

如何使用Identifiler试剂盒和毛细管电泳平台来鉴定人类细胞系

本应用说明包含如下内容：

- Applied Biosystems™ Identifiler™ Plus和Direct试剂盒用来鉴定体外生长的细胞系类型
- 只需使用低至 100 µg纯化基因组DNA (gDNA)即可鉴定细胞系
- 将细胞直接点样在Applied Biosystems™ NUCLEIC-CARD™ 样品采集装置上可鉴定细胞系
- 新型Applied Biosystem™ SeqStudio™基因分析仪在更低的基因组DNA和细胞浓度时，相较Applied Biosystems™ 3130和3500基因分析仪，性能更出色。

引言

人类疾病研究十分依赖培养物中的体外人类细胞系分析。然而，一个日益突出的问题却是，体外培养的细胞可能会出现识别错误或被其他细胞系污染[1]。细胞系识别错误会产生误导性的结果、困惑并增加研究成本[2-4]。众多期刊和资助机构目前都要求研究人员保证所用细胞系真实可靠并确定贯穿整个研究过程的维持策略(例如，Yu et al., 2015 [5]和Neimark, 2015 [6])。

鉴定细胞系的真实性或来源通常要分析几个基因座的突变或位点是否匹配预期的等位基因。目前几种分析变异位点的方法 — 如通过分析限制性片段长度多态性(RFLPs)和扩增片段长度多态性(AFLPs)进行同工酶变异的电泳分析，以及新一代测序(NGS)和MALDI-TOF质谱法 — 都存在各自的缺点，如集成性不足、成本高等。标记高度可变的STR (STR)可作为简便、价廉且高度特异性的细胞系识别“指纹”，是DNA法医学分析常用的技术。将这些高度变异基因座的等位基因谱与已知的细胞系标准化样品的基因谱进行比较，即可鉴定细胞系的真实性。诸如ATCC和莱布尼茨DSMZ-德国微生物和细胞培养物保藏中心(Leibniz-Institute DSMZ—German Collection of Microorganisms and Cell Cultures)等组织提供在线访问可搜索的数据库，以便研究人员查询已知的细胞类型。另外，研究人员可以建立实验室独特的细胞系的等位基因图谱，将细胞的等位基因图谱与自己的内部标准进行比较，以确认细胞系类型。

通常情况下，使用毛细管电泳对具有不同重复数目的微卫星基因座扩增的片段进行STR分析。我们已经建立起毛细管电泳的金标准体系，开发出适合研究人员需求的高灵敏度、高通量仪器。此外，Applied Biosystems™产品线还为使用毛细管电泳仪器进行基于PCR的STR指纹识别，提供了几款试剂盒。其中，Applied Biosystems™ Identifiler™ Plus PCR扩增试剂盒经过专门优化，能够从多种浓度的纯化gDNA制剂分析16种高度变异的人类STR。Applied Biosystems™ Identifiler™ Direct PCR扩增试剂盒最初用于以干血或颊部点样(例如NUCLEIC-CARD采样卡)或口腔拭子为起始材料，分析相同的16个STR基因座。采用NUCLEIC-CARD采样卡时，将卡片尖端1.2 mm直接放入PCR管或孔中，不需纯化即扩增。最后，Applied

Biosystems™ GeneMapper™的GeneMapper™ 5软件可通过预设的等位基因分子量标准和筛选试剂盒覆盖的各种STR等位基因的大小进行STR分析。图1显示了完整的细胞系鉴定工作流程图。

我们最近推出了毛细管电泳仪器系列的新成员SeqStudio基因分析仪，新增了多项可简化毛细管电泳分析的功能。该产品具有全新设计的用户界面、机载触摸屏和集成计算机，方便安装和运行。嵌入式运行模块和可插拔的集成化卡夹，使SeqStudio基因分析仪无需重新配置即可用于Sanger测序或片段分析。该基因分析仪具有最大限度的灵活性和易用性，配合Identifiler试剂盒成为细胞系鉴定等诸多应用的理想平台。



图1. 细胞系鉴定流程。两种方法可鉴定细胞系。(A) 将细胞点在NUCLEIC-CARD采样卡上，用Identifiler Direct试剂盒直接扩增卡片，并使用GeneMapper 5软件在Applied Biosystems™ CE仪器上进行片段分析。**(B)** 也可用Identifiler Plus试剂盒扩增从细胞系中纯化的gDNA，并使用毛细管电泳和GeneMapper 5软件分析片段。

结论

GeneMapper 5软件可将未知片段与已知等位基因大小的STR片段分子量标准(图2, 顶部)进行比对, 简化STR等位基因检出。例如, 含有来自M4A4GFP细胞gDNA的样品在D7S820基因座处有两个峰。GeneMapper 5软件将这两个峰与等位基因分子量标准中的等位基因8和10对齐(其大小分别为大约263bp和271bp), 没有任何其他明显的峰存在

