of RNA structure) 评估整个酵母转录组中的 lncRNAs 结构。PARS 技术是结合高通量测序技术和结构特异性核酸酶消化方法，通过两种核酸酶分别切割双链和单链 RNA，对随机片段化的 RNA 进行高通量测序，根据每一条片段序列在完整 RNA 序列上的比对位置来判断

此序列前一位碱基所处的结构状态, 最后整合所有测序结果从而明确 RNA 分子中每一个碱基是处于单链状态或是双链状态等结构信息。Howard Y. Chang 研究组在活细胞中利用 PARS 成功比较了不同条件下的 lncRNAs 转录组^[41]。随后, Underwood 等^[43]建立了一种类似的方法, 只用一种核酸酶的 Frag-Seq (fragmentation sequencing) 技术对整个小鼠转录组的消化片段进行分析, 成功定位了多个结构已知的 ncRNAs 的单链区域。2012 年, Howard Y. Chang 实验室设计出两种化学探针, 可以在活细胞中通过引物延伸分析选择性地对 2' 羟基酰化作用, 这种方法简称 SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension)^[44]。通过这种方法可以直接检测活体细胞内的 RNA 结构, 评估不同细胞状态下或基因敲除细胞株内的 RNA 结构的动态变化。SHAPE 方法的开发对 lncRNAs 的结构探索与功能研究都具有非常重要的实际意义。

将 lncRNA-Seq、lncRNA Microarray、ChIP-Seq 等相关高通量测序结果和数据分析结果以及相关数据库等大数据与 lncRNAs 的二级结构分析结合起来, 将这些散在的信息通过机器学习算法进行整合分析, 预测 lncRNAs 功能, 也是 lncRNAs 的生物信息学大数据分析未来的一种趋势。目前, 已有从结构入手进行 lncRNA 功能的研究, 并开发了多种 lncRNAs 二级结构和三级结构预测软件^[45-49]以及 RNA 与 RNA、RNA 与蛋白质、RNA 与 DNA 分子间相互作用的预测软件^[38]。

1.4 lncRNA 的临床应用前景

近年来, 越来越多的实验结果表明, lncRNAs 对生物体生长、发育、进化等过程至关重要, 因此, 很多科学家在一定理论研究成果的基础上将 lncRNAs 作为临床治疗的靶标。快乐木偶综合征 (Angelman syndrome, AS), 又称天使综合征, 是一种由母系单基因遗传缺陷引起的疾病, 在来自父本沉默的 *UBE3A* 等位基因存在情况下母本的 *UBE3A* 表达缺失会导致这种疾病的发生。Meng 等^[50]发现, lncRNA *UBE3A* 反义转录子的反义寡核苷酸可用来作为治疗这种疾病的一种方案, 在培养的小鼠神经元中或小鼠上, 利用 *UBE3A* 反义转录子的反义寡核苷酸使 *UBE3A* 反义转录子沉默, 激活父本的 *UBE3A* 等位基因的表达。利用这种方法, 在这种疾病的小鼠模型上, *UBE3A* 蛋白表达水平恢复并明显改善了认知缺陷。在印迹调控方面, 小鼠和人基因的组织与调控都是高度保守的, 因此, 对于快乐

木偶综合症, 将来也是一种很有前景的治疗策略。目前, 越来越多的科学家开始关注并研究如何将 lncRNA 更好地转化到治疗临床重大疾病的课题上来, 如利用与生殖^[51]、发育^[52-53]、癌症^[54-55]等重大疾病相关的临床样本进行高通量测序技术或 Microarray 分析, 筛选差异表达且潜在具有生物学意义的 lncRNA, 研究这些 lncRNA 在生殖发育以及重大疾病中的作用与功能。因此, 未来将 lncRNAs 作为临床治疗的靶标的例子还将不断增加。

