

用于基因编辑工作流程的基因分析工具

主要内容:

- 为什么基因组编辑的效率如此重要
- 不同的基因分析方法及其在评估基因组编辑效率方面的特定优势
- Thermo Fisher Scientific针对基因组编辑实验的完整工作流程方案

简介

基因组编辑技术有望彻底颠覆我们对生物系统多样性的理解和控制，由此先后催生了许多不同的系统来精确改变基因位点的序列，例如CRISPR-Cas9、TALEN和锌指核酸酶等。CRISPR-Cas9被认为是最简单、最精确和用途最广泛的基因组编辑技术，它是一种基于细菌适应性免疫应答的系统。它由两个部分组成，首先通过Cas9酶切割双链DNA，然后利用向导RNA (gRNA) 将该酶引导至基因组内特定位置。双链断裂可通过多种机制来修复。在缺少修复模板的情况下，由Cas9酶所产生的断裂将会通过一种较易出错的非同源末端连接过程来修复，导致在gRNA限定的断裂周围具有不同插入或缺失 (indel) 的异质细胞群。该过程可用来产生在特定序列周围具有随机缺失的细胞系，从而实现功能性敲除。而在修复模板存在的情况下，则可通过无错的同源重组在基因组内特定基因位点处引入用户自定义的序列改变。该过程可用于过表达新基因、创建与疾病相关的细胞模型，以及使用可报告基因标记内源基因

在任何基因组编辑实验中，修复的过程都不是完全有效或准确的。因此，为了实现具有克隆群体均一性的理想基因组编辑，需要对初级转化细胞群的单个克隆进行筛选。这需要进行两轮分析。首先，必须执行初次筛选以确定带有编辑的细胞的相对比例。了解编辑的效率将能确定那些需要进行单独扩增的单细胞克隆的数量。对于某些实验案例，如果效率足够高，则可能不需要进行二次纯细胞克隆的筛选。例如，如果研究涉及分析具有特定基因功能性丧失的细胞的反应，则可能不需要具有单一敲除的同质性群体。而对绝大多数细胞具有某种功能丧失敲除的混合细胞群进行分析可能也能提供所需的实验模型。

因此，在对工程细胞进行克隆分离之前，应确定转染细胞群中编辑效率的百分比。根据模型系统的不同，相应的转染效率将决定初次筛选的细胞群是否可以用作模型，或者确定需要检查的单细胞的大致数量，以便在二次筛选中寻找所需的编辑基因。无论如何，对于编辑效率的计算都是任何基因组编辑工作流程中的关键步骤。

赛默飞工作流程和工具简介

我们可提供用于基因组编辑实验的完整工作流程，包括靶标特异性编辑工具设计、合成及转运试剂以及用于确认和验证基因修饰的一系列分析工具（图1）。

选择gRNA序列

进行基因组编辑实验的第一步是设计gRNA。通过Invitrogen™ GeneArt™ CRISPR查找及设计工具，您可以从我们60万多种已预先设计的靶向人类和小鼠基因的CRISPR gRNA数据库中进行查找。该工具支持目标区域的FASTA序列格式并采用我们的专利算法进行全新的gRNA设计。每个基因可最多提供25种gRNA序列以及基于每种CRISPR序列潜在的脱靶效应的建议。同样地，Invitrogen™ GeneArt™ TAL Effector查找及设计工具也可用于查找由TALEN介导的基因组编辑的最佳序列。体外gRNA转录所需的带有T7启动子序列的寡核苷酸模板可直接使用设计工具进行订购。此外，也可直接通过Invitrogen™ TrueGuide™ gRNA合成网站订购即用型gRNA。在这里，可以订购用于小鼠和人类宿主的定制及预制gRNA。在大多数情况下，选择推荐排名前2或前3的设计并对每种gRNA进行单独测试有助于获取最佳的编辑效果。强烈推荐将Invitrogen™ TrueGuide™合成gRNA与TrueCut™

Cas9蛋白V2进行配套使用，以保证在对困难细胞系进行编辑时仍可保持较高的细胞存活率。

合成gRNA

一旦选择好gRNA序列，则需对gRNA进行合成。我们推荐使用Invitrogen™ GeneArt™ Precision gRNA合成试剂盒进行gRNA的合成。该试剂盒可提供gRNA快速合成的完整体系，包括从通过使用带有T7启动子的靶向特异性寡核苷酸进行的模板DNA组装，体外转录 (IVT) 反应，直到最终可产生用于转染的gRNA的快速纯化步骤。如果采用即用型TrueGuide合成gRNA，则这些合成步骤都可省去。gRNA随后将会与Cas9核酸酶发生复合并通过一系列的方法被转入至靶细胞中。例如，核酸酶可通过Invitrogen™ GeneArt™ CRISPR核酸酶载体以DNA的形式提供，或通过Invitrogen™ GeneArt™ CRISPR核酸酶mRNA以mRNA的形式提供，或直接以即用型的蛋白形式提供。对于所有的瞬时转染，我们均建议使用我们纯化后的重组Invitrogen™ TrueCut™ Cas9蛋白v2。最后，预先设计的gRNA阵列可以慢病毒储存液的形式提供，用于功能性基因组筛选。关于现有CRISPR-Cas9形式以及基于细胞类型进行推荐的完整信息，请参阅thermofisher.com/genomeedit。

基因组编辑工作流程以及我们的产品

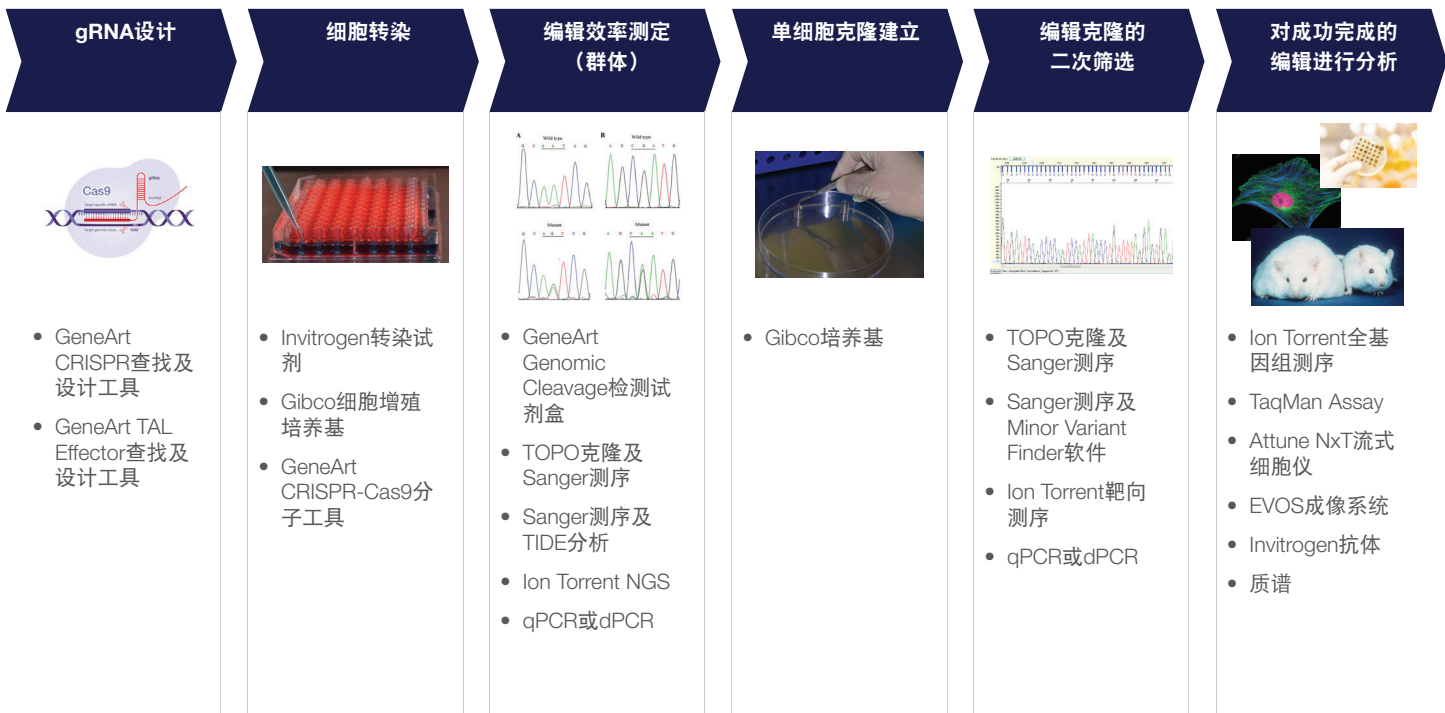


图1. Thermo Fisher Scientific用于基因组编辑的产品线。从gRNA设计和细胞转染试剂到基因分析工具和表型分析方案，我们拥有包括经优化和验证系统的一套完整的工具，可以协同使用来促进基因组编辑项目的进展。