(图2, 底部)。因此, 在这个位置, 这两个等位基因的细胞是杂合的。在D19S433位点, 等位基因14处是单峰, 表明这些细胞于该等位基因是纯合的。

等位基因的组合赋予细胞培养物独有的指纹, 并可与其他培养物或已知样品比较以鉴定此培养物的真实性。

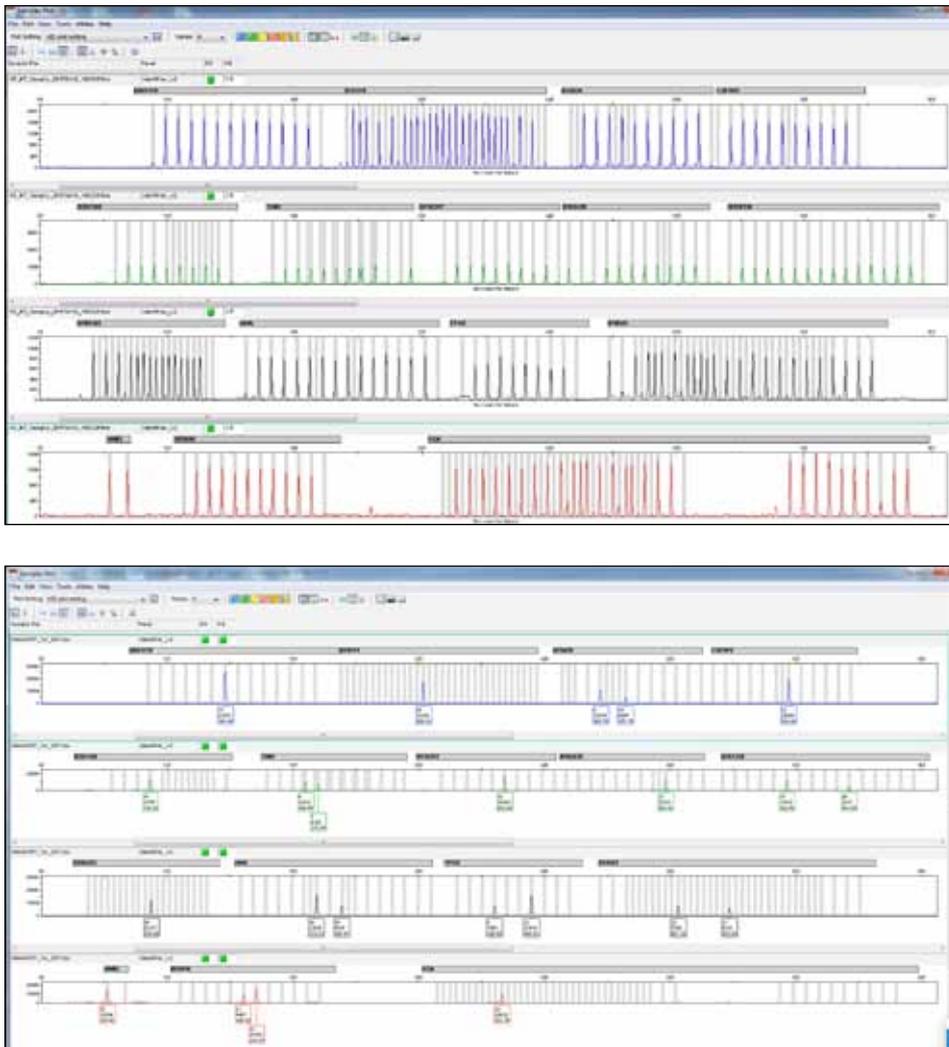


图2. GeneMapMapper 5软件分析比较Applied Biosystems™ Identifiler™ Plus等位基因分子量标准(顶部)和培养物中生长的M4A4GFP细胞的纯化gDNA (底部)。 GeneMapper 5软件使用Identifiler试剂盒中提供的等位基因分子量标准, 将未知样品中存在的等位基因分配给已知的STR等位基因。峰下方的框显示等位基因编号、峰的高度以及碱基对片段大小。