2 衰老

2.1 脑衰老

衰老是生物体发育成熟后, 随着年龄的增长, 机体在形态、结构、功能方面出现的种种不利变化, 如皮肤变薄起皱、色素形成、器官老化、智力下降等。衰老是机体在退化时期功能下降和紊乱的综合表现。普遍性、进行性、退化性和内因性是人体衰老的基本特征。近年来, 随着现代遗传学、分子生物学、细胞生物学和分子免疫学等学科的飞速发展, 人们对衰老的机制有了深层次的认识, 在大量实验证据的基础上提出了许多新的学说, 如自由基学说、端粒学说、内分泌功能减退学说、基因学说、交联学说、脂褐素累积学说等, 分别从不同的角度探讨了衰老发生的机制和对策。细胞作为生物有机体的基本单位, 也在不断地新生和衰老死亡。在细胞水平上, 衰老是正常环境条件下发生的功能减退, 同时, 体外培养的细胞在衰老过程中伴随着一些形态的变化: 细胞变得大而扁平, 出现空泡且多核; 衰老细胞内的糖、脂、蛋白质、DNA 等成分会出现损伤, 细胞内氧化蛋白增加; 重要蛋白酶出现氧化、糖基化修饰, 活性下降; 线粒体 DNA^[56]和核 DNA 出现氧化损伤、突变增多、DNA 端粒缩短等褪变现象^[57]; 细胞器数量, 特别是线粒体数量减少, 胞内出现脂褐素等异常物质沉积, 最终出现细胞凋亡和细胞坏死。细胞衰老也是细胞由正常增殖到非正常增殖的过程, 这个过程伴随着细胞表观遗传学的改变。人体的衰老是生理、心理和社会和环境变化的多维过程积累的结果, 这种积累是引起很多疾病, 包括癌症、糖尿病、心血管疾病、神经退行性疾病等的首要因素^[58]。

人的中枢神经系统衰老最主要表现在学习与记忆以及运动功能等方面^[59]。在人的正常脑衰老过程中, 程序性记忆、工作记忆、空间记忆、处理速度、内隐性记忆等都显著下降, 运动功能在正常衰老过

程中也非常敏感^[60-61]。老年人会变得动作迟缓,运动控制能力下降。同时,中枢神经系统也会发生一系列的生理学变化,神经元兴奋性降低等^[62]。造成这些改变的神经生物学机制主要有两种:衰老相关的神经元丢失和衰老相关的突触效能降低^[63]。神经元数目与性别、年龄等相关,在不同物种衰老的大脑中神经元数目都普遍显著降低^[64]。

脑是人类最先衰老的组织之一。脑与机体的感觉、运动、视觉、听觉、情绪、语言与注意、学习与记忆等多方面功能的行使紧密相关,并发挥着关键性的作用。而衰老同时也是由各种因素长时间影响产生的结果。脑被视为宇宙间最复杂的物质结构,脑的物质组成中包含着许多奇特的分子,而且它们中的许多只存在于神经系统中。这些不同的分子对于脑功能的行使,扮演着不同却又非常关键的角色。人体中的 lncRNAs 在脑中的种类和数量仅次于生殖系统^[5],同时,lncRNAs 以其功能的复杂性和灵活性与神经系统的多样性相结合,使得 lncRNAs 成为中枢神经系统研究的最新切入点,成为神经系统发育、脑衰老以及神经系统疾病等方面的分子机制研究的热点。

2.2 lncRNA在脑衰老及衰老相关的神经系统疾病等方面的研究

目前,在脑衰老研究方面,科学家们主要通过电生理实验、动物行为学实验等手段研究哺乳动物的学习、记忆、认知功能、运动功能等方面的变化。对 lncRNAs 是否参与到脑衰老这一生物学过程中的研究甚少。近些年,lncRNAs 与脑相关的研究都主要集中在 lncRNAs 对中枢神经系统发育及中枢神经系统疾病的影响等方面(表 1)。