将gRNA和Cas9转入细胞

Cas9-gRNA复合物可使用Invitrogen™ Lipofectamine™ CRISPRMAX™ Cas9转染试剂通过脂质体介导的转染被转入细胞系中。此外，也可采用Invitrogen™ Neon™转染系统通过电转将复合物转入细胞中。对于较难转染的细胞，强烈推荐使用Neon系统，以获取最大的编辑效率。一旦完成转染，细胞可进行恢复。然后，可以使用围绕gRNA的一系列切割事件的初级转化群体测定编辑效率。

测量编辑细胞的比例

在进行繁重而昂贵的实验之前，需要确定初级转化细胞群中实现了成功编辑的细胞比例。下文将介绍一些采用基因分析的最常见方法。虽然所有这些方法都能提供编辑效率的定量信息，但其中某些方法在所需编辑的性质、通量和工作流程方面优于其他方法。其他利用特定被编辑基因的表型方法（例如酶活性）也可以使用，但本文不予考虑。一旦确定了带有编辑的细胞比例，就可以考虑进行下一步。例如，为了进一步研究，可能需要获得具有确切已知编辑序列的基因同源性细胞克隆。如果初次筛选获得的已编辑细胞比例足够高，则可不进行克隆分离，直接对混合样品进行研究。如果基因编辑的效率很低，则可能需要选用不同的gRNA或转染方法重新开始实验。

对已编辑细胞进行同源性克隆分离

可通过编辑效率和预期的存活率来对需要筛选以获得具有正确敲除的克隆细胞系的单克隆数量进行预估。例如，如果需要在两个基因拷贝中都有突变的纯合子敲除，而最终的编辑检测效率为50%，则在任何单个细胞中敲除两个等位基因的概率为25% (0.5×0.5)。由indel导致移码突变的概率为2/3，则indel在每个细胞中引起的纯合子敲除概率为10-11% ($(0.5 \times 0.5) \times (0.66 \times 0.66)$)。在通过有限稀释法进行单细胞分离时，推荐使用96孔板中平均每孔0.8个细胞的密度。如果假设仅有20%的细胞可以存活，则至少需要按此方法对十块96孔板进行铺板，而仅有192个细胞可以存活下来。因此，根据提供的该编辑效率案例，在存活细胞中纯合子敲除的数量预计为19-21个细胞 ($192 \times 10-11\%$)。请注意，一些特定的细胞类型将会需要采用特定的培养基条件来维持克隆细胞的存活率。

对二次筛选克隆进行编辑序列的验证

由于从单个分离细胞生长而来的二次培养细胞应该是具有基因同源性的，用于测定编辑是否成功的方法应该是更为多样化的。在这些情况下，可采用Applied Biosystems™ TaqMan® SNP分型试剂盒以及表型分析中的任何产品，并同样需要考虑诸如编辑的性质、工作流程、通量需求以及成本等因素。在进行二次筛选时需要考虑的一些注意事项，请参阅本文的最后部分。



确定初级转化细胞群的编辑效率

利用琼脂糖凝胶进行基因组编辑检测

一种确定编辑效率的常用方法是使用Invitrogen™ GeneArt™ 基因编辑效率检测 (GCD) 试剂盒。该方法使用基因位点特异性引物扩增初级细胞群的裂解物, 然后将得到的PCR产物变性并重新退火, 促使带indel的链与不带indel或带不同indel的链退火时发生错配。随后, 检测错配并进行酶切。通过凝胶电泳和条带亮度测定法对所得片段进行分析。对于设计用于gRNA的序列, GeneArt CRISPR和TALEN设计工具均可提供用于GCD测定的建议引物选择。

图2是用GCD测定获得的琼脂糖凝胶结果的实例。其中显示了几种不同的基因位点, 包括未经处理的位点和检测酶处理的位点。两条较淡的电泳条带表明转染池中包含成功的编辑。对于处于扩增区域中间的gRNA, 可以确定预期切割产物的大小。通过比较条带亮度测定法或荧光法得到的切割片段的信号强度与测定DNA (完整野生型扩增片段+切割片段) 的总量, 可以得到基因编辑效率的估算值。在本例中, 基因编辑效率在10%到50%之间。

利用琼脂糖凝胶检测GCD分析总结:

- 利用琼脂糖凝胶检测GCD分析为检查基因编辑效率提供了一种简单方法。
- 琼脂糖凝胶的分辨率相对较低, 因此只能预估切割产物的大小。
- 准确量化需要未达到扫描系统饱和度。
- 在琼脂糖凝胶上上样及分析大量样品的过程较为繁琐。
- 与基于测序的方法不同, GCD不能识别编辑过程中可能存在的indel类型。

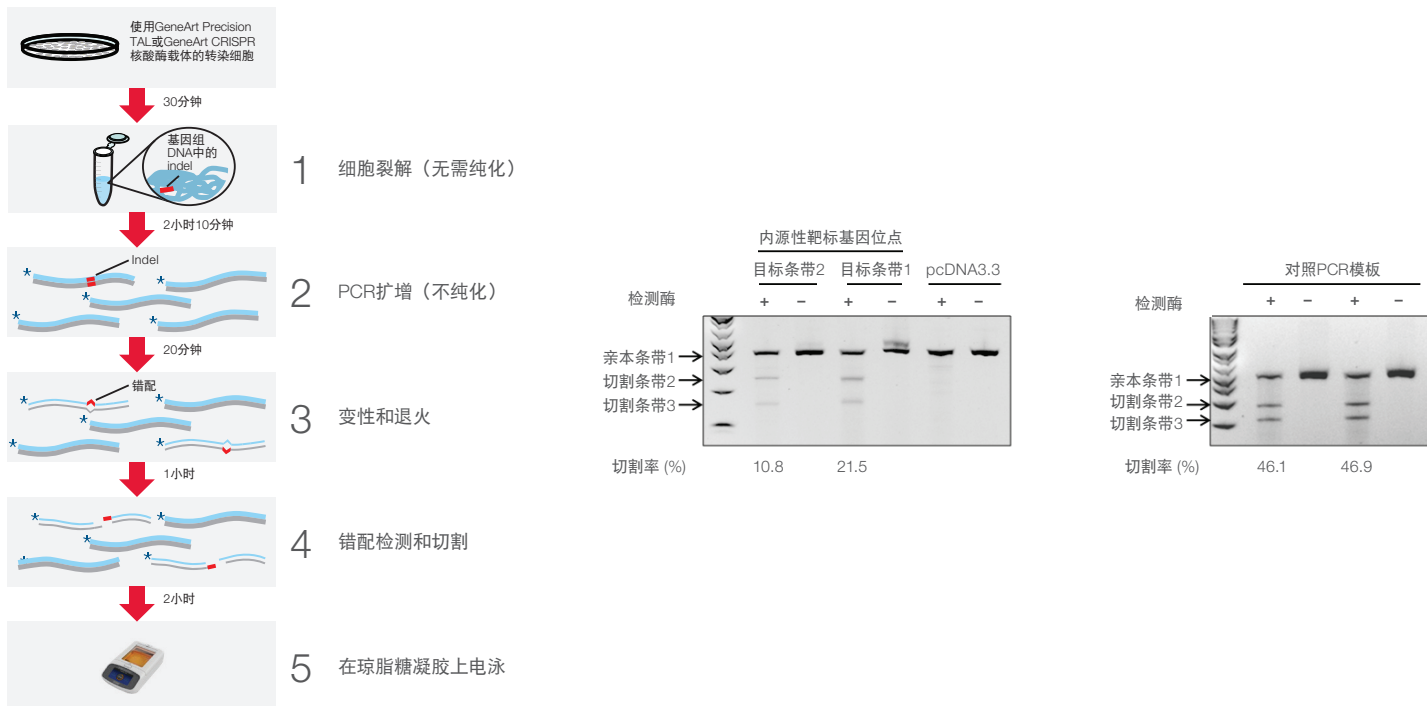


图2. 使用GCD和琼脂糖凝胶确定编辑效率。在琼脂糖凝胶中分析基因编辑效率很简单, 但对片段大小的量化可能不准确。

使用片段分析进行GCD测定

上述GCD测定也可以使用毛细管电泳 (CE) 进行分析。在该方法中，基因位点的扩增引物之一用荧光染料进行标记。首先对初级转化子的裂解物进行PCR扩增，然后用检测酶处理，该实验步骤与之前琼脂糖凝胶实验相同。然后，将裂解片段重新悬浮于Applied Biosystems™ Hi-Di™ 甲酰胺中，由毛细管电泳区分片段大小，并使用标准片段分析软件进行分析。