通过分析五种常用细胞系培养物的DNA，揭示Identifiler Plus试剂盒和SeqStudio仪器用于细胞系鉴定的实用性。使用Invitrogen™ RecoverAll™ FFPE样本总核酸分离试剂盒，从细胞沉淀纯化gDNA。由于这类样本不需要脱蜡，我们省略了对应的这些步骤。使用Applied Biosystems™ Quantifiler™ 人类DNA定量试剂盒来测定DNA量。将1 ng纯化的gDNA溶于10μL水，或从3 ng/μL浓度gDNA制备3倍连续稀释的gDNA水溶液，然后按照Identifiler Plus试剂盒用户指南(文件号4440211, 版本F)进行分析。PCR后，在Applied Biosystems™ Hi-Di™ 甲酰胺中将1 μL反应物变性，并上样毛细管电泳仪上。在GeneMapper 5软件中使用导入的人类身份鉴定(HID)分析方法(参见附录)分析片段峰。

执行初步实验。确定在不同仪器平台上进行等位基因检测所需的PCR循环次数(图3)。通过与ATCC数据库中存在的等

位基因(参见后面的章节，采用已知标准品验证细胞真实性)比较，确定这五种细胞系的真实性。在3130型仪器中，27-29个循环产生的精确检出百分比最高，而在SeqStudio和3500仪器上，27-28个循环产生的数据置信度最高。以此，我们建议找出每个实验室中使用的最佳DNA量和最佳的毛细管电泳系统循环数。

通过连续稀释纯化的M4A4GFP gDNA (数据未显示)确定Identifiler Plus试剂盒使用的最小DNA量。当在增反应中使用0.3-3 ng gDNA时结果最准确。值得注意的是，通过增加PCR循环数，我们能够准确分析少量DNA；然而，在DNA浓度较高的情况下，增加PCR循环次数也会导致杂峰增多、精确性下降。因此，为确保等位基因检出结果具有最高的置信度，我们建议在细胞系真实性分析中使用1 ng纯化gDNA。

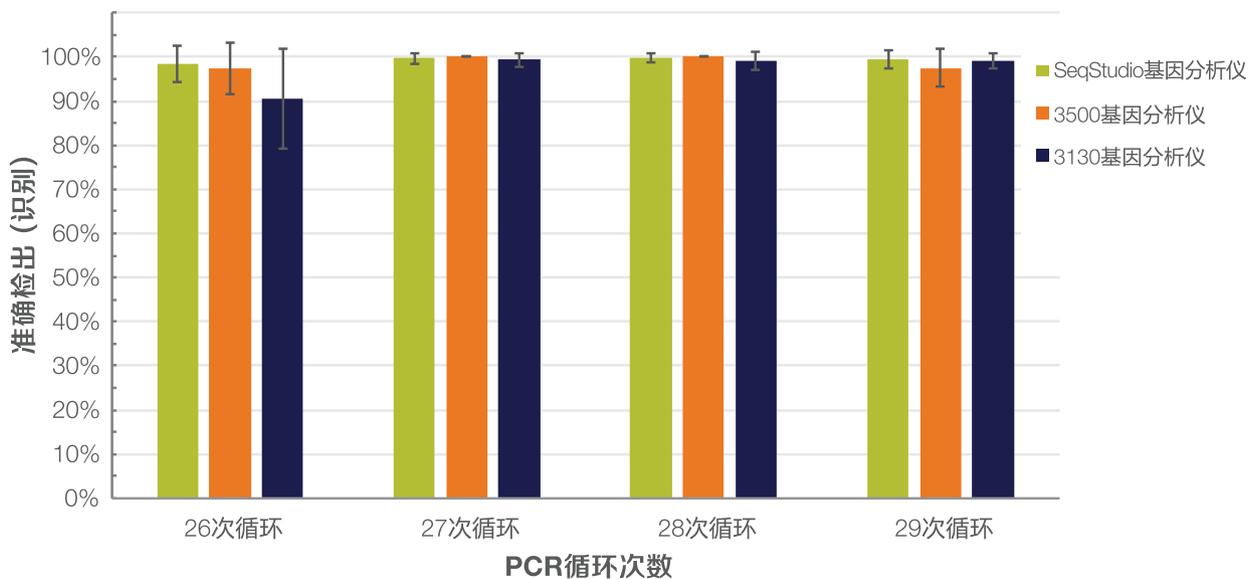


图3. 使用Identifiler Plus试剂盒分析gDNA。使用不同的PCR循环次数分析五种细胞系的1 ng DNA。27-29次循环PCR结果最精确。

Identifiler Direct试剂盒和NUCLEIC-CARD采样卡的使用

鉴定细胞培养物的另一方法是使用NUCLEIC-CARD采样卡。卡片中的基质经过化学处理，可使细胞裂解和蛋白质变性，从而将卡片上的DNA固定并在室温下长期保存。我们制备了几种不同的人细胞系PBS悬浮液(约 5×10^5 个细胞/ml；具体浓度参见图4)。将100 μ l悬浮液直接点在NUCLEIC-CARD采样卡上并干燥过夜。从沾有干燥悬浮液的区域取出单个1.2 mm孔样，放入96孔PCR板孔中。根据用户指南(文件号4415125，版本J)中提供的步骤，将Identifiler Direct试剂盒对照品和试剂添加到平板中，并扩增29次循环。如上所述，在PCR扩增之后，将1 μ l产物在Hi-Di甲酰胺中变性，并使用GeneMapper 5软件在SeqStudio、3500或3130基因分析仪上分析。