近几年,关于 lncRNAs 影响神经系统发育以及与其他神经系统疾病的关系的研究成果逐渐增加。科学家发现,海马区正常的 γ -氨基丁酸(γ -aminobutyric acid, GABA)抑制性中间神经元在胚胎和成年大脑中负责学习和记忆功能,从 *Dlx-5/6* 超保守区域转录而来的 lncRNA *Evf-2* 在 GABA 神经元发育过程中起到重要的调控作用,*Evf-2* 通过与 *Dlx-2* 转录共激活以增加 *Dlx-5/6* 和谷氨酸脱羧酶 1 的活性,而后者在谷氨酸到 GABA 形成的过程中发挥重要作用;同时,*Evf-2* 还可以调控发育中大脑的 GABA 神经元基因的表达。*Evf-2* 沉默可以导致海马和齿状回区域的 GABA 神经元的突触活动异常^[65]。此外,研究证明,lncRNAs 在诸多层面参与神经系统疾病的发生发展,广泛参与包括脑

发育、神经元分化及功能保持、神经细胞凋亡等过程。*lncRNA SOX2OT* 及 *1810014BORik* 在神经变性疾病早期和晚期都存在明显变化,可能成为潜在的生物标志物,为该类疾病的诊断提供了新思路^[66]。帕金森病(PD)是一种神经系统变性疾病,主要是由于脑内多巴胺能神经元缺失所致,常见于老年人。*lncRNA naPINK1* 通过加强其靶基因的正义转录产物 *svPINK1* 的稳定性,干扰正常线粒体呼吸链功能,提高细胞对凋亡信号的敏感性,参与帕金森病的发病进程^[67]。

MEG3 为脑膜瘤中潜在的肿瘤抑制 RNA,脑膜瘤患者缺乏 *MEG3* 表达,编码的 lncRNAs 与靶基因启动子甲基化程度有关^[68]。体外研究显示,脑膜瘤患者 *MEG3* 启动子与印记控制区 CpG 甲基化频率与肿瘤的侵袭生长进程密切相关,而去甲基剂 5 硫唑嘌呤 2 脱氧胞苷能促进 *MEG3* 重表达,这些结果均提示该 lncRNA 可能成为潜在的脑膜瘤治疗新靶点^[69]。阿尔兹海默病(AD),又称脑退化症、老年痴呆症,是一组病因未明的原发性退行性脑变性疾病,多起病于老年期,临床上以智能损害为主,其病理改变主要为皮质弥漫性萎缩、沟回增宽、脑室扩大、神经元大量减少,并可见老年斑、神经原纤维结、神经元丢失、胶质细胞增生等病变,胆碱乙酰化酶及乙酰胆碱含量显著减少。*BACE1* 基因的蛋白产物 β -分泌酶能产生 β 淀粉样蛋白,其沉积是 AD 的主要诱因。*BACE1* 基因的反义转录产物 *BACE1-AS* lncRNA 通过和 *BACE1* mRNA 以及 miR-485-5p miRNA 形成复合物,稳定 mRNA 结构,在转录后水平参与 β 淀粉样蛋白前体代谢,加速 β 淀粉样蛋白累积,表明 *BACE1-AS* lncRNA 促进 AD 的发生^[70-71]。此外,*NAT-Rad18*、*17A*、*TCONS_00021856*/*linc-SLITRK5-11*、*BC200* 等都被证明与阿尔兹海默症的发病密切相关。

本领域的最新研究显示,大量 lncRNAs 在体外培养的衰老细胞中高度差异表达^[72],而人体中的 lncRNAs 在脑中的种类和数量仅次于生殖系统^[5]。人的多个脑区的 RNA-Seq 结果表明,ncRNAs 在脑区间差异表达,且比脑区间 mRNA 更具多样性,说明 lncRNAs 很可能在调控脑区基因表达与功能分化方面具有重要作用^[73]。小鼠多个脑区的 RNA-Seq 分析也得到同样的结论^[74]。从 lncRNAs 的角度去研究脑衰老具有很大的挑战性,但挑战与机遇并存,因此,透过 lncRNAs 去探索脑衰老是非常具有理论意义与应用价值的。我们实验室目前已经完成

