

*Spatially resolved, highly
multiplexed RNA profiling in
single cells*

整理：李月、王雅文、周洋如
ppt：李月、王雅文
展示：李月、周洋如

Background

为什么要进行细胞内RNA分子的空间定位？

1. RNA的拷贝数以及其在细胞中的分布对基因表达有很重要的调节作用，测量这些特性对转录水平的基因表达和调控有很重要的意义。

2. mRNA的数量和mRNA分子在细胞中的位置影响着细胞的行为，以及细胞中的蛋白数量和类型。

3. 系统性分析单细胞的RNA丰度和空间定位，有助于我们理解细胞和发育生物学的许多方面。

4. 进行这项工作也是对基于分离细胞RNA测序的单细胞方法的一个有力补充。单细胞测序只是知道了序列信息，但是不清楚这些RNA在细胞内的空间信息。这项工作保护一个细胞内的RNA空间布局，这对于许多生物学问题的解决来说，是非常重要的。

Background

目前已有的空间定位技术：

单分子荧光原位杂交（smFISH）：是在单细胞中研究RNA拷贝数和空间定位的有力武器。

将DNA（或RNA）探针用特殊的核苷酸分子标记，然后将探针直接杂交到染色体或DNA纤维切片上，再用与荧光素分子偶联的单克隆抗体与探针分子特异性结合来检测DNA序列在染色体或DNA纤维切片上的定性、定位、相对定量分析。传统的smFISH利用一系列短的寡核苷酸来检测转录本，每个寡核苷酸上标记有一个荧光基团。这样，mRNA上就有足够浓度的荧光分子，产生可见的光点，但通常只有在60x或100x放大倍数下，使用油浸、大数值孔径的透镜才行。

Background

多路复用抗误码的出现

多路复用抗误码(multiplexed error-robust fluorescence in situ hybridization, MERFISH): 针对smFISH缺点期望作出相应改进。可以同时在各自细胞内成像的RNA种类数是有限的。这使以空间解析方法形成单个细胞的转录组分析很困难。期望在单细胞中能监测到更多的RNA种类, 并监测和纠正错误。

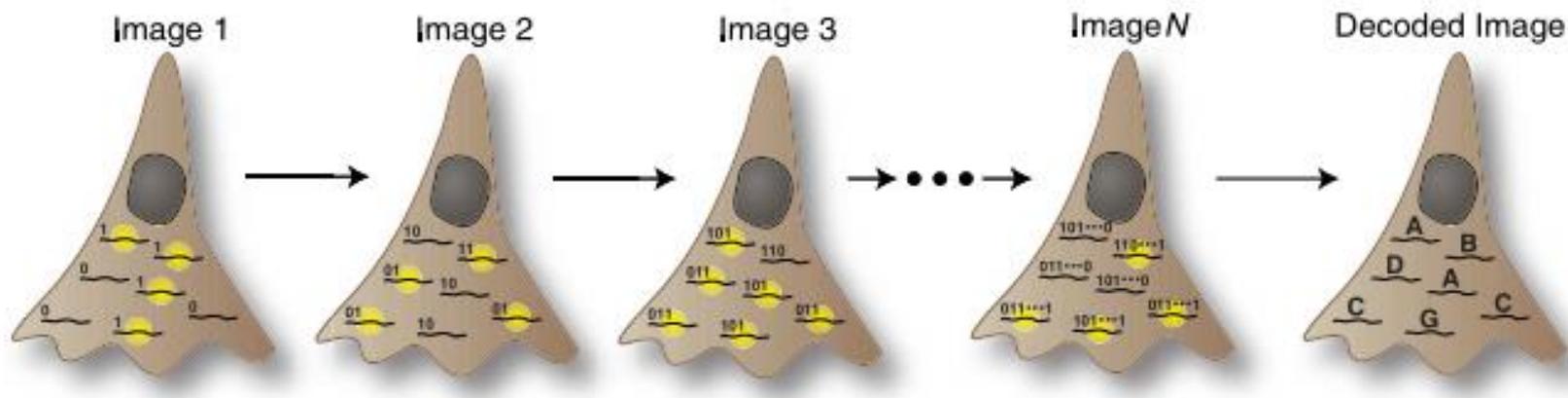
多路复用采用编码的方式相对于smFISH测得RNA种类多的原因可能是对于同样多的readout序列, 一个readout序列标记一种RNA的话, 能测得RNA种类少, 若将这些readout序列分四个一组的组合起来话可以有大约可以测 $n \times n-1 \times n-2 \times n-3 / 24$ 种RNA。

