

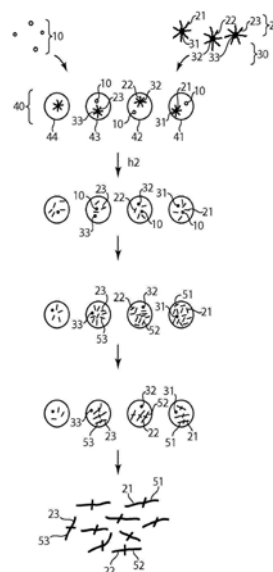


(45) 授权公告日 2021.11.16

序列表5页 附图76页

用于条形码化核酸的系统和方法

本发明总体上涉及微流体和标记的核酸。例如,某些方面总体上涉及用于标记微流体液滴内的核酸的系统和方法。在一组实施方案中,核酸可包含“条形码”或独特的序列,其可用于将液滴中的核酸与另一液滴中的核酸区分开,例如即使在将核酸合并在一起之后也如此。在一些情况下,可以使用颗粒将独特的序列掺入个体液滴中并且连接至包含在液滴内的核酸(例如,从裂解细胞释放的)。在一些情况下,条形码可用于区分数十、数百或甚至数千个核酸,例如源自不同细胞或其它来源的核酸。



1. 一种方法,其包括:

将多个细胞和多个颗粒包封在至少10,000个的多个微流体液滴内,至少一些颗粒包含共价键合到其上的寡核苷酸标签,以使得所述至少10,000个的多个液滴中的液滴含有一个或多个可与包含在所述多个液滴的其它液滴中的寡核苷酸标签区分的寡核苷酸标签;

裂解液滴内的至少一些细胞以从细胞释放核酸;

在至少一些液滴内键合释放的核酸和寡核苷酸标签;和

通过施加光从所述颗粒中释放至少一些所述寡核苷酸标签。

2. 根据权利要求1所述的方法,其包括将所述多个细胞包封在所述多个微流体液滴中,以使得至少90%的所述微流体液滴含有一个细胞或不含细胞;和/或使得至少90%的所述微流体液滴含有一个颗粒或不含颗粒。

3. 根据权利要求1-2中任一项所述的方法,其中所述颗粒以不超过1个颗粒/液滴被包封在所述液滴内。

4. 根据权利要求1-2中任一项所述的方法,其中所述颗粒以不超过0.1个颗粒/液滴被包封在所述液滴内。

5. 根据权利要求1-2中任一项所述的方法,其中所述颗粒以不超过0.01个颗粒/液滴被包封在所述液滴内。

6. 根据权利要求1-2中任一项所述的方法,其中所述细胞以不超过1个细胞/液滴被包封在所述液滴内。

7. 根据权利要求1-2中任一项所述的方法,其中所述细胞以不超过0.1个细胞/液滴被包封在所述液滴内。

8. 根据权利要求1-2中任一项所述的方法,其中所述颗粒中的至少一些是水凝胶颗粒;和/或所述颗粒中的至少一些是聚合物颗粒;和/或所述颗粒中的至少一些是微粒;和/或所述颗粒中的至少一些包含聚丙烯酰胺;和/或所述颗粒中的至少一些包含琼脂糖;和/或所述颗粒中的至少一些包含聚苯乙烯;和/或所述颗粒中的至少一些包含聚-N-异丙基丙烯酰胺;和/或所述颗粒中的至少一些是磁性的;和/或所述多个颗粒具有不超过500微米的平均直径;和/或所述多个颗粒具有至少1微米的平均直径。

9. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些通过丙烯酸亚磷酰胺键与所述颗粒共价键合;和/或所述寡核苷酸标签中的至少一些通过氨基键与所述颗粒共价键合;和/或所述寡核苷酸标签中的至少一些通过生物素-链霉亲和素键与所述颗粒共价键合;和/或所述寡核苷酸标签从其它细胞释放的核酸中独特地鉴定所述释放的核酸;和/或所述寡核苷酸标签中的至少一些包含至少一个条形码序列。

10. 根据权利要求9所述的方法,其中所述条形码序列选自条形码序列库。

11. 根据权利要求10所述的方法,其中所述条形码序列库包括至少10,000个条形码序列。

12. 根据权利要求10所述的方法,其中所述条形码序列库包括至少100,000个条形码序列。

13. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含至少两个条形码序列。

14. 根据权利要求13所述的方法,其中所述至少两个条形码序列选自不同的条形码序

列库。

15. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含引物序列。

16. 根据权利要求15的方法,其中所述寡核苷酸标签中的至少一些包含多于一个引物序列。

17. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含可切割接头。

18. 根据权利要求17所述的方法,其中所述可切割接头中的至少一些是可光切割接头或化学可切割接头或酶促可切割的接头。

19. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含多聚T序列。

20. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含随机DNA序列。

21. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含能够特异性结合基因的序列。

22. 根据权利要求1-2中任一项所述的方法,其中所述寡核苷酸标签中的至少一些包含引物,并且在至少一些液滴内键合释放的核酸和寡核苷酸标签包括使用引物扩增至少一些核酸以产生其中释放的核酸和寡核苷酸标签键合在一起的序列。

23. 根据权利要求1-2中任一项所述的方法,其包括在至少一些液滴内共价键合释放的核酸和寡核苷酸标签;和/或使用酶在至少一些液滴内键合释放的核酸和寡核苷酸标签。

24. 根据权利要求1-2中任一项所述的方法,其还包括将标记有寡核苷酸标签的释放的核酸与从其它液滴中的其它细胞释放的核酸组合。

25. 根据权利要求1-2中任一项所述的方法,其还包括将至少一些裂解的细胞暴露于抗体。

26. 根据权利要求25所述的方法,其中所述抗体是DNA标记的抗体。

27. 根据权利要求1-2中任一项所述的方法,其中至少一些细胞使用细胞裂解试剂裂解。

28. 根据权利要求27所述的方法,其包括将细胞、颗粒和细胞裂解试剂一起包封在微流体液滴内。

29. 根据权利要求1-2中任一项所述的方法,其还包括破坏至少一些液滴。

30. 根据权利要求29所述的方法,其包括在破裂所述液滴之前或之后从所述颗粒释放至少一些所述寡核苷酸标签。

31. 根据权利要求1-2所述的方法,其还包括测序至少一些释放的核酸。

32. 根据权利要求31所述的方法,其还包括使用所述寡核苷酸标签确定至少一些释放的核酸。

33. 一种制品,其包含:

至少10,000个的多个微流体液滴,所述液滴的至少一些含有包含核酸片段的细胞裂解物和多个颗粒,其中所述液滴内的多个核酸片段与寡核苷酸标签结合,其中所述液滴内的寡核苷酸标签与所述10,000个的多个微流体液滴的其它液滴内的寡核苷酸标签是可区分

的,并且其中所述寡核苷酸标签中的至少一些包含光切割接头。

34.根据权利要求33所述的制品,其包含至少100,000个的多个微流体液滴。

35.根据权利要求33所述的制品,其包含至少300,000个微流体液滴。

36.根据权利要求33所述的制品,其包含至少1,000,000个微流体液滴。

37.根据权利要求33-36中任一项所述的制品,其中至少90%的液滴包含一个颗粒或不含颗粒;和/或所述颗粒中的至少一些是水凝胶颗粒;和/或所述颗粒中的至少一些是聚合物颗粒;和/或所述颗粒中的至少一些是微粒;和/或所述颗粒中的至少一些包含聚丙烯酰胺;和/或所述颗粒中的至少一些包含琼脂糖;和/或所述颗粒中的至少一些包含聚苯乙烯;和/或所述颗粒中的至少一些包含聚-N-异丙基丙烯酰胺;和/或所述颗粒中的至少一些是磁性的;和/或所述颗粒具有不超过500微米的平均直径;和/或所述颗粒具有至少1微米的平均直径。

38.根据权利要求33-36中任一项所述的制品,其中所述寡核苷酸标签中的至少一些通过丙烯酸亚磷酰胺键与至少一些所述颗粒共价键合;和/或所述寡核苷酸标签中的至少一些通过氨基键与至少一些所述颗粒共价键合;和/或所述寡核苷酸标签中的至少一些通过生物素-链霉亲和素键与至少一些所述颗粒共价键合;和/或所述寡核苷酸标签中的至少一些包含至少一个条形码序列。

39.根据权利要求38所述的制品,其中所述条形码序列选自条形码序列库。

40.根据权利要求39所述的制品,其中所述条形码序列库包括至少100,000个条形码序列。

41.根据权利要求39所述的制品,其中所述条形码序列库包括至少1,000,000个条形码序列。

42.根据权利要求33-36中任一项所述的制品,其中所述寡核苷酸标签中的至少一些包含至少两个条形码序列。

43.根据权利要求42所述的制品,其中所述寡核苷酸标签的所述至少两个条形码序列选自不同的条形码序列库。

44.根据权利要求33至36中任一项所述的制品,其中所述寡核苷酸标签中的至少一些包含引物序列;和/或所述寡核苷酸标签中的至少一些包含化学可切割接头或酶促可切割的接头;和/或所述寡核苷酸标签中的至少一些包含多聚T序列;和/或所述寡核苷酸标签中的至少一些包含随机DNA序列;和/或所述寡核苷酸标签中的至少一些包含能够特异性结合基因的序列。

用于条形码化核酸的系统和方法

[0001] 相关申请

[0002] 本申请要求于2014年4月21日提交的美国临时专利申请序列号No.61/982,001、于2014年10月17日提交的美国临时专利申请序列号No.62/065,348、于2014年10月20日提交的美国临时专利申请序列号No.62/066,188和于2014年10月30日提交的美国临时专利申请序列号No.62/072,944的权益。这些申请中的每一个通过引用并入本文。

[0003] 政府资金

[0004] 本发明是在由国立卫生研究院授予的资金号R21DK098818的政府支持下进行的。政府对本发明具有一定的权利。

[0005] 领域

[0006] 本发明总体上涉及微流体和标记的核酸。

[0007] 背景

[0008] 后生动物的许多生理学反映在组成细胞中基因表达的时间和空间变化。这种变化中的一些是稳定的,并且已经帮助我们定义成年细胞类型,以及许多发育中的中间细胞类型。其它变化来自动态生理事件,例如细胞周期、细胞微环境的变化、发育、衰老和感染。其它表达变化在性质上似乎是随机的,并且可能具有重要的后果。为了理解在发育和生理学中的基因表达,生物学家的梦想是不仅在RNA水平上,而且在蛋白质水平上映射基因表达变化,和甚至监测每个细胞中的翻译后修饰。

[0009] 目前可用于RNA序列分析(RNA-Seq)的方法具有以高灵敏度定量细胞群中RNA分子丰度的能力。通过一些相当大的努力,已经利用这些方法来分析单个细胞中的RNA含量。限制是有效的方法来分离和处理大量的单个细胞以进行深度RNA测序,以及定量地这样做。这需要在均匀条件下分离细胞,优选具有最小的细胞损失,特别是在临床样品的情况下。对细胞数目、覆盖深度和RNA丰度的测量准确性的要求将取决于实验考虑,其将包括因素诸如获得材料的困难、材料的独特性、细胞群体的复杂性、以及细胞在基因表达空间中多样化的程度。由于缺乏当今的高容量单细胞转录组数据,很难知道所需的覆盖深度,但稀有细胞类型(例如隐藏的肿瘤细胞或组织干细胞亚群)在感兴趣的群体中的存在,结合异质性的其它独立驱动器,例如细胞周期和随机效应,表明了分析大量细胞的需要。

[0010] 虽然通过RNA-seq的RNA丰度的分析是良好建立的,但单细胞RNA-Seq的准确性相比大批测定对其酶促步骤的效率更加敏感;此外,对来自单细胞的PCR或线性扩增的需要冒着引入相当大的误差的风险。对于数千或甚至数万个细胞的并行处理以及有效地处理细胞的小样本以使得几乎每个细胞都被测量也存在较大障碍。在过去十年中,微流体已经成为具有解决这些挑战的潜力的单细胞研究的有前途的技术。然而,目前可以用微流体芯片处理的单细胞的数量仍然低于每次运行70-90个细胞,这在运行成本和细胞保持活力以用于分析的有限时间方面对大量细胞的分析设置了限制。此外,细胞在微流体室中的捕获效率通常较低,这是可用的细胞数量有限的稀有或临床样品的潜在问题。

[0011] 概述

[0012] 本发明总体上涉及微流体和标记的核酸。在一些情况下,本发明的主题涉及相关

的产品、特定问题的替代解决方案和/或一个或多个系统和/或制品的多个不同用途。

[0013] 在一个方面,本发明总体上涉及一种制品。在一组实施方案中,制品包含至少10个的多个微流体液滴,每个液滴含有包括核酸片段的细胞裂解物。在一些情况下,液滴内的多个核酸片段各自结合至寡核苷酸标签。在某些实施方案中,液滴内的寡核苷酸标签可与多个液滴的其它液滴内的寡核苷酸标签区分开。

[0014] 在另一组实施方案中,所述制品包括至少10个的多个微流体液滴,每个液滴含有细胞裂解物。在一些实施方案中,至少约90%的液滴仅含有一个颗粒。在一些情况下,所述颗粒包含与其共价结合的寡核苷酸。

[0015] 根据另一组实施方案,所述制品包含多个颗粒,至少约90%的颗粒包含与其共价键合的寡核苷酸,所述寡核苷酸包含至少2个引物位点和至少2个条形码区域。在一些实施方案中,基于寡核苷酸的条形码区域,至少约90%的颗粒与多个颗粒中的其它颗粒是可区分的。

[0016] 在一组实施方案中,制品包含至少10,000个的多个微流体液滴。在一些实施方案中,至少一些含有细胞裂解物的液滴包括核酸片段。在某些情况下,液滴内的多个核酸片段与寡核苷酸标签结合。在一个实施方案中,液滴内的寡核苷酸标签与10,000个的多个微流体液滴的其它液滴内的寡核苷酸标签是可区分的。

[0017] 在另一组实施方案中,所述制品包括至少10,000个的多个微流体液滴。在一些实施方案中,至少一些液滴含有细胞裂解物。在某些情况下,10,000个的多个微流体液滴中的至少约90%可以仅包含一个颗粒。在一些实施方案中,颗粒可以包含与其共价键合的寡核苷酸。在各种情况下,液滴内的寡核苷酸可以与10,000个的多个微流体液滴的其它液滴内的寡核苷酸区分开。

[0018] 在另一方面,本发明总体上涉及一种方法。在一组实施方案中,所述方法包括将细胞和颗粒包封在微流体液滴内的步骤,所述颗粒包含与其共价键合的寡核苷酸标签,裂解液滴内的细胞以从细胞释放核酸,以及在液滴内键合释放的核酸和寡核苷酸标签。

[0019] 在另一组实施方案中,所述方法包括提供含有细胞的多个微流体液滴的步骤,至少约90%的所述液滴含有一个细胞或不含细胞,裂解所述多个微流体液滴内的细胞以从细胞释放核酸,并将核酸键合至寡核苷酸标签,其中对于至少约90%的液滴如此。在一些情况下,液滴内的寡核苷酸标签可以与多个液滴的其它液滴内的寡核苷酸标签区分开。

[0020] 根据另一组实施方案,所述方法包括步骤:提供多个颗粒,将第一寡核苷酸连接至所述多个颗粒,以使得至少约90%的颗粒具有共价键合至其的仅一个第一寡核苷酸,其中第一寡核苷酸取自至少10个独特的第一寡核苷酸的库;和将第二寡核苷酸连接至第一寡核苷酸,以使得至少约90%的第一寡核苷酸具有共价键合至其的仅一个第二寡核苷酸,其中第二寡核苷酸取自至少10个独特的第二寡核苷酸的库。

[0021] 根据一组实施方案,所述方法包括步骤:将细胞和水凝胶微球或颗粒包封在液滴内,其中水凝胶微球或颗粒具有与其连接的条形码化核酸,裂解液滴内的细胞以释放来自细胞的RNA和/或DNA,并使RNA和/或DNA与条形码化核酸酶促反应。

[0022] 在另一组实施方案中,所述方法包括提供含有细胞的液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴内的细胞以从细胞释放RNA和/或DNA,以及用液滴特异性条形码独特地标记RNA和/或DNA。

[0023] 根据另一组实施方案,所述方法包括步骤:提供含有细胞的液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴内的细胞以从所述细胞中释放RNA和/或DNA,并用选自至少10,000个条形码的库中的条形码独特地标记RNA和/或DNA。

[0024] 在另一组实施方案中,所述方法包括步骤:提供携带核酸的多个微球或颗粒,将寡核苷酸共价连接至微球或颗粒,用随机选自预定义的第一条形码库的第一条形码酶促延伸寡核苷酸,并且用随机选自预定义的第二条形码库的第二条形码酶促延伸寡核苷酸。

[0025] 在一组实施方案中,所述方法包括将多个细胞和多个颗粒包封在至少10,000个的多个微流体液滴内,至少一些所述颗粒包含与其共价键合的寡核苷酸标签,以使得所述至少10,000个的多个液滴含有与所述多个液滴的其它液滴中包含的寡核苷酸标签可区分的一个或多个寡核苷酸标签,裂解液滴内的至少一些细胞以从细胞释放核酸,以及在至少一些液滴内键合释放的核酸和寡核苷酸标签。

[0026] 在另一组实施方案中,所述方法包括提供含有细胞的至少10,000个的多个微流体液滴,至少约90%的所述多个液滴含有一个细胞或不含细胞,裂解所述多个微流体液滴内的细胞以从细胞释放核酸,并将释放的核酸键合至寡核苷酸标签,其中对于至少约90%的液滴,液滴内的寡核苷酸标签与多个液滴的其它液滴内的寡核苷酸标签是可区分的。

[0027] 根据又一组实施方案,所述方法包括将细胞和水凝胶微球或颗粒包封在液滴内,其中水凝胶微球或颗粒可以具有与其连接的条形码化核酸,裂解液滴内的细胞以从细胞释放核酸,并使所释放的核酸与条形码化核酸酶促反应。

[0028] 根据另一组实施方案,所述方法涉及提供含有细胞的至少约10,000个的多个微流体液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴以从细胞释放核酸,以及用液滴特异性条形码独特地标记释放的核酸。

[0029] 在另一组实施方案中,所述方法包括提供含有细胞的液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴内的细胞以从细胞释放核酸,和用选自至少10,000个可区分的条形码的库的条形码独特地标记释放的核酸。

[0030] 在另一组实施方案中,所述方法包括提供含有细胞的至少约10,000个的多个微流体液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴内的细胞以释放核酸,并用液滴特异性条形码独特标记释放的核酸。

[0031] 在另一组实施方案中,所述方法包括提供含有细胞的液滴,以使得不超过10%的液滴含有两个或更多个细胞,裂解所述多个液滴内的细胞以从细胞释放核酸,和用选自至少10,000个可区分的条形码的库的条形码独特地标记释放的核酸。

[0032] 在另一方面,本发明包括制备本文所述的一个或多个实施方案的方法。在另一方面,本发明包括使用本文所述的一个或多个实施方案的方法。

[0033] 当结合附图考虑时,本发明的其它优点和新颖特征将根据本发明的各种非限制性实施方案的以下详细描述而变得明显。在本说明书和通过引用并入的文件包括冲突和/或不一致的公开内容的情况下,以本说明书为准。如果通过引用并入的两个或更多个文件包括彼此相冲突和/或不一致的公开内容,则以具有较晚生效日期的文件为准。

附图说明

[0034] 将参考附图通过示例来描述本发明的非限制性实施方案,附图是示意性的并且不

旨在按比例绘制。在图中,所示的每个相同或几乎相同的部件通常由单个数字表示。为了清楚起见,当举例说明对于本领域普通技术人员理解本发明而言是非必需的时,并非每个部件都在每个图中标出,并且并非本发明的每个实施方案的每个部件都被示出。在附图中:

[0035] 图1示出了根据本发明的一个实施方案的流程图;

[0036] 图2A-2B示出了本发明的另一个实施方案中的微流体装置;

[0037] 图3示出了本发明的另一个实施方案中的液滴内的细胞和颗粒;

[0038] 图4示出了在本发明的另一个实施方案中包含液滴内的细胞和颗粒的微流体通道;

[0039] 图5示出了在本发明的一个实施方案中作为收集时间的函数的样本计数;

[0040] 图6示出了在本发明的另一个实施方案中的读出序列的分布;

[0041] 图7示出了在本发明的另一个实施方案中寡核苷酸标签的产生;

[0042] 图8示出了本发明另一个实施方案中的微流体装置;

[0043] 图9A-9B示出了在本发明的另一个实施方案中的含有寡核苷酸标签的颗粒。

[0044] 图10A-10C示出了在本发明的另一个实施方案中的包含在液滴内的延伸的寡核苷酸标签。

[0045] 图11示出了在本发明的另一个实施方案中的DNA片段的测序;

[0046] 图12示出了根据本发明的另一个实施方案的作为液滴尺寸的函数的逆转录效率;

[0047] 图13A-13H示出了在本发明的一个实施方案中的用于对数千个细胞加DNA条形码的微流体液滴;

[0048] 图14A-14G示出了根据本发明的另一个实施方案的液滴完整性和随机条形码化;

[0049] 图15A-15G示出了在本发明的另一个实施方案中的某些ES细胞群体的异质结构;

[0050] 图16A-16C示出了根据本发明的另一个实施方案生成的基因相关网络;

[0051] 图17A-17H示出了在本发明的另一个实施方案中的分化ES细胞中的时间异质性和群体结构;

[0052] 图18示出了本发明的另一个实施方案中的微流体装置。

[0053] 图19A-19B示出了在本发明的又一个实施方案中的某些微流体装置;

[0054] 图20A-20C示出了在本发明的另一个实施方案中的用于颗粒的寡核苷酸标签的合成;