将初始细胞悬浮液用PBS连续稀释，确定可分析的最小细胞悬液量(图4)。分析细胞系的未稀释样品可获得完整和准确的图谱。悬浮液稀释10倍后点到NUCLEIC-CARD采样卡上，或造成一些等位基因丢失。等位基因丢失的数量与悬浮液的初始浓度相关，浓度较高的悬浮液丢失率较低，而浓度较低的则有较多的丢失率。稀释100倍时，约有50%的等位基因是可以检测到。我们建议首先将大约 5×10^5 个细胞/ml的悬浮液点到NUCLEIC-CARD采样卡上，以便在检测等位基因时获得最高的置信度。

细胞系	起始浓度 (细胞数/ml)
A549	2.6×10^5
M4A4GFP	4.7×10^5
U2OS	7.2×10^5
HeLa	2.1×10^5
HEK293	5.0×10^5

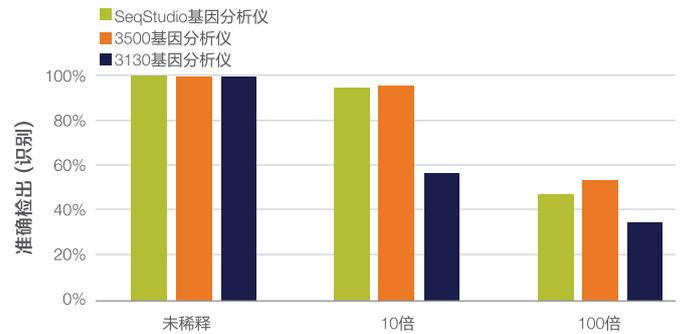


图4. 滴定NUCLEIC-CARD采样卡上干燥的细胞悬液。按照表中所示的浓度制备细胞悬浮液，然后使用NUCLEIC-CARD采样卡和Identifiler Direct试剂盒分析这些悬浮液的连续稀释液。稀释倍数下所有细胞系正确检出等位基因的平均百分比如图所示。

采用已知标准品鉴定细胞的真实性

多家研究机构已认识到需要将未知样品与已知细胞系进行匹配以确定真实性。例如，ATCC已经基于他们的细胞系STR数据库建立了一个网络查询平台(图5, atcc.org/STR_Database.aspx)。使用时只需在测试样本中输入等位基因并选择查询标准，随即返回ATCC数据库中与研究样本中STR相匹配的细胞系列表。同样，Leibniz-Institute DMSZ有一个在线STR查询系统(dsmz.de/fp/cgi-bin/str.html)。值得注意的是，这两个页面只能查询9个基因座(8个常染色体和1个性别区分基因座)，因此并非Identifiler试剂盒的所有基因座都可用此数据库比较。这些额外的基因座对于新细胞系来源的比较是非常有用的。我们已使用这两个数据库验证本应用说明所述的每个细胞系类型。因此，这些机构可以将未知样品中存在的等位基因与已知常用细胞系中的等位基因进行比较，从而简化细胞系鉴定。根据国际细胞系鉴定委员会制定的指南，80-100%等位基因匹配的细胞来自同一个供体，<50%匹配通常意味着细胞来自不同的供体或产地不同[7]。



图5.将等位基因调用上传到ATCC细胞系STR数据库可以鉴定细胞系的类型。输入每个基因座的等位基因，即显示出与STR相匹配的细胞系。值得注意的是，ATCC数据库需要查询8个位点，所有这些位点都包含在Identifiler试剂盒中。该数据库可以访问atcc.org/STR_Database.aspx。详细信息，请参阅ATCC STR数据库页面上的教程。

鉴定培养物中的污染细胞

细胞系鉴定的目标之一是确定细胞系是否被不相关的细胞污染。由于污染细胞可能具有不同于母系的STR谱。在细胞系的混合物中，最终的鉴定结果将反映所有细胞的基因组组合。例如，基因座处的单个峰可以表示：对于同一等位基因，两个细胞类型为纯合；或者一个细胞类型对于等位基因是纯合的，另一个对于除去基因座的缺失是纯合的。两个峰证明，两种细胞对于不同基因座是纯合，对于同一基因座是杂合的等等。尽管解释单个基因座上的异常峰较为困难，但通过分析16个不同基因座可以确定污染细胞系的存在——即使污染细胞系的基因组组成可能不完全可辨。