对于五个不同基因位点编辑的裂解物，使用GCD和CE方法对编辑效率进行了测定。正常细胞的提取物（未用gRNA处理）显示了扩增子的预测大小的单峰（图3，使用HPRT扩增子的示例）。片段大小不受变性以及随后的检测酶处理的影

响。相比之下，用靶标特异性的gRNA处理的细胞（预计在HPRT基因位点产生随机缺失）显示了对应完整扩增子的峰，以及当没有添加检测酶时达到的小于15nt的峰。这些较小的峰对应于扩增子中的缺失，可视为编辑的结果。当用单链核酸酶处理时，可以看到在229-251nt范围内的新峰，指示错配的切割产物。通过计算切割片段峰值的总高度与所有峰的总高度的比值，可以得到68%的编辑效率。而使用其他四个基因位点编辑的裂解物，也获得了类似的结果（图中未显示数据）。

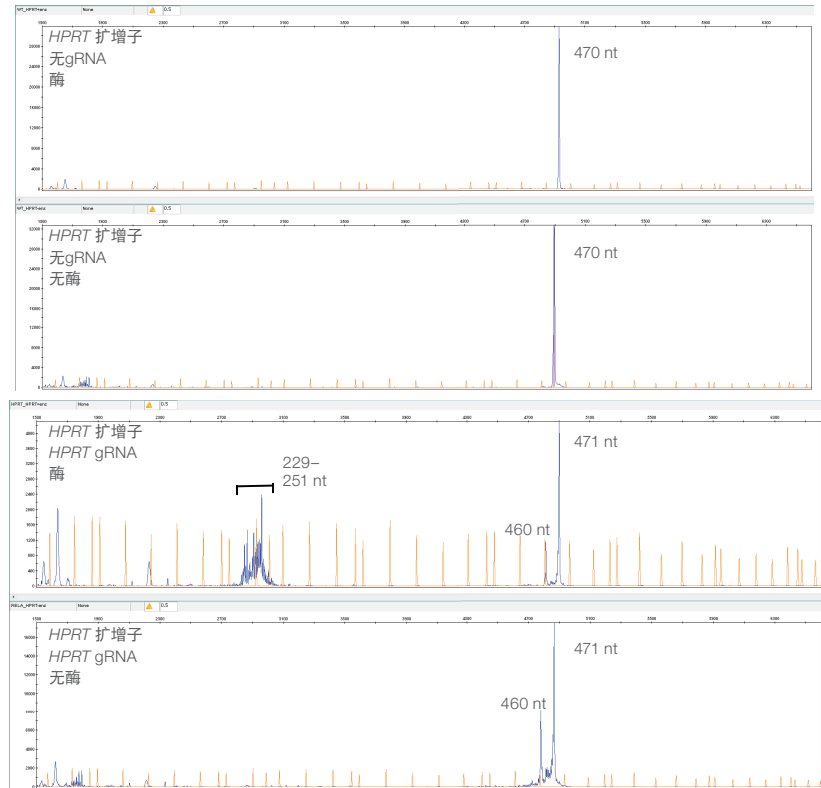
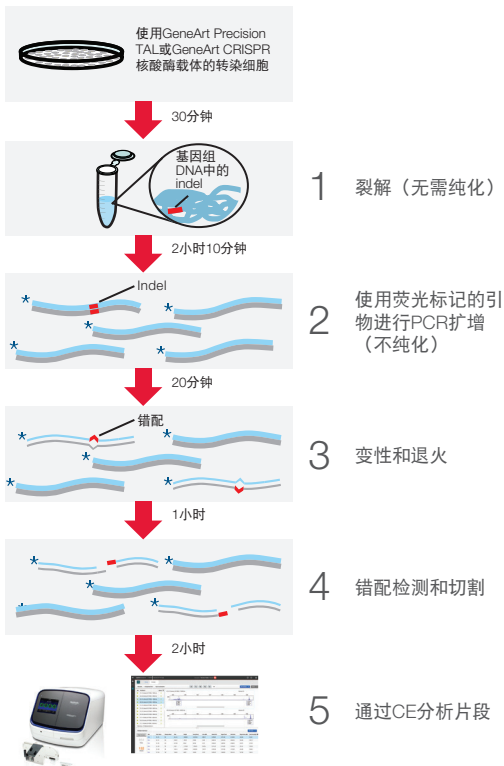
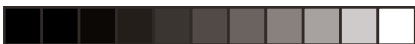


图3. 使用GCD确定编辑效率并通过CE进行片段分析。与使用琼脂糖凝胶相同的常规方法相比，可以通过CE分析编辑效率。这里，PCR引物之一用荧光标签标记。当分析缺乏gRNA的对照细胞时，在有或没有检测酶的情况下都观察到了对应于HPRT亲本基因位点的单峰。用靶标特异性gRNA处理的初级转化细胞群显示出亲本峰值 (471nt)。然而，在检测酶存在的情况下，一组在229-251nt范围内的峰表明检测到了切割产物。此外，还显示出460nt附近的一组峰，表明剪切掉的序列没有错配，并重新退火编辑。通过确定229-251nt峰中所有峰（切割片段加亲本片段）的总峰值，计算编辑效率。橙色峰是Applied Biosystems™ LIZ™染料标记的marker，用于计算片段大小。



毛细管电泳 (CE) 分析片段通常利用Applied Biosystems™ GeneMapper™ Software进行分析。然而，我们已经发布了基于云计算的片段分析工具Applied Biosystems™ Peak Scanner™ Software。该软件为GeneMapper Software提供了峰值分析功能，且界面更简洁、更易操作。Peak Scanner Software是基于云计算的应用程序，加入新功能模块更加方便。此外，研究人员通过云连接更容易分享数据。借助于Peak Scanner Software的GCD分析的结果与使用GeneMapper Software分析的结果相同（图4）。

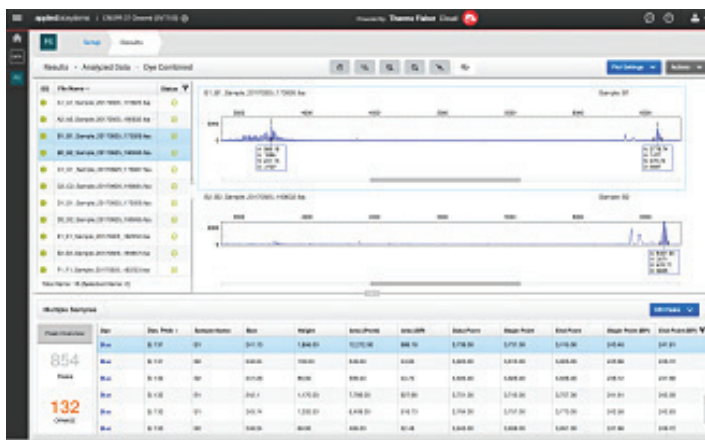


图4. 基于云计算的Peak Scanner Software方便了片段分析。可以使用免费的Peak Scanner Software分析GCD片段。该软件包含了GeneMapper Software的功能，但能够进行云存储，且用户界面简洁。图示为对从RELA gRNA处理的初级转化细胞群中得到的RELA扩增子进行的GCD分析；顶部曲线显示用检测酶处理时产生了基因片段，底部曲线表示不使用检测酶处理的情况。

通过CE进行GCD分析的总结：

- 与琼脂糖凝胶相比，通过CE进行GCD分析，基因片段大小分辨率更高。可以识别出单个碱基对的差异。
- 峰值的总和可以更准确地量化特定片段中DNA含量。与基于琼脂糖凝胶切割片段的计算相比，基于峰值大小（或面积）的计算更精确。
- 将来自不同位点的GCD分析整合到一道毛细管中，并通过使用不同的染料颜色进行引物标记，可以提高通量。尽管需要针对每个查询位点选择对应染料标记的引物，现已有采用通用染料标记的引物进行类似分析的系统 [1]。
- 使用Peak Scanner Software可以简化对GCD片段的分析。