[0055] 图21A-21H示出了在本发明的一个实施方案中的DNA的定量;

[0056] 图22A-22E示出了在本发明的另一个实施方案中的随机条形码化和独特分子标识符(UMI)过滤;

[0057] 图23A-23D示出了在本发明的一个实施方案中的mES细胞的单细胞基因表达;

[0058] 图24A-24G示出了在本发明另一个实施方案中的mES细胞群的结构。

[0059] 图25示出了在本发明的另一个实施方案中的主要基因的tSNE图谱;

[0060] 图26显示表2;和

[0061] 图27示出了本发明的另一个实施方案中的微流体装置。

[0062] 详述

[0063] 本发明总体上涉及微流体和标记的核酸。例如,某些方面总体上涉及用于标记微流体液滴内的核酸的系统和方法。在一组实施方案中,核酸可包含“条形码”或独特的序列,

其可用于将液滴中的核酸与另一液滴中的核酸区分开,例如即使在将核酸合并在一起之后也如此。在一些情况下,可以使用颗粒将独特的序列掺入个体液滴中并且连接至包含在液滴内的核酸(例如,从裂解细胞释放的)。在一些情况下,条形码可用于区分数十、数百或甚至数千个核酸,例如源自不同细胞或其它来源的核酸。

[0064] 本发明的某些方面总体上涉及用于在微流体液滴或其它合适的隔室内容纳或包封具有寡核苷酸标签的核酸并将它们共价键合在一起的系统和方法。在一些情况下,核酸可以来自裂解的细胞或液滴内的其它材料。液滴内的寡核苷酸标签可以与例如多个液滴或液滴群体内的其它液滴中的寡核苷酸标签区分开。例如,寡核苷酸标签可以含有在各种液滴之间不同的一个或多个独特的序列或“条形码”;因此,可以通过确定与核酸相关的条形码来独特地鉴定每个液滴内的核酸。例如,如果液滴被“破坏”并且来自不同液滴的核酸随后被组合或合并在一起以例如用于测序或其它分析,这可以是重要的。

[0065] 在一些实施方案中,通过最初将寡核苷酸标签连接至颗粒(例如水凝胶或聚合物颗粒),然后在颗粒掺入液滴中之后将它们从颗粒中释放,来将寡核苷酸标签引入液滴中。在一些情况下,可制备颗粒,以使得大多数或所有颗粒相对于具有其它可区分的寡核苷酸标签的其它颗粒仅具有一个独特的可区分的寡核苷酸标签。如果颗粒以1个颗粒/液滴(或更小)的密度存在于液滴内,则一旦寡核苷酸标签从颗粒中释放,则大部分或全部液滴将含有一个独特的寡核苷酸标签(或不含独特的寡核苷酸),从而允许独特地鉴定每个液滴(和其中包含的核酸)。

[0066] 现参考图1,提供了本发明的一个方面的实例。然而,应当理解,这仅是示例性的;本发明的其它实例和实施方案在下面进一步详细讨论。在图1的非限制性实例中,期望分析细胞10的群体,例如通过测序其DNA,通过鉴定可能被怀疑存在于至少一些细胞中的某些蛋白质或基因,通过确定其mRNA或转录组,等。虽然在该实例中细胞用作核酸材料的来源,但是这是举例的方式,在其它实施方案中,可以将核酸从其它来源引入液滴,或使用其它技术。

[0067] 细胞可以首先被包封在一系列微流体液滴40中。本领域普通技术人员将知道用于将细胞包封在微流体液滴内的技术;参见,例如,例如,美国专利号7,708,949、8,337,778、8,765,485或国际专利申请公开号W0 2004/091763和W0 2006/096571,其各自通过引用并入本文。在一些情况下,可以以小于1个细胞/液滴(并且在一些情况下,远远小于1个细胞/液滴)的密度包封细胞,以确保大部分或全部液滴在其中仅具有零个或一个细胞。因此,如图1所示,每个液滴41,42,43,...在其中具有零个或一个细胞。

[0068] 包封在液滴中的还有存在于颗粒30上的寡核苷酸标签20。颗粒30可以是例如微粒,并且可以是水凝胶或聚合物颗粒,或其它类型的颗粒,例如本文所述的那些。颗粒和细胞可以以任何合适的顺序同时或顺序地包封在液滴内。在一组实施方案中,每个颗粒含有独特的寡核苷酸标签,尽管在颗粒上可以存在标签的多个拷贝。例如,每个寡核苷酸标签可以具有存在的一个或多个独特的序列或“条形码”。因此,例如,颗粒31仅含有寡核苷酸标签21的拷贝,颗粒32仅含有寡核苷酸标签22的拷贝,颗粒33仅含有寡核苷酸标签33的拷贝等。在一些情况下,颗粒可以以小于1个颗粒/液滴(并且在一些情况下,远小于1个颗粒/液滴)的密度存在于液滴中,以确保大多数或所有液滴在其中仅存在零个或一个颗粒。此外,在某些实施方案中,寡核苷酸标签可以是可切割的或可以以其它方式从颗粒释放。

[0069] 应当注意,根据本发明的某些实施方案,寡核苷酸标签最初附着于颗粒以便于仅向每个液滴引入一个独特的寡核苷酸标签,如图1所示。(然而,在其它实施方案中,可以存在多个寡核苷酸标签,例如含有相同的独特条形码)。例如,如果颗粒以小于1个颗粒/液滴的密度存在于液滴中,则大多数或所有液滴将各自仅具有单个颗粒,因此仅存在单一类型的寡核苷酸标签。因此,如图1所示,寡核苷酸标签可以被切割或以其它方式从颗粒释放,例如以使得每个液滴41,42,43,...包含独特的寡核苷酸标签21,22,23,...,其不同于可能存在于其它液滴中的其它寡核苷酸标签。因此,存在于液滴内的每个寡核苷酸标签将与存在于其它液滴中的寡核苷酸标签区分开。尽管在图1中使用光(hv)来从颗粒切割寡核苷酸,应当理解,这仅是举例,并且也可以使用其它切割或释放方法,例如本文所讨论的。例如,在一组实施方案中,可以使用含有(例如,物理地)寡核苷酸的琼脂糖颗粒,并且可以通过加热琼脂糖释放寡核苷酸,例如直到琼脂糖至少部分液化或软化。

[0070] 在一些情况下,裂解细胞以从细胞释放核酸或其它材料51,52,53,...。例如,可以使用化学品或超声波裂解细胞。细胞可以释放例如DNA,RNA,mRNA,蛋白质,酶等。在一些情况下,释放的核酸可任选地经历扩增,例如通过包括对扩增方法特异的合适试剂。本领域普通技术人员已知的扩增方法的实例包括但不限于聚合酶链式反应(PCR),逆转录酶(RT)PCR扩增,体外转录扩增(IVT),多重置换扩增(MDA),或定量实时PCR(qPCR)。

[0071] 核酸或其它材料51,52,53,...中的一些或全部可以与存在于液滴中的寡核苷酸标签相关联,例如通过共价键合。例如,核酸或其它材料51,52,53可以连接或酶促附着至存在于液滴中的寡核苷酸标签。因此,如图1所示,液滴41显示连接至寡核苷酸标签21的核酸51,液滴42显示连接至寡核苷酸标签22的核酸52,液滴43显示连接至寡核苷酸标签23的核酸53等。因此,每个液滴内的核酸通过寡核苷酸标签(其在该实例中对于每个液滴是独特的)与多个液滴50的其它液滴内的核酸是可区分的。

[0072] 还应当理解,尽管图1描述了从颗粒切割寡核苷酸标签,然后裂解细胞,但在其它实施方案中,这些不一定以这种顺序发生。例如,细胞裂解可在切割后发生,或两者可同时发生。

[0073] 液滴41,42,43,...可以被“破裂”或“破坏”以释放其内含物,并且在一些情况下,存在于每个液滴中的核酸可以组合或汇集在一起,如图1所示。然而,由于核酸通过不同的寡核苷酸标签标记,来自一个液滴(即,来自一个细胞)的核酸仍然可以使用寡核苷酸标签与来自其它液滴(或其它细胞)的核酸区分开。因此,可以进行组合的核酸库的后续分析(例如,测序),并且可以通过确定不同的寡核苷酸标签来确定每个核酸的来源(例如,个体细胞)。

[0074] 因此,例如,可以以这种方式分析正常细胞和癌细胞的群体(例如,源自组织样品或活检),并且癌细胞可以被鉴定为具有异常DNA,即使存在于大的正常细胞库。例如,由于使用寡核苷酸标签跟踪细胞水平上的DNA的能力,异常DNA仍然可以被鉴定,即使存在大量的正常DNA。作为其它非限制性实例,干细胞可以从正常细胞中分离,或者可以进行感兴趣群体中稀有细胞类型的分离。

[0075] 在另一方面,本发明提供用于从大量细胞平行捕获DNA或RNA和使其条形码化的系统和方法,例如为了分析细胞群体或其它目的(例如本文所述的那些)。在一些实施方案中,这依赖于将条形码化的核酸或其它合适的寡核苷酸标签(例如,连接至颗粒或微球(例如,

水凝胶或聚合物微球))与细胞和/或可用于RNA和/或DNA捕获和/或扩增的其它试剂一起包封。

[0076] 在一组实施方案中,来自基本上每个个体细胞的内容物可以例如用独特条形码(其可以是随机确定的或如本文所讨论的确定的)来标记,这在一些情况下可以允许数百、数千、数万或甚至数十万或更多的不同细胞在单个实验中被条形码化或以其它方式标记,例如以确定或定义群体中的细胞之间的异质性或用于筛选细胞群体等。本文已描述了其它目的。

[0077] 在一组实施方案中,例如在单个反应容器中使用微流体系统将单细胞捕获在单个液滴(例如,50pL至10nL体积)中。可以裂解每个细胞,并且例如通过酶促反应,通过连接等,将其RNA和/或DNA用液滴特异性条形码独特地条形码化或标记。本文还提供了微流体系统的实例包括具有除这些之外的尺寸的微流体系统。在一些实施方案中,一些实施方案还可以用于与RNA或DNA平行地定量单个细胞中的蛋白质丰度,例如通过首先用DNA标记的抗体处理细胞,在这种情况下,DNA标签可以类似地用液滴特异性条形码条形码化。一旦液滴中的细胞组分已条形码化,则液滴可被破坏或破裂,并且样品可以被例如批量处理,以用于高通量测序或其它应用。测序后,可以根据DNA条形码分割或以其它方式分析数据。

[0078] 为了在单个细胞中进行DNA、RNA和/或DNA-抗体标签的平行条形码化,根据一组实施方案,单个水凝胶或聚合物颗粒或微球可以与生物或化学试剂和细胞一起包封在每个液滴中。携带高浓度(例如1至100微摩尔)DNA片段(下文称为“引物”)的颗粒或微球可以编码(a)条形码序列,其随机选自例如至少10,000个条形码(或至少30,000个条形码,至少100,000个条形码,至少300,000个条形码或至少1,000,000个条形码等)的库,其中在颗粒或微球上的所有核酸片段上发现相同的条形码;和/或编码(b)用于DNA或RNA的杂交和捕获的一个或多个引物序列。不同条形码的数量可以是待捕获的细胞数量的至少10倍,并且在一些情况下为至少100倍,以便降低两个或更多个细胞占据具有携带相同条形码的颗粒或者微球的不同液滴的可能性。例如,具有150,000个条形码和1,000个细胞,平均只有3个细胞将获得重复的条形码(导致997个检测到的条形码)。

[0079] 在一些实施方案中,选择包封条件使得这样的液滴包含一个颗粒(或微球)和一个细胞。空液滴和/或具有单个颗粒但不含细胞的液滴和/或具有细胞但没有颗粒的液滴的存在可能基本上不影响性能。然而,在一个液滴中存在两个或更多个颗粒或两个或更多个细胞可能导致难以控制的错误,因此这种事件的发生率在一些情况下保持最小,例如小于约10%或小于约5%。除了细胞和颗粒以外,其它生物和化学试剂可以在液滴之间平均分布。根据特定应用的目的,可以收集和处理共同包封的细胞和颗粒。例如,在一个具体实施方案中,单细胞的DNA或RNA被与颗粒一起引入的引物捕获,然后可在逆转录或其它DNA聚合反应时转化成条形码化的互补DNA。

[0080] 在纯化和任选的DNA扩增后,可以例如通过测序或其它技术来确定细胞核酸的碱基组成和条形码身份。或者,在一些实施方案中,与颗粒或微球一起引入的引物可用于扩增来自基因组的特定核酸序列。

[0081] 在一些实施方案中,使用颗粒或微球引入的条形码化的引物可以通过例如光、化学、酶促或其它技术从其中切割,例如以提高引发液滴中酶促反应的效率。然而,引物的切割可以在任何步骤或点进行,并且在某些情况下可由使用者定义。这种切割在某些情况和/

或条件下可能是特别重要的；例如，单个细胞中的一些部分的RNA和DNA分子可能非常大，或者可能缔合在复合物中，因此不能有效地扩散到颗粒或微球的表面或内部。然而，在其它实施方案中，切割不是必需的。

[0082] 诸如这些的技术可用于分析例如基因组、单核苷酸多态性、特定基因表达水平、非编码RNA、整个转录组(或其部分)、整个基因或其部分等。然而，本发明不应仅限于这些应用。

[0083] 在一个非限制性实施方案中，条形码引物的3'末端以可用于捕获细胞mRNA用于全转录组分析的多聚T序列终止。组合所有细胞的所得文库可任选地使用基于PCR的方法或使用基于杂交捕获的方法(例如Agilent SureSelect)进行富集，例如以允许仅对感兴趣的基因子集测序。在另一个实施方案中，条形码引物的3'末端可以以可用于捕获细胞中的RNA的随机DNA序列终止。在另一个实施方案中，条形码引物的3'末端可以以特定DNA序列终止，例如，所述特定DNA序列可以用于捕获目的DNA或RNA物质(“基因”)，或杂交至除了颗粒或微球之外例如与酶试剂一起被递送至液滴中的DNA探针。在另一个实施方案中，颗粒或微球可以携带多种不同的引物以靶向几种感兴趣的基因。另一个实施方案涉及液滴尺寸和液滴条形码化所需的反应组分浓度的优化。

[0084] 本发明的另一方面总体上涉及产生附着于颗粒或微球的条形码化的核酸。核酸可以附着于颗粒或微球的表面，或者在一些情况下附着在或掺入颗粒内。例如，可以在颗粒形成期间，例如物理和/或化学地将核酸掺入颗粒中。

[0085] 例如，一组实施方案通常涉及产生携带核酸片段(各自编码条形码、引物和/或可能用于核酸的捕获、扩增和/或测序的其它序列)的颗粒或微球。微球可以指1至500微米尺寸或例如本文所述的其它尺寸的水凝胶颗粒(聚丙烯酰胺，琼脂糖等)或胶体颗粒(聚苯乙烯、磁性或聚合物颗粒等)。在一些实施方案中，微球可以是多孔的。可以使用的其它合适的颗粒或微球在本文中更详细地讨论。

[0086] 在一些情况下，携带DNA的颗粒或微球的制备可以依赖于将初始DNA寡核苷酸并入颗粒或微球的共价连接或其它技术，随后通过一个或多个例如随机选自预定义库的条形码酶促延伸每个寡核苷酸。在一些情况下，可能的独特条形码的最终数量可以取决于预定义的条形码库的大小和/或延伸步骤的数量。例如，使用384个预定义的条形码和2个延伸步骤的库，每个颗粒或微球携带 $384^2=147,456$ 个可能条形码中的一个；使用3个延伸步骤，每个颗粒或微球携带 $384^3=56,623,104$ 个可能条形码中的一个；等等。在一些情况下也可以使用其它数量的步骤；此外，每个库可以具有各种数量的预定义条形码(不只是384)，并且库可以具有相同或不同数量的预定义条形码。库可以包括相同和/或不同的序列。

[0087] 因此，在一些实施方案中，所使用的可能的条形码由条形码元件的一个或多个单独“库”(随后例如使用分割-合并法将它们连接在一起以产生最终条形码)形成。库可以包含例如至少约300，至少约500，至少约1,000，至少约3,000，至少约5,000或至少约10,000个可区分的条形码。例如，第一库可以包含 x_1 个元件，第二库可以包含 x_2 个元件；形成包含来自第一库的元件和来自第二库的元件的条形码可以产生例如可使用的 x_1x_2 个可能的条形码。应当注意， x_1 和 x_2 可以相等或不相等。该过程可以重复任何次数；例如，条形码可以包含来自第一库、第二库和第三库(例如，产生 $x_1x_2x_3$ 个可能的条形码)或来自第一库、第二库、第三库和第四库的元件(例如，产生 $x_1x_2x_3x_4$ 个可能的条形码)等。还可以有5,6,7,8或任何其它合

适数量的库。因此,由于组合的潜在数量,即使相对少量的条形码元件也可用于产生数量远更多的可区分的条形码。

[0088] 在一些情况下,多个库的这种使用组合地可用于产生显著大量的可用条形码,而不必个别地单独制备和合成大量的条形码。例如,在许多现有技术系统中,需要100或1,000个条形码将需要单独合成100或1,000个条形码。然而,如果需要更大数量的条形码,例如对于待研究的更大数量的细胞,则需要合成相应地更大数量的条形码。这样的系统在较大数量(例如10,000,100,000或1,000,000个条形码)下变得不切实际且不可行。然而,通过使用条形码的单独“库”,可以实现更大数量的条形码,而不必要求每个条形码被单独合成。作为非限制性示例,可以合成1,000个可区分的条形码(或任何其它合适数量)的第一库和1,000个可区分的条形码的第二库,从而需要合成2,000个条形码(或者如果条形码在每个库中被重新使用,则仅需要1,000个),但是它们可以组合以产生 $1,000 \times 1,000 = 1,000,000$ 个可区分的条形码,例如,其中每个可区分的条形码包括从第一库取得的第一条形码和从第二库取得的第二条形码。使用3个、4个或更多个库来组装条形码可以导致可制备的甚至更大数量的条形码,而基本上不增加将需要合成的可区分的条形码的总数。

[0089] 在一些方面,可以使用多种技术(包括光,温度,化学和/或酶处理)从颗粒或微球释放DNA片段或寡核苷酸。例如,通过使用光,可以在选择的时间和/或在期望的条件下释放核酸片段,从而为其使用提供灵活性。

[0090] 在一些实施方案中,颗粒或微球可以长时间储存并作用于后续应用的试剂。

[0091] 在另一方面,本发明提供了用于从大量单细胞平行捕获、条形码化和定量一组数十至数百或更多的特异性DNA和/或RNA序列的系统和方法,例如用于分析细胞群体的目的或其它目的。某些实施方案依赖于例如连接至颗粒例如水凝胶或聚合物微球的条形码化的核酸与用于例如RNA和/或DNA捕获和扩增的细胞和/或其它试剂的包封。

[0092] 在一些情况下,用于利用独特的随机条形码标记来自个体细胞的特定组的基因(例如,数十个或数百个基因或在一些情况下更多个基因)的系统和方法,从而允许例如在单个实验中标记或条形码化数百、数千或甚至数万或更多的不同细胞,以用于定义群体中细胞之间的异质性或筛选细胞群体或用于其它目的。

[0093] 例如,在通过多路高通量测序分析大量细胞的情况下,在一些实施方案中,可能需要关注感兴趣的基因的子集(例如数十至数百个基因),而不是全转录组或全基因组捕获和测序。

[0094] 一些实施方案涉及关注细胞DNA或RNA的特定序列的细胞内容物的平行条形码化。这些可包括例如DNA条形码化的微球(或其它颗粒)的合成,和/或这种微球用于例如在单个反应容器中捕获和条形码化个体液滴(例如,体积为50pL至10nL,或本文所述的其它体积)中的单个细胞。在一些情况下,可以裂解基本上每个细胞,并且用液滴特异性核酸条形码(例如通过酶促反应)独特地条形码化(标记)其RNA和/或DNA。在一些实施方案中,修饰DNA条形码化的微球可以以这样的方式进行,即它们仅靶向特定的一组DNA序列,而不是使用一种感兴趣的序列或使用随机序列。这可以允许将高浓度的序列特异性条形码化引物递送到每个液滴中,这在一些情况下可允许互补DNA的酶促条形码化和合成主要发生在感兴趣的序列上。这可以使用例如任何酶促方法,其中一组序列特异性引物可用于捕获感兴趣的基因。

[0095] 本发明的一些实施方案可以用于与RNA或DNA平行地定量单个细胞中的蛋白质丰度,例如,通过首先用DNA标记的抗体处理细胞,在这种情况下,颗粒或微球上的一个或多个序列或寡核苷酸可以与由抗体递送的DNA标签互补。在一些情况下,一旦液滴中的细胞组分已经被条形码化,则液滴可以被破坏或破裂,并且样品可以例如被批量处理,用于诸如高通量测序的应用。在测序后,在某些实施方案中,根据DNA条形码分割数据,从而提供关于类型、序列、分子计数、感兴趣的核酸和/或蛋白质的来源等的信息。

[0096] 根据另一方面,本发明提供了优化用于在小体积内酶促处理细胞的反应条件,例如,对于其中反应的直接测试将极其缓慢(因为其将需要产生多个微流体装置,或者运行具有大量测试样品的微流体装置)的情况。在一些情况下,这还可以具体报告在将来自在小体积中裂解的单细胞的mRNA酶促逆转录为互补DNA所需的理想体积上。

[0097] 本发明的某些实施方案提供了使用体积大于5微升的反应(其可以使用常规的分子生物学试剂例如在没有微流体装置的情况下进行)优化单细胞上的微流体反应。这在某些应用中可能是有用的,例如,用于测试参数例如多个数量级的反应体积,其原本将需要多个测试微流体装置的设计和合成和这种装置的性能的并排比较。它还可用于快速优化微流体反应的条件,例如不同反应组分的最佳浓度。