Introduction

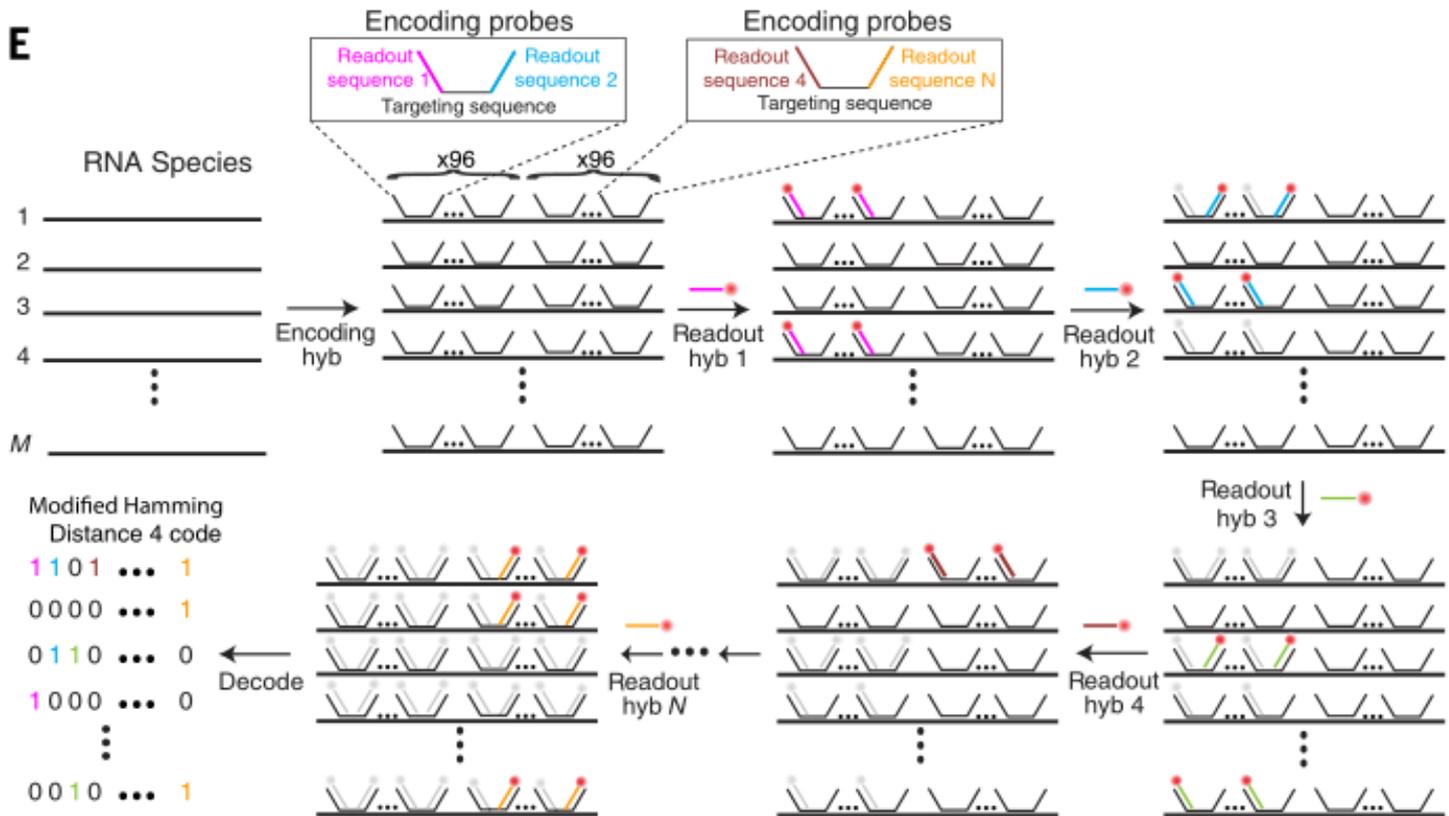
MERFISH的基本原理：

- 1.用一组编码探针标记目标RNA。使用连续轮杂交和成像,每轮都有不同的读出探针,以确定读出序列绑定到每个RNA。
- 2.原则上,组合标签允许检测RNA种类的数量成倍增长的成像,但检测错误也成倍增加。应对这样的积累误差,我们利用error-robust编码方案来设计我们编码探针,解释荧光标签测量的误差特性。