表1 脑相关lncRNA

lncRNAs	生物学功能	描述
HAR1F/HAR1A	皮层神经元的特化与迁移、调控区域脑组织的构建 ^[75]	与脑发育、分化有关
Pnky	在神经干细胞内敲降Pnky会促进神经元分化, 与mRNA剪接调控子PTBP1相互作用调控神经元分化 ^[76]	与脑发育、分化有关
HOTAIRM1	在骨髓细胞生长期间的多种HOXA基因的调控者 ^[77]	与脑发育、分化有关
RNCR2/Miat/Gomafu	调控视网膜细胞分化 ^[78]	与脑发育、分化有关
Nkx2.2 AS	Nkx2.2-AS过表达引起Nkx2.2 mRNA水平增加, Nkx2.2的上调诱导少突胶质细胞分化 ^[79]	与脑发育、分化有关
Six3OS	作为分子支架募集组蛋白修饰酶至Six3基因上, 调控Six3活性和视网膜膜细胞分化 ^[80]	与脑发育、分化有关
Evf-2	与GABA能中间神经元发育以及学习有关 ^[65]	与学习及记忆相关
Malat1	与突触形成有关 ^[81]	与突触可塑性相关
Antisense NOS pseudogene	通过调节NOSs mRNAs表达, 进而影响记忆形成	与学习及记忆相关
BC1/BC200	突触后树突局部微环境蛋白合成的调控者, 有助于长期突触可塑性的维持, 与AD和PD有关	与突触可塑性、AD及PD相关
antiBDNF/BDNF-AS	影响脑源性神经营养因子的表达水平, 与神经突的生长和成熟有关 ^[82]	与突触可塑性相关
BACE1-AS	增加BACE1 mRNA稳定性, 促进淀粉样蛋白A β 42产生, 与AD有关 ^[70]	与AD相关
17A	通过调控GPR51的可变剪接, 降低GABAB R2的转录, 影响GABAB信号通路; 增加A β 分泌, 与AD有关	与AD相关
NAT-Rad18	降低神经元抗DNA损伤应激能力, 增加神经元凋亡的敏感度, 与AD有关 ^[83]	与AD相关
GDNFOS	调控人脑内源性GDNF表达, 与神经变性、精神和神经发育障碍有关, GDNFOS多种亚型在AD中差异表达 ^[84]	与AD相关
Sox2OT	通过调控Sox2基因表达减少神经发生, 是神经退行性疾病的生物标志物, 在AD和PD中都上调 ^[66]	与AD、PD相关
naPINK1	稳定svPINK1, 干扰线粒体呼吸链, 增加细胞凋亡敏感性, 与PD有关 ^[67]	与PD相关
HAR1F	和HAR1基因重叠, 突变的huntingtin会导致纹状体内HAR1表达异常以及REST/NRSF的核质运输异常, 与HD有关, 表达水平降低	与HD相关
DGCR5	在HD中, DGCR5是REST的下游靶标, 表达降低	与HD相关
NEAT1	对nuclear paraspeckle亚结构完整性是必不可少的, 与HD有关 ^[85]	与HD相关

HAR1, human accelerated region 1; Nkx2.2-AS, Nkx2.2 antisense; Six3OS, Six3 opposite strand; NOSs, nitric oxide synthases; BDNF, brain derived neurotrophic factor; BACE1-AS, β -site amyloid precursor protein cleaving enzyme 1 anti-sense; Malat1, metastasis-associated lung adenocarcinoma transcript 1; GDNFOS, glial cell derived neurotrophic factor opposite strand; Sox2OT, Sox2 overlapping transcript; PINK1, phosphatase and tensin homologue induced putative kinase 1; REST/NRSF, RE1-silencing transcription factor/neuron-restrictive silencer factor; AD, Alzheimer's disease; PD, Parkinson's disease; HD, Huntington's disease.