利用TIDE（通过分解追踪插入/缺失）软件进行Sanger测序基因组编辑效率也可以使用Sanger测序来确定。通过设计编辑位点旁侧的引物用于扩增初级编辑细胞群中的基因位点。由GeneArt CRISPR和TALEN设计工具推荐的GCD引物可用于生成扩增子，这些扩增子将覆盖编辑区域并且可进行测序分析。此外，也可以使用与标准PCR引物相似的规则设计gRNA目标序列的旁侧的PCR引物。然后，使用Applied Biosystems™ BigDye™ Terminator或BigDye™ Direct测序试剂对所得扩增子进行测序。包含成功编辑的混合样品的测序反应曲线将包含一段分型明确的测序序列，并且在编辑位点下游含有一段混合碱基序列（图5）。通过观察这些测序峰图，通常能够说明编辑是否成功，但是，仅通过简单的观察很难确定编辑效率以及混合样品所含序列的性质。

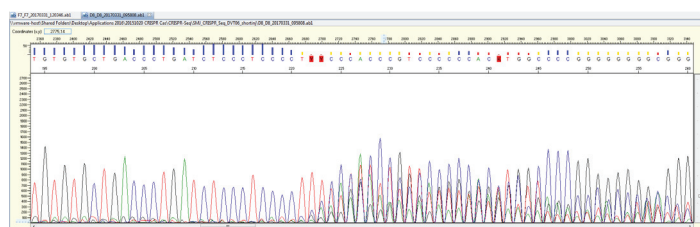


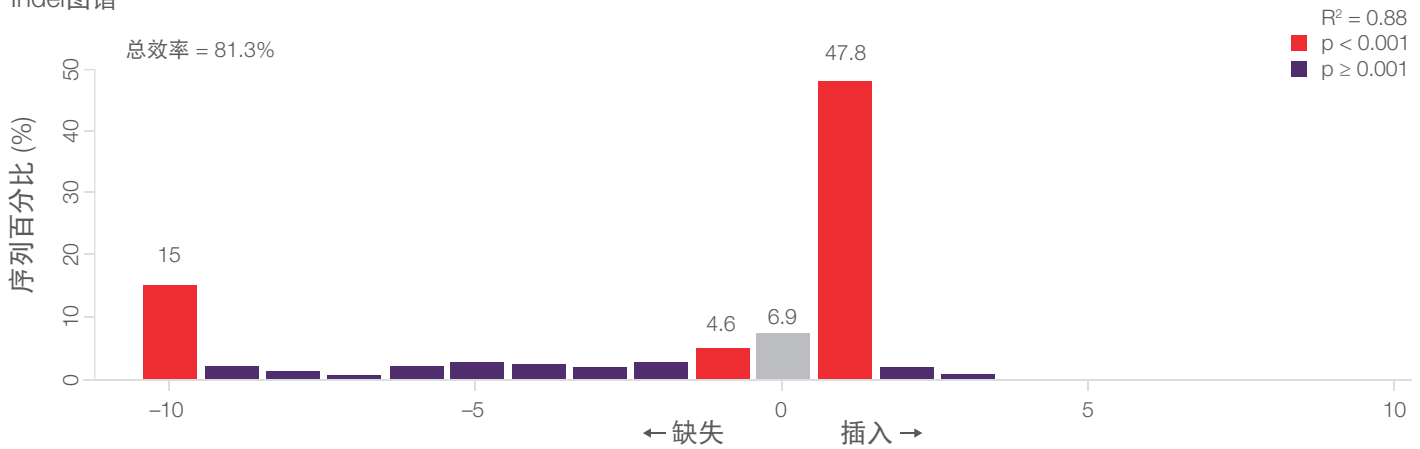
图5. 编辑细胞混合样品的Sanger测序。来自初级转化细胞群的扩增子曲线示例。应注意，红色箭头之前的序列是一致的，之后的序列有所不同。混合样品含有各种插入和缺失。此外，尽管很确定已经发生编辑，但是无法确定序列和编辑效率。



为了便于分析，Brinkman等 [2] 设计了用于对基因组编辑实验产生的混合碱基测序曲线进行去卷积的方法。他们将这种方法打包至一款名为TIDE (<https://tide.nki.nl>) 的软件工具中。TIDE软件利用gRNA序列以及未修饰和已编辑DNA的扩增子测序数据，确定不同indel长度的曲线对混合曲线的贡献度。软件将生成一张直方图，其中展示了插入和缺失（例如，+2, +1, 0, -1, -2）的各种贡献度，以及整个细胞群的总体编辑效率计算值（图6）。尽管编辑细胞群中分子的确切序列未确定，但建立了一个测量插入和缺失范围的方法，有助于确定是否继续进行下一步实验。

对于来自具有不同基因位点编辑的初级编辑细胞群提取物，使用TIDE软件分析其编辑效率（图6和7）。应注意，正向和反向测序具有相似的计算所得频率和可能突变图谱。对于 *HPRT* 编辑细胞（图6），整体效率约为80%，大部分种类为+1插入。*RELA* 编辑细胞（图7）的效率较低，约为20%。TIDE软件还提供R²值作为拟合优度的衡量标准，并强调每个indel的统计学意义。尽管 *HPRT* 基因位点的R²值较高，但 *RELA* 序列的R²值明显较低，表明这些值的置信度较低。

Indel图谱



Indel图谱

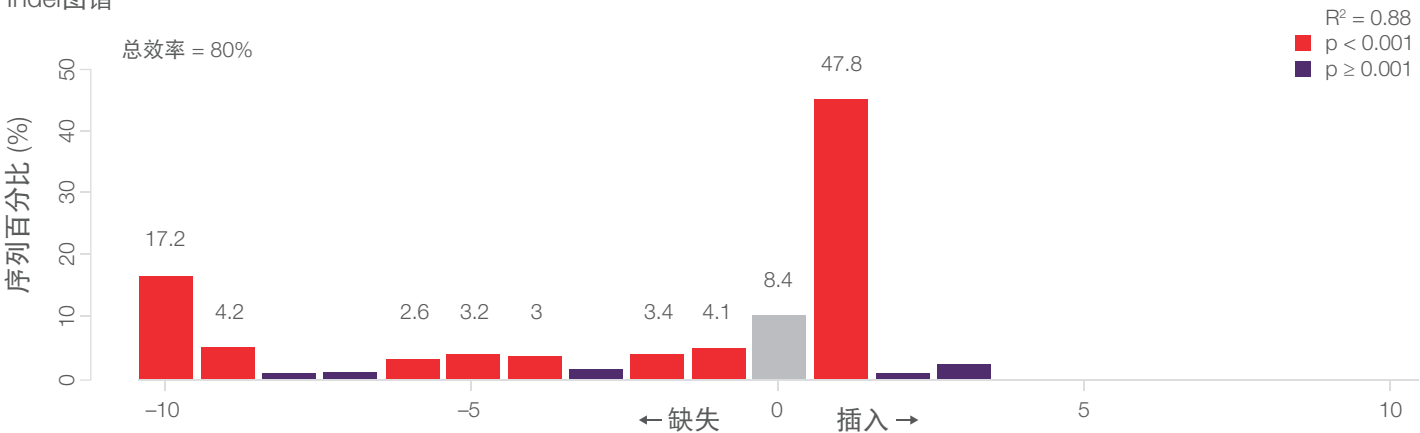


图6. TIDE软件可用于评估编辑效率。该软件能够呈现混合测序曲线（如图5所示），并预测不同indel长度的贡献度。红色数据条代表具有高置信度的计算所得百分数；紫色数据条表示具有低置信度的计算所得百分数。请注意，两条链之间的突变图谱和计算所得编辑频率是相匹配的。



很重要的一点是，TIDE软件的设置允许用户更改发生去卷积的窗口。将窗口的默认值10 nt更改为25 nt，结果表明*RELA*位点中的许多编辑是大于15 nt的删除，使用默认设置会遗漏这些编辑（图7）。随着窗口大小的增加， R^2 值也更高，并且计算所得效率（68%）与使用其他方法确定的效率更加一致（见下文）。

Sanger测序分析总结：

- Sanger测序及随后的TIDE软件分析为从初级编辑细胞群中获取信息提供了一种实用方法。
- 完善的测序工作流程为确定克隆扩增细胞中的预期编辑效率和缺失类型提供了一种简单方法。
- 使用TIDE软件时，可能需要更改查询窗口的默认值10 nt，以确定是否存在更长的缺失。

- 由于同源介导修复比双链断裂修复的发生率更低，因此，任何SNP或其他定向编辑都可能在原始样品的随机删除中丢失。然而，通过引入更新（TIDER软件）软件可以计算模板定向编辑（如SNP）的效率 [3]。
- 使用TIDE软件可以免费查询单个位点。此外，也可实现高通量研究的批量查询，但可能需要购买许可证。相关的最新信息，请访问TIDE网站 (<https://tide.nki.nl>)。

