mRNA 药物分子 poly(A)尾部结构优化

完成人

张楚玥

指导小组成员

林金钟 教授

目 录

摘要		I
Abst	ract	II
_,	前青	È 1
<u> </u>	材料	与方法
	2.1	材料3
	2.2	试剂
	2.3	仪器
	2.4	软件8
	2.5	方法
三、	研究	结果
	3.1	候选序列设计 13
	3.2	质粒稳定性检验······13
	3.3	IVT mRNA 功能检验16
四、i	讨 论	21
	4.1	质粒稳定性实验的反思
	4.2	IVT mRNA 表达水平和半衰期实验反思
	4.3	IVT mRNA 表达水平和半衰期与 poly(A)结合蛋白指纹相关 22
参考	文献	
致谢	•••	

摘要

mRNA 具有强大的药用潜力,2019 年的新冠疫情更激发了 mRNA 药物的研究热潮。Poly(A)尾是 mRNA 药物不可缺少的一部分,然而现有的加尾方式却存在各种问题。携带长链 poly(A)的质粒会在大肠杆菌中发生高频重组,最终导致poly(A)尾长度变短或丢失,限制了 mRNA 药物的工业生产和实际应用。本论文通过向 poly(A)尾中按不同频率掺入单碱基鸟苷酸 G 构建了 16 条候选序列。我利用建立的简化版质粒稳定性检验方法,计算出携带不同候选序列质粒的灰度重组率;使用流式细胞仪定量候选序列的表达水平,并定性研究最优候选序列的半衰期。从候选序列中,我筛选出了具有高质粒稳定性、原始表达量和原始半衰期的最优序列。这一新型 poly(A)尾更适于工业生产,为后续的 mRNA 药物研发与应用提供了新的素材。

关键词: mRNA 药物, poly(A)尾, 质粒稳定性

Abstract

Messenger RNA holds great potential in therapeutics, particularly with the COVID-19 pandemic fueling the research enthusiasm. Poly(A) tail is an indispensable part of mRNA medicine, but current tailing techniques have their limitations. Plasmids carrying long stretches of poly(A) are highly prone to recombination, which would result in a loss of its number, thus limiting the industrial production and application of mRNA medicine. By inserting intermittent single nucleotide guanosine into poly(A) tail, 16 candidate sequences were constructed. I established a simplified plasmid stability assessment assay, and used it to calculate the gray value-based recombination rate of different candidates. Flow cytometry is used for expression level quantification of all candidate with high plasmid stability, unchanged expression level, and unchanged half-life. This novel poly(A) tail is better suited for industrial production and provides new insights for future mRNA medicine development and application.

Key words: mRNA medicine, poly(A) tail, plasmid stability

一、前言

mRNA 药物适用于各种疾病的治疗,且相较其他手段具有强竞争力。mRNA 药物比 DNA 药物更加安全可靠,也比蛋白类药物更成本低廉、应用范围广泛。此外,mRNA 药物在精准医疗方面也具有巨大潜力。体外转录 mRNA(IVT mRNA)制备简单,编辑方便,可根据需要使用不同的核苷酸修饰。由于 mRNA 仅在靶 细胞中短暂存在,其产量和时长也能被人为控制。目前 mRNA 药物已存在疫苗 ^{[1],[2]}、免疫治疗^[3]和代谢性疾病^[4]等多方面的应用。如近期获批的新冠病毒疫苗 就有两款是 mRNA 药物,相较传统疫苗,它们研发周期短且免疫效果强。

尽管 mRNA 药物具有强大的治疗潜力,其实际应用中仍面临一些问题。 mRNA 序列在质粒上的稳定性是其中极为重要的一点。完整的 mRNA 药物一般 具有 5'帽子、5'非翻译区、编码区、3'非翻译区和 poly(A)尾这五个结构。 其中 5'加帽可通过帽子类似物实现,而 3'加尾则有酶法、PCR 法和质粒固有几 种方式^[5]。使用大肠杆菌 poly(A)聚合酶 I 的酶法加帽存在添加的碱基数量不均一 问题,无法作为药物使用。PCR 加尾则一般用于实验室制备,在进行大量生产时 因价格高昂和突变风险高而并不适用。因此仅有质粒固有这一方法因质粒扩增方 便而适于进行 mRNA 药物的工厂化生产。然而,长链的 poly(A)尾在质粒上并不 稳定,很容易因大范围的重组而丢失序列,限制了质粒法的应用。

质粒作为一个运用广泛、利于快速大量表达,且便于储存运输的载体,在实验研究和药物生产中有着重要的地位。其中,质粒的稳定性在 mRNA 药物的研究、生产和应用中至关重要。其结构稳定性弱不仅会影响质粒产量,更会导致最终产物的不均一,药物质量低与高危险性等问题。而质粒的稳定性受到质粒长度、单细胞内质粒数量、复制方式、宿主菌种类、培养条件等多方面因素的影响^[6]。如重复 DNA 结构单元会影响质粒的高级结构,出现的非沃森-克里克配对的高级结构比正常序列更容易发生删除、重复、插入、转位等基因重组^[7]。因此,提高 poly(A)在质粒上的稳定性对发展 mRNA 药物迫在眉睫。

目前解决 poly(A)稳定性的方法可主要分为利用线性质粒进行扩增和截断 poly(A)两种。尽管线性质粒能够维持长片段(~500 bp)poly(A)的稳定性^[8],但 其存在拷贝数低、质粒太大、可用的限制性酶数量少等问题,导致产率低且应用

