

人类约有2万个基因，仅占DNA的2%，剩下的98%是什么？这些区域如同基因组中的“暗物质”，有待科学家去发现。非编码RNA（核糖核酸）是基因组“暗物质”中的一类重要分子，最近十几年才被被发现。它们不仅在生命活动中发挥功能，还与许多疾病息息相关。

作为国际上较早从事长非编码RNA研究的科学家之一，中科院分子细胞科学卓越创新中心研究员陈玲玲自2011年回国后，带领团队创建并发展了新的研究技术和体系，在国际上率先发

现了多种在生命活动中发挥重要调控作用的新型长非编码RNA分子家族。

在生命的未知领域追光“解码”，这支上海科学家团队的相关成果先后发表于《科学》《细胞》等国际顶尖学术期刊。今年3月，他们又一项关于核糖体RNA的研究成果登上《自然》杂志。这些认知极大丰富了人类对生命中心法则的理解，也将揭开更多生命运行的奥秘。

——编者



上海科学家打破技术瓶颈，让人们认识到一个崭新的长非编码RNA世界

解码基因组“暗物质”，拓宽生命认知疆域

■ 乌皓 南杉

人类的基因不过2万多个，体内蛋白质的种类却超过5万种。分子生物学中心法则认为，遗传信息DNA通过转录产生RNA，RNA通过翻译产生蛋白质。作为遗传信息DNA的“信使”，RNA编码蛋白质，而蛋白质决定生物性状。随着研究的深入，科学家发现，生物体内成千上万个RNA分子中，编码蛋白质的RNA仅占不到2%，剩余高达98%都是不编码功能性蛋白质的非编码RNA。

起初，这些非编码RNA被认为是不具备生物学功能的“转录垃圾”。随着RNA测序技术的发展与遗传学研究的深入，非编码RNA分子的神秘身份逐渐被揭开。其中，一类长度大于200个核苷酸的长非编码RNA受到格外关注。

2012年，上海科学家团队通过创建新的RNA测序方法，率先发现了多种新型长非编码RNA分子家族。在这个神奇的未知世界里，它们有的在末端带有小核仁RNA，有的则呈环形，这些环形RNA还有着不同来源……近年来，越来越多研究表明，长非编码RNA在表观遗传、细胞周期以及细胞分化等众多生命活动中发挥着重要的调控作用。

占比98%的“争议神秘人”其实大有来头

1957年，DNA双螺旋结构的提出者之一、英国生物学家弗朗西斯·克里克提出了遗传学“中心法则”：DNA储存着大量遗传信息，它们通过转录产生RNA，RNA又通过翻译产生蛋白质，进一步参与到各种生物体的生命活动中并发挥多重作用。

RNA作为连接DNA和蛋白质的中间桥梁，既携带着从DNA传递而来的遗传信息，又是蛋白质合成的直接模板。因此，这些RNA也被形象地称为“信使RNA”(mRNA)——它们时刻不停地往来于DNA和蛋白质之间，完成重要的“收信”和“发信”工作。

如今，中心法则已成为现代生物学最重要、最基本的规律之一。所有细胞组成的生命体，小到细菌、大到哺乳动物乃至人类，都通过这种方式完成遗传信息的传递和表达。

在中心法则中，RNA不仅是生命遗传密码的传递者，还肩负着将DNA序列信息解码为功能蛋白质的重任。然而，科学家深入研究后发现，细胞中，承载这一重要使命的RNA占比极低，绝大部分并不参与蛋白质编码。很长一段时间以来，这些不参与蛋白质编码的RNA被视作基因组中的“暗物质”，它们的存在形式和功能不甚明确且备受争议，甚至有部分科学家认为它们就是“转录垃圾”。

但随着高通量RNA测序技术的革新与遗传学研究的不断深入，科学家们发现，在人类基因组中，98%的基因序列都是非编码序列，它们会广泛地转录并产生大量RNA，并在诸多生理和病理过程中发挥重要作用。

实际上，在动物、植物及真菌等真核生物中，非编码RNA在RNA总量中占据了相当大的比例，且存在形式多种多样。

上世纪60年代，科学家发现了一系列持家非编码RNA，包括核糖体RNA、转运RNA和小核仁RNA等。这些RNA在所有细胞中稳定而均一地表达，产生细胞生命活动所必需的蛋白质，发挥着极为重要的“持家”作用。