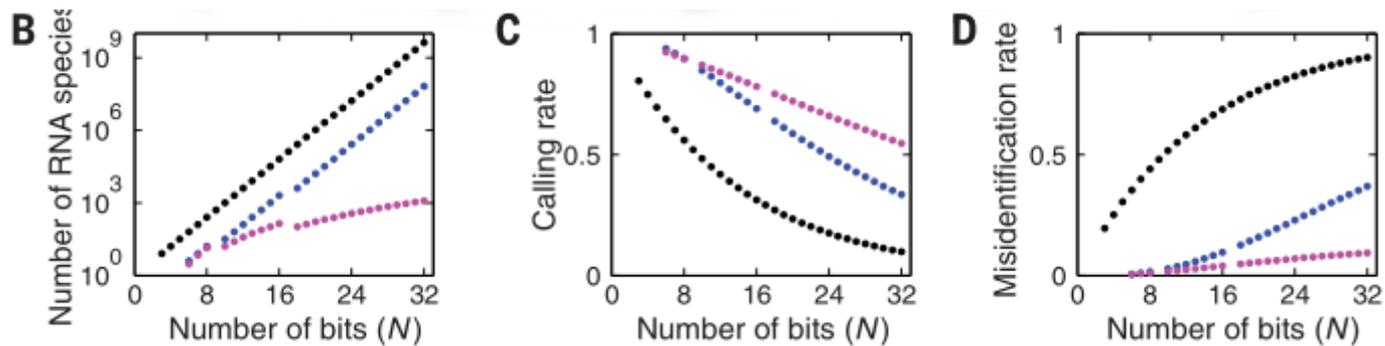
重要实验步骤:



编码探针



多轮杂交



(B) The number of addressable RNA species

(C) the rate at which these RNAs are properly identified—the “calling rate”

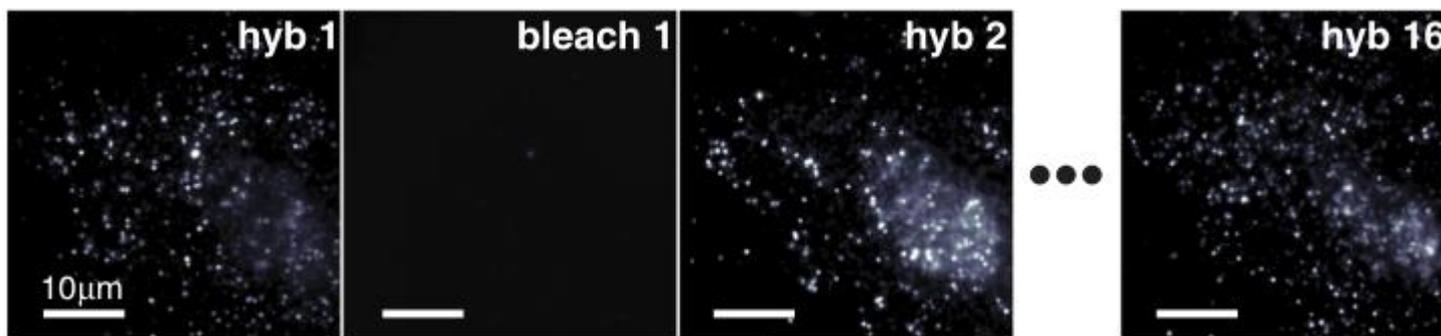
(D) and the rate at which RNAs are incorrectly identified as a different RNA species—the “misidentification rate”

Black indicates a simple binary code that includes all 2^N-1 possible binary words. Blue indicates the HD4 code in which the Hamming distance separating words is 4. Purple indicates a modified HD4 (MHD4) code where the number of 1 bits are kept at four.

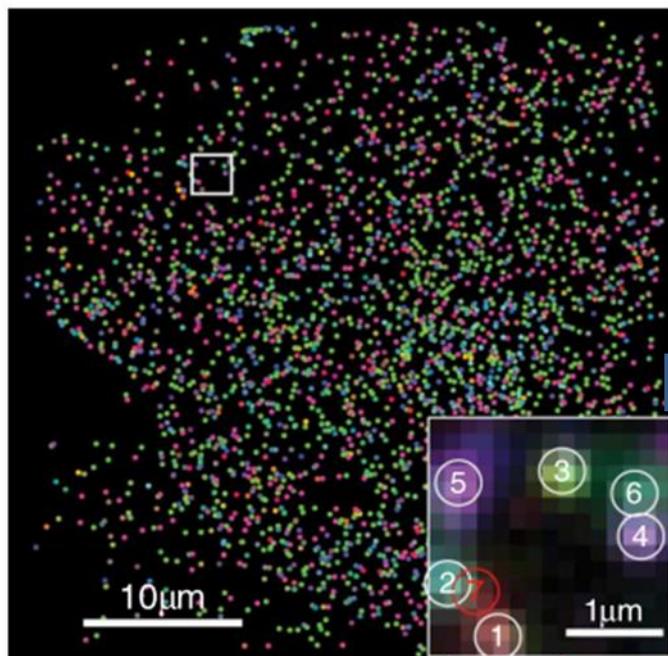
B图横坐标表示字节数，纵坐标表示所能测得的RNA种类的数目，C图横坐标依然是字节数，纵坐标表示的是正确识别率，D图的横坐标是字节数，纵坐标表示的是错误识别率。

