

手把手教你做 MeRIP-qPCR 并计算相对表达量

本文中，我们会为老师详细解析在做完 m⁶A 测序后，如何对差异甲基化的基因进行验证。简单来说就是 IP 后如何进行 qPCR 实验。文中的表述：MeRIP-qPCR 与 m⁶A-IP-qPCR 等同。

【确定样本起始量和 IgG】

在进行讲解之前，我们要回答的第一个问题就是，用多少 RNA 来做后期验证呢？举个例子，假如要验证 10 个基因，究竟需要多少 total RNA 或 polyA 的 RNA 才合适呢？

我们先进行一个简单的计算，常规的 RT-qPCR 1 个反应（1 个复孔）需要的 total RNA 至少 0.1-1μg 左右。我们按照最小反应起始量来算的话，那么 1 个反应需要至少 100ng 的 total RNA，3 个技术重复（3 个复孔），总计 total RNA 为 300ng。

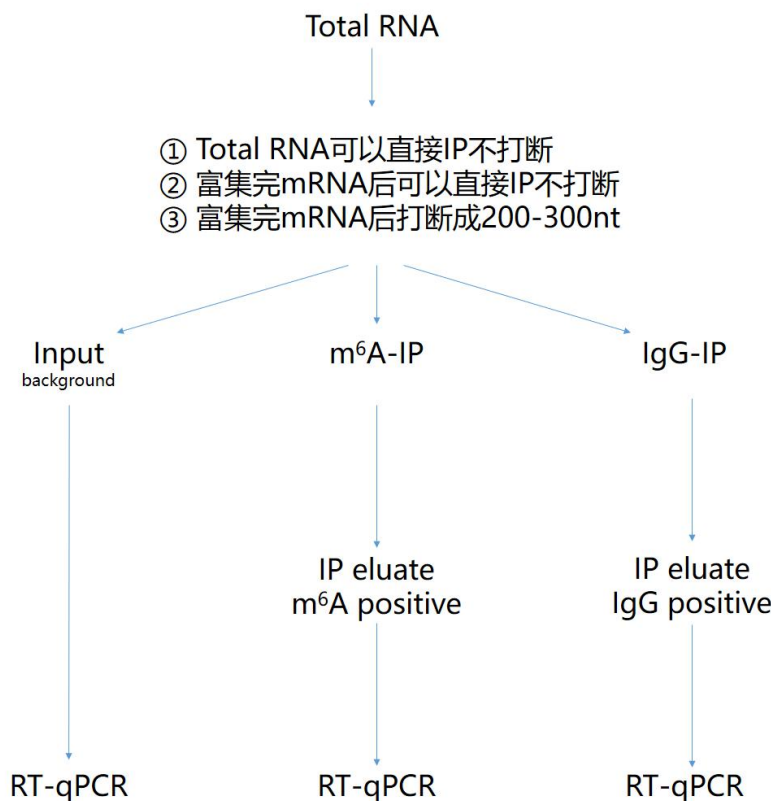
我们知道 total RNA 中绝大部分为 rRNA，那么只有不到 5% 的 RNA 为我们需要研究的对象。PCR 反应需要引物跟目标序列区域进行充分碰撞，那么如果只使用 mRNA 会大大降低一次 PCR 反应的起始量。通过计算我们认为 5-10ng 做 1 次 PCR 反应较为合理。也就是 1 个样本中验证 1 个基因不做生物学重复只做 3 个复孔需要至少 15ng 的 mRNA 或 300ng 的 total RNA。那么 3 个生物学重复就需要 9 个反应，总计 45-90ng mRNA 或 900ng 的 total RNA。

根据上面的推算，IP 下来的 mRNA 若有 500ng，可以验证 6-12 个基因左右（1 个基因 3 个生物学重复，每个重复 3 个复孔即技术重复，所以一个基因总计 9 次反应）。哺乳动物 IP 效率在 10-20% 左右，植物在 20-40% 左右（最高可达 50%）。那么假设要验证 6-10 个基因，哺乳动物需要至少 2.5-5μg 的 mRNA 和 100-200μg 的 total RNA，植物样本至少需要 1.25-2.5 μg 的 mRNA 和 50-100μg 的 total RNA。若加入 IgG 抗体这个步骤，则需要对 total RNA 的起始量再次翻倍。

