

综述



叶克穷, 中国科学院生物物理研究所研究员、课题组长、博士生导师。北京市海外高层次人才、百人计划和国家杰出青年基金获得者。主要从事RNA结构生物学研究, 包括H/ACA和C/D RNA蛋白质复合物的结构、功能和生化机制和核糖体组装机。作为课题负责人组织了多个科技部、国家自然科学基金委和中国科学院的科研项目。迄今, 在*Nature*、*Science*等杂志发表了70多篇论文。

C/D RNA的研究进展

董文杰^{1,2}, 张昱源^{1,2}, 吴松麟¹, 叶克穷^{1,2*}

(¹中国科学院生物物理研究所, 核糖核酸功能与应用重点实验室,
北京 100101; ²中国科学院大学, 北京 100049)

摘要: C/D RNA是在古菌和真核生物中普遍存在的非编码RNA, 和多个蛋白质结合形成复合物发挥功能。它的重要功能是指导rRNA和snRNA上2'-O-甲基化修饰, 还有一些特殊的成员参与核糖体组装。结构研究揭示了C/D RNA如何和底物配对并选择特定的位点进行修饰的分子机制。C/D RNA也可以靶向mRNA、tRNA和其他RNA, 通过非典型方式选择靶点, 其生物发生和活性在体内受到多种调控。这些新发现大大拓展了对C/D RNA作用机制和生物学的认识。

关键词: 向导RNA; RNA修饰; RNA蛋白质复合物

Research progress of C/D RNA

DONG Wenjie^{1,2}, ZHANG Yuyuan^{1,2}, WU Songlin¹, YE Keqiong^{1,2*}

(¹Key Laboratory of RNA Innovation, Science and Engineering, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; ²University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: C/D RNA forms a large group of non-coding RNAs universally present in eukaryotes and archaea. It assembles with multiple proteins into ribonucleoprotein (RNP) to function. The classic function of C/D RNA is to guide 2'-O-methylation in rRNA and snRNA, several special members are also involved in ribosome assembly. Structural analysis of C/D RNP has revealed the mechanism by which C/D RNA selects a specific site of substrate for modification. Recent studies show that C/D RNA can guide modification of mRNA, tRNA and other RNA and select target sites through non-canonical ways. Biogenesis and activity of C/D RNP can be regulated in cells. These developments greatly expand our understanding about the mechanism and biology of

收稿日期: 2024-07-30

基金项目: 中国科学院先导B项目(XDB0570000, XDB37010201); 中国科学院依托重大科技基础设施的建制化科研项目(JZHKYPT-2021-05); 国家自然科学基金项目(91940302, 91540201, 31430024, 31325007); 国家重点研发计划项目(2017YFA0504600)

第一作者: E-mail: dongwenjie1134@163.com

*通信作者: E-mail: yekeqiong@ibp.ac.cn

C/D RNA.

Key Words: guide RNA; RNA modification; RNA-protein complex

真核生物核糖体RNA(rRNA)上存在大量的化学修饰, 其中2'-O-甲基化修饰和假尿嘧啶是丰度最高的修饰, 各有50~100个。这两种修饰主要是通过两类RNA指导的修饰酶——C/D RNA蛋白质复合物(ribonucleoprotein, RNP)和H/ACA RNP催化形成的^[1]。复合物中的C/D RNA和H/ACA RNA成分负责底物的识别, 它们利用反义序列结合修饰位点附近的序列, 选择特异的位点呈递给复合物中的催化亚基进行修饰。由于rRNA的转录和修饰发生在核仁, 这两类RNA主要定位在核仁, 也称为小核仁RNA(small nucleolar RNA, snoRNA)。剪接体的snRNA上也有这两类修饰, 大部分也是RNA指导合成。指导snRNA修饰的RNA定位在核内的卡哈尔体(Cajal body), 也称为scaRNA(small Cajal body RNA)^[2]。C/D RNA和H/ACA RNA在古菌中也普遍存在, 介导rRNA和tRNA的修饰^[3,4]。由于古菌复合物容易通过体外重组获得, 它们的生化性质和结构最早得到深入的研究。

本文主要介绍C/D RNA的研究进展。首先介绍C/D RNA基本特点和C/D RNP的结构, 重点介绍最近发现的C/D RNA识别底物新规律、新底物和新调控机制。这些进展拓展了人们对C/D RNA生物学的认识。