1

壁垒大,限制了其在工业上的应用。截断 poly(A)则是目前 BioNTech 公司使用的 设计,即向 poly(A)中插入 10 nt 各碱基数量均匀的随机序列作为间隔,就能得到 质粒重组率低(~3%)且 mRNA 功能不变的 poly(A)结构^[9]。

mRNA 药物的基本性质包括翻译效率和半衰期^[10]。我们需要在提高 poly(A) 尾质粒稳定性的同时维持 mRNA 的正常功能。为了保证 poly(A)尾在保护 mRNA 不被降解以及起始翻译方面的关键作用不被破坏,并为了降低可能出现的免疫反 应,对 poly(A)序列的改变应尽可能小幅且最好是内源已存在有的。

除在细胞核中向 mRNA 后方添加 poly(A)尾的 poly(A)聚合酶(PAP)外,细胞质中也存在能够延伸 poly(A)的聚合酶,即末端核苷酸转移酶(TENTs)。不同的 TENT 酶具有不同的碱基偏好性,在生理状态下能够向 poly(A)中加入 U, C,G 核苷酸^[11]。其中,C 出现频率较低,U 一般会导致 mRNA 被更快速地降解,而G 则多在长链 poly(A)尾中单个出现^[12],并被证实具有抑制脱腺苷酶 CCR4-NOT 复合体的功能,能够延迟 mRNA 降解^[13]。此外,也有体外实验证据表明G 能够扰乱 poly(A)的卷积螺旋结构,并抑制脱腺苷酶 PAN2-PAN3 的降解 活性^[14]。因此,我选择了向 poly(A)尾中掺入鸟苷酸G 以稳定质粒。

本研究旨在探究作为 mRNA 药物重要结构之一的 poly(A)尾,在质粒扩增过 程中不稳定、易于重组问题的解决方案。通过在 poly(A)序列中按不同频率掺入 单碱基 G 持入能够在不同程度上提高 poly(A)序列的质粒稳定性测试结果表明, 单碱基 G 掺入能够在不同程度上提高 poly(A)序列的质粒稳定性。IVT mRNA 表 达实验结果则显示,掺入 G 对 mRNA 药物在真核细胞中的表达量无显著影响。 对其中具有高稳定性的最优候选序列,定性分析其半衰期,发现半衰期也无明显 变化。因此,可以使用掺入单碱基鸟苷酸 G 的新型 poly(A)尾构建兼具高表达量、 长半衰期与高质粒稳定性的 mRNA 药物。

综上,本研究对 IVT mRNA 的 poly(A)尾结构进行了优化,使其更适于药物 研发与工业化生产,以期为新冠病毒 mRNA 疫苗等 mRNA 药物的研发提供全新 的素材。

2

二、材料与方法

2.1 材料

2.1.1 细胞样品

293T 细胞,实验室保存,由蔡亮老师实验室杜容瑢赠予获得。

2.1.2 载体与宿主菌

所用质粒载体为 pUC57-Kan 和 pcDNA3.1。 DH5α 感受态细菌(诺唯赞 Vazyme Biotech) JM109 感受态细菌(全式金 TransGen Biotech)

2.2 试剂

2.2.1 培养基

LB 培养基(1L)

组分	含量
Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g
*Agar	15 g

*配制固体培养基时加入。

2.2.2 质粒、酶、引物

2.2.2.1 质粒

本实验所用质粒与用途如下:

质粒名称	用途
pcDNA3.1-EGFP_DC50-genscript	密码子修改的 EGFP
pcDNA3.1-EGFP_DC52-genscript	密码子修改的 EGFP
pcDNA3.1-EGFP_DC54-genscript	密码子修改的 EGFP,RNA 合成载体
pcDNA3.1-EGFP_DC56-genscript	密码子修改的 EGFP

pcDNA3.1-EGFP_DC58-genscript	密码子修改的 EGFP
pcDNA3.1-EGFP_DC60-genscript	密码子修改的 EGFP
pcDNA3.1-EGFP_DC62-genscript	密码子修改的 EGFP
pcDNA3.1-EGFPwt-TL	野生型 EGFP 对照
pcDNA3.1_mRNABB-Fluc64-TL	萤火虫荧光素酶,后续 RNA 合成载体
pUC57-kan-mBB-A10x10	质粒稳定性检验
pUC57-kan-mBB-A12x8	质粒稳定性检验
pUC57-kan-mBB-A15x6	质粒稳定性检验
pUC57-kan-mBB-A20A30A50	质粒稳定性检验
pUC57-kan-mBB-A20A40A40	质粒稳定性检验
pUC57-kan-mBB-A20×5	质粒稳定性检验
pUC57-kan-mBB-A20×2A30×2	质粒稳定性检验
pUC57-kan-mBB-A20A40A40	质粒稳定性检验
pUC57-kan-mBB-A20A30A50	质粒稳定性检验
pUC57-kan-mBB-A25×4	质粒稳定性检验
pUC57-kan-mBB-A30A30A40	质粒稳定性检验与后续实验载体质粒
pUC57-kan-mBB-A30A70	质粒稳定性检验
pUC57-kan-mBB-A40A40A20	质粒稳定性检验
pUC57-kan-mBB-A40A60	质粒稳定性检验
pUC57-kan-mBB-A50A50	质粒稳定性检验
pUC57-kan-mBB-A100GA	质粒稳定性检验