但是如果 total RNA 太少该怎么处理呢？没关系，可以直接用 m⁶A 抗体对 total RNA 进行 IP，而且不用打断。

那么接下来在进行讲解前的第二个问题就是，明明我们做高通量测序只有一个 input 文库和一个 IP 文库，为何我们做 MeRIP-qPCR 要加入 IgG 抗体呢（如下图所示）？；

那是因为高通量测序可以对所有被 m⁶A 抗体富集下来的 RNA 进行全转录层面的扫描分析，一旦 m⁶A 抗体有非特异性富集，在数据呈现上我们可以明显发现 reads 的分布比较乱且没有规律，而特异性富集则 reads 几乎富集在哺乳动物的 3' UTR 区和植物的 5' UTR 和 3' UTR 区。所以这就是测序不需要 IgG 文库，而我们做 m⁶A-IP-qPCR 需要 IgG 富集的原因。



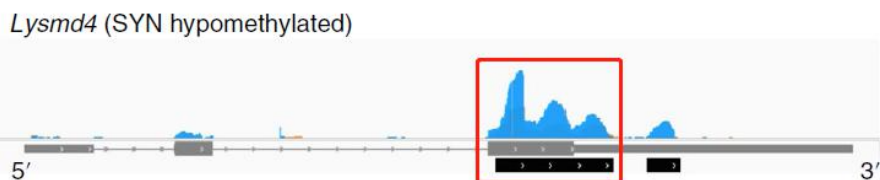
搞定了样本起始量和 IgG 的问题后，我们具体来看下步骤：

【步骤详解】

第一步，先对 RNA 进行特异性富集。至于打断或不打断，都是可行的。假如要打断的话请在打断之前，确认研究对象只是 mRNA 还是 lncRNA。若只研究 mRNA 可以先用 oligodT 磁珠进行富集后再进行打断。如果是 lncRNA 和 mRNA 都要研究，则需要先用试剂盒将 total RNA 的 rRNA 先消化干净，再进行打断。如果是进行高通量测序打断长度约为 100nt 左右，而 qPCR 则需要长度至少在 200nt 以上，部分论文甚至打断长度在 300-500nt 左右甚至都不打断。一般根据文献中的资料显示，IP 下来的片段用引物扩增出来的长度在 130-150 左右，所以 100nt 不到的长度要设计出合适的引物会非常困难。除非是用 5'RACE 扩增，否则还是建议大于 200nt。当然了，如果不打断的话设计引物就会比较方便。当然如果 total RNA 本身量比较少，建议老师直接用 m⁶A 抗体对 total RNA 进行富集。

第二步，对 RNA 进行抗体孵育和特异性富集。打断完了 RNA 我们会得到长度约为 200-300 的 RNA 片段。如果是不打断则是较为完整的 RNA 全长。这时候我们可以分出约占 m⁶A-IP 部分 RNA 只有 10% 的 RNA 作为 input。当然如果是细胞样本比较充足，Input 和 IP 的比例可以直接是 1:1。如果按照 1:1:1，简单来说用于 m⁶A IP 的 RNA 有 3μg，那么 Input、IP 和 IgG 的 RNA 都在 1μg 左右（无论是 total RNA 还是 mRNA）。

第三步，洗脱和逆转录扩增。接下来我们需要对 m⁶A 抗体和 IgG 抗体上洗脱下来的 RNA，以及 input 的 RNA，按照常规试剂盒要求进行洗脱，并用**随机引物**进行逆转录。



第四步，设计 MeRIP-qPCR 的特异性引物。由于 m⁶A 甲基化基本富集在 UTR 区和转录终止区，所以如上图所示在这些区域附近我们可以从测序结果里有明显的 peak 峰。那么我们设计引物可以考虑在基因高甲基化区域来进行。如果某几个常见的基因在不同文献里频繁出现，那么文献里列出来的引物可以被我们直接拿来引用。通常最后 PCR 产物长度最低会在 130-150bp 左右。这里需要注意的是，IP 下来的 RNA 需要用随机引物进行逆转录，而第四步中设计的引物用的是**特异性引物**，这个大家要注意区分和甄别。