[0098] 在一组实施方案中,将批量反应用于模拟微流体体积中存在的精确条件。这是一般性的,并且可以应用于优化微流体反应的其它方面或其它反应。例如,这可以应用于测试不同添加剂减轻逆转录(RT)反应的抑制的能力,并且在某些实施方案中可以定义从在小体积中裂解的单细胞进行RT反应所必需的DNA引物浓度。

[0099] 上述讨论是本发明的各种实施方案的非限制性示例。然而,其它实施方案也是可能的。因此,更一般地,本发明的各个方面涉及用于在微流体液滴内标记核酸的系统和方法的各种系统和方法,如下所述。

[0100] 在一个方面,本发明总体上涉及用于标记液滴(例如微流体液滴)群体内的核酸的系统和方法。在一些情况下,微流体液滴可以具有小于约1mm的液滴的平均直径和/或微流体液滴可以是基本上单分散的,例如,如本文所讨论的。

[0101] 在一些情况下,包含DNA和/或其它核酸的寡核苷酸标签可以附着于颗粒并递送至液滴。在一些情况下,寡核苷酸标签附着于颗粒以控制其至液滴中的递送,例如,以使得液滴将通常在其中具有至多一个颗粒。在一些情况下,在递送到液滴中时,可以例如通过裂解、通过降解颗粒等从颗粒中去除寡核苷酸标签。然而,应当理解,在其它实施方案中,液滴可以含有2、3个或任何其它数目的颗粒,其可具有相同或不同的寡核苷酸标签。

[0102] 寡核苷酸标签可以是任何合适的长度或包含任何合适数量的核苷酸。寡核苷酸标签可以包含DNA、RNA和/或其它核酸如PNA,和/或这些和/或其它核酸的组合。在一些情况下,寡核苷酸标签是单链的,尽管在其它情况下它可以是双链的。例如,寡核苷酸标签可以具有至少约10nt,至少约30nt,至少约50nt,至少约100nt,至少约300nt,至少约500nt,至少约1000nt,至少约3000nt,至少约5000nt,至少约10000nt等的长度。在一些情况下,寡核苷酸标签可以具有不超过约10,000nt,不超过约5000nt,不超过约5000nt,不超过约3000nt,不超过约1000nt,不超过约500nt,不超过约300nt,不超过约100nt,不超过约50nt等的长度。任何这些的组合也是可能,例如,寡核苷酸标签可以在约10nt至约100nt之间。寡核苷酸标签的长度不是关键的,并且在各种实施方案中可以使用多种长度。

[0103] 寡核苷酸标签可以含有多种序列。例如,寡核苷酸标签可以含有一个或多个引物序列、一个或多个独特的或“条形码”序列、一个或多个启动子序列、一个或多个间隔序列等。在一些实施方案中,寡核苷酸标签还可以包含一个或多个可切割的间隔物,例如可光切割的接头。寡核苷酸标签可以化学地(例如,通过接头)或物理地(例如,不一定需要接头)连接至颗粒,例如,以使得可以通过切割从颗粒中去除寡核苷酸标签。其它实例包括可用于增加寡核苷酸标签(例如,使用特定序列或无义序列)的体积,用于促进处理(例如,标签可包括多聚A尾),用于增加结合的选择性(例如,如下所述),用于促进通过酶(例如,合适的连接酶)的识别,用于促进鉴定等的部分。这些和/或其它序列的实例在本文中进一步详细描述。

[0104] 作为实例,在一些实施方案中,寡核苷酸标签可以包含“条形码”或独特的序列。可以选择序列,以使得一些或大多数寡核苷酸标签(例如,存在于颗粒上和/或液滴中的)具有独特的序列(或独特的序列组合),但是其它寡核苷酸标签(例如,其它颗粒或液滴上的)不具有独特的序列或序列组合。因此,例如,所述序列可以用于独特地鉴定或区分液滴或来自液滴的所包含的核酸(例如,来自裂解的细胞)与其它液滴或来自其它液滴的其它核酸(例如,从其它细胞释放的)。

[0105] 序列可以是任何合适的长度。条形码序列的长度不是关键的,并且可以是足以将条形码序列与其它条形码序列区分开的任何长度。一个、两个或更多个“条形码”序列可以存在于寡核苷酸标签中。条形码序列可以具有5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25nt的长度。在一些情况下也可存在多于25个核苷酸。

[0106] 在一些情况下,独特的或条形码序列可以取自潜在条形码序列的“库”。如果在寡核苷酸标签中存在多于一个条形码序列,则条形码序列可以取自相同或不同的潜在条形码序列库。序列库可以使用任何合适的技术(例如,随机地)选择,或者使得序列允许错误检测和/或校正,例如通过间隔一定距离(例如,汉明距离)以使得可以检测和在一些情况下校正条形码序列的读取中的错误。库可以具有任何数量的潜在条形码序列,例如至少100,至少300,至少500,至少1,000,至少3,000,至少5,000,至少10,000,至少30,000,至少50,000,至少100,000,至少300,000,至少500,000或至少1,000,000个条形码序列。

[0107] 在一些情况下,寡核苷酸标签可以含有一个或多个启动子序列,例如以允许标签的产生,允许酶促扩增等。本领域普通技术人员将知晓引物序列,例如P5或P7。许多这样的引物序列可商购获得。启动子的实例包括但不限于T7启动子、T3启动子或SP6启动子。

[0108] 在一些情况下,寡核苷酸标签可以含有一个或多个引物序列。通常,引物是用作核酸合成的起始点的单链或部分双链核酸(例如,DNA),从而允许聚合酶例如核酸聚合酶延伸引物并复制互补链。引物可以与靶核酸互补并与靶核酸杂交。在一些实施方案中,引物是合成引物。在一些实施方案中,引物是非天然存在的引物。引物通常具有10至50个核苷酸的长度。例如,引物可以具有10至40、10至30、10至20、25至50、15至40、15至30、20至50、20至40或20至30个核苷酸的长度。在一些实施方案中,引物具有18至24个核苷酸的长度。引物的实例包括但不限于P5引物、P7引物、PE1引物、PE2引物、A19引物或本文所讨论的其它引物。

[0109] 在一些情况下,寡核苷酸标签可以含有无义或随机序列,例如以增加寡核苷酸标签的质量或大小。随机序列可以具有任何合适的长度,并且可存在一个或多于一个。作为非限制性实例,随机序列可以具有10至40、10至30、10至20、25至50、15至40、15至30、20至50、20至40或20至30个核苷酸的长度。

[0110] 在一些情况下,寡核苷酸标签可以包含能够特异性结合基因或其它实体的一个或多个序列。例如,在一组实施方案中,寡核苷酸标签可以包含能够识别mRNA的序列,例如,含有多聚T序列的序列(例如,在一列中具有几个T,例如4、5、6、7、8或更多个T)。

[0111] 在一组实施方案中,寡核苷酸标签可以含有一个或多个可切割的接头,例如,其可以在施加合适的刺激后被切割。例如,可切割序列可以是可通过应用光或合适的化学物质或酶来切割的可光切割的接头。可光切割的接头的非限制性实例可见于图20A。在一些情况下,例如,可以制备多个颗粒(例如,在其表面上包含寡核苷酸标签)并将其添加到液滴中,例如,以使得平均来说每个液滴含有一个颗粒或在某些情况下更少(或更多)。在加入到液滴中之后,可以例如使用光或其它合适的切割技术从颗粒切割寡核苷酸标签,以允许寡核苷酸标签变得存在于溶液中,即在液滴内部。以这种方式,在一些实施方案中可通过将颗粒加载到液滴中来容易地将寡核苷酸标签加载到液滴中,然后将其切割掉以允许寡核苷酸标签处于溶液中,例如以与核苷酸或其它物质相互作用,例如在本文中所讨论的。

[0112] 另外,在一组实施方案中,寡核苷酸标签可以包含抗体,例如,其可以特异性结合被怀疑存在于细胞(或液滴)中的蛋白质。例如,液滴可以含有一种或多种用本文所述的寡核苷酸标签标记的抗体。

[0113] 寡核苷酸标签可以连接至颗粒,例如,如本文所讨论的。在一些实施方案中,颗粒可仅包含一种寡核苷酸标签,尽管寡核苷酸标签的多个拷贝可存在于颗粒上;其它颗粒可包含可通过例如使用本文所述的条形码序列区分的不同寡核苷酸标签。

[0114] 可以使用任何合适的方法将寡核苷酸标签连接至颗粒。连接的确切方法不是关键的,并且可以是例如化学或物理的。例如,寡核苷酸标签可以通过生物素-链霉亲和素键、氨基键或丙烯酸亚磷酰胺键与颗粒共价键合。参见例如图20A的丙烯酸亚磷酰胺键的实例。在另一组实施方案中,可以将寡核苷酸掺入(例如物理地)颗粒中,其中可以通过改变颗粒来释放寡核苷酸。因此,在一些情况下,寡核苷酸不需要具有可切割的连接。例如,在一组实施方案中,可以在形成颗粒后将寡核苷酸掺入颗粒例如琼脂糖颗粒中。在颗粒降解(例如,通过加热颗粒直到其开始软化、降解或液化)后,寡核苷酸可以从颗粒中释放。

[0115] 在本发明的某些方面,颗粒是微粒。颗粒可以是多种类型中的任何一种;如所讨论的,颗粒可以用于将特定的寡核苷酸标签引入液滴中,并且可以使用寡核苷酸标签可以与其缔合(例如,物理或化学地)的任何合适的颗粒。颗粒的确切形式不是关键的。颗粒可以是球形或非球形的,并且可以由任何合适的材料形成。在一些情况下,使用具有基本上相同的组成和/或基本上相同的平均直径的多个颗粒。多个或一系列颗粒的“平均直径”是每个颗粒的平均直径的算术平均值。本领域普通技术人员将能够例如使用激光散射、显微镜检查或其它已知技术来确定多个或一系列颗粒的平均直径(或其它特征尺寸)。在非球形颗粒中的单个颗粒的平均直径是具有与非球形颗粒相同的体积的完美球体的直径。颗粒(和/或多个或一系列颗粒)的平均直径可以例如小于约1mm,小于约500微米,小于约200微米,小于约100微米,小于约75微米,小于约50微米,小于约25微米,小于约10微米或在一些情况下小于约5微米。平均直径也可以是至少约1微米,至少约2微米,至少约3微米,至少约5微米,至少约10微米,至少约15微米或在一些情况下是至少约20微米。

[0116] 在一组实施方案中,颗粒可以是水凝胶颗粒。水凝胶颗粒包括含有DNA的水凝胶颗粒的实例参见例如标题为“Assay and other reactions involving droplet”的国际专利

申请公开号W02008/109176(通过引用并入本文)。水凝胶的实例包括但不限于基于琼脂糖或丙烯酰胺的凝胶,例如聚丙烯酰胺、聚-N-异丙基丙烯酰胺或聚-N-异丙基聚丙烯酰胺。例如,单体的水溶液可以分散在液滴中,然后聚合,例如以形成凝胶。另一个实例是水凝胶,例如可以通过加入钙离子而胶凝的藻酸。在一些情况下,可以例如通过与水相共流动、通过油相共流动或通过两种不同滴剂的聚结,将凝胶引发剂(用于丙烯酰胺的过硫酸铵和TEMED或用于藻酸盐的 Ca^{2+})加入到液滴中,例如,如下列文件中所讨论的:由Link等人在2006年2月23日提交的题为“Electronic Control of Fluidic Species”的美国专利申请系列号11/360,845,其在2007年1月4日公开为美国专利申请公开号2007/000342;或由Ahn等人于2007年1月24日提交的题为“Fluidic Droplet Coalescence”的美国专利申请系列号11/698,298;其各自通过引用将全部内容并入本文。

[0117] 在另一组实施方案中,颗粒可包含一种或多种聚合物。示例性聚合物包括但不限于聚苯乙烯(PS),聚己内酯(PCL),聚异戊二烯(PIP),聚乳酸,聚乙烯,聚丙烯,聚丙烯腈,聚酰亚胺,聚酰胺和/或这些和/或其它聚合物的混合物和/或共聚物。此外,在一些情况下,颗粒可以是磁性的,这可以允许颗粒的磁性操纵。例如,颗粒可以包括铁或其它磁性材料。颗粒也可以被官能化,以使得它们可以具有连接的其它分子,例如蛋白质、核酸或小分子。因此,本发明的一些实施方案涉及定义例如核酸、蛋白质、小分子或例如本文所述的其它物质的文库的一组颗粒。在一些实施方案中,颗粒可以是荧光的。

[0118] 在一组实施方案中,形成含有细胞或其它核酸来源和例如包含如上所述的寡核苷酸标签的颗粒的液滴。可以选择任何合适的方法来产生液滴,并且用于形成液滴的多种不同的技术将是本领域普通技术人员已知的。例如,通道的接合部可以用于产生液滴。该接合部可以是例如T形接合部、Y形接合部、通道在通道内的接合部(例如,在同轴布置中,或者包括内部通道和围绕至少一部分内部通道的外部通道)、交叉(或“X”)接合部、流动聚焦接合部或用于产生液滴的任何其它合适的接合部。参见,例如,由Link等人于2004年4月9日提交的题为“Formation and Control of Fluidic Species”的国际专利申请号PCT/US2004/010903,其于2004年10月28日公开为W02004/091763,或由Stone等人于2003年6月30日提交的题为“Method and Apparatus for Fluid Dispersion”的国际专利申请号PCT/US2003/020542,其于2004年1月8日公开为W0 2004/002627,这些专利申请通过引用整体并入本文。在一些实施方案中,接合部可以被配置和布置成产生基本上单分散的液滴。液滴也可以在流体装置上产生,和/或液滴可以单独产生,然后被送至装置。

[0119] 如果使用细胞,细胞可以来自任何合适的来源。例如,细胞可以是希望研究或测序来自细胞的核酸的任何细胞等,并且可以包括一种或多于一种细胞类型。细胞可以是例如来自特定细胞群,例如来自某一器官或组织(例如,心脏细胞,免疫细胞,肌肉细胞,癌细胞等),来自特定个体或物种的细胞(例如,人细胞,小鼠细胞,细菌等),来自不同生物体的细胞,来自天然存在的样品(例如库塘水,土壤等)的细胞等。在一些情况下,细胞可从组织分离。

[0120] 此外,本发明的某些实施方案涉及使用其它离散的隔室,例如微孔板的微孔、载玻片或其它表面上的个体点等。在一些情况下,每个隔室可以在不会意外地与其它隔室混合的特定位置中。在一些情况下,隔室可以相对较小,例如,每个隔室可以具有小于约1ml,小于约300微升,小于约100微升,小于约30微升,小于约10微升,小于约3微升,小于约1微升,

小于约500n1,小于约300n1,小于约100n1,小于约50n1,小于约30n1或小于约10n1的体积。

[0121] 在一组实施方案中,加载液滴(或其它隔室),以使得平均来说每个液滴在其中具有少于1个颗粒。例如,平均加载率可以小于约1个颗粒/液滴,小于约0.9个颗粒/液滴,小于约0.8个颗粒/液滴,小于约0.7个颗粒/液滴,小于约0.6个颗粒/液滴,小于约0.5个颗粒/液滴,小于约0.4个颗粒/液滴,小于约0.3个颗粒/液滴,小于约0.2个颗粒/液滴,小于约0.1个颗粒/液滴,小于约0.05个颗粒/液滴,小于约0.03个颗粒/液滴,小于约0.02个颗粒/液滴或小于约0.01个颗粒/液滴。在一些情况下,可以选择较低的颗粒加载率以使将在其中产生具有两个或更多个颗粒的液滴的概率最小化。因此,例如,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约95%,至少约97%,至少约98%,或至少约99%的液滴可以不含有颗粒或仅含有一个颗粒。

[0122] 类似地,在一些实施方案中,加载液滴(或其它隔室),以使得平均来说每个液滴在其中具有少于1个细胞。例如,平均加载率可以小于约1个细胞/液滴,小于约0.9个细胞/液滴,小于约0.8个细胞/液滴,小于约0.7个细胞/液滴,小于约0.6个细胞/液滴,小于约0.5个细胞/液滴,小于约0.4个细胞/液滴,小于约0.3个细胞/液滴,小于约0.2个细胞/液滴,小于约0.1个细胞/液滴,小于约0.05个细胞,小于约0.03个细胞/液滴,小于约0.02个细胞/液滴或小于约0.01个细胞/液滴。在一些情况下,可以选择较低的细胞加载率以使将在其中产生具有两个或更多个细胞的液滴的概率最小化。因此,例如,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约95%,至少约97%,至少约98%,或至少约99%的液滴可以不含细胞或仅含有一个细胞。此外,应当注意,颗粒的平均加载率和液滴内的细胞的平均加载率可以相同或不同。

[0123] 在一些情况下,可以产生相对大量的液滴,例如至少约10,至少约30,至少约50,至少约100,至少约300,至少约500,至少约1,000,至少约3,000,至少约5,000,至少约10,000,至少约30,000,至少约50,000,至少约100,000个液滴等。在一些情况下,如前所述,一些或全部液滴可以例如基于存在于至少一些液滴中的寡核苷酸标签(例如,其可以包含一个或多个独特的序列或条形码)而是可区分的。在一些情况下,至少约50%,至少约60%,至少约70%,至少约80%,至少约90%,至少约95%,至少约97%,至少约98%,或至少约99%的液滴是可区分的。

[0124] 根据本发明的某些方面,在将颗粒和细胞加载到液滴中之后,寡核苷酸标签可以从颗粒释放或切割。如上所述,可以使用任何合适的技术(例如光(例如,如果寡核苷酸标签包括可光切割的接头)、化学品或酶等)从液滴中释放寡核苷酸标签。如果使用化学品或酶,可以在形成液滴后将化学品或酶引入液滴中,例如通过皮下注射或其它方法例如题为“Fluid Injection”的国际专利申请公开号W0 2010/151776(通过引用并入本文)中讨论的那些,通过液滴与含有化学品或酶的液滴的融合,或通过本领域普通技术人员已知的其它技术。

[0125] 如所讨论的,在某些方面,液滴可以含有核酸。核酸可以来自细胞或来自其它合适的来源。在一组实施方案中,如果存在细胞,则细胞可以在液滴内裂解,例如以从细胞释放DNA和/或RNA,和/或在液滴内产生细胞裂解物。例如,可以通过暴露于裂解化学物质或细胞裂解试剂(例如,表面活性剂例如Triton-X或SDS,酶例如溶菌酶,溶葡萄球菌酶,麦芽糖酶,纤维素酶,变溶菌素,聚糖酶,蛋白酶,甘露聚糖酶,蛋白酶K等)或物理条件(例如,超声,紫

外光,机械搅拌等)来裂解细胞。如果使用裂解化学品,则可以在形成液滴后将裂解化学品引入液滴中,例如通过皮下注射或其它方法例如于2011年12月21日提交的题为“Fluid Injection”的美国专利申请系列号13/379,782(于2012年5月31日公开为美国专利申请公开号2012/0132288,其全文通过引用并入本文),通过将液滴与包含化学品或酶的液滴融合,或通过本领域普通技术人员已知的其它技术。细胞的裂解可以在寡核苷酸标签从颗粒释放之前、期间或之后发生。在一些情况下,裂解细胞将导致细胞释放其内容物,例如细胞核酸、蛋白质、酶、糖等。在一些实施方案中,一些细胞核酸也可以连接至包含在液滴内的一个或多个寡核苷酸标签(例如,如本文所讨论的)。例如,在一组实施方案中,通常在细胞内产生的RNA转录物可以被释放,然后与核酸标签连接。

[0126] 在一些实施方案中,一旦释放,来自细胞的释放的核酸(例如,DNA和/或RNA)可以例如通过引物延伸、通过连接等键合(例如共价地)到寡核苷酸标签。可以使用多种不同技术中的任一种,并且本领域普通技术人员将意识到许多这样的技术。所使用的精确连接技术不一定是关键的,并且可以在实施方案之间变化。

[0127] 例如,在某些实施方案中,可以使用连接酶将核酸与寡核苷酸标签连接。连接酶的非限制性实例包括DNA连接酶,例如DNA连接酶I,DNA连接酶II,DNA连接酶III,DNA连接酶IV,T4 DNA连接酶,T7 DNA连接酶,T3 DNA连接酶,大肠杆菌DNA连接酶,Taq DNA连接酶,等。许多这样的连接酶可以商购获得。作为另外的实例,在一些实施方案中,可以使用退火或引物延伸法将两个或更多个核酸连接在一起。

[0128] 在另一组实施方案中,核酸可以与寡核苷酸标签连接和/或使用PCR(聚合酶链式反应)或其它合适的扩增技术(包括本文所述的那些)扩增。通常,在PCR反应中,加热核酸以引起核酸解离成单链,并且使用热稳定性DNA聚合酶(例如Taq聚合酶)来扩增核酸。该过程通常重复多次以扩增核酸。

[0129] 在一组实施方案中,可以在液滴内进行PCR或核酸扩增。例如,液滴可以含有聚合酶(例如Taq聚合酶)和DNA核苷酸,并且可以处理液滴(例如,通过重复加热和冷却)以扩增液滴内的核酸。聚合酶和核苷酸可以在任何合适的点加入,例如在将编码各种条件的各种核酸加入到液滴之前、期间或之后。例如,液滴可以含有聚合酶和DNA核苷酸,其与液滴融合以允许发生扩增。本领域普通技术人员将知道合适的PCR技术和变型,例如组装PCR或聚合酶循环组件,其可以在一些实施方案中用于产生扩增的核酸。这种程序的非限制性实例也在下面讨论。此外,在一些情况下,合适的引物可用于引发聚合,引物例如P5和P7、或本领域普通技术人员已知的其它引物。在一些实施方案中,可以将引物加入到液滴中,或者引物可以存在于液滴内的一个或多个核酸上。本领域普通技术人员将知道合适的引物,其中许多可容易地商购获得。