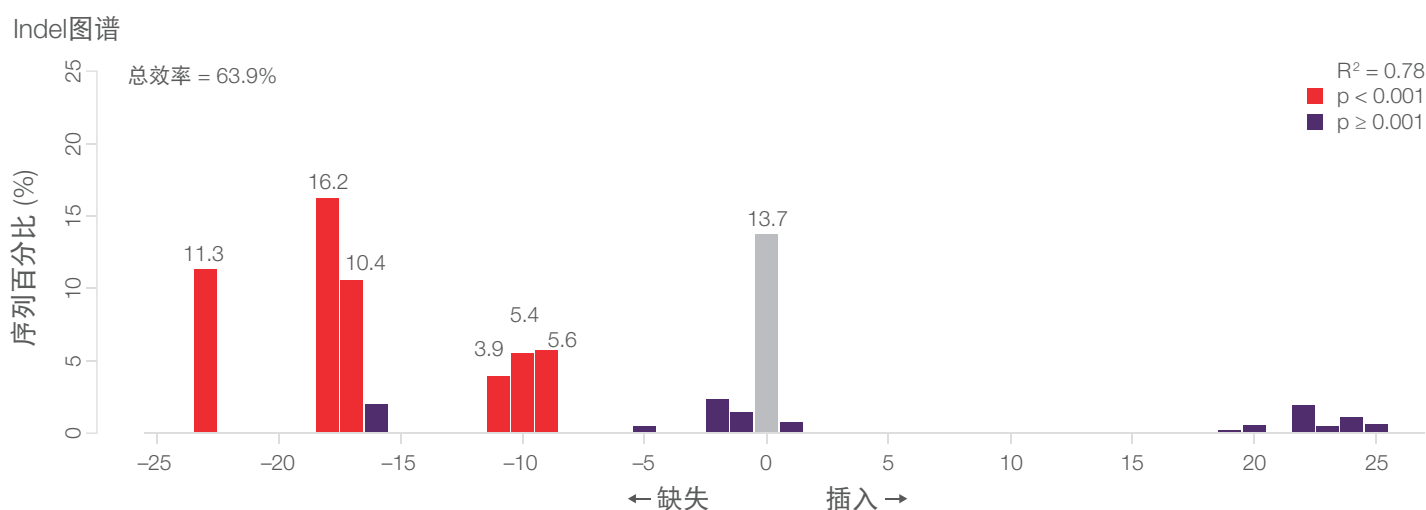
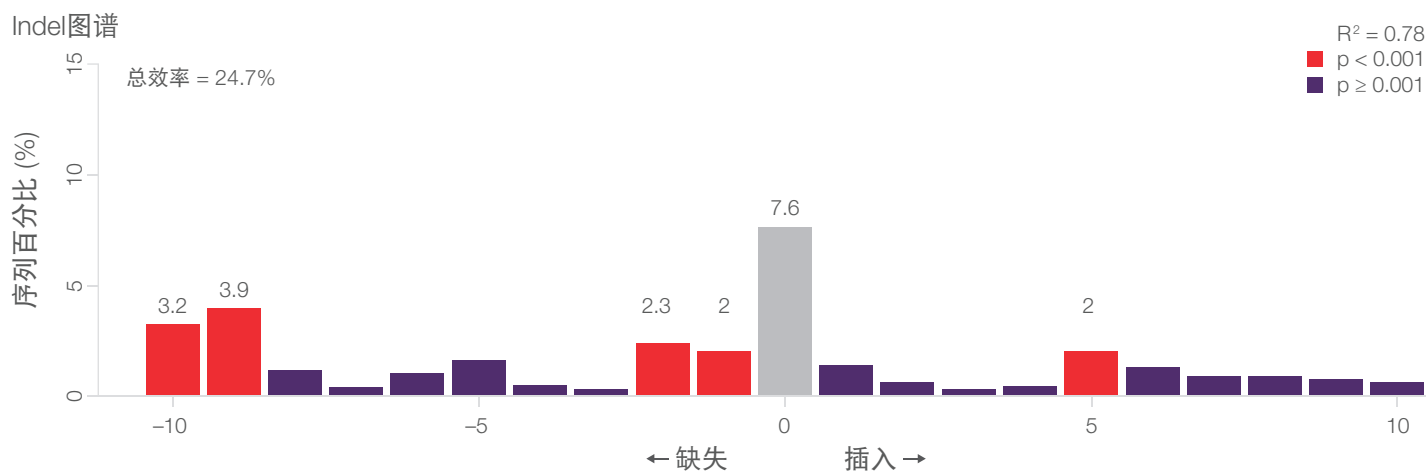


图7. 改变TIDE软件中的分析参数对编辑效率计算具有很大影响。在默认情况下，TIDE软件在gRNA起始位点的-10至+10窗口内运行。但是，更改窗口大小可以提高计算的准确性。在默认设置下，*RELA*编辑扩增子的总编辑效率约为25%，而将窗口扩大到-25至+25会产生更接近真实值的结果（效率约为64%）（见图11）。



使用TOPO克隆进行Sanger测序

通过将扩增子亚克隆到质粒中并对单个质粒进行测序，也可以确定基因组编辑实验的效率（图8）。该方法对单个分子进行测序，因此，可以提供关于所产生突变类型的准确信息。为了阐明该方法，我们使用靶标外侧引物扩增细胞裂解物中的DNA，产生长度不超过600个碱基对的PCR扩增子。随后，使用Invitrogen™ Zero Blunt™ TOPO™ PCR克隆试剂盒对PCR产物进行亚克隆，并转化到TOP10大肠杆菌细胞中。从每个转化的编辑细胞群中挑选96个细菌菌落，并使用Invitrogen™ PureLink™ 96 HQ Mini质粒DNA纯化试剂盒分离

DNA，然后进行Sanger测序。随后，分析所得测序数据，以测量含有精确编辑序列的PCR产物百分比，并挑选用于维持的克隆分离株。通过测序大量质粒，可以确定包含编辑细胞的百分比。该方法还可以初步了解总体基因敲除率或编辑效率以及已经发生的indel变化类型。