第五步，使用上一步设计好的引物，对 input、m⁶A-IP 和 IgG-IP 中的 RNA 进行 qPCR 反应并计算相应的 CT 值。

所以到这里我们的实验部分已经完成了，简单总结下就是想要做够 10 个左右基因的验证，必须至少提供 >200μg 的 total RNA 以上。富集的 RNA 类型根据自己要求出发，如 polyA RNA 或 rRNA-depleted RNA 等，最后就是打断成 200-300nt 左右的长度。用于 qPCR 的 RNA 一共分成 3 类，即 Input、m⁶A-IP 和 IgG-IP，其中 input 占 m⁶A-IP 起始量约 10%。如果用数字来表示那就是 input 需要 100ng，m⁶A 在 IP 前需要 1μg，IgG 在 IP 前也需要 1μg。如果 IP:Input:IgG=1:1:1，则不需要做特别复杂的换算。部分课题组甚至会舍弃 IgG，只有 IP 和 Input，分别为 1:1。

所以接下来我们就要详细讨论下 m⁶A-IP-qPCR 是如何计算相对表达量的。MeRIP-qPCR 是 m⁶A-IP-qPCR 的另一种说法。所以从名称上我们就能发现，m⁶A-IP-qPCR 基本和常规的 RIP-qPCR 相似。当然在讨论 RIP-qPCR 之前，我们先来看看常规的 RT-qPCR 是如何计算相对表达量的。

【基于传统 RT-qPCR 原理 MeRIP-qPCR 解读】

关于 RT-qPCR 的方法学论文很多，所以这里关于实验部分就不细讲了。计算基因的绝对定量需要加入外参基因，所以暂且我们先来看看相对表达量，也就是 RT-qPCR 传统的 2^{-ΔΔCt} 法是如何计算的。

假设验证一个基因，那么 3 个生物学重复 (biological replication) 是必须的，每个生物学重复要做 3 个复孔，我们也称之为技术重复 (technical replication)。那么一个基因总共要做 9 个反应。

每 1 个样本的 1 个基因做 3 个复孔，内参基因和目的基因的 Ct 值分别取均值，最后得到 Ct 内参和 Ct 目的。如哺乳动物会选择 GAPDH 作为内参基因。其公式为：

$$\Delta Ct = Ct_{\text{目的}} - Ct_{\text{内参}}$$

处理组和对照组的所有样本都要这样计算。例如，处理组和对照组分别有 3 个和 3 个样本（即 3 个生物学重复），你会分别得到处理组的 3 个 ΔCt 和对照组的 3 个 ΔCt 。

然后用 t 检验下 3 个处理组样本的 ΔCt 和 3 个对照组样本 ΔCt 之间是否有统计学差异。将处理组的 3 个 ΔCt 取均值得到 $\Delta Ct_{\text{处理组}}$ ，将对照组的 3 个 ΔCt 取均值得到 $\Delta Ct_{\text{对照组}}$ 。最后来计算 $\Delta\Delta Ct$ 和相对表达量（即 2 的 $-\Delta\Delta Ct$ 次方），公式分别为：

$$\Delta\Delta Ct = \Delta Ct_{\text{处理组}} - \Delta Ct_{\text{对照组}}$$

$$\text{相对表达量} = 2^{(-\Delta\Delta Ct)}$$

也就是处理组中某基因的表达量是对照组中表达量的 $2^{-\Delta\Delta Ct}$ 倍。我们假设 $\Delta Ct_{\text{处理组}}$ 和 $\Delta Ct_{\text{对照组}}$ 分别是 7 和 10，那么 $\Delta\Delta Ct = -3$ ，最后相对表达量为 8（2 的 3 次方）。