[0130] 在一些情况下,液滴可以被破裂、破坏或以其它方式破碎。本领域普通技术人员可获得用于“破坏”或“破裂”液滴的各种各样的方法,并且所选择的确切方法不是关键的。例如,可以使用诸如机械破碎或超声波的技术破碎携带流体中包含的液滴。液滴也可以使用化学试剂或表面活性剂例如1H,1H,2H,2H-全氟辛醇来破碎。

[0131] 然后可将来自不同液滴的核酸(标记有寡核苷酸标签)合并或组合在一起或分析(例如测序,扩增等)。然而,来自不同液滴的核酸可由于在破碎之前存在于每个液滴中的不同的寡核苷酸标签(例如包含不同的条形码)的存在而保持可区分。

[0132] 例如,可以使用PCR(聚合酶链式反应)或其它扩增技术扩增核酸。通常,在PCR反应中,加热核酸以引起核酸解离成单链,并且使用热稳定性DNA聚合酶(例如Taq聚合酶)来扩增核酸。该过程通常重复多次以扩增核酸。

[0133] 在一组实施方案中,PCR可用于扩增核酸。本领域普通技术人员将知道合适的PCR技术和变型,例如组装PCR或聚合酶循环组件,其可以在一些实施方案中用于产生扩增的核酸。这种程序的非限制性实例也在下面讨论。此外,在一些情况下,合适的引物可用于引发聚合,引物例如P5和P7、或本领域普通技术人员已知的其它引物。本领域普通技术人员将知道合适的引物,其中许多可容易地商购获得。

[0134] 可以使用的本领域普通技术人员已知的扩增方法的其它非限制性实例包括但不限于逆转录酶(RT)PCR扩增,体外转录扩增(IVT),多重置换扩增(MDA)或定量实时PCR(qPCR)。

[0135] 在一些实施方案中,可以使用多种技术和仪器对核酸进行测序,其中许多是可容易商购获得的。这些技术的实例包括但不限于链终止测序,通过杂交测序,Maxam-Gilbert测序,染料-终止子测序,链终止法,大规模平行签名测序(Lynx Therapeutics),polony测序,焦磷酸测序,连接测序,离子半导体测序,DNA纳米球测序,单分子实时测序,纳米孔测序,微流Sanger测序,数字RNA测序(“数字RNA-seq”)等。所选择的确切测序方法不是关键的。

[0136] 另外,在一些情况下,液滴还可以含有一种或多种DNA标记的抗体,例如,以例如通过用DNA适当标记来确定细胞中的蛋白质。因此,例如,可以使用对蛋白质特异的DNA标记的抗体在本文所讨论的多个细胞中检测蛋白质。

[0137] 关于用于在微流体系统中操纵液滴的系统和方法的其它细节遵循例如用于测定液滴(或液滴内的物质)、分选液滴等。例如,用于筛选和/或分选液滴的各种系统和方法描述于由Link等人在2006年2月23日提交的题为“Electronic Control of Fluidic Species”的美国专利申请系列号11/360,845,其在2007年1月4日公开为美国专利申请公开号2007/000342,其通过引用将全部内容并入本文。作为非限制性实例,通过施加(或移除)第一电场(或其一部分),液滴可被引导到第一区域或通道;通过将第二电场施加(或移除)到装置(或其一部分),可将液滴引导到第二区域或通道;通过将第三电场施加到装置(或其一部分),液滴可以被引导到第三区域或通道;等等,其中电场可以以某种方式不同,例如在强度、方向、频率、持续时间等方面不同。

[0138] 在本发明的某些实施方案中,提供传感器,其可以以允许确定流体液滴的一个或多个特性的方式感测和/或确定流体液滴的一个或多个特性,和/或包含流体液滴的流体系统的一部分(例如,围绕流体液滴的液体)的特性。本领域普通技术人员可以鉴定可相对于液滴确定并且可用于本发明的特性。这种特征的非限制性实例包括荧光、光谱学(例如光学,红外,紫外等)、放射性、质量、体积、密度、温度、粘度、pH、物质诸如生物物质(例如,蛋白质,核酸等)的浓度等等。

[0139] 在一些情况下,传感器可以连接至处理器,处理器进而导致对流体液滴执行操作,例如通过对液滴进行分选、从液滴添加或去除电荷、使液滴与另一液滴融合、分割液滴、引起液滴内发生混合等,例如,如前所述的。例如,响应于流体液滴的传感器测量,处理器可以使流体液滴分割、与第二流体液滴合并等。

[0140] 一个或多个传感器和/或处理器可以被定位成与流体液滴感测连通。如本文所使用的“感测连通”意味着传感器可以定位在任何地方,以使得可以以某种方式感测和/或确定流体系统内(例如,通道内)的流体液滴和/或包含流体液滴的流体系统的一部分。例如,传感器可以与流体液滴和/或包含流体液滴的流体系统的一部分流体地、光学地或视觉地、热地、气动地、电子地等感测连通。传感器可以定位在流体系统附近,例如嵌入在通道的壁内或一体地连接至通道的壁,或者与流体系统分开定位,但是与流体系统进行物理、电和/或光通信,以便能够感测和/或确定流体液滴和/或包含流体液滴的流体系统的一部分(例如,通道或微通道,包含流体液滴的液体等)。例如,传感器可以没有与包含液滴的通道的任何物理连接,但是可以定位成检测从液滴或流体系统产生的电磁辐射,例如红外线、紫外线或可见光。电磁辐射可以由液滴产生,和/或可以从流体系统的其它部分(或者流体系统的外部)产生,并且以例如通过吸收、反射、衍射、折射、荧光、磷光、极性变化、相变、相对于时间的变化等指示流体液滴的一个或多个特性的方式与流体液滴和/或包含流体液滴的流体系统的一部分相互作用。作为示例,激光可以指向流体液滴和/或围绕流体液滴的液体,并且可以确定流体液滴和/或周围的液体的荧光。如本文所使用的“感测连通”还可以是直接或间接的。作为示例,来自流体液滴的光可以被引导到传感器,或者在被引导到传感器之前首先被引导通过光纤系统、波导等。

[0141] 可用于本发明的传感器的非限制性实例包括基于光学或电磁学的系统。例如,传感器可以是荧光传感器(例如,由激光激发)、显微镜系统(其可以包括相机或其它记录装置)等。作为另一示例,传感器可以是电子传感器,例如能够测定电场或其它电特性的传感器。例如,传感器可以检测流体液滴和/或包含流体液滴的流体系统的一部分的电容、电感等。

[0142] 如本文所使用的,“处理器”或“微处理器”是能够从一个或多个传感器接收信号、存储信号和/或引导一个或多个响应(例如,如上所述)(例如,通过使用数学公式或电子或计算电路)的任何组件或设备。该信号可以是指示由传感器确定的环境因素的任何合适的信号,例如气动信号,电子信号,光信号,机械信号等。

[0143] 在一组实施方案中,可以通过在液滴上产生电荷和/或电偶极子,并且使用施加的电场操纵液滴来引导流体液滴,施加的电场可以是AC场、DC场等。作为示例,可以根据需要选择性地施加和移除电场(或者可以施加不同的电场,例如,反向电场)以将流体液滴引导到特定区域。在一些实施方案中,可以根据需要选择性地施加和移除电场,而不实质上改变包含流体液滴的液体的流动。例如,液体可以在基本上稳态的基础上(即,包含流体液滴的液体的平均流速偏离稳态流动或者液体流动相对于时间的预期值的小于20%或小于15%,并且在一些情况下,平均流速可以偏离小于10%或小于5%)或在其它预定的基础上流动通过本发明的流体系统(例如,通过通道或微通道),并且包含在液体内的流体液滴可以例如使用电场被引导到各个区域,而基本上不改变通过流体系统的液体的流动。

[0144] 在一些实施方案中,可以通过改变包含液滴的液体的流动在本发明的流体系统内对流体液滴进行筛选或分选。例如,在一组实施方案中,可以通过将围绕流体液滴的液体引导到第一通道、第二通道等中来引导或分选流体液滴

[0145] 在另一组实施方案中,可以控制流体系统内例如不同通道内或通道的不同部分内的压力,以引导流体液滴的流动。例如,液滴可以被引导朝向包括用于进一步的流动方向

(例如,被引导朝向限定任选的下游流动通道的通道中的分支或分叉)的多个选件的通道接合部。可控制一个或多个任选的下游流动通道内的压力以将液滴选择性地引导到通道中的一个中,并且可以在连续的液滴到达接合部所需的时间量级上实现压力变化,以使得可以独立控制每个连续液滴的下游流动路径。在一种布置中,液体储存器的膨胀和/或收缩可以用于将流体液滴引导或分选到通道中,例如通过引起包含流体液滴的液体的定向运动。液体储存器可以定位成使得当被激活时,由激活的储存器引起的液体的运动导致液体沿优选方向流动,从而在该优选方向上携带流体液滴。例如,液体储存器的膨胀可以引起朝向储存器的液体流动,而液体储存器的收缩可以引起液体远离储存器流动。在一些情况下,液体储存器的膨胀和/或收缩可以与其它流动控制装置和方法(例如,如本文所述的)组合。能够引起液体储存器的膨胀和/或收缩的装置的非限制性实例包括活塞和压电部件。在一些情况下,压电部件可能是特别有用的,因为它们具有相对快速的响应时间,例如响应于电信号。在一些实施方案中,流体液滴可以被分选到多于两个通道中。

[0146] 如上所述,某些实施方案总体上涉及用于分选液体中的流体液滴的系统和方法。例如,可以以某种方式(例如,如本文进一步描述的)感测和/或确定液滴的性质,然后可以将液滴引导朝向装置的特定区域,例如微流体通道,例如,用于分选目的。在一些情况下,使用本发明的某些系统和方法可以实现高分选速度。例如,在一些情况下,可以确定和/或分选至少约10个液滴/秒,并且在其它情况下,可以确定和/或分选至少约20个液滴/秒,至少约30个液滴/秒,至少约100个液滴/秒,至少约200个液滴/秒,至少约300个液滴/秒,至少约500个液滴/秒,至少约750个液滴/秒,至少约1,000个液滴/秒,至少约1,500个液滴/秒,至少约2,000个液滴/秒,至少约3,000个液滴/秒,至少约5,000个液滴/秒,至少约7,500个液滴/秒,至少约10,000个液滴/秒,至少约15,000个液滴/秒,至少约20,000个液滴/秒,至少约30,000个液滴/秒,至少约50,000个液滴/秒,至少约75,000个液滴/秒,至少约100,000个液滴/秒,至少约150,000个液滴/秒,至少约200,000个液滴/秒,至少约300,000个液滴/秒,至少约500,000个液滴/秒,至少约750,000个液滴/秒,至少约1,000,000个液滴/秒,至少约1,500,000个液滴/秒,至少约2,000,000或更多个液滴/秒,或至少约3,000,000或更多个液滴/秒。

[0147] 在一些方面,可以使用相对小的液滴的群体。在某些实施方案中,作为非限制性实例,液滴的平均直径可以小于约1mm,小于约500微米,小于约300微米,小于约200微米,小于约100微米,小于约75微米,小于约50微米,小于约30微米,小于约25微米,小于约20微米,小于约15微米,小于约10微米,小于约5微米,小于约3微米,小于约2微米,小于约1微米,小于约500nm,小于约300nm,小于约100nm或小于约50nm。在某些情况下,液滴的平均直径也可以为至少约30nm,至少约50nm,至少约100nm,至少约300nm,至少约500nm,至少约1微米,至少约2微米,至少约3微米,至少约5微米,至少约10微米,至少约15微米或至少约20微米。液滴群体的“平均直径”是液滴直径的算术平均值。

[0148] 在一些实施方案中,取决于具体应用,液滴可以具有基本上相同的形状和/或尺寸(即,“单分散”)或具有不同的形状和/或尺寸。在一些情况下,液滴可以具有横截面直径的均匀分布,即液滴可以具有这样的直径分布,以使得不超过约5%,不超过约2%或不超过约1%的液滴具有小于约90%(或小于约95%,或小于约99%)和/或大于约110%(或大于约105%,或大于约101%)的多个液滴的总平均直径的直径。用于产生液滴的横截面直径的均

匀分布的一些技术公开于由Link等人于2004年4月9日提交的题为“Formation and Control of Fluidic Species”的国际专利申请号PCT/US2004/010903,其于2004年10月28日公开为W0 2004/091763,其通过引用并入本文。

[0149] 本领域普通技术人员将能够例如使用激光散射或其它已知技术确定液滴群体的平均直径。如此形成的液滴可以是球形的,或在某些情况下是非球形的。在非球形液滴中的液滴的直径可以被认为具有与非球形液滴相同的体积的完美数学球体的直径。

[0150] 在一些实施方案中,可以通过在被液体包围的流体上产生电荷来在通道内产生一个或多个液滴,这可以使流体分离在液体内的个体液滴中。在一些实施方案中,可以向流体施加电场以使液滴形成发生。流体可以作为在液体内的一系列单独的带电和/或电诱导的液滴存在。可以使用任何合适的技术在液体内的流体中产生电荷,例如通过将流体置于电场(可以是AC、DC等)内,和/或通过引起使流体具有电荷的反应发生。

[0151] 在一些实施方案中,电场从电场发生器(即,能够产生可施加到流体的电场的装置或系统)产生。电场发生器可以产生AC场(即,相对于时间周期性变化的场,例如正弦波、锯齿波、方波等)、DC场(即,相对于时间恒定的场)、脉冲场等。用于产生合适的电场(其可以是AC、DC等)的技术是本领域普通技术人员已知的。例如,在一个实施方案中,通过在一对电极上施加电压来产生电场,所述电极可以位于通道附近,以使得电场的至少一部分与通道相互作用。电极可以由本领域普通技术人员已知的任何合适的电极材料形成,所述材料包括但不限于银、金、铜、碳、铂、铜、钨、锡、镉、镍、铟锡氧化物(“ITO”)等,及其组合。

[0152] 在另一组实施方案中,可以通过以能够引起流体形成个体液滴的方式改变通道尺寸来从由通道内的液体包围的流体产生流体液滴。通道可以例如是相对于流动方向变宽的通道,例如,以使得流体不粘附到通道壁而是形成个体液滴,或者相对于流动方向变窄的通道,例如,以使得流体被迫聚结成个体液滴。在一些情况下,通道尺寸可以导致个体液滴的形成发生的方式相对于时间改变(例如,机械地或机电地,气动地等)。例如,通道可以机械地收缩(“挤压”)以引起液滴形成,或者流体流可以被机械地破坏(例如通过使用移动挡板、旋转叶片等)以引起液滴形成。产生液滴的其它技术包括例如流体的混合或涡旋。

[0153] 某些实施方案总体上涉及用于将液滴分割成两个或更多个液滴的系统和方法。例如,可以使用施加的电场来分割液滴。液滴可以具有比周围液体更大的电导率,并且在一些情况下,液滴可以是中性电荷的。在某些实施方案中,在施加的电场中,可以促使电荷从液滴的内部迁移至待分布在其上的表面,从而可以消除液滴内部中经历的电场。在一些实施方案中,液滴表面上的电荷也可能由于所施加的电场而经受力,这导致具有相反极性的电荷在相反的方向上迁移。在一些情况下,电荷迁移可导致液滴被拉开成两个单独的液滴。

[0154] 本发明的一些实施方案总体上涉及用于将两个或更多个液滴融合或聚结成一个液滴的系统和方法,例如,当两个或更多个液滴例如由于组成、表面张力、液滴尺寸、表面活性剂的存在或不存在等而通常不能融合或聚结时。在某些情况下,相对于液滴尺寸,液滴的表面张力还可以防止发生液滴的融合或聚结。

[0155] 作为非限制性实例,两个液滴可以被给予相反的电荷(即,正电荷和负电荷,不一定具有相同的量级),这可以增加两个液滴的电相互作用,以使得液滴的融合或聚结由于它们相反的电荷而发生。例如,电场可以被施加到液滴,液滴可以通过电容器,化学反应可以使液滴变得带电等。在一些情况下,液滴可能不能够融合,即使施加表面活性剂以降低液滴

的表面张力。然而,如果液滴以相反的电荷(其可以是但不一定是相同的量级)带电,则液滴可以能够融合或聚结。作为另一个实例,液滴可以不必须被给予相反的电荷(并且在一些情况下,可以不被给予任何电荷),并且通过使用在液滴中诱导的导致液滴聚结的偶极子而融合。此外,允许聚结的两个或更多个液滴不一定需要满足“迎头相对的”。只要初始发生液滴的至少一些融合,则任何接触角度都是足够的。还参见例如由Ahn等人于2007年1月24日提交的题为“Fluidic Droplet Coalescence”的美国专利申请系列号11/698,298,其于2007年8月23日公开为美国专利申请公开号2007/0195127,通过引用将其全部内容并入本文。

[0156] 在一组实施方案中,流体可以被注入到液滴中。在一些情况下,可以例如使用微针或其它这样的装置将流体微注射到液滴中。在其它情况下,当液滴与流体通道接触时,可以使用流体通道将流体直接注入到液滴中。流体注射的其它技术公开于例如由Weitz等人于2010年6月25日提交的题为“Fluid Injection”的国际专利申请号PCT/US2010/040006,其于2010年12月29日公开为W0 2010/151776;由Weitz等人于2009年12月18日提交的题为“Particle-Assisted Nucleic Acid Sequencing”的国际专利申请号PCT/US2009/006649,其于2010年7月15日公开为W0 2010/080134,这些专利申请通过引用整体并入本文。

[0157] 根据本发明的某些方面,多种材料和方法可用于形成制品或组件(诸如本文所述的那些),例如通道诸如微流体通道、腔室等。例如,各种制品或组件可以由固体材料形成,其中通道可以通过微加工、膜沉积工艺如旋涂和化学气相沉积、激光制造、光刻技术、蚀刻法包括湿化学或等离子体工艺等形成。参见,例如,Scientific American,248:44-55,1983 (Angell等人)。

[0158] 在一组实施方案中,本文所述制品的各种结构或组件可由聚合物形成,所述聚合物例如弹性体聚合物,例如聚二甲基硅氧烷(“PDMS”),聚四氟乙烯(“PTFE”或**Teflon**[®])等。例如,根据一个实施方案,可以通过使用PDMS或其它软光刻技术单独制造流体系统来实现微流体通道(适合于此实施方案的软光刻技术的细节在Younan Xia和George M.Whitesides的题为“Soft Lithography”,公开于Annual Review of Material Science,1998,第28卷,第153-184页以及George M.Whitesides,Emanuele Ostuni,Shuichi Takayama,Xingyu Jiang和Donald E.Ingber的题为“Soft Lithography in Biology and Biochemistry”,公开于Annual Review of Biomedical Engineering,2001,第3卷,第335-373页的参考文献中讨论;这些参考文献中的每一个通过引用并入本文)。

[0159] 潜在合适的聚合物的其它实例包括但不限于聚对苯二甲酸乙二醇酯(PET),聚丙烯酸酯,聚甲基丙烯酸酯,聚碳酸酯,聚苯乙烯,聚乙烯,聚丙烯,聚氯乙烯,环烯烃共聚物(COC),聚四氟乙烯,氟化聚合物,硅酮诸如聚二甲基硅氧烷,聚偏二氯乙烯,双苯并环丁烯(“BCB”),聚酰亚胺,聚酰亚胺的氟化衍生物等。也设想了包括聚合物(包括上述的那些)的组合、共聚物或掺合物。装置还可以由复合材料(例如,聚合物和半导体材料的复合材料)形成。

[0160] 在一些实施方案中,制品的各种结构或部件由聚合和/或柔性和/或弹性体材料制成,并且可以方便地由可硬化流体形成,从而便于经由模制(例如复制成型、注射成型、铸塑成型等等)制造。可硬化流体可以基本上是可以被诱导固化或自发固化成能够容纳和/或输送预期用于流体网络或与流体网络一起使用的流体的固体的任何流体。在一个实施方案中,可硬化流体包括聚合物液体或液体聚合物前体(即“预聚物”)。合适的聚合物液体可包

括例如热塑性聚合物、热固性聚合物、蜡、金属或其在高于其熔点的温度下加热的混合物或复合物。作为另一个实例,合适的聚合物液体可以包括一种或多种聚合物在合适溶剂中的溶液,该溶液在除去溶剂(例如通过蒸发)后形成固体聚合物材料。可以例如从熔融状态固化或通过溶剂蒸发固化的这种聚合物材料是本领域普通技术人员公知的。对于其中一个或两个母模模具由弹性体材料组成的实施方案,各种聚合物材料(其中许多是弹性体的)是合适的,并且也适用于形成模具或母模模具。这种聚合物的实例的非限制性列表包括一般类别的硅酮聚合物、环氧聚合物和丙烯酸酯聚合物的聚合物。环氧聚合物的特征在于存在通常被称为环氧基团、1,2-环氧化物或环氧乙烷的三元环醚基团。例如,除了基于芳族胺、三嗪和脂环族主链的化合物之外,还可以使用双酚A的二缩水甘油醚。另一个实例包括公知的酚醛清漆聚合物。适用于根据本发明使用的硅酮弹性体的非限制性实例包括由包括氯硅烷如甲基氯硅烷、乙基氯硅烷、苯基氯硅烷、十二烷基三氯硅烷等的前体形成的那些。