质粒购自金斯瑞生物科技。

2.2.2.2 引物

本实验所用引物及其序列如下:

引物名称	序列
pcDNA3.1-T7-F	CCACTGCTTACGCCGTAATACGACTCACT
EGFP-TL-R	CTTAATTAAGCGGCCGCCTACTTGTACAGCTCGTCCATGC
TL-3UTR+T10	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
C-R	ТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТ

	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
TL-3UTR+T12	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
C-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+T15	ТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТ
C-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TCTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+T20	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
C-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T2+3	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T2+4	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T2+3	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
+5-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+T25	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
C-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T3+4	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T3+7	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T4+2	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T4+6	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT

	TTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR-T5+5	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+C-R	CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+CT-	ТСТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТ
R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+CTC	СТСТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТТ
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT
TL-3UTR+T3L	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
7-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
TL-3UTR+100T	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
	TTTTTTTTTTTTTTTCTTCCTACTCAGGCTTTAT

引物购自金斯瑞生物科技。

2.2.2.3 酶

用于 PCR 的 Phusion DNA 聚合酶, 5×Phusion Buffer 缓冲液由实验室自行 纯化或配置。dNTP 购自诺唯赞 Vazyme Biotech 公司。

用于酶切反应的 HindIII-HF, NotI-HF, CutSmart 缓冲液购自 NEB 公司。

用于体外转录 RNA 的 T7 RNA 聚合酶, 10×T7 Buffer 缓冲液由实验室自行 纯化或配置。NTP 购自诺唯赞 Vazyme Biotech 公司, ARCA 购自 APExBIO 公司。

组分	含量
0.5 M EDTA	100 mL
Tris Base	242 g

$2.2.350 \times IAE$ Buffer (1 L

调节 pH 至 8.4。

2.2.4 10×TBE Buffer (1 L)

组分	含量
Tris base	108 g
Boric Acid	55 g
EDTA (0.5 M, pH 8.0)	40 mL

2.2.5 2×RNA 上样缓冲液(20 mL)

	含量
冰系	9.6 g
$10 \times \text{TBE}$	2 mL
2.5% 溴酚蓝	200 µL
2.5%二甲苯蓝 FF	200 µL
EDTA (0.5 M, pH 8.0)	600 µL

-20℃保存。

2.2.6 Solution D (200 mL)

组分	含量	
Gudine Thiocyanate	100 g	
100 mM Sodium Citrate	53 mL	
定容至 200 mL, 0.22 µm 滤芯过滤。		
14.28 Μβ-巯基乙醇(使用前再加)	1.48 mL	

2.2.7 小抽试剂盒

购自诺唯赞 Vazyme Biotech 公司。

2.2.8 10×PBS 溶液(1 L)

组分

NaCl	80 g
KCl	2 g
Na ₂ HPO ₄ ·7H ₂ O	27.15 g
KH ₂ PO ₄	2.4 g

调节 pH 至 7.4。

2.2.9 细胞培养与转染

细胞培养液:

组分	含量
Gibco ®DMEM	500 mL
10% Gibco ®FBS	50 mL
$100 \times P/S$ (Penicillin/Streptomycin)	5 mL

DMEM 细胞培养液与 FBS 购自 Thermo Fisher Scientific 公司,抗生素 P/S 购自 Sangon Biotech 公司。

Trypsin, Opti-MEM 无血清培养基, Invitrogen Lipofectamine 2000 均购自 Thermo Fisher Scientific。多聚甲醛(PFA)购自 Sangon Biotech 公司。

2.2.10 其他试剂

RNA 纯化柱购自 Solarbio 索莱宝科技有限公司。

2.3 仪器

Eppendorf Mastercycler[®] nexus 进行 PCR Olympus 倒置荧光显微镜 IX73 进行显微成像 BD FACSJazz[™] Cell Sorter 流式细胞仪进行流式细胞术 Bio-Rad ChemiDoc[™] Touch 进行凝胶成像

2.4 软件

SnapGene[®] 4.2.1 查看质粒 ImageJ 1.53c 处理胶图 FlowJo v10.6.2 处理流式数据 GraphPad Prism 7 做数据图

2.5 方法

2.5.1 Phusion PCR (50 μL)

PCR 反应体系:

组分	含量
Phusion DNA Polymerase	1 µL
5×Phusion Buffer	10 µL
25 mM dNTP	0.4 μL
10 µM Primer (Forward + Reverse)	$1 + 1 \ \mu L$
DNA Template	5 ng (0.5 μL)
DEPC-H ₂ O	36.1 µL

PCR 反应程序:

程序	参数	循环数
预变性	95 °C, 2 min	1 cycle
变性	95 °C, 20 s	
退火	58 °C, 20 s	32 cycles
延伸	72 °C, 1 kb/30 s	
终延伸	72 °C, 4 min	1 cycle

产物在1%琼脂糖凝胶中进行电泳(180 V, 20 min)验证,用 GelRed 燃料 染色 DNA 并观察。

2.5.2 酶切反应(NotI-HF 与 HindIII-HF)

组分	含量
Plasmid	400 ng
10×CutSmart 缓冲液	2 µL
NotI-HF + HindIII-HF	$0.4 + 0.4 \ \mu L$
ddH_2O	补足至 20 μL