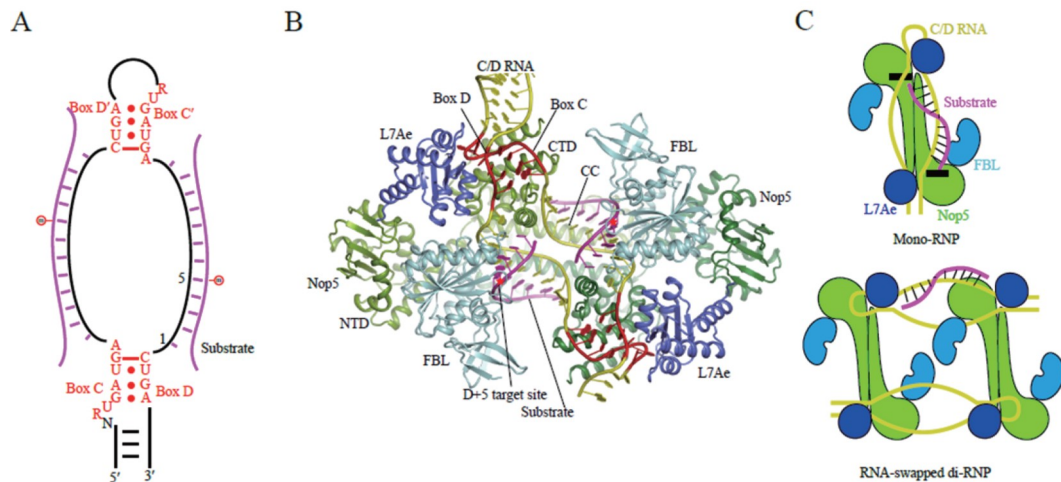
1 C/D RNP的组成和结构

1.1 C/D RNA

C/D RNA由于含有保守的C盒(RUGAUGA, 其中R代表嘌呤)和D盒(CUGA)基序而得名, 它们分别位于RNA的5'和3'末端, 相互结合形成K转角(K-turn)结构^[5](图1A)。K转角结构的一端是包含两个保守的G•A和A•G错配的非典型双链, 另一端是沃森克里克(Watson-Crick)配对形成的经典双链, 中间是3个不参与配对的碱基。C/D RNA的内部还包含类似的C'盒与D'盒基序, 它们通常形成缺乏经典双链区的K环(K-loop)结构。连接C与D'基序的间隔序列或C'与D基序的间隔序列可以结合底物, 称为D'向导序列与D向导序列。C/D RNA通过向导序列和底物互补序列形成10~20个碱基对(base pair, bp), 选择底物上和D或D'基序上游第5位碱基配对的位置作为修饰位点, 称为“D+5”规则^[6-8]。

1.2 C/D RNP结构

在古菌中, C/D RNA和甲基转移酶Fibrillarin (FBL)、骨架蛋白质Nop5和RNA结合蛋白质L7Ae结合形成C/D RNP复合物。真核生物C/D RNP的核心蛋白质组分包括FBL(酵母Nop1)、Nop56、Nop58以及L7Ae的同源蛋白质(人NHP2L1/15.5K、酵母Snu13), 其中Nop56和Nop58是古菌Nop5的同



A: C/D RNA二级结构以及底物识别示意图; B: 古菌C/D RNP结合底物的活性结构(PDB: 3PLA); C: C/D RNP的单体模型和交叉二体模型

图1 C/D RNA和C/D RNP的结构

源蛋白质。

C/D RNP最早在真核生物中被发现,其结构和生化研究主要来自于古菌复合物。2002年开始,通过体外重组获得来源于多种古菌的C/D RNP,它们具有催化特异位点修饰的活性,这为后续的结构和生化研究提供了基础^[9-11]。C/D RNP的结构研究始于对一些亚复合物结构的解析^[12-15]。结合底物的C/D RNP活性构象在2011年得到解析^[16],揭示了其利用向导RNA指导底物特异位点修饰的结构基础。

在复合物的活性结构中,一条C/D RNA结合两份Nop5、FBL和L7Ae蛋白质,整体结构呈现双重对称性(图1B)。Nop5由氨基端结构域(N-terminal domain, NTD)、卷曲螺旋(coiled-coil domain, CCD)结构域和羧基端结构域(C-terminal domain, CTD)构成,在复合物结构中起骨架作用。Nop5的CCD形成同源二聚体,构成复合物的中央平台。Nop5的CTD和L7Ae分别结合在C/D RNA K转角的两侧,组成RNA结合模块。Nop5的NTD结合FBL,构成催化模块。Nop5 NTD和CC结构域之间的连接是活动的^[12,14],有利于催化模块在底物的结合和解离过程中调整位置。晶体结构中C/D RNA的间隔区为12个核苷酸(nucleotide, nt),底物和间隔区的2-11位碱基形成10 bp双链,间隔区的1位和12位碱基没有和底物匹配。底物和向导RNA形成的双链躺在Nop5的CCD二聚体平台之上,双链两头分别和两个Nop5的CTD结合。FBL利用S-腺苷甲硫氨酸(S-adenosine methionine, SAM)作为甲基供体,对底物RNA核糖2'-OH进行甲基化修饰。FBL结合在底物和向导区形成的RNA双链的小沟,没有改变双螺旋的结构。同时,FBL卡在L7Ae和Nop5 CCD上的突起区域,使得底物第五位的核糖正好位于FBL的催化口袋,该构型揭示了D+5靶向规则背后的结构机制。

C/D RNA具有二重性的结构特点,它包含两个C/D模体和两个向导序列,并和两份蛋白质结合形成对称的复合物结构。虽然C/D RNA具有二重性,体外实验表明,古菌和真核生物的C/D RNA的两部分功能是相对独立的^[17]。一个K转角突变只影响和其D盒相邻的向导的功能,另外只含有一个C/D模体和一条向导序列的半体C/D RNA也能指导

位点特异的修饰。那么天然C/D RNA结构的二重性有什么意义呢?和半体RNA相比,天然C/D RNA在遇到弱结合底物和C/D模体突变时具有优势,结构二重性增加了C/D RNA的鲁棒性^[18]。另外,具有二重性的C/D RNA可以协同结合两个底物,促进弱底物的结合和修饰^[19]。

C/D RNA标志性的结构特征是两对保守C/D基序,它们形成K转角结构,并结合L7Ae蛋白质。我们设计构建了一种新型甲基化向导RNA,它们无C/D基序,也不需要L7Ae进行组装和行使功能^[18]。晶体结构研究显示,该新型向导RNA中的一段RNA螺旋结构可以替代C/D基序结合Nop5,从而精确放置底物以进行修饰。该RNA只能与古菌蛋白质组装成有活性的酶,无法与真核生物蛋白质组装。这种无C/D盒的向导RNA比天然C/D RNA更加简单,可能代表原始的甲基化向导RNA,但是目前还没有在古菌体内发现类似结构的RNA。

1.3 真核生物C/D RNP

和古菌相比,C/D RNP在真核生物中的结构和功能都变得更加复杂。古菌中的Nop5在真核生物中分化成Nop56和Nop58两个蛋白质,它们形成异源二聚体。Nop58特异结合位于末端的C/D基序,而Nop56特异结合内部的C'/D'基序^[20,21]。但是在体外重组的C/D RNP的Nop58和Nop56不能特异结合这两个C/D基序,说明真核生物C/D RNP的组装特异性和其体内生物发生过程有关^[17]。FBL的N端含有称为GAR结构域的无结构区(intrinsic disordered region, IDR),GAR结构域能介导液液相分离。FBL形成的凝集体(condensate)和核仁的分层结构密切相关^[22,23]。