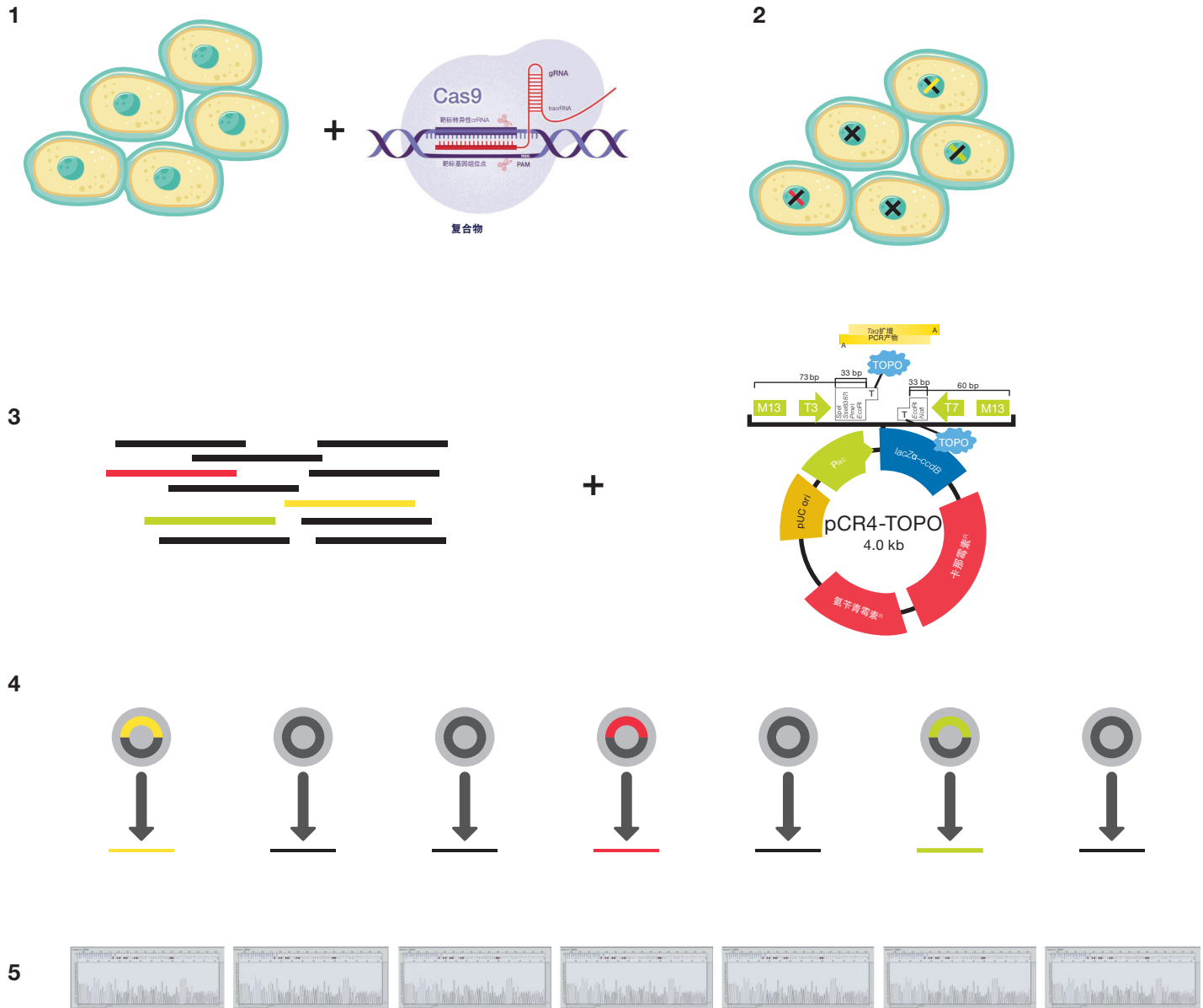


图8. 通过质粒亚克隆和测序，计算编辑效率。 第1步：用gRNA和Cas9 mRNA转染细胞。第2步：孵育细胞，以完成基因组变化。第3步：从细胞培养物中纯化基因组DNA，对同源培养物中的基因工程位点进行PCR扩增，然后将PCR片段克隆到TOPO载体中。第4步：从单个菌落中分离质粒，并对插入片段进行PCR扩增。第5步：对插入片段进行测序。编辑效率是指具有基因工程变化的被测序列插入片段总数的百分比。

我们先使用gRNA和Cas9 mRNA转染细胞，然后使用TOPO克隆对初级转化子中的位点进行亚克隆，再通过Sanger测序对96个克隆进行分析。在被分析的96个克隆中，有84个与目标序列一致。其中，只有12个克隆在扩增区域无编辑事件；有72个克隆与野生型序列相比至少有一个序列偏差，总体编辑效率约为86%。有关详情，请参阅参考文献4。

TOPO克隆分析总结：

- 亚克隆及随后的Sanger测序可以估算初级转化细胞群中基因组变化的确切性质。
- 测序更多质粒可以提供更深入的见解，因为测序的菌落越多，发现细胞群中稀有物种的可能性越大。
- 测序的质粒越多，编辑效率越准确。
- 亚克隆和测序的工作流程可能非常耗时且繁琐。

dPCR

数字PCR (dPCR) 通过将PCR反应分成数千个亚微米级反应室来精确测定含有编辑的比例。通过计算阳性反应室的数量，可以计算出原始样品中含有这些序列的分子数量。该方法可以用于实验室测定基因组编辑的效率。在使用的自定义Applied Biosystems™ TaqMan® 探针中，其中一个探针置于待检测的编辑序列上，而第二个探针则置于不易受编辑影响的扩增子的区域中（图9A）。当对分区样品进行PCR时，大多数分区对于未编辑区域的探针应为阳性，而未经过编辑的分区对于以编辑区域为目标的探针也为阳性。计算只有一个探针为阳性的孔的比例即可得到编辑效率。

为了显示这种方法的实用性，我们从含有多种人类基因的随机编辑的细胞获得裂解物。使用1微升裂解液与自定义设计的Applied Biosystems™ TaqMan® Assay和QuantStudio™ 3D数字PCR系统，遵循dPCR的标准操作方案和热循环程序。然后使用Applied Biosystems™ QuantStudio™ 3D Analysis-Suite™云软件分析数据（图9B）。例如，来自CDK4基因位点编辑的细胞的裂解物产生3,827个对于两种探针均为阳性的孔和1,082个仅对对照探针为阳性的孔。这相当于28%的编辑效率。对来自在其他四种基因编辑的细胞的裂解物得到了类似的测定结果（图11）。

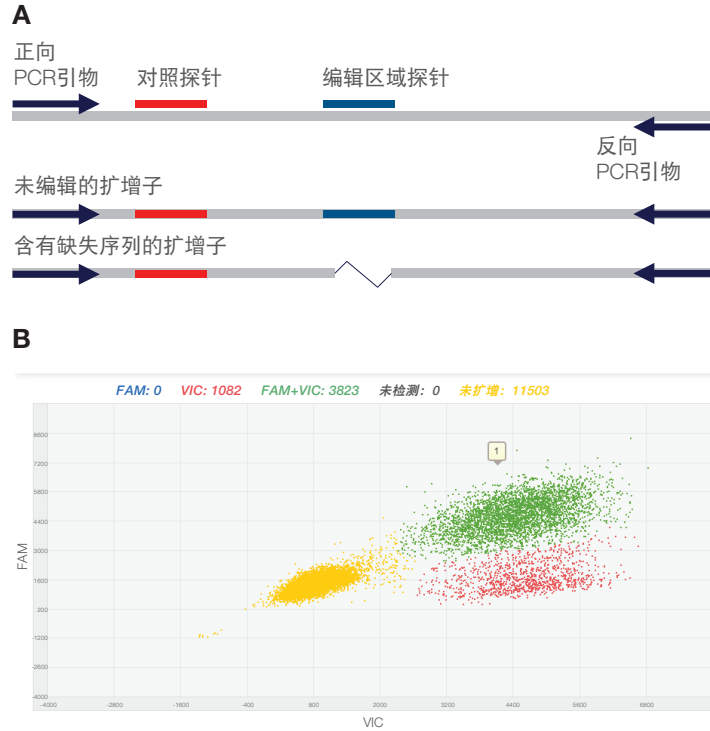


图9. 通过dPCR计算编辑效率。 (A) 通过自定义定制TaqMan SNP Assay，可以通过dPCR估计编辑效率。为此，选择一段不易受编辑和修复活动影响的探针（用VIC™染料标记探针，图示为红色），以及一种覆盖目标区域中用于编辑的原始的未改变序列的探针（用FAM™染料标记探针，蓝色）。(B) 发生编辑的细胞则仅为VIC探针阳性（红点）；未发生编辑的细胞则FAM和VIC探针均为扩增阳性（绿点）根据仅对VIC探针呈阳性的数量与总数（对于VIC探针以及FAM + VIC探针均呈阳性的孔）的比值得到编辑效率。请注意，可能需要对探针配比进行一些优化。

dPCR分析总结：

- 编辑效率可以通过dPCR来测定。
- 数据采集的数字特性意味着可以获得编辑效率的精确值。
- 可以修改dPCR方法以测定更多靶向编辑的效率，如SNP或易位。例如，可以将编辑特异性探针定位到所需的单核苷酸变异，并计算对于该探针阳性的比例。

- 在基因组编辑实验中成功使用dPCR可能需要优化实验所需Assay，包括PCR条件，引物和探针的混合比例。

新一代测序

使用Ion Torrent™新一代测序 (NGS) 方法也可以测定编辑效率。通常，对从初级编辑细胞群中的目标基因位点得到的扩增子进行测序。收集每个池中的细胞并在GeneArt基因编辑效率检测试剂盒的缓冲液中裂解。然后使用位于编辑位点外围的引物扩增目标区域。选择能产生与预期长度相符的扩增子的引物是十分重要的。我们还建议使用高保真PCR酶来避免可能影响数据质量的任何序列错误。可以使用如Ion Chef™系统、Ion Xpress™ BarcodeAdapter 1-96试剂盒、Ion Xpress™ Plus片段文库试剂盒和Ion PGM™ Hi-Q™ Chef试剂盒从PCR样品构建条形码文库。构建完成后，文库可以在Ion PGM™或Ion S5™仪器上进行测序。制备和运行Ion Torrent™文库的详细方案请参见相应产品的用户手册。

NGS是一种分析初级转化细胞群的可靠方法（图10）。可以通过比较编辑基因的reads数与该基因位点的总reads数来计算编辑效率。此外，可以通过分析返回的序列来测定目标位点的编辑性质。测序深度和获得单个扩增子的多个reads数的能力有助于提高该方法的准确性和可靠性。由于拥有多达384个条形码，该方法适用于高通量工作流程。然而，NGS也有缺点，那就是缺乏自动化数据分析工具。自定义的数据分析方法可以辅助自动数据分析，但必须自主开发。此外，可能并不需要NGS生成的大量数据来确定编辑实验的效率。同时基于化学反应和数据分析的复杂性可能会影响结果获得的时间，除非进行多样本检测，否则对于单个样本而言，NGS相对昂贵。