2.5.3 体外转录 RNA(20 µL)

组分	含量
T7 RNA Polymerase	2 μL
$10 \times$ T7 Reaction Buffer	2 µL
100 mM ARCA	1.2 μL
NTP mix (25 mM ATP, 18.75 mM	16.11
GTP,25 mM CTP, 25 mM UTP)	1.0 μΕ
PCR Template	2.5 μL
DEPC-H ₂ O	10.7 µL

37℃水浴/金属浴2h。

2.5.4 体外转录 RNA 纯化

 消化 DNA 模板: 向 20 μL 体外转录 RNA 体系中加入 1 μL DNaseI, 37 ℃ 水浴/金属浴 15 min。

向 20 μL 体系中加入 80 μL DEPC-H₂O, 350 μL Solution D 与 250 μL 100%
乙醇并混匀。装到 RNA 纯化柱上,静置 2 min, 11,000 g 离心 1 min。

3) 加入 500 µL 75% 乙醇, 11,000 g 离心 1 min, 重复 2 次。

4) 将 RNA 柱转移至干净 EP 管中,加入 70 μL 1 mM 柠檬酸钠,65℃静置
10 min, 11,000 g 离心 2 min。

转录获得的 RNA 于-80℃储存,使用时稀释至~100 ng/ µL 并保存于-20℃。

2.5.5 RNA 跑胶验证

取 2.5 µL RNA 产物加入 2.5 µL 2×RNA 上样缓冲液,混合,70℃金属浴 7-10 min,迅速放到冰上冷却,1%琼脂糖凝胶跑胶验证。

2.5.6 细胞传代培养(10 cm 培养皿)

1) 吸去培养皿内的培养液,加入 1-2 mL PBS 润洗细胞后移除。

2) 加入1 mL 胰酶消化至部分细胞脱落后,加入2 mL DMEM 培养基中和, 轻吹细胞使其从培养基表面全部脱落。

3) 将培养皿内液体全部转移至 15 mL 移液管中, 1000 rpm 离心 1 min, 吸去上清。

4) 向新的细胞培养皿中加入 7-10 mL DMEM 培养液。

5) 加入 1-2 mL DMEM 培养基重悬细胞,吸取适量比例后以悬滴形式加入新的细胞培养皿中, 37℃ 5% CO₂培养。

2.5.7 细胞瞬时转染(以24孔板为例)

1) 按照 1.5×10⁵ 个细胞/孔的数量将细胞均匀地铺在 24 孔板内。

2) 培养 24 h 后进行转染。

将 mRNA 统一稀释至~100 ng/µL。

A 管	250 ng mRNA	20 µL opti-MEM
B 管	0.75 μL lipo 2000	20 µL opti-MEM

将A管和B管分别配置后室温静置5min。混合A、B管,移液器轻轻吹打 混匀后室温静置15min,以悬滴形式加入24孔板内。

3) 培养16h后观察或进行后续实验。

2.5.8 流式细胞样品准备(以24孔板为例)

1) 吸去表面培养液,加入100 µL Trypsin,震荡4-5 下后迅速吸去。

2) 加入 200-400 μL PBS 吹打至培养及表面细胞完全脱落,吸出细胞悬液加入等体积 PFA 溶液中。

3) 4℃避光保存,上样前吸去上清约 200 μL,将细胞充分吹打重悬后使用流 式细胞仪进行检测。

2.5.9 TBE-PAGE 跑胶

1) 制备 TBE-PAGE 胶(8%, 15 mL)

组分	含量	
30% 丙烯酰胺/甲叉双丙烯酰胺(29:1)	4 mL	
10×TBE 缓冲液	1.5 mL	
ddH ₂ O补齐至	15 mL	
混匀后, 倒胶前加入		
10% APS	75 μL	
TEMED	7.5 μL	

2) 每孔上样约 100 ng, 160 V 跑 37 min。

3) 将胶取出放在~50 mL 自来水中,加入 5 μL 10000×GelRed 核酸染料,静置 3-5 min,于照胶仪中拍摄胶图。

三、研究结果

3.1 候选序列设计

参照 poly(A)结合蛋白(PABP)的指纹图谱,共设计了 16 种按照不同频率 掺入 G 的 poly(A)尾,并使用 100 nt poly(A)和 BioNTech 公司开发的 A₃₀LA₇₀尾 部^[9]作为阳性对照(图 1)。



图 1. 在 100 nt 的 poly(A)尾中按不同频率插入鸟苷酸 G

蓝色代表 A,粉色代表 G, An 表示该段 poly(A)共 n 个碱基,绿色代表 BioNTech 公司所用 10 bp 连接序列(GCATATGACT);两条阳性对照序列被高亮标出

3.2 质粒稳定性检验

3.2.1 简化版质粒稳定性检验实验方法

由于实验中候选质粒数量较多,采用常规的转化、挑单菌落、摇菌、抽提质 粒、酶切含 poly(A)片段、跑胶验证该菌落是否发生重组的方式(图 2A)来计算 每个质粒的重组率太过繁琐复杂,可操作性较差,我对实验过程进行了一定的简 化,以通过所获半定量的结果对各质粒重组率进行估测。