有活性的真核生物C/D RNP一直难以通过体外重组获得,直到最近利用嗜热毛壳菌(*Chaetomium thermophilum*)中具有优异生化性质的蛋白质才解决了这个问题^[17]。有修饰活性的真核生物C/D RNP结构还没有被解析,目前只有结合在核糖体小亚基90S前体的U3 RNP的结构被解析^[24-26]。U3是特殊的C/D RNA,它并不指导修饰,而是在18S rRNA的组装和加工中发挥关键作用。在90S前体中,U3 RNP的核心部分和古菌的C/D RNP结构类似,Nop58和Nop56通过CCD形成异源二聚体,分别结

合C/D和C'/D'基序。结构还显示, U3特有的5'结构域和pre-18S上4个位点结合, 发挥核心的结构组装作用。

1.4 有争议的C/D RNP交叉双体模型

古菌C/D RNP在体外重组条件下还能形成二聚体(di-RNP), 二聚体和单体一样能催化特异位点的修饰^[16,27]。但二聚体的结构模型和生物学意义存在巨大争议。有观点认为, 二聚体中的C/D RNA和两个蛋白质复合物交叉结合, 并介导二聚体的形成^[27,28]。该“交叉双体”模型和晶体结构中观察到的C/D RNP单体模型是不一致的, 这将导致底物结合方式和活性构象也明显不同(图1C)。C/D RNP二聚体的结构曾通过电镜和核磁共振技术进行分析, 但是这些结果都不能确定C/D RNA在复合物中的走向^[28-30]。还有其他证据不支持“交叉双体”模型。C/D RNA半体只含有一个C/D模体和一条向导区, 它们在体外和蛋白质组装时也能形成二聚体RNP^[17]。由于半体RNA是单价的, 它只能结合一个Nop5, 无法连接两个复合物。这说明二聚体的形成是由复合物中的蛋白质介导的, 而不是C/D RNA介导的。“交叉双体”模型中, 底物和向导形成的双链将横跨两个复合物, 不受单体模型中底物结合通道对其长度的限制。这和C/D RNA最长结合10个碱基底物的性质并不相容^[17]。Nop5蛋白质在溶液中具有形成高聚体的倾向, 很可能介导C/D RNP进一步聚合。因此, 更合理的解释是单体C/D RNP通过蛋白质聚合形成二聚体, 这样每个RNP的内部结构还和单体模型一致, 底物结合方式和活性结构也不受影响。

2 C/D RNA识别底物的模式

2.1 C/D RNA和底物配对的长度

在C/D RNP活性结构中, 底物和向导RNA形成的双链的两端分别和Nop5二聚体的两个CTD结合, 这种结合方式将限制底物和向导的配对长度^[16]。在另一个解析的C/D RNP结构中, 虽然向导链和底物按照设计能形成11 bp的双链, 但是在结构中它们只形成了10 bp的双链^[31]。这说明C/D RNP底物结合通道的长度是固定的, 因而限制底物最多和C/D RNA形成10 bp的配对。体外生化实验也显示, 将底物和向导的配对延长到12 bp后, 修

饰水平大大降低^[31]。真核生物C/D RNP也具有类似的特点, 长于10 bp的配对会抑制体外的修饰反应^[17]。因此, C/D RNA在修饰反应中只能和底物形成最多10 bp的双链结构, 过长的双链结构需要在一端解旋后才能进入长度固定的蛋白质通道进行修饰反应。

在古菌中, C/D RNA的向导间隔区通常只有12 nt, 和底物形成10 bp左右的配对, 因此没有长配对带来的问题。但是在真核生物中, C/D RNA向导序列的长度是可变的, 根据预测向导序列可以和底物形成10~21 bp长度的双链^[31]。那么在真核生物体内, 底物和向导序列会不会形成长配对? 长配对会影响修饰效率吗? 细胞是否利用RNA解旋酶来消除过长的配对? 这些问题还需要进一步研究。

2.2 选择靶点的非经典方式

大部分C/D RNA通过向导序列识别底物并选择D/D'盒上游的第5位进行修饰。在酵母中的早期研究还发现了4条特殊的C/D RNA, 它们能利用D'向导同时指导两个位点的修饰, 这一现象被称为“双重特异性”^[7,32,33]。其中U18、snR13和U24靶向底物的第5和6位进行修饰, snR48靶向间隔一个碱基的两个位点。snR48由于缺乏保守的D'盒, 无法确定确切的相对于D'盒的靶点位置。最近在拟南芥中也发现了5条具有双重特异性的C/D RNA, 其中U24、SnoR58Y、SnoR29和SnoR10靶向相邻的两个位点, SnoR1靶向相间的两个位点^[34]。酵母U24利用D'向导靶向25S rRNA的A1449和G1450, 还利用D向导靶向上游的C1437。有意思的是, 这三个位点的修饰以及它们的向导U24在酵母、拟南芥、果蝇和人中都是保守的^[34,35]。酵母snR48利用D'向导靶向25S rRNA的G2791和G2793, 这两个位点的修饰以及它们的向导在酵母、拟南芥和果蝇中也是保守的^[34,35]。U24和snR48的家族成员在多种生物中不仅靶向保守的位点, 而且都采用非经典的靶向方式, 提示这种特殊的底物识别机制具有某种优势从而在进化上被保留下来。

真核生物C/D RNA的双重特异性只存在于D'向导。由于C'/D'模体经常出现突变, 双重特异性可能和C'/D'模体结构和组装的多样性有关, 但是确切的分子机制并不清楚。最近, 在古菌冰岛硫化叶菌(*Sulfolobus islandicus*)中也发现了具有双重特