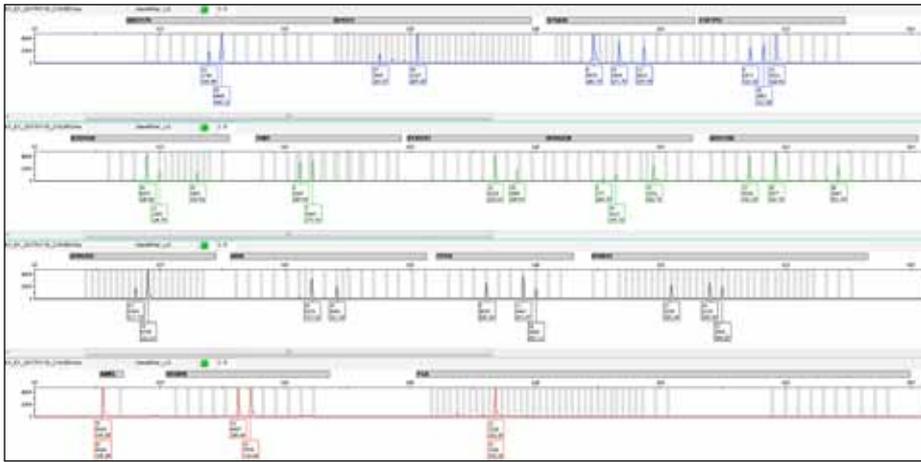
为了测试培养物中污染细胞的检测极限，我们制备了来自M4A4GFP细胞和HeLa培养物的 5×10^5 个细胞的细胞悬浮液，然后混合悬浮液使得它们含有10%，15%，20%，25%，30%和50%的HeLa细胞；将100 ul的每种混合细胞悬浮液点在NUCLEIC-CARD采样卡上。待悬浮液干燥，将卡片(1.2 mm)用Identifiler Direct试剂扩增29次PCR循环，并在三个不同的仪器平台上进行分析。同样，我们将HeLa gDNA与1%，2%，3%，4%，5%和10%比例的M4A4GFP gDNA混合。每种混合物各取1 ng，使用Identifiler Plus试剂盒PCR扩增29次循环后，在SeqStudio基因分析仪上分析。

M4A4GFP和HeLa细胞的等量混合物产生具有明显异常峰的谱(图6A)。例如,在D7S820位点,三个等位基因位于对应位置,而理论上在同质二倍体细胞群中最多具有两个等位基因。其他几个基因座也出现了三个等位基因。仅从单峰和双峰位点推论结论并不可信,但在多个位点同时存在三个等位基因是,便可判断出现异质细胞群体或DNA混合物。在所有的平台上,我们都可以清楚地检测到含有20% HeLa细胞的混合物的异质性(图6B)。随着混合物中HeLa细胞的百分比下降,检测到的HeLa特异性等位基因的数量也下降。然而,在SeqStudio和3500平台上,即使在10%的

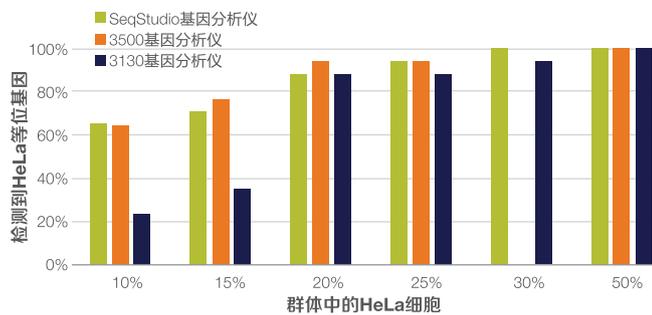
细胞混合物中,这种下降也不太明显,大约65%的HeLa特异性等位基因仍可检测到。使用混合有M4A4GFP DNA的HeLa细胞的较低百分比的gDNA进行类似的分析(图6C)。超过一半的HeLa特异性等位基因可以在仅有4%的HeLa细胞的混合物中检测到,表明Identifiler Plus试剂盒具有高灵敏度。