简化后的实验其余步骤不变,将第二步挑单菌落更换为将平板上半数菌落 (>100)刮下进入同一摇菌管中进行摇菌;而重组率的计算从

简化为灰度重组率

(上样 - 未上样)全条带灰度值 - (上样 - 未上样)未重组条带灰度值 × 100%

(上样 – 未上样)全条带灰度值

即通过胶图上各条带的灰度值对片段在质粒上的重组率进行测量。

3.2.2 宿主菌种筛选

为了让筛选出的 poly(A)尾能适于 mRNA 药物的工业化生产,我选择了 DH5α 和 JM109 这两种工业上最常用的质粒生产菌作为候选载体菌株。它们的 recA 与 endA 基因均被敲除,从而能最大程度地保证质粒在生产过程中的稳定性^[15]。其中, 相较 DH5α, JM109 曾被报导具有更低的重组率并更适宜用于重复性 DNA 的克隆^{[16],[17]}。

在对 DH5α 和 JM109 转化了含 100 nt poly(A)的质粒后,我分别从平板上挑 了 5,10,15 个菌进行重组率测定(图 2B-C)。灰度重组率随着摇菌数量的升高 而略有上升。同时,三组实验所得的灰度重组率并未出现较大的偏移,且与此前 报导的大肠杆菌中重组率结果相近^[5](图 2C),这一结果也论证了灰度重组率计 算的可行性。相较 DH5α~65%的灰度重组率,JM109的灰度重组率更低(~55%), 从胶图直观感受其重组条带也更淡(图 2B)。最终我选择了表现更为优异的 JM109 作为后续稳定性实验的菌株。



图 2. A100 在 JM109 中的重组率低于 DH5a

A. 质粒稳定性检验流程,常规方法将在第二步挑菌并通过胶图中重组质粒数量计算重组率, 简化版方法将在第二步刮板并通过胶图中重组条带相对全部条带的亮度,利用灰度值计算灰 度重组率; B. D 代表 DH5α, J 代表 JM109,数字代表共挑取几个菌落一同摇菌,右侧两泳 道为 DNA 指示条带与其对应长度,1-6 泳道最上方为未重组条带,下方为各重组条带;C. 根 据 B 图原始数据计算得到的各泳道灰度重组率,从左到右依次与条带相对应

3.2.3 掺入鸟苷酸的候选序列质粒稳定性检验

对所有候选质粒同时进行简化版稳定性检验,并计算其灰度重组率(图3)。 从 3-5 次独立实验所得的结果可以看出,13 个候选序列均表现出了灰度重组率降 低的现象,相比 A₁₀₀ 至少下降了 50%(图 3B)。其中(A₂₀G)₂A₃₀GA₃₀、(A₃₀G)₂A₄₀ 和 A₄₀GA₆₀ 三个候选序列表现出了与 A₃₀LA₇₀ 对照相近的较低重组率。而计算得 到的 A₁₀₀ 质粒灰度重组率接近 50%,与此前结果相符(图 2B),进一步证明了 灰度重组率实验方法的可行性。

实验发现,各质粒均发生了重组,但不同质粒的重组条带位置和亮度存在差 异(图 3A)。在全组实验中表现优异的(A₂₀G)₂A₃₀GA₃₀、(A₃₀G)₂A₄₀和 A₄₀GA₆₀ 三个候选序列,其重组条带亮度也最低。同时,与此前摇 5-15 个菌时出现几条 稳定重组带的情况(图 2B)不同,含 A₁₀₀的质粒在大量菌落摇菌后出现了一系 列长度不同的重组条带,暗示其重组过程发生频率高且稳定性差,从最低条带的 长度来看,直到仅剩约 30-40 bp 时重组过程才停止。这一长度与三个表现优秀的 候选序列 poly(A)片段长度相吻合。而间隔序列长度更短的 10 bp、12 bp、15 bp 的重组效率反而相对更高,可能是因为重复的 A_xG 序列(X=10, 12, 15)形成 了一个结构单元而整体发生重组。出于同样的原因,可以观察到(A₂₀G)₄A₂₀和 (A₂₅G)₃A₂₅ 候选序列的重组频率也相对更高。此外,BioNTech 公司发布的数据显 示,对于 120 nt 的 poly(A)序列,其中第 30-50 位碱基对重组最为敏感^[9],因此该 公司也选择在 30 位处截断 poly(A)插入连接序列。这也和实验中三条表现优异的 序列结果相符。

15



图 3. 掺入鸟苷酸 G 能够提高质粒稳定性

A. 对所有候选序列进行简化版稳定性检验所得的一张典型胶图,其最左侧泳道为 DNA 指示条带与其对应长度,右侧各泳道对应的候选序列在泳道上方标出;由于 A₁₀₀G, A₁₀₀GA, A₁₀₀GAG 序列均含有 100 bp poly(A),认为它们的重组率与 A₁₀₀相等; B. 根据 3-5 组独立实验得到的胶图计算所得灰度重组率,误差线为 SEM 值

3.3 IVT mRNA 功能检验

3.3.1 功能基因选择

为了能够对携带不同候选 poly(A)序列的 IVT mRNA 功能进行测试,以 EGFP 绿色荧光蛋白作为其中的功能基因。实验室原本使用的 EGFP 具有荧光强度高和 半衰期长的特点。为了能够更方便地获取表达量和半衰期数据,本实验从密码子 调整后的 EGFP 序列中,选择发光强度中等的作为后续实验的蓝本。