异性的2条RNA——sR4和sR46，它们同时靶向底物的第5和6位，这说明双重特异性向导起源于古菌。另外，sR46的双重特异性向导位于D向导，打破了真核生物C/D RNA的双重特异性向导只存在于D'向导的规律。

在冰岛硫化叶菌中还首次发现了5条C/D RNA能靶向单个非常规位点，打破了对于“D+5”规则的认知^[19]。其中，sR19和sR22选择第4位碱基进行修饰，sR9和sR31靶向第6位，sR48靶向第7位。它们是如何选择非常规位点进行修饰的，需要进一步的生化和结构研究。

冰岛硫化叶菌含有46条tRNA，其中22条在第34位含有2'-O-甲基化修饰，至少有9条C/D RNA指导该位点修饰^[19]。tRNA第34位是反密码子的摆动(wobble)位，经常被化学修饰以加强与密码子的第3位结合。有意思的是，有些C/D RNA(1条被实验证实和4条被预测)可以指导2条同工受体tRNA(isoacceptor)的C34和U34修饰。这些C/D RNA向导区第5位A和一条tRNA U34发生完全配对，但和另一个同工受体tRNA的C34会形成AC错配。通常认为C/D RNA和底物配对时，修饰位点是不可以错配的。该发现说明，在修饰位点可以存在AC错配，尚不清楚这是属于该生物的特殊现象还是普遍的规律。

2.3 额外配对和双重靶向

在酵母和人中发现，一些C/D RNA在结合底物时，除了和修饰位点序列形成主要配对，还能和修饰位点附近序列发生额外的配对^[36]。该现象在拟南芥中也广泛存在，43.3%的rRNA底物和C/D RNA结合时存在额外配对^[34]。拟南芥snoR28 RNA指导25S G2396的修饰，它和底物形成7 bp的主配对和8 bp的额外配对。在体内突变破坏额外配对会显著降低该位点修饰水平，说明额外配对对底物加载的重要性^[37]。还没发现古菌C/D RNA和底物形成额外配对，但是在冰岛硫化叶菌中，C/D RNA广泛利用两条反义序列同时靶向rRNA上相邻17~50 nt的位点，以该方式指导rRNA上79%的修饰。这种双重靶向可以看成一种特殊的额外配对。在拟南芥中具有双重靶向的C/D RNA很少见，108条C/D RNA中只有6条能指导相近位点的修饰。在真核生物中和底物附近序列的额外配对

并不指导修饰，在冰岛硫化叶菌中双重靶向的配对都能指导修饰。C/D RNA在古菌和真核生物中具有不同的长度，可能导致它们采用不同策略用以加强底物的结合。古菌C/D RNA的长度非常紧凑(~60 nt)，没有多余的序列用于额外配对，而采取双重靶向加强底物的结合，提高修饰效率。而真核生物C/D RNA长度更长而多变(80~200 nt)，可以有多余序列介导额外配对。

冰岛硫化叶菌C/D RNA如果靶向相近两个位点的修饰，其中一个位点通常和向导只形成很弱的配对，提示双重靶向能加强弱结合底物的修饰^[19]。体外生化实验也证实，单独的弱结合底物无法被修饰，而和强结合底物相连后就能被高效修饰。与强结合底物结合后，可以促进相连的弱结合底物和C/D RNA的接触机会。有意思的是，强结合底物即使和弱结合底物没有共价连接，作为分离的底物也能促进弱结合底物的修饰。这说明C/D RNA的一个向导序列结合底物后还会改变酶的构象，促进第二个底物的结合和修饰。两个位点结合底物的协同性可能是C/D RNA结构双重性的一个重要意义。

3 C/D RNA的底物和2'-O-甲基化修饰的生物学功能

3.1 rRNA

在一些模式生物中，rRNA上2'-O-甲基化修饰得到了系统鉴定，指导修饰的C/D RNA也有较全面和准确归属。酿酒酵母(*Saccharomyces cerevisiae*)是最早研究C/D RNA的模式生物，其rRNA上含有55个2'-O-甲基化修饰位点，除了25S G2922由单独蛋白质Spb1催化，其他所有位点由44条C/D RNA指导修饰^[32,38-40]。拟南芥表达108条C/D RNA，它们来自184个变体基因^[34]。拟南芥rRNA上鉴定到111个2'-O-甲基化修饰位点，其中109个位点可以找到高度可信的C/D RNA指导修饰，1个位点(25S G2923)由蛋白质Spb1催化，另外一个25S U2922位点的修饰酶还不清楚^[34]。人rRNA上有112个2'-O-甲基化位点，大部分可以找到指导修饰的C/D RNA^[41,42]。果蝇rRNA中有61个高可信2'-O-甲基化位点，其中58个位点已经有对应的C/D RNA^[35]。这些系统的研究表明，真核生物rRNA上

2'-O-甲基化修饰除了1或2个位点由Spb1催化, 剩下的很可能都是C/D RNP合成的。

rRNA上的化学修饰(包括2'-O-甲基化修饰)大多数分布在核糖体的重要功能区, 如解码中心、肽键形成中心和亚基界面。酵母遗传实验表明, 去除单个位点修饰通常对生长没有影响, 去除多个修饰会有累计效应影响核糖体的功能^[43,44]。也有研究表明, 修饰有更加特异的生物学功能。热精胺(thermospermine)是一种多胺类分子, 其缺失能影响拟南芥高度和引发根部瘤状结构^[45]。利用抑制子筛选发现, Nop56以及H/ACA RNP的组分Nap57的突变能抑制热精胺缺失的表型, 说明rRNA上2'-O-甲基化修饰和假尿嘧啶修饰与热精胺生物功能有关, 但具体的分子机制尚不清楚。