我们从传统的 RT-qPCR 会发现，每个样本都要做一个内参基因，如上面提到的哺乳动物会选择 GAPDH，而 miRNA 则会选择 U6。那么我们的 m⁶A-IP-qPCR 是不是也需要内参基因，根据我们查阅的大量文献来看，绝大部分是没有的，少量文献也会使用到内参基因或者我们称之为 negative control gene。

那么是不是 m⁶A-IP-qPCR 就真的没内参了呢？其实也不是！实际上我们会发现整个 input 充当了内参基因的概念，一个基因的甲基化水平是无法直接用 RIP 下来的 RNA 在做 qPCR 后直接用 Ct 值高低来衡量的，需要 input 中的 RNA 作为背景。而 IgG 抗体的加入则是为了排除 m⁶A 抗体非特异性结合情况发生。由于有时候 m⁶A 抗体不同厂家、保质期、蛋白性能等因素可能会产生非特异性结合的结果，这时候就需要用 IgG 来排除这部分的干扰。目前市面上使用的绝大部分 m⁶A 抗体为兔抗，那么尽量在使用 IgG 做对照时也使用兔抗来源的 IgG。

那么我们首先依旧是与传统的 RT-qPCR 相似，用相应的软件实时监测扩增曲线。溶解曲线必须是单峰，以保证引物特异性扩增。然后就是校准每一个样本的平均 Ct 值，以消除样本准备过程中的差异。

接下来我们就要开始对每个样本中，RIP 和 input 的 3 个技术重复得到的 Ct 值计算平均值，再对 3 个生物学重复样本中的 Ct 平均值再去计算其 Ct 平均值，得到 Average Ct_{RIP} 和 Average Ct_{input}。最后计算稀释系数，即 input dilution factor，也就是 input 的 RNA 占 m⁶A-IP 的 RNA 有多少比例。根据上面的示例，也就是 input RNA=100ng 而 m⁶A-IP RNA=1μg，那此处的稀释系数=10%。最后根据这些数值，我们可以计算出归一化后的 ΔCt 值，也叫 $\Delta Ct_{\text{normalized RIP}}$ 。

$$\Delta Ct_{\text{normalized RIP}} = [\text{Average Ct}_{\text{RIP}} - \text{Average Ct}_{\text{input}} - \log_2(\text{input dilution factor})]$$

这里为什么归一化后的 $\Delta Ct_{\text{normalized RIP}}$ 还要减去 $\log_2(\text{input dilution factor})$ 呢？那是因为需要减去 input 的噪音，排除其干扰。由于 RNA 样本特别稀少，所以在做实验的时候往往会把 input 的使用量按照一定比例减少。如本文中 input 的占比就只有 10%，那么稀释倍数 (Input Dilution Factor) 就是 0.1。最极端的情况，就是 RNA 十分的富余，所以当 input 的比例理论上和 m⁶A-IP 及 IgG-IP 的比例一样时，那么 $\log_2(\text{input dilution factor}) = 0$ ，也就无需相减了。所以加入稀释系数的校正，就是为了达到 input:m⁶A-IP:IgG-IP=1:1:1 的效果。同样如果前期 IP 和 Input 比例为 1:1，则不需要考虑这些复杂的换算了。

下一步就是要计算在 RIP 下来的 RNA 中某一个基因占 input 中这个基因有多少的百分比。换句话说就是有 input 中的某个基因究竟有百分之多少发生了甲基化，所以这里我们用 %input 来表示。基本上到这一步，部分老师不想做 IgG 验证的，会把这个数据直接放在论文里，但是我们建议还是要加入 IgG 的部分。当然这一步需要进行线性归一化处理，方法与上面的 RT-qPCR 类似，其公式为：

$$\%input = 2^{(-\Delta Ct_{\text{normalized RIP}})}$$

由于无法判断 m⁶A 抗体是否存在非特异性富集的情况，所以要过滤这些噪音还需要使用兔抗 IgG 再对 RNA 进行一次富集。这时候就需要对 ΔCt 再次进行背景去除。其中需要检测的目的基因在 IgG 中的 ΔCt 值就是 $\Delta Ct_{\text{negative control}}$ 。其公式为：

$$\Delta Ct_{\text{normalized RIP/negative control}} = \Delta Ct_{\text{normalized RIP}} - \Delta Ct_{\text{negative control}}$$

