

ICS 35.240.80
C 07

团 体 标 准

T/CHIA 41.1-2023

非编码 RNA 注释标准 第 1 部分：分类信息

Specifications for annotation of non-coding RNA
Part 1: Classification information

2023-11-14 发布

2024-02-01 实施

中国卫生信息与健康医疗大数据学会 发布

目 次

前 言.....	I
引 言.....	II
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 缩略语.....	3
5 非编码 RNA 的分类.....	3

前 言

T/CHIA 41-2023《非编码RNA注释标准》分为以下3个部分：

- 第1部分：分类信息；
- 第2部分：基本信息注释；
- 第3部分：功能及疾病注释。

本部分为T/CHIA 41-2023的第1部分。

本部分按照 GB/T 1.1-2020给出的规则起草。

本部分由中国科学院生物物理研究所提出，由中国卫生信息与健康大数据学会归口。

本部分主要起草单位：中国科学院生物物理研究所、中国科学院北京基因组研究所（国家生物信息中心）、浙江大学、复旦大学、清华大学、中国人民解放军总医院、北京蛋白质组研究中心、中国科学院微生物研究所、北京大学人民医院、中国科学院上海营养与健康研究所、中南大学、空军军医大学（第四军医大学）和北京睿博解码生物科技有限公司。

本部分的主要起草人：陈润生、何顺民、宋廷瑞、张鹏、周红红、王晓娜、方向东、金力、何昆仑、李亦学、张学工、段会龙、周水庚、渠鸿竹、赵思琪、钱颖、王霞、吕旭东、朱云平、马俊才、杨忠、石乐明、吴松峰、吴林寰、王振、陈先来、贾志龙、张昭军、娄晓敏、阮修艳、单广乐、乔楠、刘登辉、丁子建。

引 言

《非编码 RNA 注释标准 第 1 部分：分类信息》旨在为非编码 RNA 的注释提供一套术语规范、条理清晰、意义明确、语义语境无歧义的分类标准，防止注释者在注释过程中错误认定非编码 RNA 或非编码 RNA 的类型，并保证分类信息描述的一致性，便于对非编码 RNA 已知知识的获取，保证信息的有效交换、有效挖掘分析和共享。

非编码 RNA 注释标准

第 1 部分：分类信息

1 范围

本标准规定了非编码RNA注释中相关的分类信息。

本标准适用于对非编码 RNA 的注释。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 29859-2013 生物信息学术语

3 术语和定义

GB/T 29859-2013 中界定的以及下列术语和定义适用于本标准。

3.1

非编码 RNA non-coding RNA

非编码 RNA 是一类不具有蛋白编码能力的 RNA 转录本。这里的“非编码”是指转录本中缺乏有效的可以编码并翻译为蛋白质的开放阅读框和/或保守密码子。值得注意的是，近年来有研究表明某些长非编码 RNA 可以翻译产生小蛋白（小肽），但通常认为这些小肽不是传统意义上有功能的蛋白质，因此不足以认定这些 RNA 属于编码的 RNA。总的来说，如果一个 RNA 被认定为编码的 RNA（信使 RNA，mRNA），那它就不属于非编码 RNA，否则就属于非编码 RNA。