[0161] 在某些实施方案中使用硅酮聚合物,例如,硅酮弹性体聚二甲基硅氧烷。PDMS聚合物的非限制性实例包括由Dow Chemical Co., Midland, MI以商品名Sylgard销售的那些,特别是Sylgard 182, Sylgard 184和Sylgard 186。包括PDMS的硅酮聚合物具有几种使本发明的各种结构的制造简化的有益性质。例如,这种材料便宜、容易获得,并且可以通过用热硬化而从预聚物液体固化。例如,PDMS通常通过将预聚物液体暴露于约例如约65°C至约75°C的温度进行例如约1小时的暴露时间而硬化。此外,硅酮聚合物例如PDMS可以是弹性体的,因此可用于形成具有相对高的纵横比的非常小的特征,这在本发明的某些实施方案中是必需的。在这方面,柔性(例如,弹性体)模具或母模可以是有利的。

[0162] 由硅酮聚合物(例如PDMS)形成结构诸如微流体结构或通道的一个优点是这些聚合物被氧化的能力,例如通过暴露于含氧等离子体如空气等离子体,以使得氧化结构在其表面含有能够与其它氧化的硅酮聚合物表面或与各种其它聚合和非聚合材料的氧化表面交联的化学基团。因此,可以制造结构,然后氧化并且基本上不可逆地密封至其它硅酮聚合物表面,或者密封至与氧化的硅酮聚合物表面反应的其它基底的表面,而不需要单独的粘合剂或其它密封剂。在大多数情况下,可以简单地通过使氧化的硅酮表面与另一表面接触来完成密封,而不需要施加辅助压力以形成密封。也就是说,预氧化的硅酮表面用作针对合适的配合表面的接触粘合剂。具体地,除了不可逆地密封自身之外,氧化的硅酮例如氧化的PDMS还可以不可逆地密封至除了自身之外的一系列氧化材料,包括例如玻璃、硅、氧化硅、石英、氮化硅、聚乙烯、聚苯乙烯、玻璃碳和环氧聚合物,它们以与PDMS表面类似的方式氧化(例如,通过暴露于含氧等离子体)。在本发明的上下文中有用的氧化和密封方法以及整体模制技术描述于本领域中,例如描述于题为“Rapid Prototyping of Microfluidic Systems and Polydimethylsiloxane” Anal. Chem., 70:474-480, 1998 (Duffy等人)的文章中,其通过引用并入本文。

[0163] 因此,在某些实施方案中,制品的设计和/或制造可以相对简单,例如通过使用相对公知的软光刻和其它技术诸如本文所述的那些。另外,在一些实施方案中,制品的快速和/或定制设计(例如,在几何形状方面)是可能的。在一组实施方案中,制品可以被制成是一次性的,例如在制品与放射性、毒性、有毒、反应性、生物危害性等的物质一起使用的实施方案中,和/或当物质的性质(例如,毒理学性质、放射性性质等)是未知的时。从氧化的硅酮聚合物形成通道或其它结构(或内部、流体接触表面)的另一个优点是这些表面比典型的弹

性体聚合物的表面亲水得多(当需要亲水的内表面时)。因此,与由典型的未氧化的弹性体聚合物或其它疏水性材料构成的结构相比,这种亲水性通道表面可以更容易地用水溶液填充和润湿。

[0164] 为了所有目的,以下文献通过引用整体并入本文:Weitz等人的题为“Methods and Systems for Droplet Tagging and Amplification”的美国专利申请系列号61/980,541; Bernstein等人的题为“Systems and Methods for Droplet Tagging”的美国专利申请系列号61/981,123; Link等人的题为“Formation and Control of Fluidic Species”的国际专利申请公开号W0 2004/091763; Stone等人的题为“Method and Apparatus for Fluid Dispersion”的国际专利申请公开号W0 2004/002627; Weitz等人的题为“Method and Apparatus for Forming Multiple Emulsions”的国际专利申请公开号W0 2006/096571; Link等人的题为“Electronic Control of Fluidic Species”的国际专利申请公开号W0 2005/021151; Weitz等人的题为“Droplet Creation Techniques”的国际专利申请公开号W0 2011/056546; Weitz等人的题为“Creation of Libraries of Droplets and Related Species”的国际专利申请公开号W0 2010/033200; Weitz等人的题为“Fluid Injection”的美国专利申请公开号2012-0132288; Agresti等人的题为“Assay And Other Reactions Involving Droplets”国际专利申请公开号W0 2008/109176; 和Weitz等人的题为“Fluid Injection”的国际专利申请公开号W0 2010/151776。

[0165] 还通过引用并入本文的是于2014年4月21日提交的美国临时专利申请系列号61/982,001; 于2014年10月17日提交的美国临时专利申请系列号62/065,348; 于2014年10月20日提交的美国临时专利申请系列号62/066,188; 和于2014年10月30日提交的美国临时专利申请系列号62/072,944。

[0166] 此外,以下通过引用将其全部内容并入本文:于2014年4月17日提交的美国专利申请系列号61/981,123; 于2015年4月17日提交的题为“Systems and Methods for Droplet Tagging”的PCT申请; 于2014年4月17日提交的美国专利申请系列号61/981,108; 于2015年4月17日提交的题为“Methods and Systems for Droplet Tagging and Amplification”的PCT申请; 于2015年4月17日提交的题为“Immobilization-Based Systems and Methods for Genetic Analysis and Other Applications”的美国专利申请; 于2015年4月17日提交的题为“Barcoding Systems and Methods for Gene Sequencing and Other Applications”的美国专利申请; 和于2014年10月30日提交的美国专利申请系列号62/072,944。

[0167] 以下实例旨在举例说明本发明的某些实施方案,但不例示本发明的全部范围。

[0168] 实施例1

[0169] 本实施例使用水凝胶或聚合物微球,其各自携带浓度为1-100微摩尔的DNA片段(引物)。这些引物可以通过化学品或光从微球切割,其中每个DNA片段编码(a)从至少10,000个条形码(但通常超过100,000个条形码)的库中随机选择的条形码序列,其中在每个微球上的所有核酸片段发现相同的条形码;和(b)用于DNA或RNA的杂交和捕获的一个或多个引物序列; (c) 任选地,另外的DNA序列,例如用于条形码化每个分子的随机核苷酸序列,或用于扩增或捕获条形码产物的序列。这些微球的合成在下面更详细地描述。

[0170] 在用于液滴生产的本实施例中,使用通过软光刻制备的微流体装置。其示意图如

图2所示,但例如还可以使用其它工具例如毛细管或管道执行乳化。还可以使用其它微流体构造。通过使用该微流体装置,产生~4nL体积的液滴(图3和4),但是可以基于酶促条形码化反应的要求容易地调整液滴的尺寸。

[0171] 在本实施例中使用的微流体装置具有用于液滴载体油的一个入口和用于液滴水相的组分的附加入口(图14)。对于载体油,使用含有~0.75% (v/v) 表面活性剂(含有两个全氟聚醚嵌段(PFPE)和一个聚乙二醇(PEG)嵌段的PFPE-PEG-PFPE三嵌段共聚物)的氟化油(例如HFE-7500)。表面活性剂用于防止液滴聚结,并且可以例如基于其物理化学性质来调节量。用于乳化的载体油不限于氟化液体,并且可以成功地使用替代流体例如基于烃(例如矿物油,己烷等)、硅油和其它类型的油。在本实施例中使用的三个入口递送以下组分:(1)解离的细胞的悬浮液;(2)细胞裂解试剂;(3)携带条形码化的引物的水凝胶或聚合物微球的悬浮液;和(4)用于酶促产生与捕获的DNA或RNA互补的条形码化DNA的反应混合物。在一些情况下可以预组合这些组分中的一些,例如(2)和(4)。

[0172] 在本实施例中根据以下考虑制备细胞悬浮液。如果细胞是贴壁的或来自组织,可首先将细胞解离并任选过滤或离心以除去两个或更多个细胞的团块。调节细胞悬浮液缓冲液(通常为PBS)的质量密度,以使注射过程中的细胞沉淀最小化,例如通过加入~16% (v/v)的Optiprep。可以调节细胞数密度(每单位体积的细胞)以使两个或更多个细胞在同一液滴中被捕获的发生率最小化。正确数量密度的精确计算例如取决于因素诸如可以耐受的多细胞事件的量以及取决于液滴体积和由细胞悬浮液贡献的相对液滴体积。例如,对于具有由细胞悬浮液贡献的50%液滴体积的4nL液滴,可以使用50个细胞/微升的数量密度以导致0.1个细胞/液滴的平均占据率,从而导致约5%的具有多于一个细胞的含细胞的液滴。如果需要,可以将小的磁力搅拌棒引入细胞注射器中,以允许细胞悬浮液的连续或临时组合。在注入微流体装置期间,可使用冰袋或其它合适的冷却技术将细胞保持冷却。

[0173] 在本实施例中制备酶反应混合物和/或裂解试剂,以使得其在与细胞悬浮液和与微球混合后的最终浓度适合于细胞裂解和进行酶促反应(例如逆转录反应)。

[0174] 如果使用水凝胶微球,这些球体可以被装填(浓缩)以使得其至液滴中的递送变得有序和同步,从而确保大多数液滴精确地容纳一个微球。当使用刚性聚合物微球时,这些可以通过例如使用流动而是有序的。在一些情况下,目的是确保具有单个微球的液滴的数量相对高,并且具有0或2个微球的液滴的数量是稀少的或甚至可忽略的。

[0175] 为了进一步增加细胞和微球的共包封事件,例如,每个液滴一个细胞和一个DNA条形码化的微球,可以在包封之前对细胞进行排序。

[0176] 作为一个非限制性实施例,将水相以100微升/小时、100微升/小时和10-15微升/小时的流速递送到装置中,分别用于细胞悬浮液、裂解/反应混合物和浓缩的水凝胶微球悬浮液。例如,可以将细胞悬浮液的数量密度调节至50,000/mL,以使得在乳化的1小时内捕获5,000个细胞用于条形码化。然而,根据具体应用,可以在1和10,000微升/小时之间独立地调节所有相的流速。

[0177] 在包封步骤之后(或期间),可以裂解细胞,并且使用例如光、化学、酶促或其它技术将附着于微球表面的DNA片段释放在液滴内。

[0178] 释放的DNA片段可以用作用于细胞编码核酸扩增的引物。例如,可以使用逆转录将来自细胞的mRNA转化为cDNA,或者在另一个实例中,可以使用DNA聚合酶合成编码细胞蛋白

的基因。

[0179] 为了将合成的核酸(DNA或RNA)释放到混合物中,在一些情况下可以例如通过化学或物理技术破坏液滴。可以收集释放的DNA,并且如果需要,可以扩增或进一步加工。例如,可以通过首先在液滴破坏之前将一部分液滴乳液转移到新的反应管中来调节待分析的细胞的数量(图5)。例如,在收集含有5,000个细胞的200微升液滴乳液之后,可以首先将乳液分成五个40微升的管,每个管含有约1,000个细胞。如果需要,可以单独处理这些样品。另外,在其它实施方案中可以执行其它调整。

[0180] 可以通过DNA测序或其它技术确定核酸的碱基组成,包括条形码和捕获的序列(图6)。

[0181] 诊断测试可以例如使用定量实时PCR(qPCR)进行,以将捕获的在液滴中条形码化的DNA或RNA的丰度与当在受控条件下(例如在合并的批量反应中的液滴的外部)进行酶促反应时或使用来自相同数量的细胞的纯化的DNA或RNA获得的丰度进行比较。qPCR使用两种引物,一种与由微球递送的条形码化DNA片段的末端杂交;另一种与待捕获的靶DNA或RNA序列杂交。

[0182] 作为非限制性实例,上述可用于分析基因组、单核苷酸多态性、特异性基因表达水平、非编码RNA、全转录组、全基因或其部分等。

[0183] 图2示出了根据本发明的实施方案的一个实例的微流体装置的示意图和操作。其它微流体装置设计也是可能的,例如,如本文所讨论的。图2A示出了指示系统的操作的示意图。使用微流体装置将细胞、条形码化的微球(条形码化的珠粒)和试剂包封在液滴中。图2B示出了具有三个入口和一个出口的装置。入口用于引入i)细胞,ii)DNA条形码化的微球,iii)生物和/或化学试剂和iv)载体油。凝胶、细胞和试剂可以通过三个入口I、II、III中的任一个引入装置中。在流动聚焦接合处发生包封,然后在出口处收集包封的样品。可以调节每个入口的流速以便获得细胞和DNA条形码化的微球共包封的最佳条件。图3显示了共包封在一起的细胞和DNA条形码化的微球的数字图像。上箭头表示细胞,下箭头表示微球。指示了从第一帧开始的时间。图4示出了显示微球和细胞共包封的装置出口的实例。图5显示根据本发明的一个实施方案产生的条形码化样品的数量相对乳液体积和包封(收集)时间。图6显示了在本发明的一个实施方案中显示出主要均匀条形码化的测序读数/丰富条形码的分布。

[0184] 实施例2

[0185] 本实施例举例说明了用于产生附着至微球的条形码化核酸的某些技术。首先,合成将DNA引物(P1)掺入到水凝胶中的微球(图7)。可以使用制备微球或各种类型的水凝胶颗粒的几种技术。在本实施例中描述的微球使用聚丙烯酰胺(pAc)水凝胶,但是也可以使用替代的水凝胶材料(例如琼脂糖、聚-N-异丙基丙烯酰胺(pNIPAM)等)。

[0186] 在一个实施方案中,制备将个体组分混合在一起的含有丙烯酰胺(Ac)、N,N'-亚甲基双丙烯酰胺(bis-Ac)和丙烯酸亚磷酰胺修饰的DNA(Ac-DNA)和/或过硫酸铵(APS)的水溶液。

[0187] 基于水凝胶网孔的孔径和后续应用所需的引物浓度,可以相应地调节Ac和bis-Ac组分的量以及Ac-DNA浓度。例如,在一种情况下,通过含有0.1-0.6% (v/v) 聚合诱导剂(N,N,N',N'-四甲基乙二胺,被称为TEMED)的载体油乳化~0.0258% 丙烯酰胺、~0.036% (v/

v) N,N'-亚甲基双丙烯酰胺、1-50微摩尔Ac-DNA和~0.2%APS的混合物用于生产水凝胶微球。作为载体油,可以使用含有~1.5% (v/v) 表面活性剂(含有两个全氟聚醚嵌段(PFPE)和一个聚乙二醇(PEG)嵌段的PFPE-PEG-PFPE三嵌段共聚物)的氟化油(例如HFE-7500)。表面活性剂可用于例如防止液滴聚结。在一些实施方案中,其量应基于其物理化学性质进行调节。用于乳化的载体油不限于氟化液体,并且在其它实施方案中可以使用基于烃的替代流体(例如矿物油,硅油,己烷等)。

[0188] 在用于液滴生产的本实施例中,使用通过软光刻制备的微流体装置。其示意图如图8所示,但是例如也可以使用其它工具例如毛细管或管道进行乳化。此外,也可以使用具有不同示意图的微流体装置。

[0189] 通过使用这种微流体装置,产生直径约62微米的液滴(图9),但是可以基于其它应用的要求调节液滴的尺寸。

[0190] 在本实施例中,将液滴收集到管中,然后在~65℃孵育>2小时以诱导聚丙烯酰胺的聚合。聚合所需的孵育期和温度可以相应地变化。液滴的聚合也可以通过光或各种化学方法诱导。

[0191] 聚合后,可以通过诸如化学(例如全氟辛醇)或物理技术(例如电场)的技术破坏乳液,这可以使乳液的内容物(例如微球)释放到本体溶液中。然后将释放的微球在己烷和水性缓冲液中洗涤。在典型的示例程序中,用含有1% (v/v) Span 80的己烷处理微球,然后用水性缓冲液(例如10mM Tris-HCl (pH 7.0), 10mM EDTA和0.1% (v/v) Tween-20)洗涤3次,然后悬浮在具有所需组成(例如10mM Tris-HCl (pH 7.0), 0.1mM EDTA和0.1% (v/v) Tween-20)的缓冲液中。微球的最终体积可以不同于合成期间观察到的体积,并且随着水凝胶悬浮缓冲液的条件而变化。

[0192] 微球可以储存延长的时间段,例如在4℃下在含有10mM EDTA的溶液中,或在-20℃下在含有5mM EDTA和50%甘油的溶液中。

[0193] 核酸或引物在微球中或其表面上的掺入取决于例如存在于引物和/或构成微球的材料上的官能团。作为非限制性实例,在聚合过程中,可以将在其5'端含有丙烯酸亚磷酰胺的核酸掺入某些微球的聚丙烯酰胺网孔中。作为另一个实例,亚磷酰胺(acrydite)修饰的寡核苷酸可以与硫醇基团共价反应,并因此,具有巯基的微球将结合亚磷酰胺修饰的寡核苷酸。在另一个实例中,具有氨基的寡核苷酸可以共价结合到某些微球的羧基上。在另一个实例中,具有生物素基团的寡核苷酸可以连接至链霉亲和素包被的微球。在另一个实例中,颗粒可以包括能够识别标签上存在的某些寡核苷酸序列的抗体或抗体片段。[移动到文本的主体]因此,核酸至微球中/上的不同类型的掺入是可能的。

[0194] 在一个实施方案中,可以使用含有用于捕获靶核酸(例如RNA或DNA)、扩增(例如携带T7启动子序列或用于PCR引物的杂交位点)和/或测序的序列的引物P1。在另一个实施方案中,P1引物具有可光切割的位点。

[0195] 在一个实施方案中,P1引物的结构如下(从5'至3'的方向):丙烯酸亚磷酰胺-光可裂解的间隔基-T7启动子的核苷酸序列-用于测序的核苷酸序列(PE1)。

[0196] 在另一个实施方案中,来自DNA引物的第一库的DNA引物N₁的结构如下(从5'至3'的方向):衔接子序列(P2)-条形码序列-与PE1互补的核苷酸序列。

[0197] 在另一个实施方案中,来自DNA引物的第二库的DNA引物N₂的结构如下(从5'到3'

的方向):目的序列(P3)-条形码序列-与P2互补的核苷酸序列。

[0198] 在一些情况下,携带P1引物的微球可以平等地分成 N_1 个库,并且每个库可以与 N_1 个不同的DNA模板中的一个杂交,所述DNA模板具有(依次从3'端到5'端):与部分或全部的P1引物互补的DNA序列,从而允许与DNA P1引物形成双链体;由超过6个确定的核苷酸组成的 N_1 个独特的核酸条形码之一,所述核苷酸对于相同库中的所有分子是相同的;任选地,由多于5个随机核苷酸组成的随机核酸序列,所述核苷酸在具有相同库的分子之间是不同的;和可以用作后续条形码化的杂交位点的DNA序列(P2)。P2序列可以含有用于在随后的步骤中引发测序反应的序列。

[0199] 可以对 N_1 个库中的每一个进行酶促反应,从而导致P1核酸片段通过每个库中的模板DNA片段的复制延伸。在一些情况下,可以使用连接反应,例如,代替聚合反应。

[0200] 酶促反应可以通过加入抑制剂如EDTA、钒或通过其它方法停止。

[0201] 微球可以合并在一起并任选地洗涤以除去酶或任何过量的模板分子。

[0202] 微球上的DNA片段可以转化为单链DNA,例如通过经由变性除去模板分子,例如通过在0.1M氢氧化钠中重复洗涤微球或,通过其它技术。

[0203] 现在以P2序列结束的微球可以平等地分成 N_2 个库(通常 $N_2=N_1$,尽管这不是必需的),并且每个库可以与 N_2 个不同的DNA模板之一杂交(如步骤(2)中),所述模板具有(从3'端到5'端):与部分或全部的P2引物互补的DNA序列,从而允许与DNA P2引物形成双链体;由超过6个确定的核苷酸组成的 N_2 个独特的核酸条形码之一,所述核苷酸对于相同库中的所有分子是相同的;由多于5个随机核苷酸组成的随机核酸序列,所述核苷酸在相同库中的分子之间是不同的;和可以用作后续延伸的杂交位点或用作单细胞分析操作(例如DNA或RNA分子的捕获)的引物序列的DNA序列位点(P3)。

[0204] 在一些实施方案中,可以重复这些步骤中的一些。

[0205] 在一些情况下,可以产生携带编码引物P1、随后第一条形码、随后序列P2、随后第二条形码、随后序列P3的单链DNA片段的微球。独特微球库的数目是 $N_1 \times N_2$ (也参见图10和11)。

[0206] 在一些情况下,制备的微球可以储存延长的时间段,并在随后的应用中用作试剂。

[0207] 如果需要,可以用 $N_{3,4}, \dots$ 个条形码模板的额外库进行额外的重复,每个添加条形码和序列P4、P5等。独特的微球库的数量可以随着每个步骤而增加至 $N_1 \times N_2 \times N_3 \times \dots$ 。

[0208] 任选地,所有微球可以一起杂交到M个DNA模板的混合物,所述DNA模板具有(连续地从3'端至5'端):与部分或全部的最终P3(或P4、P5等)引物互补的DNA序列,从而允许与DNA P3(或P4、P5等)引物形成双链体,以及将用作用于单细胞分析操作(例如特异性DNA或RNA分子的捕获)的特异性引物序列的M个序列 S_1, \dots, S_M 之一。这些步骤可以重复。这可以产生相同数量的 $N_1 \times N_2 \times N_3 \times$ 个微球库,每个携带上述序列但是现在DNA片段属于M个种类的分子,其在最终M个可能的序列 S_1, \dots, S_M 之外是相同的。

[0209] 在一些情况下,这可能产生用ssDNA片段包被的微球,其中每个ssDNA片段以如下顺序(从5'到3')编码:P1引物,例如含有可以用作用于但不限于用于核酸扩增和测序的位点的引物位点PE1和T7启动子位点;两个或更多个DNA条形码(每个由6个或更多个核苷酸组成),其对于包被单个珠粒的所有引物是相同的但在珠粒之间不同;任选地,分子特异性DNA条形码(由5个或更多个随机核苷酸组成);P2“引物位点2”,其可以用作用于单细胞中的DNA