体外转录并纯化密码子调整后的 EGFP 序列 DC50-62 的 mRNA,转染细胞 并置于荧光显微镜下观察。在一系列调整序列中,DC54 和 DC56 表现出了相较 野生型 EGFP 更低的荧光强度,其中 DC54 的表现更为适中(图 4A)。流式细胞 术实验结果也表明 DC54 的荧光水平低于野生型(图 4B)。因此我选择了 DC54 作为功能基因进行后续的实验。



图 4. DC54 相较野生型 EGFP 表达水平适中

A. 野生型 EGFP(wtEGFP)与密码子调整后的系列 EGFP(DC50-52)的 IVT mRNA 在荧光显微镜下观察结果; IVT mRNA 使用 ARCA 加帽, 100 nt poly(A)尾,转染 293T 细胞后 12 h 观察荧光强度; B. DC54 密码子调整型和野生型 EGFP 的流式细胞术实验结果,纵轴为 相对荧光强度,转染后 16 h 测定

3.3.2 掺入鸟苷酸的候选序列 IVT mRNA 表达水平检验

由于细胞内 EGFP 蛋白的荧光强度由其含量决定,而蛋白含量由其翻译、成 熟和降解速率共同决定;对同种 EGFP 可认为其成熟和降解速率是常数,若进一 步假设不同组细胞内 IVT mRNA 含量一致,则可以通过测量某一时刻 EGFP 的 荧光强度来推断并比较其翻译速率。为了保证组间 IVT mRNA 的含量相对一致, 我选择了荧光强度最高的 16 h (未展示数据)对细胞内荧光强度进行测量。

将候选 poly(A)序列通过 PCR 连接到 DC54-EGFP 基因后方,经过体外转录 和 293T 细胞转染,在荧光显微镜下观察亮度并用流式细胞术测量其荧光强度, 以反映各候选序列对应的 mRNA 在细胞中的表达水平。在全部 16 条候选序列中, 10 条表现出与 A₃₀LA₇₀ 对照相近或更高的表达量,其中 A₅₀GA₅₀ 的表达水平甚至 略高于 A₁₀₀ 对照(图 5、6)。在此前表现优异的三条序列(A₂₀G)₂A₃₀GA₃₀、 (A₃₀G)₂A₄₀ 和 A₄₀GA₆₀ 的表达量均与 A₃₀LA₇₀ 对照基本一致,其中(A₃₀G)₂A₄₀ 和 A₄₀GA₆₀ 表达量和 A₁₀₀ 并无显著差异(t 检验, P>0.05)。

在三个表现优异的序列中,(A₃₀G)₂A₄₀综合表现最佳。在最重要的稳定性实验中,(A₃₀G)₂A₄₀始终拥有与A₃₀LA₇₀对照接近的低重组率(图3)。相较A₄₀GA₆₀,(A₃₀G)₂A₄₀中插入G的数量适中,各片段均在最重组敏感的30-50 bp处^[9];而相较(A₂₀G)₂A₃₀GA₃₀,其表达量与A₁₀₀更为接近,对IVT mRNA 正常功能的影响更小。最终综合考虑质粒稳定性和 IVT mRNA 的表达效率,我选择了性能最优的

(A₃₀G)₂A₄₀进行后续实验。



图 5. 荧光显微镜观察各候选序列荧光强度

在 DC54-EGFP 尾部添加各候选序列,制备 IVT mRNA,使用 ARCA 加帽,转染后 18 h 显 微镜下观察荧光;各图左上角依次标出了所用的 poly(A)尾组成,右下角为 50 μm 标尺



图 6.10/16 个候选序列具有与 A30LA70 对照接近或更高的表达效率

横轴为阴性对照(空转)、两阳性对照与各候选序列,纵轴为相对荧光强度;与图5使用同一批制备的 IVT mRNA,转染后 16 h 收集细胞,进行流式细胞术测定荧光强度;图中各组实验数据用点标出,方框顶端为4组独立实验荧光强度的平均值,误差线为 SD 值;虚线代表 A₃₀LA₇₀ 对照荧光强度的平均值,认为平均值接近或高于该虚线的候选序列具有与A₃₀LA₇₀ 接近或更高的表达效率

3.3.3 (A30G)2A40 随时间的表达量变化测定

从 3.3.2 首段论述可见,在 EGFP 蛋白的翻译效率一致的前提下,某一时间 点处的荧光强度取决于胞内 IVT mRNA 的数量。因为 IVT mRNA 的初始量一定, 其胞内含量由转染率和 mRNA 降解速率决定。其中前者可通过阳性细胞比例反 映,后者体现了该 mRNA 的功能半衰期。因此,可认为其在某一时间点处的荧 光强度取决于 IVT mRNA 的转染率和降解速率。

在分别转染了具有(A₃₀G)₂A₄₀和 A₃₀LA₇₀尾部的 DC54-EGFP 后的 0.5-96 h 分 别取样,通过流式细胞术测量细胞荧光强度,并计算各组中阳性细胞所占的比例。 在各时间点,(A₃₀G)₂A₄₀和 A₃₀LA₇₀对照的荧光强度和阳性细胞比例均基本一致 (图 7)。综合上文结果可见,(A₃₀G)₂A₄₀中鸟苷酸的掺入对 IVT mRNA 表达量 和半衰期均无显著影响。