上世纪90年代，科学家在研究线虫和牵牛花的过程中，发现了一些长度只有20-22个核苷酸且具有调控功能的小非编码RNA。本世纪初，更多具备调控功能的小非编码RNA陆续被发现，它们可以让基因表达“沉默”，进而抑制蛋白质的表达。

与此同时，又有一类庞大的长非编码RNA分子家族被发现。这类RNA分子长度大于200个核苷酸，却不编码任何功能性蛋白质或多肽。于是，它们被归为无用的“转录垃圾”。

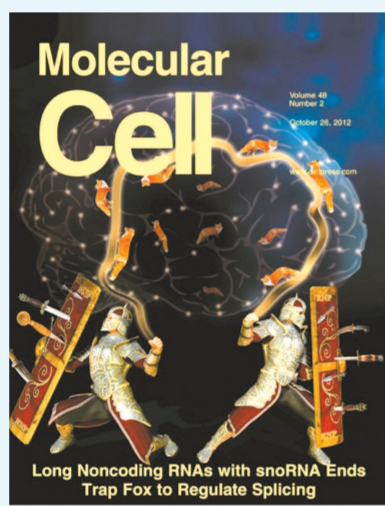
直到2010年，科学家们发现，由增强子转录形成的一种长非编码RNA——增强子RNA，能够调控基因表达。尽管科学家并不清楚它是如何进行调控的，但这一发现扭转了人们对于非编码RNA的认知。

中国科学家让“转录垃圾”实现逆袭

数量巨大、种类繁多、形态多样，这表明长非编码RNA可能具有很多潜在生物学功能，值得进一步挖掘和探索。凭借自身强大的调控潜力，被长期视为“转录垃圾”的长非编码RNA实现“逆袭”，成为炙手可热的研究热点。

不过，在探索长非编码RNA这个未知世界的过程中，由于技术所限，科学家始终无法一窥全貌。一位中国女科学家通过十年多努力，打破了这一僵局，拓宽了人类对长非编码RNA的认知。

在人体细胞中，经过前期加工的成熟mRNA为了保持自身稳定并顺利指导蛋白质合成，首尾两端会分别被特殊的甲基化“帽子”和长长的多聚腺苷酸“尾巴”保护起来。这一现象吸引了中国科学院分子细胞科学

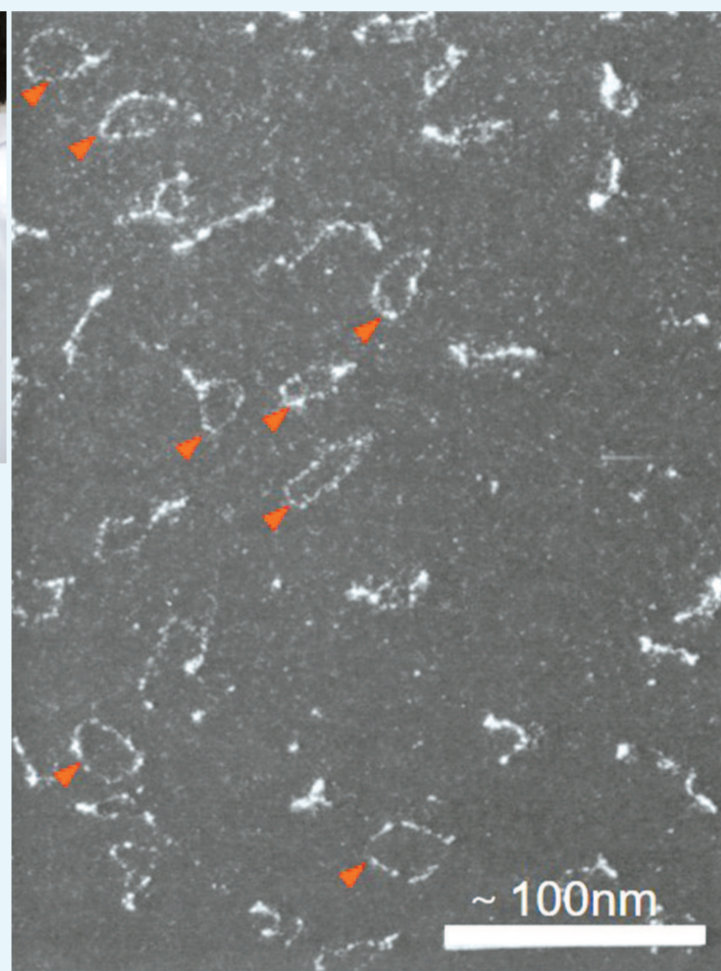


▲陈玲玲(左一)和中国科学院分子细胞科学卓越创新中心实验室的学生们。 新华社发

▶电子显微镜下观察到变性后的类病毒RNA呈现为由单链RNA构成的环形共价闭环结构。(摘自Heinz L. Sanger等人发表于1976年的《美国科学院院报》文章)