因此,如果具有完整谱的污染细胞(最高置信度)的数量超过总细胞数的20%,则可以在群体中检测到。如果群体仅有4%的污染细胞,则可检测指示等位基因。

A



B



C

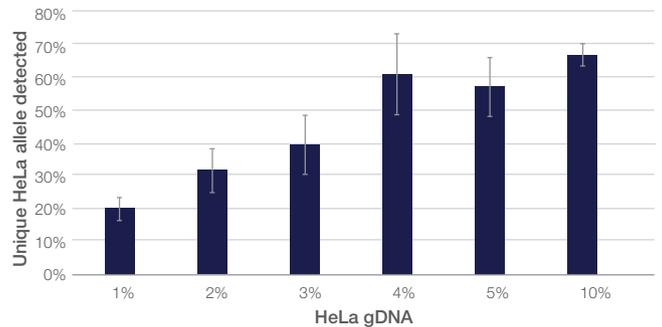


图6. 检测培养物中的污染细胞。(A) 培养物中的外源峰可能预示存在基因组异质性或污染细胞。使用NUCLEIC-CARD采集卡和Identifiler Direct试剂盒分析M4A4GFP和HeLa细胞的1:1混合悬浮液。根据宿主和污染细胞的等位基因组成,受污染的培养物的图谱可能变化。**(B)** 分析NUCLEIC-CARD采集卡上的混合样品。将HeLa细胞和M4A4GFP细胞悬浮液稀释至 5×10^5 个细胞/ml,按指定的比例混合并在NUCLEIC-CARD采集卡上。在指定的仪器上分析相同的均分试样。请注意,3500平台会有30%的细胞未分析。**(C)** 分析混合样品中的gDNA。将来自HeLa细胞的gDNA与来自M4A4GFP细胞的指定比例的gDNA混合。超过一半HeLa细胞特有的等位基因可以在仅含4% HeLa gDNA的混合物中检测到。

细胞系中等位基因的不平衡

使用Identifiler试剂盒分析时，基因型呈现为单峰(纯合)或两个平衡高峰(杂合)的等高高峰。但体外传代的细胞系显示出基因组的不稳定性，并且由于基因座、染色体部分或整个染色体的重复或缺失，姊妹等位基因的峰高可能不同。例如，本研究中使用的HEK293细胞在D8S1179位点处是平衡的，但在CSF1P0和D13S317位点处显示出显著不同的峰高(图7)。而且，当扩增反应中加入的DNA量为100 ng或更少时，随机效应可能会引起峰高不平衡。

为了证明Identifiler试剂盒检测到真正的等位基因不平衡，我们分析了来自正常人类供体的8种不同的标准化gDNA制备物(1 ng)。

我们通过计算每个基因座检测到的等位基因的峰高比来确定等位基因平衡。对于杂合基因座，这个比例是1.0；对于纯合基因座，我们将该比率定义为1.0。我们同时分析了上述细胞系中1 ng gDNA。

标准化的人类gDNA的大多数基因座显示出等位基因平衡：使用1 ng DNA时，其比率非常接近1.0 (图8)。有一些例外情况，两个个体捐赠者有一个比率为1.5的位点(这表明它们可能包含重复的等位基因(3:2))，并且一个个体具有TH01等位基因座的多重重复。然而，细胞系显示出等位基因比率从0.2到4.5不等。该试剂盒能够准确地测量正常的等位基因比率，可能揭示多年传代的细胞系的基因组不稳定性。相反，长期未传代的细胞(如原代细胞或未诱导的干细胞)，不可能表现出等位基因不平衡。



图7. 细胞系中的等位基因不平衡。正常细胞中，杂合子的峰高非常相似；在细胞系中，峰高可能在一个轨迹(红色圆圈)上不同。这可能是由于细胞系的基因组不稳定，导致基因座的重复和缺失，因此显示出不同的等位基因丰度。