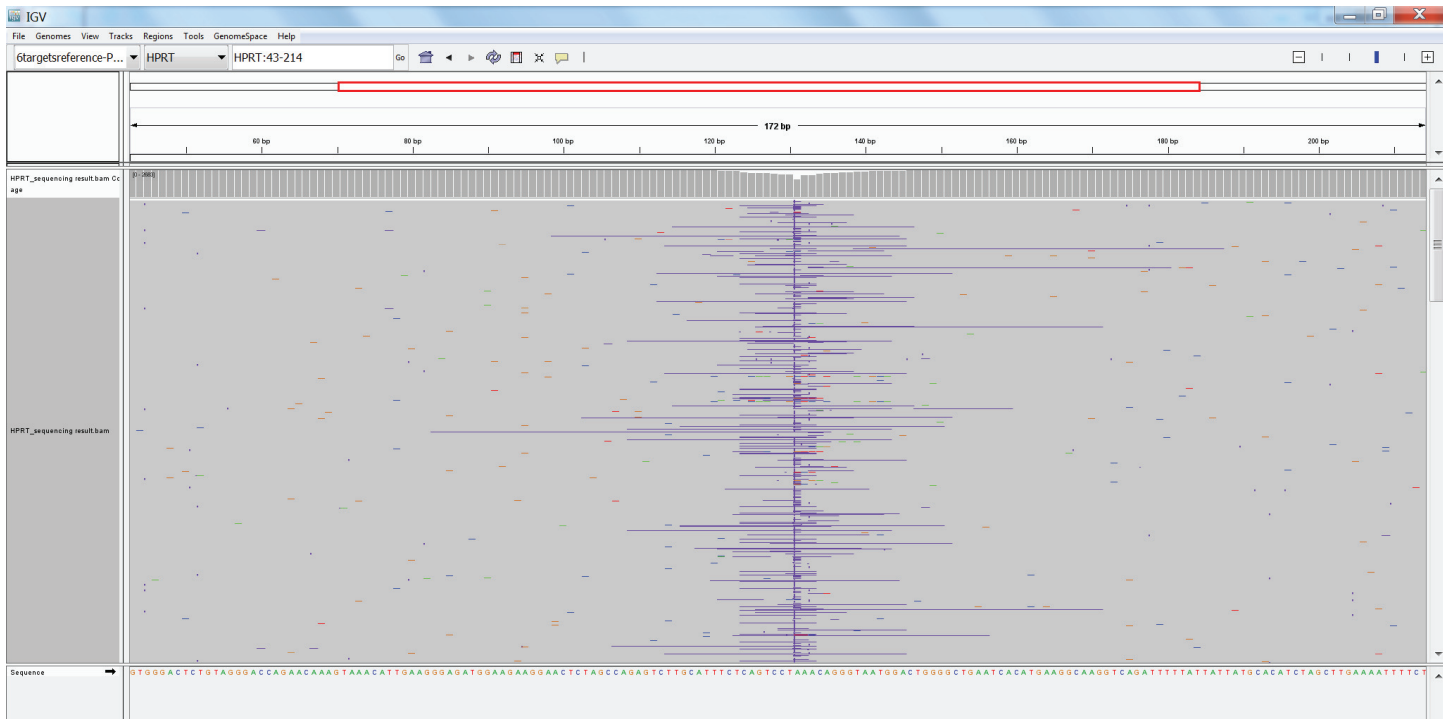


图10. 在HPRT基因位点编辑的混合细胞群的NGS分析。将Ion Torrent NGS分析结果导入Integrative Genomics Viewer (IGV) 并在gRNA序列上对齐。每行表示一个单独的读数；紫色线代表缺失的序列。其他修饰的基座通过不同颜色的短棒显示。读数面板上方的灰色条显示该位置未修饰碱基的相对数量；请注意，大多数变异都围绕在gRNA位点。通过分析IGV的数据，可以轻松确定混合池的编辑效率和类型。

NGS分析总结:

- 初级细胞群中的基因组编辑可以通过下一代测序进行分析。
- 增加测序深度可提供细胞群中编辑的更多细节并提高编辑效率测定的准确性。
- NGS解决方案适用于高通量工作流程。
- 生成的复杂数据可能需要自定义数据分析工具。
- 除非大量样品多重检测，否则NGS平台相对昂贵。

方法比较

本文介绍了许多检测初次编辑效率的方法（还有其它方法尚未涉及）。为了比较通过上述方法获取的数值，我们通过五个不同基因位点的细胞群裂解物来测定编辑效率（图11）。为了确保比较的更加直观的比较，每次测定使用相同裂解产物的稀释样品。总的来说，观察到的效率是相似的，只有几个异常值。例如，琼脂糖分析的GCD获得的*HPRT*编辑效率高于片段分析或Sanger测序的GCD获得的*HPRT*编辑效率；这可能是由于测量琼脂糖凝胶条带卡亮度的精确性不够导致。再如，由dPCR获得的*CDK4*编辑效率低于其他方法的值；这可能是由于测定设计得不理想。此外，使用*CHUK*裂解物的TIDE频率测定值低于其他方法的值；这是由于测序数据的质量较低导致（图中未显示数据）。

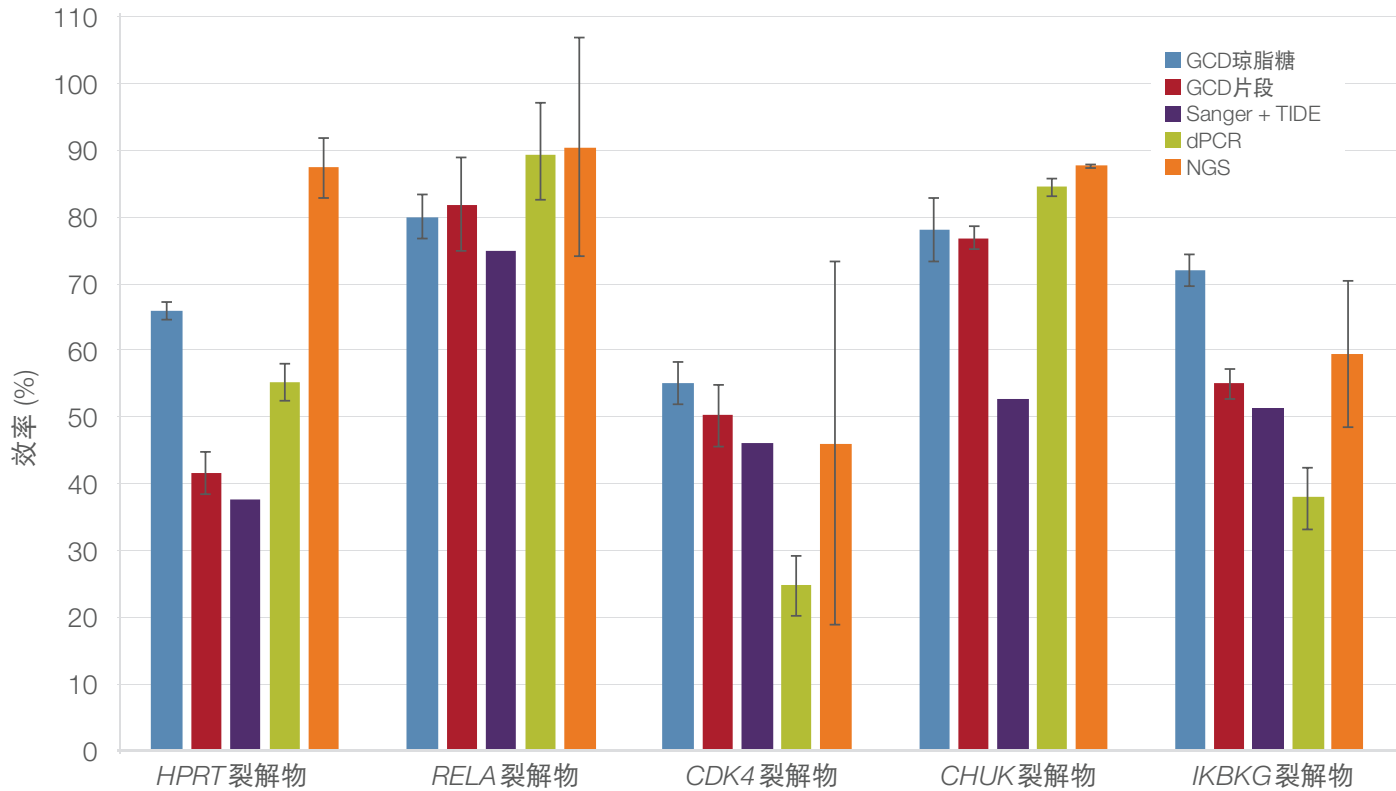


图11. 比较不同技术方法获取的编辑效率。使用所示技术分别分析A549细胞中指示基因敲除编辑的初级细胞群的裂解物。通过电泳方法（GCD和Sanger测序）测定的编辑效率基本相似，但dPCR和NGS差异较大。