◀2012年，中国科学家团队率先发现两端带有小核仁RNA的长非编码RNA(sno-lncRNA)的研究成果登上当年10月26日出版的《分子细胞》期刊封面。

(本版图片除注明外，均由陈玲玲研究组提供)



非编码RNA对中心法则的再认识

■ 陈玲玲

1953年，DNA双螺旋模型的提出标志着近代生物学的诞生。1957年，弗朗西斯·克里克进一步提出了经典的遗传学中心法则，开启了分子生物学时代。

经过几十年的研究，科学家对于非编码RNA的认知极大丰富了人类对中心法则的理解。从广义上讲，非编码RNA在真核生物转录组中占据了非常大的部分，包括小RNA、长非编码RNA、环形RNA以及高度重复序列的RNA等多种类别。更重要的是，尽管非编码RNA不具备编码蛋白质的能力，但可以直接调控生物性状。

我和非编码RNA的渊源可以追溯到我在博士阶段接触的课题。当时，我就对非编码RNA的研究产生了浓厚兴趣。回国建立自己的实验室后，我提出了一个猜想：并非所有长非编码RNA都具有当时科学界所认为的、类似信使RNA的结构，可能存在一些

卓越中心研究员陈玲玲的注意。在早期研究中，她发现，很多类似mRNA结构的长非编码RNA在体内广泛表达。在此基础上，她提出了一个关键问题：是不是所有的长非编码RNA都具有和mRNA相似的结构？

想要解答这个问题，陈玲玲遇到了一个关键的技术难题：传统的RNA测序技术并不适用于这些未知的RNA分子。原来，传统RNA测序技术往往需要借助于磁珠来实现对RNA的富集，进而对RNA进行测序。这种磁珠能够特异性地捕获那些带有“尾巴”结构的mRNA，但也势必会遗漏掉那些可能不带“尾巴”的长非编码RNA。

为了解决这个问题，陈玲玲带领团队对传统RNA测序技术进行了优化，创建了新的RNA分离与测序方法。相较于传统RNA测序技术，这种方法

新型长非编码RNA家族有待发现。

基于这个猜想，我们实验室创建了针对无多聚腺苷酸尾RNA的研究新体系，为发现新型长非编码RNA分子奠定了基础。通过对人类细胞中无多聚腺苷酸尾转录本的深入分析，我们相继发现了来源于内含子的带有小核仁RNA的长非编码RNA家族、多基因连续转录加工产生的SPA长非编码RNA家族，以及环形长非编码RNA(简称环形RNA)，并证明这些RNA分子具有全新的生成途径和重要的生物学功能。

近些年来，一系列新兴的研究工具和技术极大推动了非编码RNA领域的研究，包括三代RNA测序、化学探针、超高分辨率显微成像，以及基于CRISPR-Cas9的基因编辑技术等。这些技术有利于我们更快地发现细胞、组织以及物种特异性的RNA异构体，也使我们能从分子二级结构、精准亚细胞定位等多个角度去深入解析长非编码RNA的作用模式及时空表达图谱。

仍然借助磁珠来富集带“尾巴”的mRNA，但关注的焦点则转移到那些剩余的、未被捕获到的RNA中。

在新型RNA测序技术的助力下，科学家们先后发现了多种新型长非编码RNA分子。

2012年，陈玲玲团队与合作者一起，率先发现并鉴定出一类新型长非编码RNA分子。这类RNA分子两端既没有“帽子”，也没有“尾巴”，而是带有一种能形成特定结构的小核仁RNA。这些小核仁RNA的存在，使得中间的RNA序列不会被降解。形象地说，小核仁RNA就像分别驻守在RNA分子两端的“护卫”，守护“中间领地”免受外界攻击。

进一步研究发现，这类RNA分子的表达缺失与小胖威利综合征的发生及发展密切相关。小胖威利综合征是一种由神经发育异常导致的罕

见遗传性疾病，但其具体病因至今仍未明晰。这种新型长非编码RNA分子的发现为阐明小胖威利综合征的病理机制提供了新的研究思路。此外，这类RNA分子对于调节细胞核仁结构与功能也至关重要。