了对青年和老年小鼠脑组织的 lncRNA-Seq 测序及 Microarray 分析工作, 发现了大量显著差异表达的 LncRNAs, 并进行了初步分类分析。后续, 我们将对其进行深入的功能和机理研究。

3 总结与展望

lncRNAs 已被公认为基因调控家族的新成员, 与小非编码 RNA 相比, 人们对 lncRNAs 的认识还非常有限。2016 年, NONCODE 数据库已整合了 141 353 个人 lncRNA 与 117 405 个鼠 lncRNA^[86], 这些 lncRNAs 是生物大数据研究的重要源泉, 为科技发展提供了广阔的空间。但是, 在 lncRNAs 数据库

中人工注释过的人 lncRNAs 和鼠 lncRNAs 分别只有 166 个和 108 个, lncRNAs 的序列、结构和功能等方面仍有很多未解决的科研问题需要去探索。目前, lncRNAs 的功能研究过程主要在体外培养的细胞水平和分子水平上开展, 难以模拟体内的真实情况。因此, 后续研究需要进一步在生物体内的系统化背景下, 在动物模型中研究 lncRNAs 生物学功能与作用机制。lncRNAs 序列保守性差, 但不同序列的 lncRNAs 并不代表没有功能或功能不同, lncRNAs 的结构研究仍处于起始阶段, 解析 lncRNAs 结构对后续功能研究具有非常重要的帮助。此外, 近几年, lncRNAs 的研究手段日新月异,

各种方法优势各异,然而,一些RNA研究方法的局限性也阻碍了lncRNAs功能的深入研究,还需要广大科研工作者不断地优化与完善。关于脑衰老的研究目前主要集中在电生理和动物行为学实验等方面,关于lncRNAs对正常脑衰老影响的分子机制研究还很少。lncRNAs对脑衰老影响的研究将有助于我们进一步了解复杂有机体的进化、生长发育并加深对衰老的认识,并带来新的见解,同时为人类衰老相关疾病的诊断和治疗、全新的药物设计和研发以及延缓衰老提供全新的探索方向。

[参 考 文 献]