随着转染时间延长,1h后随着 EGFP 的表达与成熟,细胞开始发出荧光; 随后细胞荧光强度逐步上升,在24h附近趋于峰值;之后逐渐下降。结合表达 荧光的阳性细胞比例可以看出,早期1-4h的荧光强度上升主要是由细胞逐渐接 收 IVT mRNA 并开始表达荧光,此阶段荧光强度和阳性细胞比例都基本呈线性 增长。中期 8-24h 荧光强度上升主要是由于 EGFP 表达效率达到最大和表达的 EGFP 逐渐积累,阳性细胞比例仍有增多但低于荧光强度增长幅度。后期24-96h 荧光强度开始减弱,主要是由于 IVT mRNA 被降解,细胞无法再生产新的 EGFP 蛋白,而此前生产的 EGFP 逐渐被降解或失去功能无法发光,因此虽然阳性细胞 比例仅略有减少,但荧光强度已经很低,趋近0.5h时的本底强度。



图 7. (A₃₀G)₂A₄₀ 尾部不影响对 IVT mRNA 的表达效率与半衰期 蓝线代表具有(A₃₀G)₂A₄₀ 尾部的 DC54-EGFP 体外转录 mRNA, 橙线代表 A₃₀LA₇₀ 对照; 主

图横轴为转染后时间,纵轴为相对荧光强度;副图横轴同样为转染后时间,纵轴为发出荧光的阳性细胞比例;转染后 0.5, 1, 2, 4, 8, 24, 48, 96 h 分别收集细胞,之后进行流式细胞术测定 其荧光强度;图上各点为三组独立实验数据的平均值,误差线代表 SD;使用的 IVT mRNA 与图 5、6 为同一批制备

四、讨 论

4.1 质粒稳定性实验的反思与展望

实验中使用了简化版质粒稳定性检验方法,通过灰度值测量计算细胞的灰度 重组率。这一方法虽然简化了实验流程,但同时也存在不足:

其一是各质粒条带灰度值测量的不精确性,在独立实验组间同一质粒的灰度 重组率结果可能出现较大差异,导致了较高的误差线(图 3B)。

其二是胶图只能显示出一定范围内条带的长度差异。尽管使用了分辨率更高的聚丙烯酰胺凝胶,但受限于所用凝胶的大小,对于 A₁₀₀ 这种高重组率、多重 组条带的质粒,会出现重组条带重叠现象(图 3A)。对于仅发生少数碱基丢失的 重组,也有重组条带和未重组条带难以分开的可能。

其三是胶图无法识别碱基突变。目前的实验中,对于在质粒复制过程中可能 发生的碱基突变未加考虑。

实验中另一个值得注意的点在于(A₁₀G)₉A₁₀的未重组条带始终高于其他候选 序列(图 3A),尽管其与旁边的(A₁₂G)₇A₁₆在长度上仅相差 2 bp。这可能是因为 非变性胶保留了 DNA 片段的二级结构,而(A₁₀G)₉A₁₀由于结构更为松散,体积 更大,导致其迁移速率更低。

尽管简化版质粒稳定性检验方法帮助我筛选出了三个表现优异的候选序列, 但对于最终决定用于后续 mRNA 药物研发的(A₃₀G)₂A₄₀ poly(A)尾,还需要对其 稳定性有更精细的实验数据。计划下一阶段将通过单菌落摇菌来计算常规重组率, 并通过测序获得少数碱基重组和突变的概率。

4.2 IVT mRNA 表达水平和半衰期实验反思与展望

实验仅对一种功能蛋白的表达水平进行了检验,其结果可能具有一定的偶然性。计划下一阶段将更换 IVT mRNA 中的功能蛋白以验证目前的实验结果。

实验通过表达量随时间的变化对 IVT mRNA 的功能半衰期进行了定性预测, 但这种间接的预测方式并不准确,所得结果可能存在其他的解读。计划后续实验 将通过 qPCR 来定量研究(A₃₀G)₂A₄₀ 尾对 IVT mRNA 半衰期的影响。

4.3 IVT mRNA 表达水平和半衰期与 poly(A)结合蛋白指纹相关

Poly(A)结合蛋白(PABP)共有 4 个 RNA 识别基序(RRM),其中靠近 N 端的 RRM 1 和 RRM 2 与 mRNA 的结合更为紧密且对腺苷酸选择性更强,RRM 4 则是 PABP 与其他蛋白相互作用的主要位点^[18]。此前的研究发现,在脱腺苷酶降 解 poly(A)的过程中,PABP 的四个 RRM 是逐步从 mRNA 上释放的,每个 RRM 结合~8 nt 的 poly(A)^[18]。同时,在蛋白翻译过程中,与 5'帽结合的 eIF4F 复合体 可以与 PABP 相互作用形成闭环结构,这一结构可以促进核糖体循环并调控翻译 的起始效率^[19]。