2016年，陈玲玲团队又发现了另一种新型长非编码RNA分子。与“掐头去尾”的小核仁RNA-长非编码RNA不同，这种RNA分子带有和mRNA末端相同的“尾巴”结构，但头部的“帽子”则被小核仁RNA所替代，研究者将其命名为“SPA”。他们发现，在小胖威利综合征患者基因的关键缺失区域，存在两条SPA长非编码RNA。这在解开小胖威利综合征致病机理谜团的同时进一步提示，非编码RNA在人类疾病中扮演重要角色。

2017年，陈玲玲研究员带领团队发现了一条

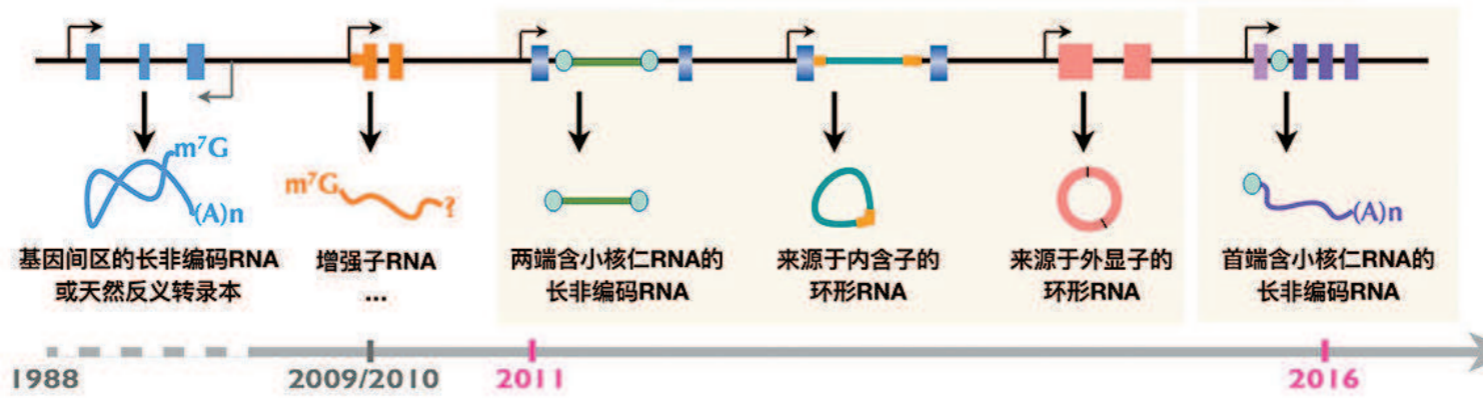
见遗传性疾病，但其具体病因至今仍未明晰。这种新型长非编码RNA分子的发现为阐明小胖威利综合征的病理机制提供了新的研究思路。此外，这类RNA分子对于调节细胞核仁结构与功能也至关重要。

2016年，陈玲玲团队又发现了另一种新型长非编码RNA分子。与“掐头去尾”的小核仁RNA-长非编码RNA不同，这种RNA分子带有和mRNA末端相同的“尾巴”结构，但头部的“帽子”则被小核仁RNA所替代，研究者将其命名为“SPA”。他们发现，在小胖威利综合征患者基因的关键缺失区域，存在两条SPA长非编码RNA。这在解开小胖威利综合征致病机理谜团的同时进一步提示，非编码RNA在人类疾病中扮演重要角色。

2017年，陈玲玲研究员带领团队发现了一条

见遗传性疾病，但其具体病因至今仍未明晰。这种新型长非编码RNA分子的发现为阐明小胖威利综合征的病理机制提供了新的研究思路。此外，这类RNA分子对于调节细胞核仁结构与功能也至关重要。

针对无多聚腺苷酸尾的新RNA测序技术



定位于核仁的全新小核仁RNA-长非编码RNA，并根据其结构特征和功能，将其命名为“SLERT”。这是科学家首次在人类细胞中发现可调控RNA聚合酶转录的长非编码RNA。

由此可见，此前被认为是“转录垃圾”的长非编码RNA实际上蕴藏着巨大的功能潜力。越来越多研究表明，长非编码RNA能够以不同的分子机制参与增殖、分化、代谢、自噬、衰老和凋亡等多种生命过程，发挥重要的调控作用。

近年来，针对长非编码RNA的研究快速增长。这些研究表明，RNA不仅是遗传信息的传递者，还是遗传过程的调节者，由此对遗传学中心法则进行了有力补充。与此同时，一系列新发现揭开了许多困扰人类多年的遗传疾病的病因，为相关疾病的诊疗带来新希望。