图中黑色颜色是简单的二进制编码方式，蓝色表示汉明编码方式，紫色是经过改进的汉明编码方式。

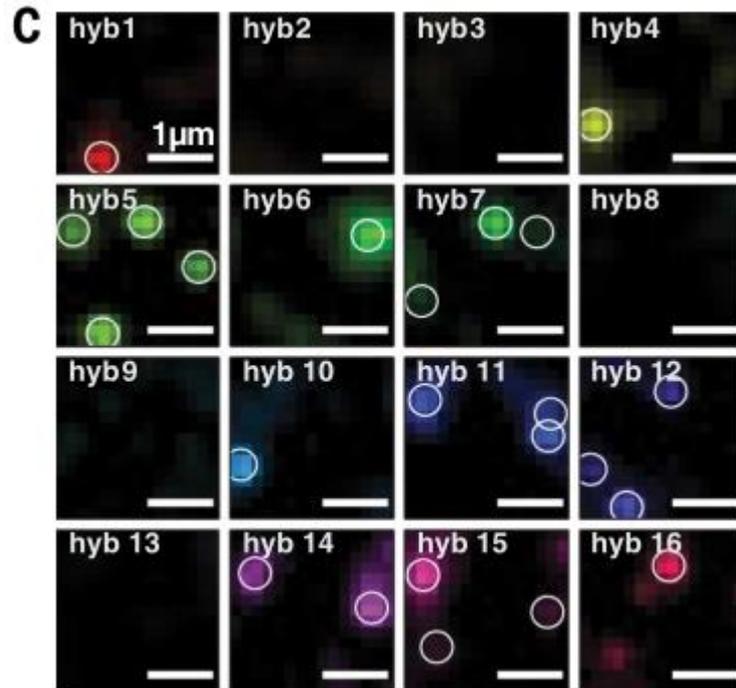
成像结果:



bleach1图表示的是在每轮杂交过程中经过漂白后的图像。



此图是经过多轮杂交后得到的，其中的颜色都是经过叠加的，不是本身探针荧光的颜色。



D

Spot number	Hybridization round																Gene
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1	1	0	0	0	1	0	0	0	0	0	0	1	0	0 1	0	0	DYNC1H1
2	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0	0	EGFR
3	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	FLNA
4	0	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	TLN1
5	0	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	TLN1
6	0	0	0	0	0	1	1	0	0	0	1	0	0	0 1	0	0	LRP1
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	unidentified

(C) Fluorescent images from each round of hybridization for the boxed subregion in (B), with circles indicating potential RNA molecules. (D) Corresponding words for the spots identified in (C). Red crosses represent the corrected bits.

两种不同编码方式的比较：

two different error-robust encoding schemes

	速度	正确识别率	错误识别率	纠错机制
MHD4	较慢	较高	较低	识别纠正错误
MHD2	较快	较低	较高	只识别错误

应用：

- 高通量分析细胞之间基因表达的不同
- 分析不同基因间的共同变异
- RNA空间定位

Discussion

- 使用我们目前的成像和分析方法，经过32轮杂交后，允许表达程度前10%的基因同时成像。通过使用更先进的图像分析算法，能允许几乎整个转录组成像。但是，根据理论预测，RNA在密集压缩结构仍然可能不被检测出来。

Discussion

- 因为MHD4和MHD2的错误率随着比特的增加而增长的速度较慢，所以可以使用这两种编码方式中的任意一个来增加编码的比特数，可以进一步提高检测的RNA种类的数目。但是通过增加杂交次数来增加比特数会加长数据收集的时间，以及造成样本的退化。这个问题可以通过在每一轮杂交过程中使用多种颜色读出多个比特来解决。

Thank You!