3.2 snRNA

人主要剪接体的snRNA(U1、U2、U4、U5、U6)上有22个非末端位点的2'-O-甲基化修饰, 其中15个位点有预测的指导C/D RNA^[46]。拟南芥主要snRNA上有19个2'-O-甲基化修饰位点, 其中14个位点可以找到对应的C/D RNA^[34]。没有发现酵母C/D RNA中能指导snRNA上的2'-O-甲基化修饰^[46]。

最近发现了一类新型和15.5K结合的C/D RNA变体, 其末端形成反方向的K转角结构^[47]。其中, bktRNA1能指导次要剪接体的U12 snRNA上A8位点的修饰, 该修饰对于次要剪接体的组装和U12-型内含子的剪接至关重要。U6上含有8个位点2'-O-甲基化修饰。敲降LARP7可以抑制U6上2'-O-甲基化修饰, 在小鼠中影响雄性生殖细胞的mRNA的剪接和精子发生, 在人细胞中显著影响mRNA剪接位点选择^[48,49]。Nopp140对于snRNA定位于卡哈尔体至关重要, 敲降Nopp140能抑制U1、U2、U5和U12上的2'-O-甲基化修饰, 也能改变剪接位点的选择^[50]。这些研究说明, snRNA的2'-O-甲基化修饰对于剪接体功能的重要性。

3.3 tRNA

古菌C/D RNA能指导tRNA修饰早已知晓, 但是直到最近才发现真核生物C/D RNA也可以指导tRNA修饰。据Vitali等^[51]报道, 两条人C/D RNA(位于核仁的SNORD97和位于卡哈尔体的SNORD133/SCARNA97)能指导延伸tRNA-Met上

C34位点的2'-O-甲基化修饰。这两条C/D RNA在脊椎动物中普遍存在, 在非脊椎动物和开花植物中也能找到一条同源RNA, 说明该修饰机制的普遍性。最近通过改进的PARIS2 RNA-RNA交联技术, 发现C/D RNA和大量tRNA之间存在广泛的交联, 交联涉及很多以前没有发现底物的孤儿C/D RNA^[52]。但是这些新发现的C/D RNA和tRNA之间的配对通常比较弱, 利用RiboMeth-seq技术鉴定tRNA预测靶点的修饰显示, 修饰程度普遍很低。由于RiboMeth-seq难以检测低水平的修饰, 需要用更可靠的质谱技术确定tRNA预测靶点的修饰水平。该研究还发现, 敲除SNORD97和SNORD133会降低延伸tRNA-Met的水平, 影响密码子偏好基因的表达, 对细胞生长和分化产生复杂的影响^[52]。

3.4 mRNA

通过高通量测序技术在哺乳动物mRNA上发现了上千个2'-O-甲基化修饰位点, 修饰分布多在CDS和3'UTR区^[53-56]。很多位点的修饰依赖于FBL, 说明它们可能是C/D RNP催化的^[53,54]。mRNA上2'-O-甲基化修饰的生物功能还不清楚。在单独基因上的研究显示, 修饰能影响mRNA的表达水平和破坏tRNA解码过程^[54,57]。在整体水平, mRNA的2'-O-甲基化修饰程度与其丰度和稳定性呈正相关^[53,54]。FBL的表达水平在癌细胞中经常增加, 有可能通过提高mRNA稳定性从而改变细胞的基因表达^[53]。

3.5 古菌C/D RNA的底物

古菌C/D RNA的经典底物包括rRNA和tRNA^[3]。最近在冰岛硫化叶菌中系统鉴定了C/D RNA和2'-O-甲基化修饰位点, 揭示了C/D RNA底物的全局谱图^[19]。该菌中存在61条C/D RNA, 其中36条指导rRNA上65个位点修饰, 只有3个rRNA修饰位点没有找到合适的向导RNA。另外有18条C/D RNA指导tRNA上至少55个位点修饰, 还有1条C/D RNA指导RNase P的修饰, 有9条C/D RNA没有找到合适的底物。tRNA的第34和50位是C/D RNA的主要靶点。其中, 丰度最高的sR14 RNA指导18条tRNA第50位的修饰。有22条tRNA第34位发生修饰, 由于该位点处于反密码环, 周围序列各异, 至少有9条C/D RNA指导第34位修饰。

4 真核生物C/D RNP的生物发生

在真核生物里, C/D RNA可以作为独立的基因、基因簇或者存在于其他基因的内含子。C/D RNA在酵母中主要存在于独立基因和基因簇, 在植物中主要以基因簇表达, 在哺乳动物中主要分布在内含子区域。所有C/D RNA是通过RNA聚合酶 II (RNA polymerase II, Pol II) 转录形成的。C/D RNP组装起始于转录位点, 然后要进入卡哈尔体成熟, 最后被运输到核仁发挥功能^[58]。

通过遗传和生化研究发现, 一些因子参与C/D RNP的组装, 这些因子从酵母到哺乳动物都是保守的^[58]。其中Rvb1(人RUVBL1)和Rvb2(RUVBL2)是两种AAA+ ATP酶, 它们形成环状结构的异源六聚体, 还参与许多其他复合物的组装^[59]。它们和Tah1(RPAP3)、Pih1(PIH1D1)形成R2TP复合物, 参与C/D RNP的组装。R2TP复合物和热休克蛋白质Hsp90一起招募Nop58, Nop58通过C末端无结构区域和Pih1(PIH1D1)结合^[60-64]。另外, Rsa1(NUFIP1)和Hit1(ZNHIT3)形成的二聚体复合物招募Snu13^[65-67]。在这些组装因子的帮助下, Snu13和Nop58组装到C/D RNA上的C/D模体。随后, Nop56和第二个Snu13组装到C'/D'模体, FBL也加入复合物。C/D模体是先于C'/D'模体组装的, 而且它们的重要性也完全不同。C/D模体对于C/D RNA的稳定性是必需的, 但是C'/D'的缺失不影响C/D RNA的积累^[8,68,69]。