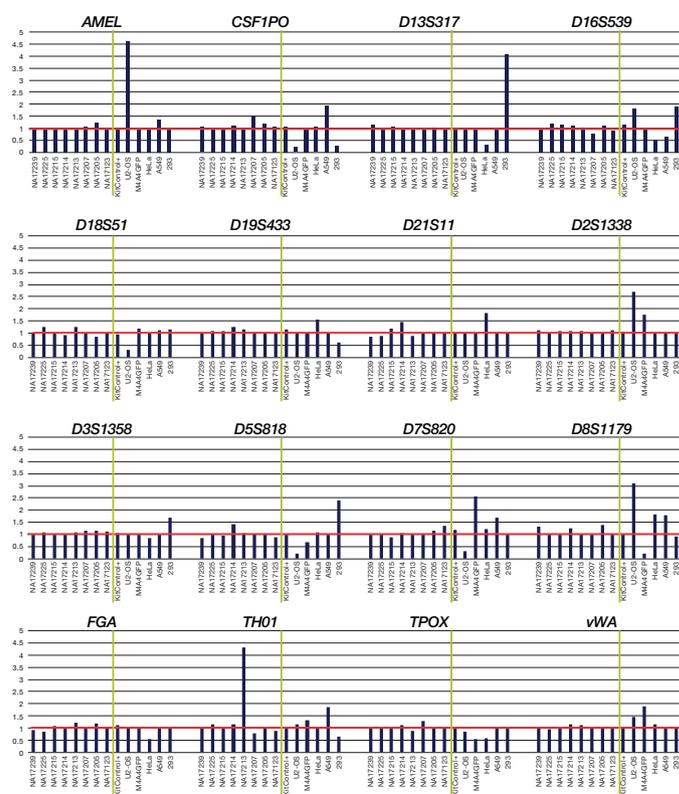


图8. 细胞系显示可测量的等位基因不平衡。用Identifiler Plus试剂盒分析来自人类供体的八个标准gDNA样品(绿线左侧)和五个细胞系样品(绿线右侧)。等位基因比例计算使用杂合子峰值的高度计算，或在纯合子(红线)中定义为1.0。值得注意的是，来自正常人类供体的大多数样品具有接近1.0的等位基因比例，而细胞系样品通常有偏差。

结论

在本应用说明中，我们描述了用于细胞系鉴定的简单且经济的工作流程，证明Identifiler Plus和Identifiler Direct试剂盒可用于区分体外培养细胞中的STR等位基因“指纹”。实际上，由于Identifiler试剂盒可检测15个不同的染色体上的16个STR基因座，因而在从降解样本中检测DNA时，可灵活检测细胞系中的多倍突变，并日益充实细胞系衍生物和数据库。我们证明，仅需0.1 ng gDNA即可获得鉴定结果。但为了确保最高的置信度，我们建议使用1 ng gDNA。此外，我们还可从固定在NUCLEIC- CARD采集卡上100 μ L的 5×10^5 个细胞/ml悬液获得鉴定结果。当gDNA或细胞浓度相对较高时，我们的毛细管电泳仪器性能相当。最后，我们发现与3130和3500型基因分析仪相比，SeqStudio基因分析仪所需的gDNA和细胞浓度更低。值得注意的是，整个毛细管电泳仪器系列的初始样品制备、PCR扩增、毛细管电泳制备和下游分析方法几乎完全相同。这便于研究人员选择最适合其实验室的通量和预算的平台。

参考文献

1. Lorsch JR et al. (2016) Fixing problems with cell lines. *Science* 6216:1452–1453.
2. Huang Y et al. (2017) Investigation of cross-contamination and misidentification of 278 widely used tumor cell lines. *PLoS One* 12(1):e0170384. doi: 10.1371/journal.pone.0170384.
3. He Y et al. (2016) Retracted: Knockdown of tumor protein D52-like 2 induces cell growth inhibition and apoptosis in oral squamous cell carcinoma. *Cell Biol Int* 40:361.
4. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-017.html>
5. Yu M et al. (2015) A resource for cell line authentication, annotation and quality control. *Nature* 520:307–311.
6. Neimark J (2015) Line of attack. *Science* 347: 938–940.
7. http://standards.atcc.org/kwspub/home/the_international_cell_line_authentication_committee-iclac_/Authentication_SOP.pdf

附录

分析所需的GeneMapper模块和设置

在分析细胞系认证项目的FSA文件之前，必须将相应的BIN文件导入GeneMapper 5软件。导入所需文件时，请按照以下说明进行操作。

- AmpFISTRBins_v2.txt和AmpFLSTR_Panels_v2.txt 可以在这里下载：thermofisher.com/us/en/home/technical-resources/software-downloads/genemapper-id-software.html
- 启动GeneMapper 5软件。从菜单栏中选择“Tools” (工具)。
- 选择“Panel Manager” (面板管理器)。

安装“Panel File” (面板文件)方法:

- 在左侧窗口中单击“Panel Manager” (面板管理器)。
- 选择“File” (文件)并从菜单栏导入“Panels” (面板)。
- 随即出现一个对话框。找到计算机上“Panels” (面板)文件的路径。将其放置在默认的“Panels” (面板)文件夹中，对话框随即打开正确的文件夹。
- 选择标题为AmpFISTR_Panels_v2.txt的文件
- 在左窗格中安装“Panel” (面板)。