或RNA的测序和/或与其杂交的引物,用于逆转录或PCR扩增;P3“引物位点3”,其可以用作用于单细胞中的DNA或RNA的测序和/或与其杂交的引物,用于逆转录或PCR扩增。一个P3片段也可以编码多个基因特异性引物(GSP)中的一个,因此用多个ssDNA片段包被的每个珠粒将包含所有的GSP。

[0210] 携带条形码化DNA引物的微球可用作用于单细胞中的DNA或RNA的测序和/或与其杂交的试剂,用于逆转录或PCR扩增以及涉及DNA捕获、扩增和测序的其它应用。

[0211] 图7示出了携带与携带条形码序列和引物位点PE1*和P2*的单链DNA(ssDNA)引物库杂交的PE1引物的微球。在本实施例中,随后使用DNA聚合酶延伸引物。然后将延伸的引物转化成ssDNA(例如使用升高的温度或碱性溶液)。然后可将获得的ssDNA引物与携带第二条形码序列和引物位点P2*和P3*的第二引物库杂交。在引物延伸和转化成ssDNA后,微球可用于不同的应用,例如,旨在捕获和测序样品中的核酸的应用。

[0212] 图8示出了根据本发明的一个实施方案的用于生产携带DNA的微球的微流体装置的示意图和设计。本实施例中的装置包括用于载体油的一个入口和用于试剂的一个入口。在两相相遇的流动聚焦接合处产生液滴。在收集出口处收集液滴。

[0213] 图9A和9B是根据本发明的一个实施方案生产的携带DNA的微球的明视场图像。在本实施例中,由聚丙烯酰胺水凝胶和DNA引物组成的微球连接至聚合物网孔。比例尺为50微米。

[0214] 图10示出了在另一个实施方案中的携带条形码化DNA引物的微球的DNA延伸效率的评估。图10A显示具有以FAM荧光探针杂交至PE1位点的DNA的微球,图10B显示具有以FAM荧光探针杂交至P2位点的DNA的微球。图10C显示具有以荧光探针杂交至P3位点的DNA的微球。这些结果显示DNA延伸可以在水凝胶微球中进行。

[0215] 图11示出了在另一个实施方案中的来自11个个体微球的DNA片段的高通量测序。从每个微球测序平均140,000个分子。该图显示在 384^2 种可能的条形码中在每个微球上携带相同条形码的这些(“引物”)的分数。条形码的识别对于每个微球是不同的。每条线对应一个微球。在理想条件下,每个微球上的100%的DNA片段将携带相同的条形码,并且0%将携带第二种、第三种或其它条形码。在该样品中实现的平均值是92%的携带相同条形码的DNA片段。

[0216] 实施例3

[0217] 本实施例使用如上所述合成的DNA条形码化的微球,从而得到携带具有以下序列元件(5'至3')的以下单链DNA片段的微球:P1引物,例如含有可以用作用于但不限于用于核酸扩增和测序的位点的引物位点PE1和T7启动子位点;两个或更多个DNA条形码(每个由6个或更多个核苷酸组成),其对于包被单个珠粒的所有引物是相同的但在珠粒之间不同;任选地,分子特异性DNA条形码(由5个或更多个随机核苷酸组成);P2“引物位点2”,其可以用作用于单细胞中的DNA或RNA的测序和/或与其杂交的引物,用于逆转录或PCR扩增;P3“引物位点3”,其可以用作用于单细胞中的DNA或RNA的测序和/或与其杂交的引物,用于逆转录或PCR扩增。

[0218] 在本实施例中,在合成DNA条形码化的微球之后,将微球合并,然后与M个DNA模板的单一混合物杂交,所述DNA模板具有(连续地从3'端至5'端):与部分或全部的最终P3引物互补的DNA序列,从而允许与DNA P3引物形成双链体,以及将用作用于单细胞分析操作(例

如特异性DNA或RNA分子的捕获)的特异性引物序列的M个序列 S_1, \dots, S_M 之一。这些步骤可以重复,从而产生相同数量的 $N_1 \times N_2$ 个微球库,每个携带在1中指定的序列,但是现在DNA片段属于M个种类的分子,其在最终M个可能的序列 S_1, \dots, S_M 之外是相同的。

[0219] 在一些情况下,可根据其它方法合成DNA微球,以产生具有上述相同序列的所得微球。

[0220] 实施例4

[0221] 本实施例显示,从裂解的细胞逆转录(RT) mRNA成互补DNA对于小于3nL/细胞的反应体积变得强烈抑制,特别是随着反应产率Y遵循利用液滴体积V的一阶抑制,即 $Y = 1 / (1 + K_{50} / V)$,其中 $K_{50} = 1 - 3.3\text{nL}$ 对于测试的至少三种不同细胞培养系(MCF7, K562和THP-1细胞)发生50%抑制时的体积。相比之下,目前使用液滴微流体的许多工作集中在将细胞包封在体积为10-100pL体积的液滴中。在这样的体积下,逆转录反应将被严重抑制。

[0222] 图6示出了根据本发明的一个实施方案的用于最佳液滴体积的批量测试。

[0223] 在本实施例中,可以通过模拟包含5微升或更多反应混合物的反应孔中的液滴条件(其可以被调节以模拟单个液滴内的条件)来进行对微流体体积中的单细胞的反应效力的测试。

[0224] 为了模拟大小为V的微流体体积,将完整细胞以1个细胞/体积V的终浓度加入到本体反应中。因此,4nL液滴内的单个细胞对应于用浓度为250个细胞/微升的细胞裂解物运行反应。

[0225] 此外,为了模拟大小为V的微流体体积,例如通过使用各自携带m摩尔试剂的微球,将任何试剂谨慎地施用到液滴中,将试剂以m摩尔/体积V的终浓度加入本体反应中。例如,如果1毫微微摩尔的DNA片段在微球上递送到4nL液滴中,则相同的DNA片段将以0.42微摩尔的浓度加入到本体反应中。

[0226] 本体反应可以在12孔、96孔或384孔形式中平行进行,以鉴定最佳反应条件。

[0227] 这种诊断测试可以提供用于优化用于条形码化的液滴尺寸和组成的快速方法。例如,在测试三个细胞系时,观察到当制备小于3nL体积的液滴时对条形码反应的强烈抑制。参见图12。

[0228] 实施例5

[0229] 为了解释健康和患病组织的基因表达,生物学家的梦想是对每个细胞中的基因表达变化作图。通过使用这样的数据,可以期望鉴定和跟踪异质细胞亚群,并推断基因和通路之间的调节关系。已经利用诸如RNA测序的“组学(Omics)”方法来分析单个细胞,但是限制是常规地分离和处理大量个体细胞以用于深度测序的有效方式,以及定量地这样做。本实施例举例说明了用于通过下一代测序对数千个个体细胞进行平行条形码化以用于后续分析的液滴-微流体方法。这显示低噪声特征谱并且容易适用于其它基于测序的测定。这些实例将技术应用于小鼠胚胎干(ES)细胞以定义ES细胞群体结构和ES细胞分化通过LIF撤除的异质性起始。这些结果表明了液滴条形码化用于解构细胞群和推断基因表达与高通量单细胞数据的关系的应用。

[0230] 这些实例利用液滴微流体来开发一种新的技术用于并行条形码化成千上万的个体细胞以用于通过下一代测序(drop-seq)的后续分析。在这些实例中使用的工具具有在单次运行中对成千上万个细胞进行条形码化的理论能力,尽管实际上一些实验集中于每次运

行数百至数千个细胞,因为测序深度在非常高的细胞计数下变得有限。这些实例通过分析LIF撤除之前和之后的小鼠胚胎干(ES)细胞来评估drop-SEQ。总共分析了超过10,000个条形码化的细胞和对照液滴,其中约3,000个ES和分化细胞在更大的深度下测序以用于后续分析。以下分析鉴定了表达不同谱系的标记物的稀有亚群(其难以根据数百个细胞的分析进行分类)的存在。还发现关键的多能性因子以相关的方式在整个ES细胞群体中波动,并且探索了这种波动可能用于将新因子与多能状态相关联的可能性。在分化后,观察到基因表达波动的相关性结构中的显著变化,其由多能性因子的非同步失活引起,并出现新的细胞状态。总之,这些结果显示drop-SEQ解构细胞群体和推断单个实验中的基因表达关系的潜力。

[0231] 设计和实施用于单细胞的分析和液滴条形码化的微流体平台。使用用于RNA测序(RNA-Seq)的方案,其中mRNA在逆转录反应期间被条形码化,随后合并细胞并进一步处理用于测序(图13A)。为此,drop-SEQ平台(图13A-13E和图18)用裂解缓冲液、逆转录(RT)试剂和条形码化的寡核苷酸引物将细胞包封在液滴中。从每个裂解细胞释放的mRNA保持陷入在相同的液滴中,并在互补DNA(cDNA)的合成期间被条形码化。条形码化后,来自所有细胞的物质通过破坏液滴而组合,并且处理cDNA文库用于测序(图13A)。

[0232] 实施该策略的一个挑战是确保每个液滴携带编码相同独特条形码(其应当不同于其它液滴中的条形码)的引物。通过合成与细胞共同包封的条形码化的水凝胶微球(BHM)的文库来克服这一挑战(图13B)。每个水凝胶携带编码 384^2 (即147,456)个预定义条形码之一的共价偶联的光可释放引物。该库大小允许以99%的独特标记随机标记3,000个细胞,和以97%的独特标记随机标记10,000个细胞(见下文)。图19-21描述了用于使用分割库方法合成BHM的方法;见下文。这可以以直接的方式扩展以产生更大数量的条形码用于更大规模的细胞捕获,例如用于靶向测序应用。

[0233] 为了共包封BHM和细胞,使用具有用于i) BHM, ii) 细胞, iii) RT/裂解试剂和iv) 载体油的四个入口和用于液滴收集的一个出口的微流体装置(图13C-13D)。该装置产生单分散液滴,其以每秒~10-50滴的速率在1-5nL的范围内变化,从而同时混合来自入口的等分试样(图13E)。芯片内的紧密装填的可变形水凝胶的流动可以有效地同步化,从而允许接近100%的水凝胶液滴占有率。这个特征确保了到达液滴中的随机分布的细胞几乎总是暴露于BHM。在典型的条件下,细胞浓度被设置为仅占据10%的液滴,以确保两细胞事件的低概率(图13E)。在这些实验中,液滴含有至少一个细胞和一个凝胶以产生用于测序的条形码化文库。通常,超过90%的生产性液滴恰好包含一个细胞和一个凝胶(图13F)。RT反应的效率也使用溶液或仍然结合至BHM的引物测试,并且发现引物释放对于来自液滴中的裂解细胞的有效RT反应是重要的(图13G)。因此,在RT反应之前,通过暴露于UV光将BHM结合的引物光释放到液滴中(图13A)。

[0234] 在这些实例中,对数百至数千个细胞的样品进行测序以避免极浅的测序深度,但是这也可以用于容易地捕获和条形码化更高的细胞数量,例如,具有4,000-6,000个细胞/小时的通量。实际上,在测序后,发现条形码化的样品的数量通常与收集的乳液体积成线性比例(图13H),其中对于每100微升乳液(约30分钟收集时间)约2,000-3,000个细胞或对照液滴被条形码化。

[0235] 图13示出了用于对数千个细胞进行DNA条形码化的液滴微流体平台的实例。图13A

示出drop-SEQ工作流的概况。芯片上操作发生在前三个框中,芯片外操作发生在接下来的三个框中,并且测序/数据分析发生在最后两个框中。图13B是用于将DNA条形码化的水凝胶微球(BHM)(大圆圈)与细胞(小圈)和RT/裂解混合物组合的微流体装置的示意图。BHM引物说明:PC=可光切割的接头;T7=T7RNA聚合酶启动子;PE1=测序引物;BC=BHM特异性条形码;UMI=独特的分子标识符。图13C示出了微流体装置设计。图13D示出了用于包封(右)和收集(左)的微流体模块的快照。细胞和BHM分别用下部和上部箭头注释。其它箭头表示流动的方向。比例尺为100微米。图13E示出了液滴占有率随时间的统计。图13F示出了细胞和DNA-条形码化的珠粒共包封事件的统计。超过90%的细胞用单个DNA条形码化的珠粒包封。图13G显示了使用从珠粒光释放的引物制备的文库在逆转录之前(上曲线)或之后(下曲线)的BioAnalyzer轨迹。图13H显示了对于纯RNA和mES细胞检测的条形码的数目,每100微升乳剂收集了2,159个细胞(30分钟收集时间)。

[0236] 图18示出了在本实施例中使用的液滴微流体装载的设计。该装置包括用于RT和裂解试剂混合物(1)、细胞悬浮液(2)、DNA条形码化的珠粒(3)的三个入口和用于连续相(4)的一个入口。流体电阻器掺入至由于注射器泵的机械不稳定性而引起的装置阻尼波动中。将等分样品通过60微米宽的通道汇集到70微米宽的主通道中,其中它们在流动聚焦接合处包封成液滴之前层流流动(虚线框)。在出口(5)以乳液的形式收集液滴。

[0237] 图19显示了用于生产DNA条形码化的水凝胶珠粒的液滴微流体装置的设计。图19A显示了装载的示例设计。该装置包括用于水相(试剂)的一个入口和用于连续相(载体油)的一个入口。在流动聚焦喷嘴处产生单分散的水凝胶液滴,这些尺寸在右边的虚线框中指示。液滴通过表面活性剂在2000微米长的通道中稳定,并在出口处以乳液的形式收集。图19B显示水凝胶液滴生产、稳定化和收集的数字图像。微流体通道为50微米深。比例尺表示100微米。

[0238] 图20显示了DNA条形码化的珠粒的合成。图20A显示了连接至携带T7RNA聚合酶启动子(左侧序列)和PE1引物位点(序列右侧)的DNA引物的5'端的含有丙烯酸亚磷酰胺部分(左)和可光切割间隔物(右)的结构DNA寡核苷酸。图20B显示了条形码化的水凝胶微球的合成示意图。在第一步中,连接至聚丙烯酰胺水凝胶的ssDNA引物与携带PE1*和W1*引物位点和DNA条形码的前半部分的互补DNA寡核苷酸杂交。得到的DNA异源双链通过Bst 2.0DNA聚合酶转化为dsDNA(其中虚线表示新合成的DNA链),并通过碱处理变性回ssDNA形式。在第二步中,用携带W1*序列、DNA条形码的后半部分、独特分子标识符(UMI)和多聚A序列的第二DNA寡核苷酸重复该过程。引物延伸和变性后,DNA条形码化的珠粒含有T7启动子、PE1引物位点、DNA条形码、W1位点、UMI和多聚T序列。图20C显示了完全装配的引物的DNA序列(SEQ ID NO:6)。突出显示的文本表示具有T7启动子(TAATACGACTCACTATAGGG)(SEQ ID NO:2)、PE1引物位点(CTCTTCCCTACACGACGCTCTTC)(SEQ ID NO:3)、两个DNA条形码([条形码1]和[条形码2])、W1衔接子位点(AAGGCGTCACAAGCAATCACTC)(SEQ ID NO:4)、UMI(NNNNNN)和多聚T尾(TTTTTTTTTTTTTTTTTTV)(SEQ ID NO:5)的寡核苷酸的不同部分。丙烯酸亚磷酰胺和可光切割间隔物的化学部分分别表示为/5Acryd/和/iSpPC/。[条形码1]和[条形码2]的DNA序列各为8个核苷酸长。

[0239] 图21显示了掺入条形码化的水凝胶微球(BHM)中的DNA引物的定量。图21A-21D显示了合成后BHM的iImaging,显示大小为63微米的BHM的明视场图像(图21A),和与靶向PE1

序列(图21B)、W1序列(图21C)、和多聚T序列(图21D)的互补DNA探针杂交后的荧光共聚焦成像。比例尺为100微米。图21E显示在从BHM光切割后的DNA引物的BioAnalyzer电泳图,显示全长条形码(最大峰)以及合成中间体(两个较小峰)的存在。在35和10380碱基对处的峰是凝胶迁移标记。峰上方的数字表示碱基对的理论片段大小,但这些对于单链DNA产物不准确。图21F-21H显示来自11个个体BHM的条形码化的产物的深度测序的结果。图21F显示每个凝胶上的条形码丰度的秩图;图21G和21H显示了在每个BHM上由图21G和21H中详细描述的最丰富和第二丰富的条形码占据的分数的直方图。完美的合成将导致最高的条形码占据100%,所有其它条形码占据0%。观察到连接至每个BHM的所有引物的平均约92%携带相同的显性条形码。

[0240] 实施例6

[0241] 随机条形码化和液滴完整性的验证。在本实施例中,通过将来自小鼠和人来源的细胞的混合物(小鼠ES细胞和K562红白血病细胞)以大致相等的比例应用于drop-Seq平台来测试drop-Seq平台有效地分隔和条形码化细胞的能力(图14A)。在该测试中,每个条形码应该与小鼠或人作图的转录物完全关联,只有小部分的2-细胞事件导致与小鼠和人两者关联的条形码的出现。测序后,图14A显示drop-SEQ提供复合细胞混合物中细胞的明确鉴定:96%的条形码标记的读取映射到具有超过99%纯度的小鼠或人转录组,并且仅4%的条形码显示两种生物的混合物。通过稀释细胞悬浮液以减少共包封事件,或通过收集之前在芯片上分选液滴以消除多细胞事件,可以进一步降低该已很低的错误率。

[0242] 还明确测试的是细胞条形码从 384^2 种可能条形码的预期库中随机取样以确保极低概率的重复条形码。跨八个独立运行(覆盖总共11,085个对照液滴和细胞)的条形码身份的比较一致地显示与来自 384^2 个条形码的库的随机取样的良好一致性(图22A)。

[0243] 图14显示了液滴完整性和随机条形码化的测试。图14A显示了液滴完整性控制实验的示意图和结果:小鼠和人细胞被共包封以允许明确鉴定在多个细胞间共有的条形码;4%的条形码共有混合的小鼠/人读数。

[0244] 图22显示了随机条形码化和独特分子标识符(UMI)过滤。图21A显示了覆盖140-2,930个细胞或纯RNA对照液滴的8次drop-Seq运行的随机条形码化的成对测试。上方的三角形显示对于使用3842个随机条形码化的每对运行的共有条形码的观察(左)和预期(右)数量。下方的三角形显示p值(假设来自3842个条形码的库中的均匀随机条形码化),其预测所观察到的共有条形码的数量应当围绕期望值超几何分布。p值未针对多重假设检验进行校正。图22B-22D显示了UMI过滤。图21B显示作为检测到的mRNA分子的数目的函数的观察到的UMI的预期数目(黑色曲线)可以显示为具有以下形式: $n_{obs}(m) = N_{UMI} \left(1 - e^{-\frac{m}{N_{UMI}}}\right)$, 其中m是检测到的mRNA分子的数目, $N_{UMI} = 4,096$ 是可用的UMI库的总大小。该函数与理想的线性关系(近似直线)形成对比,从而显示饱和点。图22C是来自单个mES细胞的数据中的每个基因的映射读数数量相对于不同UMI的数量的实例;数据点对应于独特的基因符号。该曲线表示没有扩增偏差,即其中每个映射的读数对应于不同的UMI。大多数基因显示一定的扩增偏差。图22D和22E显示了在没有UMI过滤的情况下(图22D)和在UMI过滤之后(图22E)作为在mES细胞群体中检测到的基因的平均转录物丰度的函数的细胞间CV(标准偏差/平均值)的双对数图。每个数据点对应于单个基因符号。

[0245] 实施例7:

[0246] drop-SEQ平台的基线技术噪声。单细胞RNA-Seq中的技术噪声的两个来源是(a) mRNA捕获效率中细胞之间的变异性, (b) 由每个细胞中捕获有限数目的mRNA转录物产生的固有取样噪声。据报道, CEL-Seq方案具有~4%或更低低捕获效率, 并且当在微量滴定板中进行时, 具有对于纯RNA对照的约25%和对于细胞的~50%的捕获效率的变化性(样品之间的变异系数)。生物信息学分析对单细胞测序数据的影响了解甚少, 但潜在的问题可能来自将不确定的读数归因于多个基因, 从而导致假的基因对相关性。在文库扩增期间也可能产生技术噪声, 但是这种噪声源通常通过使用随机的独特分子标识符(UMI) 序列(其允许重复读取的生物信息移除) 来消除。本实施例举例说明了在所有实验中使用随机六聚体的基于UMI的过滤的实现, 从而导致方法噪声的显著降低(图22)。

[0247] 为了测试该系统中的技术噪声与之前应用的CEL-Seq如何比较, 分析了技术对照样品, 其包括与已知浓度的ERCC RNA加标对照混合的稀释至单细胞浓度(每滴10pg)的总RNA(图14B)。在单次运行中对953个条形码化的对照液滴进行测序, 每个液滴具有平均 30×10^3 (+/-21%) 个UMI过滤映射(UMIFM)的读数(图14B)。在每个液滴中鉴定出5,000至15,000个独特基因符号(总共检测了25,209个), 其中数量与UMIFM计数强烈相关(图14C)。这显示与低至平均0.5分子/液滴的浓度的ERCC加标输入浓度相比, UMIFM计数的优异的线性读数(图14D); 低于该限度, 存在观察到的转录物的数目过度计数的轻微倾向。

[0248] 方法性能的另一个重要测量是其灵敏度, 即检测表达的基因的可能性。灵敏度几乎完全是转录物丰度的函数(图14E), 并且基于mRNA分子的总体捕获效率(参见下文)(从ERCC加标物测量为7.1%) 对所有基因非常好地预测灵敏度(图14D)。利用这种捕获效率, 当存在10个转录物时在50%的液滴中检测到基因, 当存在>45个转录物时, 在>95%的液滴中检测到基因(图14E)。该灵敏度和捕获效率高于先前对于CEL-Seq所测量的(3.4%)。