Bcd1(ZNHIT6)也是C/D RNP早期组装的关键因子^[70,71], 但是它的具体作用还不清楚。Bcd1影响Nop58的装配^[72], 它N端的一段保守序列能结合Snu13和C/D RNA^[73], 还能与组蛋白分子伴侣Rtt106结合^[74]。

5 真核生物C/D RNP的调控

5.1 C/D RNP的组装调控

C/D RNP的组装在体内受到其他因子的调控。EZH2是PRC2复合物中的组蛋白甲基转移酶, 有抑制转录的作用。研究发现, EZH2还存在独立于PRC2的功能, 它结合FBL, 促进FBL和Nop56的结合, 以及促进C/D RNP的组装, 从而增强rRNA 2'-O-甲基化修饰^[75]。EZH2的缺乏能全局抑制癌细胞

中的蛋白质翻译水平, 降低内部核糖体进入位点(internal ribosome entry site, IRES)介导的翻译起始^[75]。

INHEG是在神经胶质瘤干细胞中高表达的长非编码RNA, 它可以促进SUMO2 E3连接酶TAF15与Nop58相互作用, 增强Nop58的类泛素化修饰, 促进C/D RNP的组装^[76]。研究发现, 激活INHEG能增强rRNA 2'-O-甲基化修饰, 从而增加胶质瘤细胞中一些致癌蛋白的表达^[76]。鉴于C/D RNP也可以修饰mRNA, 促进C/D RNP组装的因子也有可能通过增加mRNA的稳定性而改变基因表达^[53]。

5.2 C/D RNP的活性调控

U6上有8个2'-O-甲基化修饰位点, 它们由多个C/D RNP催化形成。最近发现, LARP7在C/D RNP修饰U6时发挥重要的激活作用, LARP7可以促进U6和C/D RNP结合^[48,49]。在利用重组C/D RNP的体外修饰反应中, LARP7能增加U6的修饰效率^[48]。LARP7可以直接介导U6和C/D RNA的结合, 一方面利用La-RRM1结构域结合U6, 另一方面利用RRM2结构域识别指导U6修饰的多个C/D RNA上的一段保守的CAGGG序列^[49]。NPM1是核仁内高丰度的RNA结合蛋白质, 它能结合C/D RNA和FBL^[77]。敲降NPM1导致28S rRNA上5个位点的2'-O-甲基化修饰水平降低, 并影响IRES介导的翻译和造血干细胞的维持, 引起骨髓衰竭。NPM1的具体作用并不清楚, 由于敲降NPM1并不影响C/D RNA水平, 它可能促进了C/D RNA与底物的结合。

在修饰完成后, C/D RNA需要从底物释放以催化下一个底物。在体外的修饰反应中底物能够自然释放, 但有证据表明, 细胞内有因子催化底物的解离。酵母RNA解旋酶Prp43和Dbp3的缺失会导致rRNA上大量2'-O-甲基化位点的修饰水平下降^[78]。它们的缺失还会引起一部分C/D RNA在核糖体上滞留, 这可能减少C/D RNP的有效水平, 从而降低修饰程度^[78,79]。

6 结语与展望

C/D RNA的经典靶向规律已经有很好的结构基础, 但是我们不清楚C/D RNA如何实现双重特异性靶向和选择单个非经典位点进行修饰。这些非经典的靶向机制还需要进一步的结构分析。新的

研究不断揭示C/D RNP在体内的组装和活性受到多重调控, 我们期望未来对C/D RNP体内调控过程有更多认识。C/D RNA的底物范围也在不断扩展, 探索这些新2'-O-甲基化修饰的效应分子和生物学功能也是令人期待的未来研究方向。

参考文献

- [1] Kiss T. Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J*, 2001, 20(14): 3617-3622
- [2] Darzacq X. Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J*, 2002, 21(11): 2746-2756
- [3] Omer AD, Lowe TM, Russell AG, et al. Homologs of small nucleolar RNAs in archaea. *Science*, 2000, 288 (5465): 517-522
- [4] Tang TH, Bachellerie JP, Rozhdestvensky T, et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc Natl Acad Sci USA*, 2002, 99(11): 7536-7541
- [5] Huang L, Lilley DMJ. The kink-turn in the structural biology of RNA. *Quart Rev Biophys*, 2018, 51: e5
- [6] Tycowski KT, Smith CM, Shu MD, et al. A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci USA*, 1996, 93(25): 14480-14485
- [7] Kiss-László Z, Henry Y, Bachellerie JP, et al. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, 1996, 85(7): 1077-1088
- [8] Cavaillé J, Nicoloso M, Bachellerie JP. Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. *Nature*, 1996, 383(6602): 732-735
- [9] Tran EJ. Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs. *EMBO J*, 2003, 22(15): 3930-3940
- [10] Bortolin ML. *In vitro* RNP assembly and methylation guide activity of an unusual box C/D RNA, cis-acting archaeal pre-tRNA^{Trp}. *Nucleic Acids Res*, 2003, 31(22): 6524-6535
- [11] Omer AD, Ziesche S, Ebhardt H, et al. *In vitro* reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex. *Proc Natl Acad Sci USA*, 2002, 99(8): 5289-5294
- [12] Ye K, Jia R, Lin J, et al. Structural organization of box C/D RNA-guided RNA methyltransferase. *Proc Natl Acad Sci USA*, 2009, 106(33): 13808-13813
- [13] Moore T, Zhang Y, Fenley MO, et al. Molecular basis of box C/D RNA-protein interactions. *Structure*, 2004, 12(5): 807-818
- [14] Aittaleb M, Rashid R, Chen Q, et al. Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Mol Biol*, 2003, 10(4): 256-263
- [15] Wang H. Crystal structure of a fibrillarin homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 Å resolution. *EMBO J*, 2000, 19(3): 317-323
- [16] Lin J, Lai S, Jia R, et al. Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature*, 2011, 469(7331): 559-563
- [17] Yang Z, Wang J, Huang L, et al. Functional organization of box C/D RNA-guided RNA methyltransferase. *Nucleic Acids Res*, 2020, 48(9): 5094-5105
- [18] Wang J, Yang Z, Ye K. Methylation guide RNAs without box C/D motifs. *RNA*, 2022, 28(12): 1597-1605
- [19] Wang J, Wu S, Ye K. Complicated target recognition by archaeal box C/D guide RNAs. *Sci China Life Sci*, 2024, 67(4): 631-644
- [20] Cahill NM. Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *EMBO J*, 2002, 21(14): 3816-3828
- [21] Granneman S, Kudla G, Petfalski E, et al. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci USA*, 2009, 106(24): 9613-9618
- [22] Feric M, Vaidya N, Harmon TS, et al. Coexisting liquid phases underlie nucleolar subcompartments. *Cell*, 2016, 165(7): 1686-1697
- [23] Yao RW, Xu G, Wang Y, et al. Nascent pre-rRNA sorting via phase separation drives the assembly of dense fibrillar components in the human nucleolus. *Mol Cell*, 2019, 76 (5): 767-783
- [24] Kornprobst M, Turk M, Kellner N, et al. Architecture of the 90S pre-ribosome: a structural view on the birth of the eukaryotic ribosome. *Cell*, 2016, 166(2): 380-393
- [25] Sun Q, Zhu X, Qi J, et al. Molecular architecture of the 90S small subunit pre-ribosome. *Elife*, 2017, 6: e22086
- [26] Chaker-Margot M, Barandun J, Hunziker M, et al. Architecture of the yeast small subunit processome. *Science*, 2017, 355(6321): eaal1880
- [27] Bleichert F, Gagnon KT, Brown Ii BA, et al. A dimeric structure for archaeal box C/D small ribonucleoproteins. *Science*, 2009, 325(5946): 1384-1387
- [28] Lapinaite A, Simon B, Skjaerven L, et al. The structure of the box C/D enzyme reveals regulation of RNA methylation. *Nature*, 2013, 502(7472): 519-523
- [29] Bower-Phipps KR, Taylor DW, Wang HW, et al. The box C/D sRNP dimeric architecture is conserved across domain Archaea. *RNA*, 2012, 18(8): 1527-1540
- [30] Bleichert F, Baserga SJ. Dissecting the role of conserved