- [1] Rinn JL, Kertesz M, Wang JK, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 2007, 129: 1311-23
- [2] Khalil AM, Guttman M, Huarte M, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA*, 2009, 106: 11667-72
- [3] Tsai MC, Manor O, Wan Y, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, 2010, 329: 689-93
- [4] Chodroff RA, Goodstadt L, Sirey TM, et al. Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes. *Genome Biol*, 2010, 11: R72
- [5] Necsulea A, Soumillon M, Warnefors M, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature*, 2014, 505: 635-40
- [6] Anguera MC, Ma W, Clift D, et al. Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. *PLoS Genet*, 2011, 7: e1002248
- [7] Guttman M, Amit I, Garber M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, 2009, 458: 129-40
- [8] Johnsson P, Lipovich L, Grandér D, et al. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim Biophys Acta*, 2014, 1840: 1063-71
- [9] Necsulea A, Soumillon M, Warnefors M, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature*, 2014, 505: 635-40
- [10] Magistri M, Faghihi MA, St Laurent G III, et al. Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet*, 2012, 28: 389-97
- [11] Chu C, Qu K, Zhong FL, et al. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell*, 2011, 44: 667-78
- [12] Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*, 2012, 338: 1435-9
- [13] Bierhoff H, Dammert Marcel A, Brocks D, et al. Quiescence-induced lncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol Cell*, 2014, 54: 675-82
- [14] Hung T, Wang Y, Lin MF, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet*, 2011, 43: 621-9
- [15] Chen LL, Carmichael GG. Decoding the function of nuclear long non-coding RNAs. *Curr Opin Cell Biol*, 2010, 22: 357-64
- [16] Luteijn MJ, Ketting RF. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet*, 2013, 14: 523-34
- [17] Sarma K, Cifuentes-Rojas C, Ergun A, et al. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell*, 2014, 159: 869-83
- [18] Dimitrova N, Zamudio JR, Jong RM, et al. LincRNA-p21 activates p21 in cis to promote polycomb target gene expression and to enforce the G₁/S checkpoint. *Mol Cell*, 2014, 54: 777-90
- [19] Shen Y, Yue F, McCleary DF, et al. A map of the cis-regulatory sequences in the mouse genome. *Nature*, 2012, 488: 116-20
- [20] Kudla G, Granneman S, Hahn D, et al. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci USA*, 2011, 108: 10010-5
- [21] Helwak A, Kudla G, Dudnakova T, et al. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell*, 2013, 153: 654-65
- [22] Ule J, Jensen K, Mele A, et al. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*, 2005, 37: 376-86
- [23] Ule J, Jensen KB, Ruggiu M, et al. CLIP identifies Nova-regulated RNA networks in the brain. *Science*, 2003, 302: 1212-5
- [24] Darnell RB. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip Rev RNA*, 2010, 1: 266-86
- [25] Fusco D, Accornero N, Lavoie B, et al. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol*, 2003, 13: 161-7
- [26] Rook MS, Lu M, Kosik KS. CaMKIIalpha 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage. *J Neurosci*, 2000, 20: 6385-93
- [27] Hocine S, Raymond P, Zenklusen D, et al. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods*, 2013, 10: 119-21
- [28] Paige JS, Wu KY, Jaffrey SR. RNA mimics of green fluorescent protein. *Science*, 2011, 333: 642-6
- [29] Bratu DP. Molecular beacons: Fluorescent probes for detection of endogenous mRNAs in living cells. *Methods Mol Biol*, 2006, 319: 1-14
- [30] Cleard F, Karch F, Maeda RK. DamID as an approach to studying long-distance chromatin interactions. *Methods Mol Biol*, 2014, 1196: 279-89