实验中观察到间隔 10 nt 或 12 nt 掺入 G 会导致 IVT mRNA 翻译效率降低, 而间隔 15 nt 的(A₁₅G)₅A₂₅则恢复了与 A₃₀LA₇₀对照相近的表达水平(图 6)。由 此推断, 15 nt 能够让 PABP 的 RRM 1 和 RRM 2 同时与 poly(A)互作,从而保证 PABP 与 mRNA 的稳定结合并发挥正常功能,这与上文的研究结果相符^[18]。而 12 nt 的 poly(A)则不足以保证 PABP 的稳定结合与功能发挥。同时,要让 PABP 的 4 个 RRM 都与 poly(A)结合,需要~30 nt A,这与含有不少于 30 nt poly(A)片 段的候选序列能表现出较好翻译效率的现象是一致的(图 6)。而在 A₁₀₀ 后添加 G/GA/GAG 的三条候选序列都表现出了相当低的表达水平,猜测其原因可能是 位于末尾的 G 影响了 poly(A)尾、PABP 与 eIF4F 复合体的相互作用,从而导致 低翻译效率;或者该序列可能激活了胞质中某些单链 RNA 识别蛋白,如 NOD2^[10] 或 TLR7^[20],引发的免疫反应导致了 IVT mRNA 的降解和翻译抑制。

此外,有研究推测 PABP 因为与 eIF4F 复合体和 eRF3 蛋白相互作用,占据 了其与脱腺苷酶 CCR4-NOT 复合体相互作用的位点,让最前端的这部分 poly(A) 序列免受脱腺苷酶的降解,从而使 mRNA 保持着一段较短的 poly(A)尾并持续翻 译^[21]。因此,尽管实验所用的候选序列中添加了 G,在 PABP 能够与最前端 poly(A) 稳定结合的情况下,(A₃₀G)₂A₄₀ 尾的半衰期不变(图 7)。

4.4 总结

本实验在 poly(A)序列中间隔性掺入单碱基鸟苷酸 G,通过测定 16 种候选序 列的质粒稳定性和 IVT mRNA 表达水平,选出了表现最佳的(A₃₀G)₂A₄₀ 尾。该序 列具有高质粒稳定性和与 A₃₀LA₇₀ 对照一致的 IVT mRNA 翻译效率和半衰期,能 够用于后续的 mRNA 药物研发,为 mRNA 药物的大规模工业化生产创造条件。

参考文献

- 1. Polack, F. P. et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* (2020)383, 2603–2615.
- Baden, L. R. et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N. Engl. J. Med. (2021)384, 403–416.
- 3. Schlake, T., Thess, A., Thran, M. & Jordan, I. mRNA as novel technology for passive immunotherapy. *Cell. Mol. Life Sci. CMLS* (2019)76, 301–328 (2019).
- 4. Berraondo, P., Martini, P. G. V., Avila, M. A. & Fontanellas, A. Messenger RNA therapy for rare genetic metabolic diseases. *Gut* (2019)68, 1323–1330.
- Trepotec, Z., Geiger, J., Plank, C., Aneja, M. K. & Rudolph, C. Segmented poly(A) tails significantly reduce recombination of plasmid DNA without affecting mRNA translation efficiency or half-life. *RNA N. Y. N* (2019)25, 507–518.
- Silva, F., Queiroz, J. A. & Domingues, F. C. Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. *Biotechnol. Adv.* (2012)30, 691–708.
- Oliveira, P. H., Prather, K. J., Prazeres, D. M. F. & Monteiro, G. A. Structural instability of plasmid biopharmaceuticals: challenges and implications. *Trends Biotechnol.* (2009)27, 503–511.
- Grier, A. E. *et al.* pEVL: A Linear Plasmid for Generating mRNA IVT Templates With Extended Encoded Poly(A) Sequences. *Mol. Ther. - Nucleic Acids* (2016)5, e306.
- 9. EBERLE, F., Sahin, U., Kuhn, A., Vallazza, B. & Diken, M. Stabilization of poly(A) sequence encoding DNA sequences. (2020).
- Linares-Fernández, S., Lacroix, C., Exposito, J.-Y. & Verrier, B. Tailoring mRNA Vaccine to Balance Innate/Adaptive Immune Response. *Trends Mol. Med.* (2020)26, 311–323.
- 11. Yu, S. & Kim, V. N. A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat. Rev. Mol. Cell Biol.* (2020)21, 542–556.
- 12. Lim, J. et al. Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* (2018)361, 701–704.
- Nicholson, A. L. & Pasquinelli, A. E. Tales of Detailed Poly(A) Tails. *Trends Cell Biol.* (2019)29, 191–200.
- Tang, T. T. L., Stowell, J. A. W., Hill, C. H. & Passmore, L. A. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nat. Struct. Mol. Biol.* (2019)26, 433–442.
- Phue, J.-N., Lee, S. J., Trinh, L. & Shiloach, J. Modified Escherichia coli B (BL21), a superior producer of plasmid DNA compared with Escherichia coli K (DH5alpha). *Biotechnol. Bioeng.* (2008)101, 831–836.
- Ribeiro, S. C., Oliveira, P. H., Prazeres, D. M. F. & Monteiro, G. A. High Frequency Plasmid Recombination Mediated by 28 bp Direct Repeats. *Mol. Biotechnol.* (2008)40, 252.
- 17. Ferenc, M. Plasmids 101: Common Lab E. coli Strains.

https://blog.addgene.org/plasmids-101-common-lab-e-coli-strains.

- Webster, M. W. et al. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol. Cell* (2018)70, 1089-1100.e8.
- 19. Tomek, W. & Wollenhaupt, K. The "closed loop model" in controlling mRNA translation during development. *Anim. Reprod. Sci.* (2012)134, 2–8.
- Karikó, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* (2005)23, 165–175.
- 21. Lima, S. A. et al. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.* (2017)24, 1057–1063.