- box C/D sRNA sequences in di-sRNP assembly and function. *Nucleic Acids Res*, 2010, 38(22): 8295-8305
- [31] Yang Z, Lin J, Ye K. Box C/D guide RNAs recognize a maximum of 10 nt of substrates. *Proc Natl Acad Sci USA*, 2016, 113(39): 10878-10883
- [32] Lowe TM, Eddy SR. Unusual C'/D' motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. *Science*, 1999, 283(5405): 1168-1171
- [33] van Nues RW, Watkins NJ. Unusual C'/D' motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. *Nucleic Acids Res*, 2016, 45(4): 2016-2028
- [34] Wu S, Wang Y, Wang J, et al. Profiling of RNA ribose methylation in *Arabidopsis thaliana*. *Nucleic Acids Res*, 2021, 49(7): 4104-4119
- [35] Sklias A, Cruciani S, Marchand V, et al. Comprehensive map of ribosomal 2'-O-methylation and C/D box snoRNAs in *Drosophila melanogaster*. *Nucleic Acids Res*, 2024, 52(6): 2848-2864
- [36] van Nues RW, Granneman S, Kudla G, et al. Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *EMBO J*, 2011, 30(12): 2420-2430
- [37] Cao Y, Wang J, Wu S, et al. The small nucleolar RNA SnoR28 regulates plant growth and development by directing rRNA maturation. *Plant Cell*, 2022, 34(11): 4173-4190
- [38] Yang J, Sharma S, Watzinger P, et al. Mapping of complete set of ribose and base modifications of yeast rRNA by RP-HPLC and mung bean nuclease assay. *PLoS One*, 2016, 11(12): e0168873
- [39] Taoka M, Nobe Y, Yamaki Y, et al. The complete chemical structure of *Saccharomyces cerevisiae* rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Res*, 2016, 44(18): 8951-8961
- [40] Lapeyre B, Purushothaman SK. Spb1p-directed formation of Gm2922 in the ribosome catalytic center occurs at a late processing stage. *Mol Cell*, 2004, 16(4): 663-669
- [41] Taoka M, Nobe Y, Yamaki Y, et al. Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res*, 2018, 46(18): 9289-9298
- [42] Jorjani H, Kehr S, Jedlinski DJ, et al. An updated human snoRNAome. *Nucleic Acids Res*, 2016, 44(11): 5068-5082
- [43] Decatur WA, Fournier MJ. rRNA modifications and ribosome function. *Trends Biochem Sci*, 2002, 27(7): 344-351
- [44] Liang X, Liu Q, Fournier MJ. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA*, 2009, 15(9): 1716-1728
- [45] Li X, Yan Z, Zhang M, et al. SnoRNP is essential for thermospermine-mediated development in *Arabidopsis thaliana*. *Sci China Life Sci*, 2023, 66(1): 2-11
- [46] Karijolic J, Yu YT. Spliceosomal snRNA modifications and their function. *RNA Biol*, 2010, 7(2): 192-204
- [47] Li B, Liu S, Zheng W, et al. RIP-PEN-seq identifies a class of kink-turn RNAs as splicing regulators. *Nat Biotechnol*, 2024, 42(1): 119-131
- [48] Wang X, Li ZT, Yan Y, et al. LARP7-mediated U6 snRNA modification ensures splicing fidelity and spermatogenesis in mice. *Mol Cell*, 2020, 77(5): 999-1013.e6
- [49] Hasler D, Meduri R, Bık M, et al. The Alazami syndrome-associated protein LARP7 guides U6 small nuclear RNA modification and contributes to splicing robustness. *Mol Cell*, 2020, 77(5): 1014-1031
- [50] Bizarro J, Deryusheva S, Wacheul L, et al. Nopp140-chaperoned 2'-O-methylation of small nuclear RNAs in Cajal bodies ensures splicing fidelity. *Genes Dev*, 2021, 35(15-16): 1123-1141
- [51] Vitali P, Kiss T. Cooperative 2'-O-methylation of the wobble cytidine of human elongator tRNA^{Met} (CAT) by a nucleolar and a Cajal body-specific box C/D RNP. *Genes Dev*, 2019, 33(13-14): 741-746
- [52] Zhang M, Li K, Bai J, et al. A snoRNA-tRNA modification network governs codon-biased cellular states. *Proc Natl Acad Sci USA*, 2023, 120(41): e2312126120
- [53] Li Y, Yi Y, Gao X, et al. 2'-O-methylation at internal sites on mRNA promotes mRNA stability. *Mol Cell*, 2024, 84(12): 2320-2336
- [54] Chen L, Zhang LS, Ye C, et al. Nm-Mut-seq: a base-resolution quantitative method for mapping transcriptome-wide 2'-O-methylation. *Cell Res*, 2023, 33(9): 727-730
- [55] Zhang P, Huang J, Zheng W, et al. Single-base resolution mapping of 2'-O-methylation sites by an exoribonuclease-enriched chemical method. *Sci China Life Sci*, 2023, 66(4): 800-818
- [56] Dai Q, Moshitch-Moshkovitz S, Han D, et al. Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat Methods*, 2017, 14(7): 695-698
- [57] Choi J, Indrisiunaite G, DeMirci H, et al. 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nat Struct Mol Biol*, 2018, 25(3): 208-216
- [58] Massenet S, Bertrand E, Verheggen C. Assembly and trafficking of box C/D and H/ACA snoRNPs. *RNA Biol*, 2017, 14(6): 680-692
- [59] Dauden MI, López-Perrote A, Llorca O. RUVBL1-RUVBL2 AAA-ATPase: a versatile scaffold for multiple complexes and functions. *Curr Opin Struct Biol*, 2021, 67: 78-85