- [31] Dostie J, Richmond TA, Arnaout RA, et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res*, 2006, 16: 1299-309
- [32] Fullwood MJ, Liu MH, Pan YF, et al. An oestrogen-receptor- α -bound human chromatin interactome. *Nature*, 2009, 462: 58-64
- [33] Handoko L, Xu H, Li G, et al. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet*, 2011, 43: 630-8
- [34] Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 2009, 326: 289-93
- [35] Harismendy O, Notani D, Song X, et al. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. *Nature*, 2011, 470: 264-8
- [36] Li Y, Chen J, Zhang J, et al. Construction and analysis of lncRNA-lncRNA synergistic networks to reveal clinically relevant lncRNAs in cancer. *Oncotarget*, 2015, 6: 25003-16
- [37] Chen X, Yan CC, Luo C, et al. Constructing lncRNA functional similarity network based on lncRNA-disease associations and disease semantic similarity. *Sci Rep*, 2015, 5: 11338
- [38] Li JH, Liu S, Zhou H, et al. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*, 2014, 42: D92-7
- [39] Yang F, Huo XS, Yuan SX, et al. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol Cell*, 2013, 49: 1083-96
- [40] Loewer S, Cabili MN, Guttman M, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet*, 2010, 42: 1113-17
- [41] Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, 2011, 472: 120-4
- [42] Kertesz M, Wan Y, Mazer E, et al. Genome-wide measurement of RNA secondary structure in yeast. *Nature*, 2010, 467: 103-7
- [43] Underwood JG, Uzilov AV, Katzman S, et al. FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. *Nat Methods*, 2010, 7: 995-1001
- [44] Spitale RC, Crisalli P, Flynn RA, et al. RNA SHAPE analysis in living cells. *Nat Chem Biol*, 2013, 9: 18-20
- [45] Popenda M, Szachniuk M, Antczak M, et al. Automated 3D structure composition for large RNAs. *Nucleic Acids Res*, 2012, 40: e112
- [46] Jonikas MA, Radmer RJ, Laederach A, et al. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA*, 2009, 15: 189-99
- [47] Parisien M, Major F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature*, 2008, 452: 51-5
- [48] Sharma S, Ding F, Dokholyan NV. iFoldRNA: three-dimensional RNA structure prediction and folding. *Bioinformatics*, 2008, 24: 1951-2
- [49] Das R, Baker D. Automated *de novo* prediction of native-like RNA tertiary structures. *Proc Natl Acad Sci USA*, 2007, 104: 14664-9
- [50] Meng L, Ward AJ, Chun S, et al. Towards a therapy for Angelman syndrome by targeting a long non-coding RNA. *Nature*, 2015, 518: 409-12
- [51] Lv M, Tian H, Cao YX, et al. Downregulation of miR-320a/383-sponge-like long non-coding RNA NLC1-C (narcolepsy candidate-region 1 genes) is associated with male infertility and promotes testicular embryonal carcinoma cell proliferation. *Cell Death Dis*, 2015, 6: e1960
- [52] Bao J, Wu J, Schuster AS, et al. Expression profiling reveals developmentally regulated lncRNA repertoire in the mouse male germline. *Biol Reprod*, 2013, 89: 107
- [53] Grote P, Wittler L, Hendrix D, et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell*, 2013, 24: 206-14
- [54] Wang L, Shen X, Xie B, et al. Transcriptional profiling of differentially expressed long non-coding RNAs in breast cancer. *Genom Data*, 2015, 6: 214-6
- [55] Xue X, Yang YA, Zhang A, et al. LncRNA *HOTAIR* enhances ER signaling and confers tamoxifen resistance in breast cancer. *Oncogene*, 2015, 34: 1608
- [56] Nagley P, Wei YH. Ageing and mammalian mitochondrial genetics. *Trends Genet*, 14: 513-17
- [57] Rizvi S, Raza ST, Mahdi F. Telomere length variations in aging and age-related diseases. *Curr Aging Sci*, 2014, 7: 161-7
- [58] Niccoli T, Partridge L. Ageing as a risk factor for disease. *Curr Biol*, 2012, 22: R741-R52
- [59] Yeoman M, Scutt G, Faragher R. Insights into CNS ageing from animal models of senescence. *Nat Rev Neurosci*, 2012, 13: 435-45
- [60] Mattay VS, Fera F, Tessitore A, et al. Neurophysiological correlates of age-related changes in human motor function. *Neurology*, 2002, 58: 630-5
- [61] Hedden T, Gabrieli JD. Insights into the ageing mind: a view from cognitive neuroscience. *Nat Rev Neurosci*, 2004, 5: 87-96
- [62] Chinta SJ, Woods G, Rane A, et al. Cellular senescence and the aging brain. *Exp Gerontol*, 2015, 68: 3-7
- [63] Burke SN, Barnes CA. Neural plasticity in the ageing brain. *Nat Rev Neurosci*, 2006, 7: 30-40
- [64] Pakkenberg B, Gundersen HJ. Neocortical neuron number in humans: effect of sex and age. *J Comp Neurol*, 1997, 384: 312-20
- [65] Feng J, Bi C, Clark BS, et al. The *Evf-2* noncoding RNA is transcribed from the *Dlx-5/6* ultraconserved region and functions as a *Dlx-2* transcriptional coactivator. *Genes Dev*, 2006, 20: 1470-84