生物界“莫比乌斯环”掀起研究热潮

在非编码RNA的未知世界里，环形RNA因其特殊结构而引起了科学家们的极大兴趣。这是一类由单链RNA分子构成、呈现封闭环状结构的非编码RNA分子，犹如一个个玲珑的镯子。细胞内，由DNA转录出的只是前体mRNA，它必须经过复杂的剪接和加工，才能变成成熟的mRNA，成为合成蛋白质的模板。一般线性RNA所经历的剪接过程，被称为“经典剪接”。而环形RNA则由前体mRNA通过“反向剪接”，首尾相互连接，进而形成一个闭合环状的RNA分子。

由于具有完整的环形结构，环形RNA不易被识别线性RNA的核糖核酸酶“切割”而降解，因此比线性RNA更加稳定。

目前研究发现，环形RNA主要来自基因外显子，但多项研究显示其他多种基因结构也可产生环形RNA。这说明，环形RNA的种类及生物合成可能比目前所知的更为复杂。尽管环形RNA有很多不同来源，但当我们提到环形RNA时，通常指的是来源于基因外显子的环形RNA，这些环形RNA主要存在于细胞质中。

外显子来源的环形RNA由单链RNA分子通过共价键形成。1976年，科学家利用电镜技术，首次在植物感染类病毒中观察到了这种闭合环形结构。然而，由于环形RNA结构独特，且在细胞内表达水平较低，此后的几十年中，仅有少数环形RNA被发现或鉴定。

很久以来，环形RNA都被认为是转录过程中错误剪接的产物，并不具备生物学功能。直到1993年，科学家发现，小鼠Sry基因转录出的环形RNA可能在成年小鼠的睾丸中发挥特定功能。由此，环形RNA才在非编码RNA领域被广泛关注，并逐渐跃升为新的“明星分子”，一股环形RNA的研究热潮开始在世界范围内兴起。

近年来，随着新一代高通量RNA测序技术及多种特异性针对环形RNA的生物化学实验和计算生物学方法的快速发展，在植物、线虫、斑马鱼、果蝇、小鼠和人类等众多物种中，数以千万计的环形RNA得以被发现和识别。

科学家目前认为，环形RNA是一类表达相对保守，且长期存在于不同生物体内的生物大分子。越来越多研究表明，环形RNA在细胞内扮演着多种重要角色。例如，环形RNA不仅可以作为“分子海绵”吸附微小RNA并调控其活性，还可与核糖核蛋白结合形成复合物，调节特定的信号通路。

此外，近期研究发现，环形RNA可通过特殊途径自身翻译产生多肽。这表明，即使是非编码RNA，也可能具有潜在的编码功能，这无疑打破了人们的固有认识。而且，环形RNA比线性RNA能更稳定地存在于组织、细胞和体液中，半衰期也比线性RNA长。很多环形RNA只有在某个组织或某个发育阶段才能被检测到，在表达上呈现特异性。

仅由一条单链RNA分子首尾相连而产生的环形RNA具有诸多独特性质，这不仅让人联想到莫比乌斯环——一个永无止境的神秘之环。只需要将一根纸条翻转180°，再将两端粘在一起，就可以形成这样一个重要的拓扑学结构。环形RNA好似生物世界中的莫比乌斯环，简单的结构却承载着神奇的特质，等待我们去发现和探索。

近年来，在临床方面，环形RNA与疾病的联系也受到越来越多的关注。目前，已有大量研究揭示环形RNA在癌症中发挥重要作用，而环形RNA的高度稳定性和表达特异性，使其有望成为潜在的癌症生物标志物和治疗靶点。

尽管长非编码RNA分子天然的柔性、折叠和结构异质性，为RNA疗法的研究与开发带来极大挑战，但mRNA疫苗、反义寡核苷酸以及小分子干扰RNA等核酸药物的快速发展和应用，让人们看到了RNA的独特优势和临床应用潜力。

如今，第三代长读码RNA测序技术、超高分辨率显微成像技术以及基因编辑和碱基编辑器技术的发展，为长非编码RNA研究带来新机遇，也为相关领域研究提出新方向。未来，围绕环形RNA及其他非编码RNA分子，科学家有望开发出全新的RNA治疗策略。

(作者分别为中国科学院分子细胞科学卓越创新中心副研究员、中国科学院)