3.2

核糖体 RNA ribosomal RNA

含量最多的一类分子量较大的 RNA，结构相当复杂，与多种蛋白质结合形成核糖体，是蛋白翻译的核心组件。

3.3

转运 RNA transfer RNA

具有携带并转运氨基酸功能的一类小分子 RNA，长约 70-90 个核苷酸，折叠成三叶草形状，为蛋白翻译活动提供氨基酸原料。

3.4

小 RNA, 微 RNA microRNA

真核生物中一类长度在 19-24 nt 的内源性具有调节基因表达功能的单链非编码 RNA 分子, 由发夹结构的前体加工形成。

3.5

长非编码 RNA long non-coding RNA

一类转录本长度大于 200 nt 的, 和信使 RNA 类似但不编码蛋白质的长链非编码 RNA 分子。也称为长链非编码 RNA。

3.6

小核 RNA small nuclear RNA

位于真核细胞核中, 主要存在于核浆的可溶性部分, 是真核生物转录后加工过程中 RNA 剪接体的主要成分, 在哺乳动物中长度约为 100-215 nt。

3.7

小核仁 RNA small nucleolar RNA

一类广泛存在于真核生物细胞核仁中的由内含子编码的小分子非编码 RNA, 长度 60-300 nt, 具有保守的结构元件。

3.8

小干扰 RNA small interfering RNA

一类长度约为 20-25 nt 的非编码小 RNA 分子, 由双链 RNA 加工形成, 通过与靶基因的碱基互补配对, 指导基因沉默。

3.9

Piwi 相互作用 RNA Piwi-interacting RNA

一类与 Piwi 蛋白相互作用的长约 26-31 nt 的小非编码 RNA, 主要在动物生殖细胞中表达。

3.10

环状 RNA circular RNA

一类以共价键形成闭合环状结构的非编码 RNA 分子, 没有 5'帽子结构和 3' poly(A)尾巴结构。

3.11

小胞质 RNA small cytosol RNA

也叫 7SL RNA，一类主要存在于细胞质中的长约 300 nt 的非编码 RNA。

4 缩略语

下列缩略语适用于本文件。

RNA：核糖核酸（Ribonucleic Acid）。

DNA：脱氧核糖核酸（Deoxyribonucleic Acid）。

nt：核苷酸（nucleotide）。

mRNA：信使 RNA（messenger RNA）。

rRNA：核糖体 RNA（ribosomal RNA）。

tRNA：转运 RNA（transfer RNA）。

miRNA：小 RNA，或微 RNA（microRNA）。

lncRNA：长非编码 RNA（long non-coding RNA）。

siRNA：小干扰 RNA（small interfering RNA）。

piRNA：Piwi 相互作用 RNA（Piwi-interacting RNA）。

snRNA：小核 RNA（small nuclear RNA）。

snoRNA：小核仁 RNA（small nucleolar RNA）。

circRNA：环状 RNA（circular RNA）。

scRNA：小胞质 RNA（small cytosol RNA）。

dsRNA：双链 RNA（double-stranded RNA）。

SO：序列本体论（sequence ontology）。

5 非编码 RNA 的分类

非编码 RNA 分类描述见表 1。

表 1 非编码 RNA 分类

名称	英文名称	SO 识别号	描述
核糖体 RNA	rRNA	SO:0000252	构成核糖体的 RNA，提供结构支架和催化活性，在信使 RNA 的指导下将氨基酸合成为肽链。rRNA 的含量是所有 RNA 分子中最高的。真核生物 rRNA 包括：5S rRNA、5.8S rRNA、18S rRNA、28S rRNA。
转运 RNA	tRNA	SO:0000253	长约 70-90 个核苷酸，折叠成三叶草形状，把氨基酸转运到核糖体，参与多肽链的合成。
小 RNA/微 RNA	miRNA	SO:0000276	来源于一些转录本形成的独特发夹结构，被称为 pre-miRNA，加工成约 22nt 成熟的 miRNA，参与 RNA 诱导的沉默复合体（RISC），通过与 mRNA 转录本上的序列互补配对，导致靶基因沉默。具

名称	英文名称	SO 识别号	描述
			有高度保守性、时序表达和组织表达特异性。
长非编码 RNA	lncRNA	SO:0001877	一类转录本长度大于 200 nt 的不编码 100 氨基酸以上长度蛋白质的长链 RNA 分子，在细胞核及细胞质中都有分布，可能由 RNA 聚合酶二或 RNA 聚合酶三转录，根据基因组位置可分为多种类型，如基因间 lncRNA、反义链 lncRNA、假基因等等。许多 lncRNA 具有重要的调控功能，包括表观遗传调控、转录调控、转录后调控等。
小核 RNA	snRNA	SO:0000274	一直存在于细胞核中，与 40 种左右的核内蛋白质共同组成 RNA 剪接体，有 U1、U2、U4、U6 等，主要功能是处理细胞核中的 mRNA 前体，参与 RNA 的剪接，并在将 mRNA 从细胞核转运到细胞质的过程中起着十分重要的作用。
小核仁 RNA	snoRNA	SO:0000275	与核仁核糖核蛋白结合形成 snoRNP 复合物；主要存在于蛋白编码基因或非蛋白编码基因的内含子区域；按保守的结构元件可以分为 box C/D snoRNA、box H/ACA snoRNA 和 MRP RNA；主要参与 rRNA 的加工处理、RNA 剪接和翻译过程的调控以及氧化应激反应，还参与到遗传性疾病、造血、代谢以及癌症的过程中。
Piwi 蛋白相互作用 RNA	piRNA	SO:0001035	与 Piwi 蛋白质结合的一类小 RNA，， 5' 端第一个碱基通常为尿苷，且为单磷酸核苷，3' 端有 2'-O-甲基修饰，对于维持生殖细胞基因组完整、抑制转座子转录、抑制翻译、参与异染色质的形成、执行表观遗传调控和生殖细胞发生等有重要作用。
小干扰 RNA	siRNA	SO:0000646	由长链 dsRNA 经过 Dicer 加工成长度为 20-25 nt 的成熟 siRNA，加载到 RISC 后，通过与靶基因的碱基互补配对，指导基因沉默，该过程通常被称为 RNA 干扰 (RNAi)。
环状 RNA	circRNA	SO:0002291	以共价键形成闭合环状结构的 RNA 分子，大致分为四类：全外显子型的 circRNA，内含子和外显子组合的 EIcircRNA，内含子组成的套索型 ciRNA，由病毒 RNA 基因组、tRNA、rRNA、snRNA 等环化产生的 circRNA。主要位于细胞质或储存于外泌体中，不受 RNA 外切酶影响，表达更稳定且不易降解，广泛存在于多种真核生物体内。
小胞质 RNA	scRNA	SO:0000013	是蛋白质在粗面内质网定位合成所需的信号识别体的组成成分。
其他	others	-	其他不常见或未分类的非编码 RNA 分子。