- [66] Arisi I, D'Onofrio M, Brandi R, et al. Gene expression biomarkers in the brain of a mouse model for Alzheimer's disease: mining of microarray data by logic classification and feature selection. *J Alzheimers Dis*, 2011, 24: 721-38
- [67] Sai Y, Zou Z, Peng K, et al. The Parkinson's disease-related genes act in mitochondrial homeostasis. *Neurosci Biobehav Rev*, 2012, 36: 2034-43
- [68] Qin R, Chen Z, Ding Y, et al. Long non-coding RNA MEG3 inhibits the proliferation of cervical carcinoma cells through the induction of cell cycle arrest and apoptosis. *Neoplasma*, 2012, 60: 486-92
- [69] Balik V, Srovnal J, Sulla I, et al. MEG3: a novel long noncoding potentially tumour-suppressing RNA in meningiomas. *J Neurooncol*, 2013, 112: 1-8
- [70] Faghihi MA, Modarresi F, Khalil AM, et al. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med*, 2008, 14: 723-30
- [71] Liu T, Huang Y, Chen J, et al. Attenuated ability of BACE1 to cleave the amyloid precursor protein via silencing long noncoding RNA *BACE1AS* expression. *Mol Med Rep*, 2014, 10: 1275-81
- [72] Abdelmohsen K, Panda A, Kang MJ, et al. Senescence-associated lncRNAs: senescence-associated long noncoding RNAs. *Aging Cell*, 2013, 12: 890-900
- [73] Webb A, Papp AC, Curtis A, et al. RNA sequencing of transcriptomes in human brain regions: protein-coding and non-coding RNAs, isoforms and alleles. *BMC Genomics*, 2015, 16: 990
- [74] Kadakkuzha BM, Liu XA, McCrate J, et al. Transcriptome analyses of adult mouse brain reveal enrichment of lncRNAs in specific brain regions and neuronal populations. *Front Cell Neurosci*, 2015, 9: 63
- [75] Pollard KS, Salama SR, Lambert N, et al. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*, 2006, 443: 167-72
- [76] Ramos AD, Andersen RE, Liu SJ, et al. The long noncoding RNA *Pnky* regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell*, 2015, 16: 439-47
- [77] Lin M, Pedrosa E, Shah A, et al. RNA-Seq of human neurons derived from iPS cells reveals candidate long non-coding RNAs involved in neurogenesis and neuropsychiatric disorders. *PLoS One*, 2011, 6: e23356
- [78] Rapicavoli NA, Poth EM, Blackshaw S. The long noncoding RNA *RNCR2* directs mouse retinal cell specification. *BMC Dev Biol*, 2010, 10: 49
- [79] Tochitani S, Hayashizaki Y. *Nkx2.2* antisense RNA overexpression enhanced oligodendrocytic differentiation. *Biochem Biophys Res Commun*, 2008, 372: 691-6
- [80] Rapicavoli NA, Poth EM, Zhu H, et al. The long noncoding RNA *Six3OS* acts in *trans* to regulate retinal development by modulating *Six3* activity. *Neural Dev*, 2011 6: 32
- [81] Bernard D, Prasanth KV, Tripathi V, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J*, 2010, 29: 3082-93
- [82] Modarresi F, Faghihi MA, Lopez-Toledano MA, et al. Inhibition of natural antisense transcripts *in vivo* results in gene-specific transcriptional upregulation. *Nat Biotechnol*, 2012, 30: 453-9
- [83] Parenti R, Paratore S, Torrisi A, et al. A natural antisense transcript against *Rad18*, specifically expressed in neurons and upregulated during β -amyloid-induced apoptosis. *Eur J Neurosci*, 2007, 26: 2444-57
- [84] Airavaara M, Pletnikova O, Doyle ME, et al. Identification of novel GDNF isoforms and *cis*-antisense GDNFOS gene and their regulation in human middle temporal gyrus of Alzheimer disease. *J Biol Chem*, 2011, 286: 45093-102
- [85] Johnson R. Long non-coding RNAs in Huntington's disease neurodegeneration. *Neurobiol Dis*, 2012, 46: 245-54
- [86] Zhao Y, Li H, Fang S, et al. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. *Nucleic Acids Res*, 2016, 44: D203-8