为本研究中使用的每个仪器运行模块。

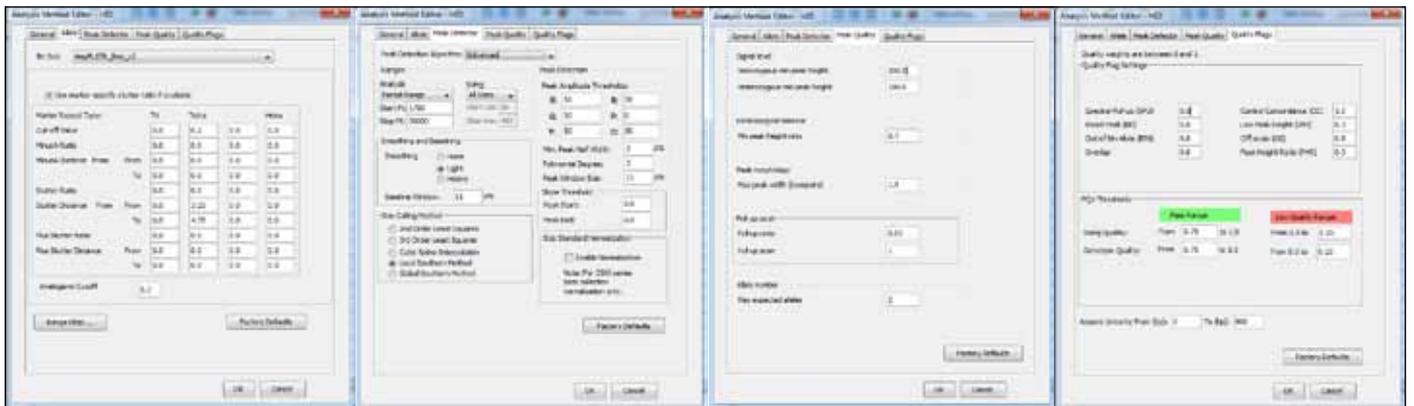
仪器	所用聚合物	加样电压	加样时间	运行电压	运行时间	毛细管长度	运行模式
3130xl	POP-7™	1.2	23	15.0	1,200	36 cm	FragmentAnalysis36_POP7
3500xL	POP-7™	1.6	15	19.5	1,330	50 cm	FragmentAnalysis50_POP7xl
SeqStudio	POP-1™	2.0	10	9.0	1,440	28 cm	FragAnalysis

要安装“Bin文件”：

- 在左侧窗口中单击“ AmpFISTR_Panels_v2”。
- 选择“File”(文件)并从菜单栏导入“Bin Set”。
- 随即出现一个对话框。找到计算机上“Bin”文件的位置。如果您将其放置在默认的“Panels”文件夹中，对话框应打开到正确的文件夹。
- 选择标题为AmpFISTR_Bins_v2.txt的文件
- 在标题为“Bin Set”的顶部的下拉菜单中安装“Bins”。

- 点击“Apply”(应用), 然后点击“OK”(确定)返回到 GeneMapper 5软件。在执行分析时, 可以在“Panel”(面板)下选择Identifiler_v2 panel和相关的“Bin Set”。
- 如要设置细胞系鉴定分析方法, 可以使用图9A中的值作为指导, 修改现有的方法或创建新方法。
- 如要设置绘图设置以显示检测到的等位基因, 可创建一组新设置并输入图9B所示的参数。
- 在基因分型表中查看两个以上的等位基因, 请更改“Table Setting Editor”(表格设置编辑器, 图9C)的“Genotypes”(基因型)选项卡中的“Allele Settings”(等位基因设置)。

A



B



C

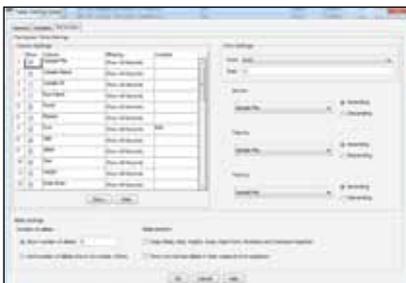


图9. GeneMapper 5软件设置：(A) 细胞系认证分析方法；(B) 细胞系认证图设置；(C) 在基因分型表中查看两个以上等位基因。可以通过打开“Manager Tool”(管理器工具)并选择相应的选项卡来访问这些页面。

订购信息

产品	数量	货号
SeqStudio基因分析仪	1台	A34274
SeqStudio分析软件	1套	4443764
SeqStudio卡夹	500次反应	A33671
SeqStudio起始试剂盒	1个	A35000
SeqStudio SmartStart一天培训	1人次	A34684
3500基因分析仪	1套	4440466
3500xL基因分析仪	1套	4440467
Identifiler Plus PCR扩增试剂盒	100次反应	A26364
Identifiler Direct PCR扩增试剂盒	200次反应	4467831
NUCLEIC-CARD采集卡, 1点样采集	100张	4473973
NUCLEIC-CARD彩色采集卡, 4点样采集	50张	4473978
RecoverAll FFPE样本总核酸分离试剂盒	40次反应	AM1975
GeneMapper 5软件	1份许可证	4475074

更多有关SeqStudio基因分析仪的信息请见 thermofisher.com/seqstudio



赛默飞
官方微信



赛默飞
Applied Biosystems
官方微信

免费服务电话: 800 820 8982/400 820 8982
销售服务信箱: sales.china@thermofisher.com
技术咨询信箱: lifescience-cnts@thermofisher.com

ThermoFisher
SCIENTIFIC