最后一步就是最激动人心的一刻，那就是计算甲基化富集倍数，也叫 Fold Enrichment，其公式为：

$$\text{Fold Enrichment} = 2^{(\Delta\text{Ct}_{\text{normalized RIP/negative control}})}$$

那么不加 IgG 可以直接将结果用于论文中的数据呈现吗？其实在部分的高分文章中我们也发现不用 IgG 直接进行验证的，即用 %input 数据直接用于表示甲基化比例。但是为了保险起见，我们还是强烈推荐老师加入 IgG 组的数据用于后期校正。

那么还有没有更暴力的方法呢？答案是肯定的！但是前提条件就是 Input 和 IP 比例保持一致即 1:1，且前期直接对 total RNA 进行富集（不需要额外富集一次 mRNA 或 lncRNA），使用以下简便的计算公式：

$$\text{Control: } \Delta\text{CT} = \text{CT}(\text{IP}) - \text{CT}(\text{Input})$$

$$\text{Treatment: } \Delta\text{CT} = \text{CT}(\text{IP}) - \text{CT}(\text{Input});$$

$$2^{(\Delta\text{CT Treatment} - \Delta\text{CT Control})}$$

你 Get 到了 MeRIP-qPCR 正确的打开方式了吗？

本文方法参考来自以下几篇文献，若要了解详细的方法，老师可以自行下载对应文献查看。

注意事项：

1. m⁶A 抗体富集这部分的 protocol，请参考 m⁶A 研究方法的权威，m⁶A-seq 的发明人以色列特拉维夫大学医学院的 Gideon Rechavi 教授发表在 Nature Protocols (Pubmed id: 23288318) 的文章。这里面会有非常详细的步骤教你如何做 m⁶A 抗体去 IP 带有甲基化修饰的 RNA。到时候注意打断长度为 200-300nt 即可，打断需要做几次条件实验。也就是说除了打断长度转换成 200-300nt，富集到的 RNA 从上机测序变为 PCR 外，其他所有步骤全部一样。当然如果没有打断条件，可以直接对 total RNA 进行富集且不打断。文章中的生物信息部分可以忽略不计；
2. 在实际操作中，也有人试验过在不打断 RNA 的情况下进行 m⁶A IP，所得到的结果也符合要求。若要不打断，则对 m⁶A 抗体质量要求较高，建议选择 Synaptic Systems m⁶A antibody No.202003 同类级别的抗体，用 DEPC 水稀释；
3. IP 下来的产物可使用体积比为 25:24:1 的酚：氯仿：异戊醇抽提 RNA，也可直接加入 Trizol 试剂进行提取；
4. 实验过程中注意在 RNase free 的实验台上操作；
5. 订购以上提到的相关产品，请+VX:hugasis。

参考文献：

- Castellanosrubio, A, *et al.* A novel RT-qPCR-based assay for the relative quantification of residue specific m⁶A RNA methylation. **Scientific Reports** 9.1 (2019).
- Weng, HY., *et al.* METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m⁶A Modification. **Cell Stem Cell** 22.2 (2017).
- Ma, CH., *et al.* RNA m⁶A methylation participates in regulation of postnatal development of the mouse cerebellum. **Genome Biology** 19.1 (2018): 68.
- Liu, J., *et al.* m⁶A mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. **Nature Cell Biology** 20 (2018): 1074
- Weng, YL., *et al.* Epitranscriptomic m⁶A Regulation of Axon Regeneration in the Adult Mammalian Nervous System. **Neuron** 97.2 (2018): 313.
- Gagliardi, M., *et al.* RIP: RNA Immunoprecipitation. **Methods in Molecular Biology** 1480 (2016): 73.
- Dominissini, D, *et al.* Transcriptome-wide mapping of N⁶-methyladenosine by m⁶A-seq based on immunocapturing and massively parallel sequencing. **Nature Protocols** 8.1 (2013): 176-189.