[0249] 在准确性上, 这显示出非常低水平的技术噪声, 其可以通过将细胞群体间每个基因的变异系数(CV=标准差/平均值) 与其平均表达水平进行比较来评估(图14F)。在仅由采样噪声限制的系统中, 所有基因应该在功率定律曲线 $CV = (\text{平均值})^{-1/2}$ 周围窄分布(图14F)。这确实被观察到了。更正式地, 在归一化后, 99.5%的检测到的基因(N=25,209) 具有与泊松分布一致的分布, 基线技术噪声为5-10% (图14F, 虚线曲线)。

[0250] 图14B显示了RNA-Seq技术对照实验的UMI-过滤作图(UMIFM) 读数的实验示意图和直方图。图14C显示作为每个液滴的UMIFM读数的函数检测的数量独特基因符号。图14D显示了与其输入浓度线性相关的加标分子的平均UMIFM读数, 捕获效率 β (beta) = 7.1%。图14E显示作为输入RNA丰度的函数的方法灵敏度; 曲线显示理论预测 $1 - e^{-x \cdot (1 - e^{-\beta})}$, 通过假定只有固有采样噪声导出。图14F显示了归一化后针对平均UMIFM计数作图的加标和纯RNA转录物的变异系数(CV)。实线曲线显示了采样噪声限制; 虚线曲线显示了具有5%的捕获效率上的残留液滴间变异性的采样噪声限制。

[0251] 实施例8

[0252] 单细胞数据的噪声建模。在预测单个ES细胞数据时, 本实施例显示了技术噪声模型, 以更好地理解当与批量测量相比时在每个细胞的基础上测量时转录物的低采样效率的影响。低效率对细胞之间观察到的基因表达的变异性以及细胞间基因表达的共变均具有影响。三个特征促成该影响: 在所有细胞间平均的转录物的捕获效率; 捕获效率上的细胞间变

化;和归一化方案的选择。通过改进先前的噪声模型,对于跨细胞的基因丰度的CV、基因Fano因子(方差/平均)和对于基因之间的成对相关性导出了生物量和观察量之间的关系(图14G;参见下文)。Fano因子是通常用于测量噪声基因表达的度量,但它对捕获效率非常敏感。该分析显示,技术噪声不仅引入了广泛认可的基线噪声,而且还虚假地放大了现有的生物变化(图14G,等式1)。图14G显示对于CV、Fano因子和基因成对相关性的观察的和潜在的生物量之间的关系概述。

[0253] 还显示低采样效率以可预测的方式显著地降低基因对之间的相关性,从而设置在数据中发现相对弱但显著的相关性的期望(图14G,等式2-3)。知晓相对弱的相关性是真实的并且是单细胞测量的统计的预期结果有助于从包括高度可变基因的测试的数据中获得有用的信息(参见下文)。这些结果还提供了用于基于基本计数统计从生物测量开发用于形式上去卷积噪声的方法的基础。

[0254] 此外,出乎意料地遇到并消除的是来自映射到两个或更多个基因转录物的读数的异常基因表达相关性的重要来源。旨在用于批量(非单细胞)应用的序列分析流水线以可能杂乱地偶联原本独立表达的基因的方式在概率性地映射模糊读数。这个问题对于单细胞的3'-测序可能特别严重,因为UTR区域可以在多个基因间是相似的;并且在相对均匀的细胞群体例如ES细胞中也如此,其特征在于变得与虚假生成的那些相当的弱基因表达偶联的广泛网络。然而,当采样效率低时,该问题更为普遍,因为这些用于削弱甚至强的生物相关性(图14G,等式3)。这些实例显示,使用利用跨不同读数的重复UMI标签以最小化映射中的模糊性的新型生物信息流水线(见下文)克服了读数映射问题。

[0255] 实施例9

[0256] 小鼠ES细胞的单细胞分析。单细胞分析能够从不同谱系鉴定分化的细胞类型,即使具有非常低的测序深度。不太清楚的是可以从研究经受随机波动或动态环境的相对均匀的人口获得的信息类型。为了探索可从我们的新方法获得的信息的类型,本实施例研究了维持在血清中的小鼠ES细胞,因为这些细胞被充分研究并且表现出良好表征的波动,但是与分化的细胞类型相比它们仍然是均匀的,并且可以提出用于高通量单细胞测序的挑战。

[0257] 为了探索drop-SEQ的行为,通过收集不同的乳液体积,对于每个ES细胞运行在不同的测序深度收获不同数目的细胞。收集935个ES细胞用于深度测序;LIF撤除后2天收集145、302和2,160个细胞;4天后收集683细胞;以及7天后收集169和799个细胞。在这些运行中每个细胞获得的平均读数数量高达 208×10^3 ,平均UMIFM计数高达 29×10^3 。运行统计详见表1。

[0258] ES细胞群的结构。对于935个ES细胞,鉴定了1,507个基因,其比从泊松统计学所预期的(10%FDR,参见下文和表2)显著更可变,并且还在至少一个细胞中在至少10个UMIFM计数的水平上表达细胞(图15A,15B)。在1,507个丰富和可变的基因中,发现先前报道在ES细胞中波动的多能性因子(Nanog,Rex1/Zfp42,Dppa5a,Sox2,Esrrb)。值得注意的是,最高度可变的基因包括原始内胚层命运的已知标记物(Col4a1,Col4a2,Lamb1,Lama1,Sox17,Sparc),外胚层命运的标记物(Krt8,Krt18,S100a6)和ES细胞状态的表现遗传调节物(Dnmt3b),以及与ES细胞调节具有未知关联性的基因,例如干细胞抗原Sca-1/Ly6a,其可在调节成体干细胞命运中起作用。其它基因显示非常低的噪声特征谱,与泊松统计学一致(例如Ttn,图15B)。对选择的基因组(图15C)评价上述泊松噪声,被定义为 $\eta(\text{eta}) = \text{CV}^2 - 1/\mu$ (μ

或 μ 是平均UMIFM计数), 并发现其与以前的报道定性一致。与CV或Fano因子不同, 即使对于低采样效率, η (eta) 也于其真实生物学值线性地相关 (图14G, 等式 (1))。

[0259] 图15显示drop-SEQ分析显示ES细胞群体的异质结构。图15A显示针对mES细胞转录组 (中间的点和上方的点) 和纯RNA技术对照 (下方的点) 的平均UMIFM计数绘制的CV。标记为黑色的基因被鉴定为比技术对照显著更可变。实线和虚线曲线如图14F所示, 但在细胞实验中具有20%的残留方法噪声。注释了可变基因的子集。图15B显示了显示适应至泊松和负二项分布的低 (Ttn)、中等 (Trim28, Ly6a, Dppa5a) 和高 (Sparc, S100a6) 表达变异性的举例说明性基因表达分布。图15C显示针对多能性因子作图并与其它因子比较的上述泊松 (a.p.) 噪声, $CV^2 - 1/\text{平均值}$ 。

[0260] 实施例10

[0261] 为了测试ES细胞在多能性ICM样状态和更分化的外胚层样状态之间表现出异质性的观点, 本实施例研究了在单一ES细胞中对比候选多能性和分化标记物的表达。基因对相关性 (图15D) 首次表现出与离散的双态视图一致, 因为外胚层标记物Krt8和原始内胚层标记物Col4a1仅在Pou5f1 (显示的) 和其它多能性标记物 (未显示) 低的细胞中表达。与多能状态相比, 分化易发状态是罕见的。相关性还显示了ES细胞中其它已知的调节性相互作用, 例如BMP信号传导的已知阴性靶标Sox2与BMP靶Id1反相关。然而, 更令人惊讶的是发现多个多能性因子 (Nanog, Trim28, Esrrb, Sox2, Klf4, Zfp42) 在大多数细胞群体中串联波动 (图15D, 23、24)。这些观察一起没有被简单的双态模型解释, 因为它们表明多能性因子独立于外胚层基因表达而保持相关; 相反, 它们暗示以不同多能性为特征的状态连续体。然而, 并非所有多能性因子显示出显著的相关性, 然而: Oct4/Pou5f1与其它核心多能性因子和其它因子远更弱地相关, 并且替代地与细胞周期蛋白D3而非其它细胞周期蛋白强烈相关 (图15D和24), 暗示违背特定的调节起点的波动。

[0262] 那么, 从数据推断的ES细胞群的结构是什么? 对高度可变基因的ES细胞群体进行主成分分析 (PCA), 并且发现多个非平凡维数的异质性 (12个维数, 95%置信度), 这对应于不能由表达中的固有噪声来解释的数据中的主成分 (PC) 的数目 (参见图15E)。这一观察结果证实了在ICM-外胚层轴以外的额外的异质性来源的存在。检查前四个主成分及其负载 (图15F), 揭示了存在至少三个小的但不同的细胞亚群: 一个稀少的群体 (6/935个细胞) 表达非常低水平的多能性标记物和高水平的Col4a1/2、Lama1/b1/c1、Sparc和Cd63, 其明确鉴定原始内胚层 (PrEn) 样细胞; 第二细胞群体 (15/935个细胞) 表达高水平的Krt8、Krt18、S100a6、Sfn和上皮细胞谱系的其它标记物。第三群体表现出似乎无特征的状态, 由热休克蛋白Hsp90、Hspa5和其它ER组分如二硫化物异构酶Pdia6的表达标记。该群体可代表来自解离的应激下的ES细胞。

[0263] PCA分析是用于可视化细胞群体的有力工具, 其可以仅用基因表达的两个或三个主轴分级。然而, 当存在多于三个非平凡主成分时, 存在用于代表高维数据的局部结构的降维的更适合的技术。此实施例应用被称为t分布随机邻近嵌入 (t-SNE) 的降维方法 (图15G)。图谱显示没有如预期的大的ES细胞亚群, 但是显示从高多能性到低多能性的连续谱, 具有由位于群体边缘的PCA鉴定的离群点群体。该图还显示了分别由Prdm1/Blimp1、Lin41/Trim71和SSEA-1/Fut4的高表达表征的三个额外的边缘亚群。与Hsp90-hi群体一样, 仍然需要看到这些群体是否代表由不同功能行为赋予的不同细胞状态, 或者这些群体是否仅仅是

访问ES细胞基因表达的离群但正常的状态的细胞。因此,尽管鉴定了ES细胞群体中经充分研究的外胚层样状态,并且发现ICM至外胚层样状态之间的集体波动的证据,但这些波动不是ES细胞群体中转录异质性的唯一轴。

[0264] 图15D显示了举例说明成对基因相关性的热图。图15E显示从mES细胞群的主成分分析获得的特征值分布,揭示数据中可检测的细胞异质性(箭头)的非平凡模式。平滑曲线显示基因表达谱的随机置换的典型特征值分布;锯齿曲线显示了随机矩阵的预测Marcenko-Pastur特征值分布。只有超过曲线的特征值才是显著的。图15F显示mES细胞主成分及其负载,显示异质性的主要不相关模式并揭示三种稀少的ES细胞亚群。图15G是mES细胞群的tSNE图,揭示了额外的边缘亚群和多能性至外胚层轴。

[0265] 图23显示mES细胞的单细胞基因表达。主要在0、2、4和7天可变的基因的基因表达(分别为图23A-23D)。每个基因的表达是z-得分标准化的。

[0266] 图24显示mES细胞群的结构。图24A显示了跨935个mES细胞的选择基因的成对相关性。如本文所讨论的,Oct4/Pou5f1与细胞周期蛋白D3更强烈相关,与Sox2、Klf4和其它多能性因子更弱相关。这里报告的相关性是在没有子采样校正的情况下观察到的(参见图14G,等式(3))。图24B-24G显示mES细胞群的3维tSNE图谱的不同投影揭示不同的细胞亚群;每个图中的细胞根据指定标记物的聚集表达而着色。

[0267] 实施例11

[0268] 来自基因表达共变的推定多能性因子。基因在群体中共变的观察提出了相关性是否可能揭示基因调节或功能中的共同性的问题。在细胞的复杂混合物中,这样的推断的尝试可能被混淆,因为基因间相关性可主要来自细胞类型之间的微小差异,其反映大规模表观遗传学改变而不是特定的调节程序。在仅由单个细胞类型组成的群体中,情况是不同的:在这里,可能更乐观的是细胞状态的波动可以揭示功能依赖性。mES细胞群满足这个要求,因为它显示相对少的离散结构,超出上述小亚群的存在。

[0269] 为了测试基因表达共变是否可能包含调节信息,本实施例使用定制的网络邻域分析(NNA)方案研究了已知的多能性因子的共变伴侣(图16,参见下文)。该方案定义与感兴趣的给定的基因最密切相关并且也彼此最紧密相关的基因集。考虑到相关性对采样效率的敏感性(图14G,等式(3)),NNA分析(其仅对相关性网络拓扑敏感)将比简单地关联高度相关的基因更加稳健。值得注意的是,应用于多能性因子Nanog和Sox2,NNA方案强烈富集其它多能性因子:在Nanog的20个最接近毗邻物中,11个被记录为多能性因子,另外三个与多能性相关,一个(Slc2a3)与Nanog同线。只有一个基因(Rbpj)已被显示是多能性所不必要的,剩下的四个基因没有以前记录的与ES细胞的相关性。吸引人的是预测这些基因在维持多能性状态中也发挥功能性作用。类似地,Sox2的整个邻域包括与多能性直接或间接相关的因子-包括核心多能性因子(6/9个基因);苏氨酸分解代谢酶Tdh,其最近被显示在ICM中以高水平表达并且是维持多能性状态所需的;Pcbp1被显示为多能性因子Ronin/Thap11的结合伴侣,并且翻译起始因子亚基Eif2s2被显示响应于Stat3过表达而上调。有趣的是,相同的分析可以提供对其它生物学通路的洞察。例如,细胞周期蛋白B(Ccnb1)的邻域很小,但含有其它核心细胞周期基因Cdk1、Ube2c和Plk1。

[0270] 然而,该方案通常不适用于所有调节功能:发现似乎独立于mES细胞生物学的许多其它通路似乎没有有意义的NNA关联。这表明单细胞共变可以捕获对所研究的细胞的生物

学最具特异性的波动,并且可以更一般地用于通过弱的通路特异性扰动人为地产生波动来鉴定其它生物学通路组分。

[0271] 实施例12

[0272] 分化ES细胞的群体动力学。在LIF撤除后,mES细胞通过异质的但不良表征的过程分化,从而最终导致主要体细胞(外胚层)谱系的形成。预先存在的PrEn细胞的命运是不清楚的,也不清楚其它细胞谱系是否可能会暂时出现然后消失的问题。在单细胞分析中,在LIF撤除后,分化ES细胞群经历群体结构的显著变化,其可以根据高度可变基因的表达从分层聚类细胞定性地理解(图17A)。这些变化和以下分析反映了无指导的分化方案;这会对未来将相同的方法应用于指导的分化方案以鉴定中间细胞类型的固有异质性和变异如何取决于信号传导是指导性的。

[0273] 为了分析细胞群体中发生的变化并验证数据的质量,本实施例首先检查多能性因子和分化标记物的基因表达动力学(图17B,17C)。多能性因子Rex1/Zfp42和Esrrb水平的平均表达迅速下降;Pou5f1和Sox2逐渐下降;外胚层标记物Krt8稳定增加;和Otx2(从ICM转移到外胚层状态所需的转录因子)瞬时增加到第2天,然后减少。然而明显的是,平均表达不代表每个细胞中的动力学:一些细胞不能表达外胚层标记物,并且甚至在LIF撤除后7天,一部分细胞继续表达多能性因子(图17C),表明ES细胞分化的时间设置本身是异质性的。

[0274] 对从所有时间点聚集的细胞进行PCA分析以鉴定这种异质性是否反映了全局趋势(图17D),并且发现即使在LIF撤除后7天后,一部分(5%, $N=799$)细胞与mES细胞群体重叠。最大时间异质性在LIF后4天是明显的,其中细胞沿着第一主成分广泛分布在mES细胞和分化状态之间。PCA分析还显示在第2天和第4天富集与从多能状态出现时发生的代谢变化一致的强代谢标记(顶部GO注释:细胞代谢过程, $p=1.4\times 10^{-8}$)。

[0275] 除了由分化中的不同步产生的异质性之外,在四和七天后,存在具有不同基因表达模式的新出现亚群的证据,并非所有这些都可以立即归因于已知的细胞类型。通过t-SNE在这些时间点可视化群体结构(图17G和图25),并在表3中列出不同的亚群标记物。在LIF撤除后两天和四天,鉴定稀少的Zscan4⁺细胞,先前被鉴定为稀少的滋养外胚层形成细胞(REF);到第7天不再检测到该群体。在4天和7天时,出现表达通常母本印记的基因H19、Rhox6/9、Peg10、Cdkn1等的另一个更少稀少的群体,提示可能与早期原生殖细胞分化相关的广泛的去甲基化。

[0276] 除了这些群体,可以在所有时间点检测到驻留的PrEn细胞(图17F,17G),其中PrEn群体在LIF撤除后两天和四天显示扩增,但在LIF后七天停滞。总体而言,分析揭露了ES细胞分化的时间异质性和不同的和新的ES细胞亚群的动力学。

[0277] 图17显示了分化ES细胞中的时间异质性和群体结构。图17A显示了在每个时间点通过对高度可变基因的细胞间相关性的热图进行分层聚类定性地看到LIF撤除后全局群体结构的变化。图17B显示在LIF撤除后基因表达的平均动力学与已知的分化模式一致。图17C显示了通过表达给定数量的计数的细胞的分数的概率密度(小提琴)图显示的图17B中的基因的动力学。数据点显示前5%的细胞。图17D和17E显示来自显示快速瞬时变化(PC 2)和分化不同步(PC 1)的多个时间点的3,034个细胞的前两个主成分(PC)(图17D)和PC负载(图17E)。图17F显示作为LIF后时间的函数的外胚层和PrEn细胞的分数的动力学。图17G显示了在LIF后4、7天后分化ES细胞的tSNE图谱揭示瞬时和新出现的群体亚结构,以及在LIF后4天

后的基因的tSNE图谱(右图)揭示推定的群体标记物。

[0278] 图25显示了在LIF撤除后4天的主基因的tSNE图。此图再现具有全基因注释的图17G。

[0279] 实施例13

[0280] mES细胞分化期间混杂基因表达波动的减少。本实施例解决了以下假设:mES细胞的特征在于混杂基因表达,牵涉许多基因的弱偶联表达,其在分化过程中变得精细。在基因表达更混杂的情况下,可以预期细胞占据基因表达的更大的子空间,如通过其中分布细胞的独立维数的数量所测量的。相比之下,更受控的模式基因表达(即使是多种细胞类型的混合物)会将细胞限制至反映基因表达的一种或多种相干态的较低维流形。本实施例评价了ES细胞和分化细胞的内在维数。发现基因表达的内在维数在分化后减少(图17H),而随机数据和纯RNA的维数显著高于ES细胞。该分析支持ES细胞异质性与混杂的弱偶联基因表达相关的假设,其与由分化中的不同步和细胞类型的发散引起的Lif撤除后的异质性形成对比。

[0281] 图17H显示了在LIF后7天和mES细胞的基因表达可变性的内在维数的估计,从而显示在分化期间波动的收缩子空间。结果与纯RNA形成对比,纯RNA应缺乏相关性并因此显示最大波动子空间。

[0282] 实施例14

[0283] 这些实例显示用于单细胞捕获、条形码化和转录组分析的平台的建立,而对可处理的细胞数没有物理限制。这些实例显示了高捕获效率、快速收集时间、非常低的液滴间CV和接近采样统计所施加的限制的技术噪声。这些在不同的实验、装置、BHM批次和乳液体积中是可重复的(表1)。这些可以容易地应用于小的临床样品(包括肿瘤样品和组织微活检)的单细胞转录组学,从而给出组织异质性的定量图片。取决于所需的应用,这允许通过收集不同的乳液体积来交换测序深度与细胞群体的大小。这些实例允许基于基因表达常规鉴定细胞类型,甚至稀少的亚群。由于低测量噪声,这些允许区分离散细胞类型与基因表达中的连续波动,如在ES细胞中的情况。除了分类细胞之外,这种类型的数据还对于基于协方差鉴定基因之间的推定的调节联系是有价值的,例如通过利用群体中个体细胞之间的天然和可能的微妙变化。这些实例仅突出显示了这种推断的几个简单示例(图16),但是这种类型的单细胞数据适合于更正式的逆向工程方法。

[0284] 这些实例可以为许多生物学问题提供准确的信息。这通过ES细胞异质性及其在早期分化期间的动力学的复杂和挑战性问题来举例说明。ES细胞不支持不同细胞类型的大亚群,因此,它们的异质性的分析需要灵敏的方法。为了解释来自这些细胞的数据,开发了单细胞噪声的统计模型,其解决了生物基因间相关性如何受到低捕获效率和液滴之间的技术可变性的影响的问题,并且我们利用来自机器学习的工具以可视化基因表达“空间”中的细胞的更高维度组织和该组织的动力学。该分析提供了证据来支持ES细胞在血清和LIF中维持时在更多的多能状态和更多的分化状态之间波动的假说。然而,除了表达分化标记物例如Krt8/18、S100a6和Fgf5的ICM样群体以外,还鉴定了与原始内胚层命运相关的其它ES细胞亚群、表达原始生殖细胞标记物Blimp1/Prdm1的另一亚群以及具有由高水平的ER相关蛋白或E3连接酶Lin-41/Trim71标记的不太明显的命运关联的亚群。这些小细胞亚群的无偏鉴定需要由drop-Seq实现的规模。这通过原始内胚层样细胞举例说明,其在所有时间点代表小于约1%的细胞群体,并且太稀少以致于我们在仅仅100-200个细胞的较小样品中确信