	初次分析 (编辑效率)	初次分析 (获取的测序 信息)	二次分析 (编辑确认)	最高相对通量	最低相对成本
敲除 (随机删除)	<ul style="list-style-type: none"> GCD Sanger测序与TIDE软件 	<ul style="list-style-type: none"> NGS TOPO克隆与Sanger测序 	<ul style="list-style-type: none"> Sanger测序 	<ul style="list-style-type: none"> GCD NGS 	<ul style="list-style-type: none"> GCD
SNP	<ul style="list-style-type: none"> dPCR Sanger测序与TIDE软件 	<ul style="list-style-type: none"> NGS TOPO克隆与Sanger测序 	<ul style="list-style-type: none"> Taqman分型试剂盒 Sanger测序 	<ul style="list-style-type: none"> Taqman分型试剂盒 NGS 	<ul style="list-style-type: none"> Taqman分型试剂盒
插入	<ul style="list-style-type: none"> dPCR TOPO克隆与Sanger测序 	<ul style="list-style-type: none"> NGS TOPO克隆与Sanger测序 	<ul style="list-style-type: none"> Taqman分型试剂盒 Sanger测序 	<ul style="list-style-type: none"> Taqman分型试剂盒 NGS 	<ul style="list-style-type: none"> Taqman分型试剂盒
定点删除	<ul style="list-style-type: none"> TOPO克隆与Sanger测序 Sanger测序与TIDE软件 	<ul style="list-style-type: none"> NGS TOPO克隆与Sanger测序 	<ul style="list-style-type: none"> Sanger测序 	<ul style="list-style-type: none"> GCD NGS 	<ul style="list-style-type: none"> Sanger测序

图12. 测定切割效率的方法。选择测定编辑效率的方法取决于所需的编辑类型、所需的通量（分析大量样本的能力）以及工作流程的复杂程度。表格中展示了针对各种编辑和实验变量的推荐方法。所有的方法都可以提供定量信息，但某些方法相对具有优势。

每种方法对于分析初级转化子都有各自的优缺点。虽然这些方法都可以得到关于编辑效率的信息，但许多变量会决定哪种方法最适合特定的实验体系。实验中应当考虑的变量包括所需的编辑类型，是否需要确切的序列信息以及实验工作流程。图12显示了实验初始应考虑的一些常见和重要的因素。对于给定的实验，可能不需要精确测定编辑效率，而是在进入实验的下一个阶段之前，“满足”能够选择出性能最佳的基因编辑所需gRNA的方式就可以。

同质细胞群的二次验证注意事项

分析初级细胞群只是基因组编辑实验的一部分。某些研究人员可能会对分离菌落进行二次筛选，以获得基因同质的细胞培养物。在这样的二次筛选中，不需要测定编辑效率。不过，需要测定派生克隆的克隆性和纯度。此外，还需要检查靶向和脱靶效应。上述许多基因组分析工具对这些筛选都有帮助。



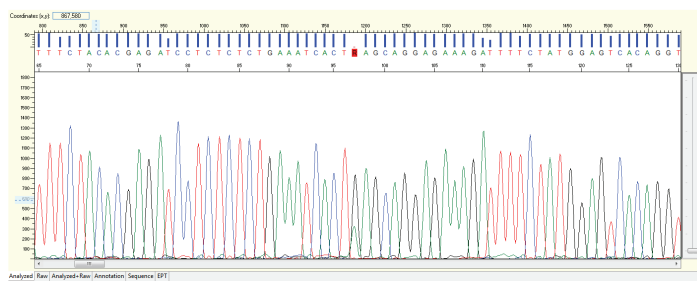


例如，Sanger测序是验证二次克隆的实用工具。如前所述，在相对简单的工作流程中，培养物中的裂解物可以被扩增和测序。同质的细胞培养会得到单一的测序图，而不会有杂峰。因此，除了确认培养物的单克隆性之外，还可以获得该位点编辑的确切序列。如果测序数据显示异质性，则可能需要新一轮的单细胞选择。在存在人工引入SNP的情况下，使用Applied Biosystems™ Minor Variant Finder软件进行分析可能会有用（图13）。这款免费的软件工具可用于检测变异SNP，检测灵敏度低至5% [5]。虽然Minor Variant Finder软件只能检测SNP的变异频率，但TIDER软件 [3] 可用于分析混合种群中其他序列的频率。

此外，基于TaqMan Assay的基因分型方法在验证过程中非常实用。在初级细胞群中，存在不同种类的序列，只有一小部分可能含有所需的编辑。这可能导致初次筛选中的反应效率和数据解释复杂化。相反，在单克隆培养物的二次分析中，可以更有效地筛选期望的编辑。由于TaqMan分型试剂盒可以查询特定的序列，因此可用于鉴别含有所需编辑的培养物。

最后，NGS可以用于分析已编辑细胞的二次克隆。例如，在单克隆细胞培养物中对编辑的基因位点进行测序可以提供变异核苷酸的准确信息，以及对不纯培养物中其他物种的定量。对靶基因或甚至整个基因组有关的基因位点进行序列测定也十分有用，可以测定在编辑过程中是否发生了脱靶效应。

A



B



图13. 检测混合群体测序数据中的SNP编辑。 Minor Variant Finder是一款使用方便的桌面性软件，专门用于检测和报告Sanger测序数据中的单核苷酸变异 (SNV)。(A) 含有次要SNP的序列数据的示例。(B) Minor Variant Finder软件可以 (1) 减少背景噪音，(2) 鉴别变异，和 (3) 测定正向和反向链中的等位基因频率。使用此软件可以检测出频率低至5%的变异。在图示示例中，正向变化频率为12.9%，反向为10.8%。





结论

Thermo Fisher Scientific为基因组编辑工作流程的每一步都开发了工具和解决方案。这些产品组合（包括Applied Biosystems™公司的基因分析工具）均基于30年的创新经验打造，能随着客户的研究需求而不断发展。我们的工具系列包括经过优化和验证的操作系统，它们可以协同使用，更快、更轻松地解答您的重要科学问题。由于每项研究都是独一无二的，所以我们提供了一系列基因组编辑解决方案，满足您的特殊实验和研究需求。

参考文献

1. Yang Z, Steentoft C, Hauge C et al. (2015) Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res* 43:e59.
2. Brinkman EK, Chen T, Amendola M et al. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* 42:e168.
3. Brinkman EK, Kousholt AN, Harmsen T et al. (2017) Easy quantification of template-directed CRISPR-Cas9 editing. Preprint from bioRxiv: <http://dx.doi.org/10.1101/218156>
4. Thermo Fisher Scientific Inc. (2016) Using Sanger sequencing to facilitate CRISPR- and TALEN-mediated genome editing workflows. Application note COL31142.
5. Thermo Fisher Scientific Inc. (2016) Low-level somatic variant detection in tumor FFPE samples by Sanger sequencing. Application note COL31176.



订购信息

产品名称	数量	目录号
GeneArt Genomic Cleavage Detection Kit	20次反应	A24372
Zero Blunt TOPO PCR Cloning Kit for Sequencing	25次反应	450031
SeqStudio Genetic Analyzer System with SmartStart Orientation	1个系统	A35644
3500 Genetic Analyzer	1个系统	4440466
3500xL Genetic Analyzer	1个系统	4440467
GeneMapper Software 5	1个许可	4475073
Peak Scanner Software v1.0	1个许可	4381867
BigDye Terminator v3.1 Cycle Sequencing Kit	100次反应	4337455
BigDye Direct Cycle Sequencing Kit	100次反应	4458687
GeneScan 500 LIZ Dye Size Standard	800次反应	4322682
QuantStudio 7 Flex Real-Time PCR System, 384-well, desktop	1个系统	4485701
Custom TaqMan SNP Genotyping Assays	1,500次反应	4331349
QuantStudio 3D Digital PCR Instrument with Power Cord	1个设备	4489084
QuantStudio 3D Digital PCR 20K Chip Kit v2 and Master Mix	8 x 12芯片	A26317
Ion S5 System	1个系统	A27212

更多信息请访问thermofisher.com/genomeedit

ThermoFisher
SCIENTIFIC

仅供研究使用。不能用于诊断用途。©2018 Thermo Fisher Scientific Inc.保留所有权利。
除非另有说明，所有商标均为Thermo Fisher Scientific及其子公司的财产。
TaqMan是Roche Molecular Systems, Inc.的注册商标，未经许可不得使用。COL32200 0318