- [60] Prieto MB, Georg RC, Gonzales-Zubiate FA, et al. Nop17 is a key R2TP factor for the assembly and maturation of box C/D snoRNP complex. *BMC Mol Biol*, 2015, 16(1): 7
- [61] Kakiyama Y, Makhnevych T, Zhao L, et al. Nutritional status modulates box C/D snoRNP biogenesis by regulated subcellular relocation of the R2TP complex. *Genome Biol*, 2014, 15(7): 404
- [62] Bizarro J, Charron C, Boulon S, et al. Proteomic and 3D structure analyses highlight the C/D box snoRNP assembly mechanism and its control. *J Cell Biol*, 2014, 207(4): 463-480
- [63] Zhao R, Kakiyama Y, Gribun A, et al. Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol*, 2008, 180(3): 563-578
- [64] Yu G, Zhao Y, Tian S, et al. Yeast R2TP interacts with extended termini of client protein Nop58p. *Sci Rep*, 2019, 9(1): 20228
- [65] Rothé B, Saliou JM, Quinternet M, et al. Protein Hit1, a novel box C/D snoRNP assembly factor, controls cellular concentration of the scaffolding protein Rsa1 by direct interaction. *Nucleic Acids Res*, 2014, 42(16): 10731-10747
- [66] Rothé B, Back R, Quinternet M, et al. Characterization of the interaction between protein Snu13p/15.5K and the Rsa1p/NUFIP factor and demonstration of its functional importance for snoRNP assembly. *Nucleic Acids Res*, 2014, 42(3): 2015-2036
- [67] Boulon S, Marmier-Gourrier N, Pradet-Balade B, et al. The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol*, 2008, 180(3): 579-595
- [68] Watkins NJ, Leverette RD, Xia L, et al. Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA*, 1996, 2(2): 118-133
- [69] Cavaillé J, Bachellerie JP. Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie*, 1996, 78(6): 443-456
- [70] Peng WT, Robinson MD, Mnaimneh S, et al. A panoramic view of yeast noncoding RNA processing. *Cell*, 2003, 113(7): 919-933
- [71] Hiley SL, Babak T, Hughes TR. Global analysis of yeast RNA processing identifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Res*, 2005, 33(9): 3048-3056
- [72] Paul A, Tiotiu D, Bragantini B, et al. Bcd1p controls RNA loading of the core protein Nop58 during C/D box snoRNP biogenesis. *RNA*, 2019, 25(4): 496-506
- [73] Khoshnevis S, Dreggors RE, Hoffmann TFR, et al. A conserved Bcd1 interaction essential for box C/D snoRNP biogenesis. *J Biol Chem*, 2019, 294(48): 18360-18371
- [74] Bragantini B, Charron C, Bourguet M, et al. The box C/D snoRNP assembly factor Bcd1 interacts with the histone chaperone Rtt106 and controls its transcription dependent activity. *Nat Commun*, 2021, 12(1): 1859
- [75] Yi Y, Li Y, Meng Q, et al. A PRC2-independent function for EZH2 in regulating rRNA 2'-O methylation and IRES-dependent translation. *Nat Cell Biol*, 2021, 23(4): 341-354
- [76] Liu L, Liu Z, Liu Q, et al. LncRNA INHEG promotes glioma stem cell maintenance and tumorigenicity through regulating rRNA 2'-O-methylation. *Nat Commun*, 2023, 14(1): 7526
- [77] Nachmani D, Bothmer AH, Grisendi S, et al. Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. *Nat Genet*, 2019, 51(10): 1518-1529
- [78] Aquino GRR, Krogh N, Hackert P, et al. RNA helicase-mediated regulation of snoRNP dynamics on pre-ribosomes and rRNA 2'-O-methylation. *Nucleic Acids Res*, 2021, 49(7): 4066-4084
- [79] Bohnsack MT, Martin R, Granneman S, et al. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell*, 2009, 36(4): 583-592