地检测到。

[0285] 在技术方面,drop-SEQ平台被开发用于全转录组RNA测序,但是该技术是高度灵活的,并且应该容易适应于需要RNA/DNA分子条形码化的其它应用,例如其它RNA-Seq方案、聚焦于小基因组的靶向测序方法、ChIP-Seq、基因组测序或染色质邻近分析(Hi-C)。一个实施方式使用非常简单的液滴微流体芯片,其仅具有单个流动聚焦接合部(图13C)以组合细胞、条形码和RT试剂。平台的其它形式可以进一步利用液滴微流体功能,以允许通过试剂皮注射(pico-injection)到现有液滴中的多步酶促反应,或者在通过芯片上分选液滴进行测序之前执行靶细胞富集。此外,drop-SEQ应该能够容易地结合靶向相对低捕获效率的生物化学创新。

[0286] 图16显示基因相关性网络邻域揭示多能性相关因子。通过选择具有至少三个相互毗邻物的网络毗邻物生成Nanog(图16A)、Sox2(图16B)和细胞周期蛋白B(图16C)的连接的相关性网络邻域(参见下文)。在图16A和16B中,灰色框表示先前验证的多能性因子;Calcoco2、Eif2s2和Igfbp2框指示先前报道与多能状态相关的因子。

[0287] 实施例15

[0288] 本实施例举例说明了上述实施例中使用的各种方法和系统。

[0289] 微流体装置设计与操作。在这些实施例的一些中使用的微流体装置的设计在图18中示出并集成了几个特征。如上所述,其包含用于i)条形码化水凝胶微球(BHM)、ii)细胞悬浮液、iii)逆转录(RT)和裂解试剂混合物以及iv)载体油的四个入口,以及用于液滴收集的一个出口。为了减少由于注射泵的机械结构可能引起的流动波动,流体电阻器以蛇形通道的形式被并入,而每个入口处的无源过滤器防止通道堵塞。该装置包括两个接合部,一个用于将三个水性输入物合并在一起,以及第二接合部用于样品包封,其中水相和油相遇并且产生液滴。为了稳定液滴以防聚结,使用溶解在HFE-7500(3M)氟化流体中的0.75%(w/w)EA-表面活性剂(RAN Biotechnologies Inc.,)。仔细选择微流体通道的尺寸以最大化BHM和细胞共包封事件的数量。BHM再注入通道的宽度(60微米)被设计为使得通过该通道的BHM(直径为63微米)变得轻微挤压,从而便于它们紧密装填和排列成单个队列。进入主通道(70微米宽)的BHM然后可以在被包封成个体液滴之前在流动的下流自由移动。由于它们的紧密装填,BHM的到达变得高度规则,从而允许每个液滴几乎100%的单珠粒加载。这确保i)包封成液滴的几乎每个细胞暴露于一个条形码化的引物,和ii)非条形码化的细胞的损失最小。

[0290] 软光刻。根据既定的方案制造具有80微米深的矩形微流体通道的微流体装置。简言之,用均匀的80微米膜厚度的SU-8 3050光致抗蚀剂(MicroChem)涂覆3英寸尺寸的硅晶片,在65℃下烘焙20分钟,并通过具有图18中所示的相应设计的掩模暴露于365nm UV光40秒(在 $\sim 8\text{mW cm}^2$),以及在95℃下烘烤5分钟。将未聚合的光致抗蚀剂用丙二醇单甲醚乙酸酯溶解,将硅晶片用异丙醇冲洗并在95℃热板上干燥1分钟。将PDMS碱和交联剂(Dow Corning)以10:1的比例混合,并将 $\sim 30\text{mL}$ 倒入含有显影硅晶片的培养皿中,脱气并在65℃孵育过夜。然后将PDMS层剥离,并用1.2mm活检穿孔器(Harris Uni Core)穿刺入口-出口端口。然后用氧等离子体处理PDMS的图案化侧面并结合至干净的载玻片。用防水Aquapel(PPG Industries)处理微通道,然后将该装置用于上述实验中。

[0291] 微流体装置操作。在装置操作期间,用冰冷夹套冷却细胞悬浮液和RT/裂解混合物,并将液滴收集在置于冰冷架(IsoTherm System,Eppendorf)上的单个1.5mL管(DNA

LoBind, Eppendorf) 中。为了防止在RT孵育期间由于蒸发而导致的水分损失,将200微升矿物油层(Sigma)置于乳液的顶部。在整个实验中,使用细胞悬浮液的100微升/小时、RT/裂解混合物的100微升/小时、BHM的10-20微升/小时和载体油的80微升/小时的流速,以在15个液滴/秒的频率下产生4nL滴剂。将每个水相通过连接至放置在注射泵(Harvard Apparatus, PC2 70-2226)上的无菌1mL注射器(Braun)的针头的聚乙烯管(ID 0.38×OD 1.09mm, BB31695-PE/2)注射到微流体装置中。

[0292] 加载条形码化的水凝胶微球(BHM)到微流体装置中。合成后,将BHM储存在含有10mM Tris-HCl (pH 8.0)、10mM EDTA、0.1% (v/v) Tween-20的 $T_{10}E_{10}T_{0.1}$ 缓冲液中。在加载到微流体芯片上之前,在含有10mM Tris-HCl (pH 8.0)、0.1mM EDTA和0.1% (v/v) Tween-20的 $T_{10}E_{0.1}T_{0.1}$ 缓冲液中洗涤BHM,然后重悬于补充有0.5% (v/v) IGEPAL CA-630的1X RT缓冲液(Invitrogen Superscript III缓冲液),并通过在5000rpm离心2分钟浓缩。在除去上清液后,将BHM第二次浓缩以实现紧密装填,并最终直接加载到连接至充油注射器的管中,以用于注射到微流体装置中。BHM样品的组成是100微升浓缩的BHM、20微升10% (v/v) IGEPAL CA-630、40微升5×第一链缓冲液和40微升无核酸酶的水(总等分试样体积200微升)。

[0293] 细胞制备和注射。细胞包封过程依赖于细胞至设备中的随机到达。为了使两个或更多个细胞进入相同液滴的情况最小化,使用稀释的细胞悬浮液(~50-100,000个细胞/mL),以获得5-10个液滴中1个细胞的平均占有率。为了防止细胞在注射器或系统的其它部分中沉降,将细胞悬浮在具有16% (v/v) 密度梯度溶液Optiprep(Sigma)的1X PBS缓冲液中。通常使用20,000个细胞,悬浮在总体积200μl的160微升5X PBS(Lonza 17-516F)、32微升Optiprep(Axis-Shield 1114542)和8微升1% (v/v) BSA(Thermo Scientific B14)微升中。

[0294] 反转录/裂解混合物。RT/裂解混合物包括25微升5X第一链缓冲液(18080-044 Life Technologies)、9微升10% (v/v) IGEPAL CA-630(#18896Sigma)、6微升25mM dNTP(Enzymatics N2050L)、10微升0.1M DTT(#18080-044, Life Technologies)、15微升1M Tris-HCl (pH 8.0) (51238Lonza)、10微升鼠RNA酶抑制剂(M0314, NEB)、15微升SuperScript III RT酶(200U/微升, #18080-044, Life Technologies)和60微升不含核酸酶的水(AM9937Ambion),总体积为150微升。

[0295] 用于生产液滴的表面活性剂和载体油。载体油是具有1% (w/w) EA表面活性剂(RAN Biotechnologies)的HFE-7500氟化流体(3M)。EA-表面活性剂是平均分子量为~13.000g mol⁻¹的三嵌段共聚物。其具有通过聚乙二醇(M_w ~600g mol⁻¹)头基连接的两个全氟聚醚尾(M_w ~6.000g mol⁻¹)。表面活性剂高度可溶于氟化流体中,并且几乎不溶于水相,从而提供~2mN/m的平衡界面张力。

[0296] 液滴内的条形码化。在细胞包封后,通过将含有乳剂液滴的管在冰上暴露于紫外光(~10mW/cm²下的365nm, BlackRay汞灯)从BHM中释放引物。接下来,将管加热至50℃并孵育2小时以使cDNA合成发生,然后通过70℃加热15分钟终止。然后将乳液在冰上冷却1分钟,并通过加入1体积的PFO溶液(20% (v/v) 全氟辛醇和80% (v/v) HFE-7500)去乳化。将来自破裂液滴的水相转移到单独的DNA Lo-Bind管(Eppendorf)中,根据CEL-SEQ方案用文库制备部分所述的修饰进行处理。

[0297] 条形码化的水凝胶微球的合成和质量控制。BHM合成依赖于补充有丙烯酸酯改性

的DNA引物(其在丙烯酸酰胺聚合时被掺入至水凝胶网孔中)的丙烯酸酰胺:双丙烯酸酰胺溶液的微流体乳化。聚合后,BHM从液滴中释放,洗涤几次并通过分割库合成进行处理以用于组合条形码化。下面概述了进行这种水凝胶珠粒合成随后组合条形码化的详细方案。

[0298] BHM合成开始于使用图19所示的微流体芯片将凝胶前体溶液乳化成62微米大小的液滴。分散相的组成为10mM Tris-HCl (pH7.6)、1mM EDTA、15mM NaCl,含有6.2% (v/v) 丙烯酸酰胺、0.18% (v/v) 双丙烯酸酰胺、0.3% (w/v) 过硫酸铵和50微摩尔丙烯酸酯修饰的DNA引物(IDT,序列参见图20A)。作为连续相,使用含有0.4% (v/v) TEMED和1.5% (w/w) EA-表面活性剂的氟化流体HFE-7500。水相的流速为400微升/小时,油相的流速为900微升/小时。将液滴收集到1.5mL管中处于200微升矿物油之下,并在65℃孵育12小时以允许珠粒发生聚合。将所得的固化的珠粒用在HFE-7500油中的1mL 20% (v/v) 的1H,1H,2H,2H-全氟辛醇(B20156, Alfa Aesar)洗涤两次,用在己烷(BDH1129-4LP, VWR)中的1mL 1% (v/v) Span 80 (S6760, Sigma)洗涤两次,在每个步骤之间孵育0.5-1分钟,最后在5000rcf离心30s。最终离心后,抽吸己烷相,将得到的BHM沉淀物溶解于1mL TEBST缓冲液(10mM Tris-HCl (pH8.0), 137mM NaCl, 2.7mM KCl, 10mM EDTA和0.1% (v/v) Triton X-100)。为了除去痕量的己烷,将珠粒用1mL TEBST缓冲液以5000rcf洗涤3次,每次30s,最后重悬于1mL TEBST缓冲液中并在4℃下储存。这些BHM含有尺寸为~100nm的孔。此外,具有~1kPa的弹性模量的珠粒是“湿软的”,这允许它们被装填成浓缩的凝胶块而不损失其完整性。

[0299] BHM分割库组合条形码化。为了在水凝胶微球上制备条形码化的引物,使用总结在图20B中的两步酶促延伸反应。开始时,使用用9微升的15微摩尔引物5'-W1*-bc1-PE1*,其编码条形码的前半部分(其中“bc1”表示每个孔的独特序列,核苷酸序列信息也参见表4)预先加载的384孔板。向每个孔中加入6微升反应混合物,含有约40,000个水凝胶珠粒(携带5'-Ac-PC-T7p-PE1引物)、2.5x等温扩增缓冲液(NEB)和0.85mM dNTP (Enzymatics)),总计~10⁷个珠粒。在85℃变性2分钟并在60℃杂交20分钟后,加入5微升Bst酶混合物(在1X等温扩增缓冲液中的1.8U Bst 2.0和0.3mM dNTP),得到每孔20微升的最终体积。在60℃孵育60分钟后,通过向每个孔中加入20微升终止缓冲液(100mM KCl, 10mM Tris-HCl (pH 8.0), 50mM EDTA, 0.1% (v/v) Tween-20)终止反应,并在冰上孵育30分钟以确保EDTA螯合镁离子并使Bst酶失活。接下来,将珠粒收集到50mL Falcon管中,以1000rcf离心2分钟,并用50mL含有10mM EDTA的STOP缓冲液洗涤三次。为了除去第二链,将凝胶悬浮于20mL 150mM NaOH、0.5% (v/v) Brij 35P中,并用10mL的100mM NaOH、0.5% (v/v) Brij 35P洗涤两次。然后用缓冲液100mM NaCl、100mM Tris-HCl (pH 8.0)、10mM EDTA、0.1% (v/v) Tween-20中和碱性溶液,并在10mL T₁₀E₁₀T_{0.1}缓冲液(10mM Tris-HCl (pH 8.0), 10mM EDTA, 0.1% (v/v) Tween-20)中洗涤一次,并在10mL T₁₀E_{0.1}T_{0.1}缓冲液(10mM Tris-HCl (pH 8.0), 0.1mM EDTA, 0.1% (v/v) Tween-20)中洗涤两次,最后将珠悬浮在1.3mL DST缓冲液中。

[0300] 对于第二个条形码化步骤,制备第二个384-微量滴定板,其预先加载有9微升的15微摩尔引物5'-T19V*-UMI-bc2-W1*(其中“bc2”表示每个孔的独特的序列,UMI是随机六核苷酸,对于序列信息也参见表4),并如对于第一384孔板所述重复程序。

[0301] 珠粒上ssDNA引物的定量。为了定量每个BHM的ssDNA引物的量,使用靶向未延伸的DNA“短线”的互补DNA探针(PE1)、一个延伸步骤后的条形码化引物(W1)和两个延伸步骤后的引物(T₁₉V) (关于序列信息参见表1)进行荧光原位杂交(FISH)。通过将约4000个DNA条形

码化的珠粒与10微摩尔FAM标记的探针一起悬浮在杂交缓冲液(1M KCl, 5mM Tris-HCl (pH 8.0), 5mM EDTA, 0.05% (v/v) Tween-20)中, 在室温下以40微升体积进行20分钟杂交。高盐浓度用于避免靶向 T_{19} V的探针(dA20-FAM) (其即使在室温下也具有弱结合)的熔化。在缺少DNA引物的微球中验证背景荧光的缺失。孵育后, 将珠粒用1.4mL杂交缓冲液洗涤三次, 重悬于40微升中, 并在共焦显微镜(Leica)下记录荧光强度。具有PE1*-FAM、W1*-FAM和dA20-FAM的珠粒的平均荧光强度分别为 2286 ± 271 、 1165 ± 160 和 718 ± 145 (图21A-21D)。这对应于W1/PE1的 $\sim 50\%$ 和多聚T/W1的60%的掺入效率, 其给出31%的最终效率或每个珠粒 ~ 15 微摩尔的完全条形码化的ssDNA引物。计算BHM体积, 这相当于每个珠粒 $\sim 10^9$ 个拷贝的完全延伸的ssDNA引物。

[0302] 为了验证引物从水凝胶网孔的释放, 将约4000个珠粒悬浮于20微升DST缓冲液中并暴露于UV光($\sim 10\text{mW}/\text{cm}^2$ 处的365nm) 8分钟。使用BioAnalyzer高灵敏度DNA分析试剂盒(Agilent Technologies)的1微升上清液的凝胶电泳图确认存在三条DNA条带(图21E), 这与上述FISH结果一致。

[0303] 来自单个BHM的引物的单分子测序。为了在合成后测试BHM的组成, 随机挑取10个BHM并使用Illumina MiSeq测序平台进行测序。为此目的, 首先将BHM与如上所述的荧光FISH探针(PE1-FAM)杂交, 并在荧光照明下使用解剖显微镜(Nikon)手动挑取, 并转移至预填充有5微升DNA悬浮(DS)缓冲液(10mM Tris-HCl pH 8.0, 0.1mM EDTA)的0.2mL PCR管中。然后将管暴露于UV光($\sim 10\text{mW}/\text{cm}^2$) 15分钟, 同时将其保持在冰上。在UV暴露后, 向管中加入0.5微升的5微摩尔PE2-(条形码) $_n$ -A19引物(这里, n 表示10个不同的条形码), 并与4.5微升的Bst 2.0即用型反应溶液混合。然后将具有10微升最终体积的样品在室温下孵育10分钟, 在95°C下灭活3分钟, 并在冰上冷却。接下来, 将含有50% (v/v) Kapa HiFi HotStart预拌混合物(2X, KK2601)、15% (v/v) PE1/PE2引物和35% (v/v) 不含核酸酶的水的20微升主混合物加入每个管, 并用PCR扩增DNA(95°C 5分钟, 30个循环的98°C 20秒、60°C 15秒、72°C 30秒, 最后步骤为72°C 5分钟)。通过凝胶电泳评估PCR产物的大小, 用GenElute PCR CleanUp试剂盒(Na1020-1KT, Sigma)纯化, 并将所有样品稀释至10ng/微升。在最终步骤中, 将所有样品汇集在一起, 并使用MiSeq Illumina平台按照制造商的建议进行测序。来自10个个体珠粒的引物的测序结果示于图21F-21H中。

[0304] 每次测序运行的细胞数量的限制。对于条形码化水凝胶微球(BHM)的大型库, 每个微球携带 N 个条形码中的一个, 在两个或细胞携带相同条形码之前可以捕获的细胞的最大数量是多少? 这个问题类似于所谓的生日问题, 条形码类似于一年中的天, BHM类似于一个房间里的人。从采样 n 个BHM中观察到的条形码的预期数量是 $n_{\text{obs}} = N(1 - e^{-n/N})$ 。因此, 预期的多条形码化错误(定义为携带相同条形码的细胞的分数的分数)近似为 $f_{\text{err}} = 1 - n_{\text{obs}}/n$ 。当 $n \sim N$ 时, 错误增多, 因此实际上采样细胞的数量必须远小于条形码的数量, 即 $n \ll N$, 因此获得条形码化的单细胞的极限是 $f_{\text{err}} \approx n/2N$ 。条形码化的单细胞 n 的数量取决于容许的错误, 例如, 允许小于 $f_{\text{err}} = 1\%$ 的错误需要上限 $n = N/50$ 。因此, 对于在我们的实验中对应于两个384孔板的 $N = 384^2$ 的值, 在极限 $n = 2,949$ 个细胞处产生1%的多条形码化错误。在实践中, 可以使用更少的细胞来产生可忽略的多条形码化错误。

[0305] 细胞培养物制备。在37°C、5% CO_2 和60-80%湿度、密度为 $\sim 3 \times 10^5$ 个细胞 mL^{-1} 的条件下, 在用明胶预包被的培养瓶内的ESC基础培养基中维持小鼠胚胎干(mES)细胞。ESC培养

基含有补充有15% (v/v) 胎牛血清 (Gibco)、2mM L-谷氨酰胺、1×MEM非必需氨基酸 (Gibco)、1% (v/v) 青霉素-链霉素抗生素、110微摩尔β-巯基乙醇、100微摩尔丙酮酸钠的无酚红DMEM (Gibco)。为了引导分化,ESC基础培养基补充有终浓度为1000U/mL的白血病抑制因子 (LIF),对于非引导的mES分化,培养基不含LIF。在LIF撤除的2天内,培养物经历显著的形态学变化,表明mES细胞的分化。

[0306] 在包封前,用1×PBS (不含Mg²⁺和Ca²⁺离子) 洗涤烧瓶并在37℃用1×胰蛋白酶/EDTA溶液处理3分钟。通过加入等体积的ESC基础培养基淬灭胰蛋白酶。将分离的细胞在260g离心3分钟,并重悬于~3mL新鲜ESC基础培养基中。在通过40微米大小的过滤器后,用血细胞计数器计数细胞,并在补充有0.04% (v/v) BSA和16% (v/v) OptiPrep溶液的0.5x PBS中稀释以获得期望量的细胞 (通常为200微升中20,000个细胞)。将悬浮液转移到与微流体装置连接的1mL注射器中,并以100微升/小时的流速注射。按照该程序,在第1天具有LIF并且在第2天早些时候、第2天晚些时候、第4天和第7天无LIF的情况下制备mES细胞。

[0307] 将K-562细胞系 (ATCC, CCL-243) 在补充有10% (v/v) 胎牛血清和1% (v/v) 青霉素-链霉素的DMEM中在37℃、5%CO₂和60-80%湿度气氛下以~3×10⁵个细胞mL⁻¹的密度维持。对于包封实验,如上所述制备K-562细胞,但使用DMEM培养基,并以1:1的比例与mES细胞混合。

[0308] DNA文库制备。文库制备基于修改的CEL-Seq方案。DNA文库制备的工作流程可以总结如下:RT-->ExoI-->SPRI纯化 (SPRIP) -->SSS-->SPRIP-->T7体外转录线性扩增-->SPRIP-->RNA片段化-->SPRIP-->引物连接-->RT-->文库富集PCR。

[0309] 参考Jaitin DA等人 (2014) Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343 (6172):776-779中详述的方案,对该方案进行了以下修改:RT引物包括P5/PE1衔接子,而连接引物包括P7/PE2衔接子,与该方案中的取向相反;在ExoI处理之前,将来自破碎液滴的水相在4℃下以14krcf离心15分钟以沉淀细胞碎片和凝胶;在ExoI处理期间,将10U HinFI加入以消化可能在RT反应期间形成的引物二聚体;在线性扩增后省略原有的DNA酶消化步骤;在进行线性扩增之后进行下一步之前,在Agilent BioAnalyzer上分析所得扩增的RNA文库;在引物连接之前,将样品用虾碱性磷酸酶处理30分钟。最终文库富集PCR所需的最终PCR循环数目为10-13个循环。其余步骤保持不变。

[0310] 生物信息学分析。如表1所详述,在Illumina MiSeq、HiSeq 2500和NextSeq机器上进行配对末端测序。使用读数1来获得样品条形码和UMI序列;然后将读数2映射到如下所述的参考转录组。首先基于在读数1中由W1衔接子序列 (参见图20和表4) 分开的两个样品条形码组分的存在过滤读数。然后使用Trimomatic (5) (版本0.30;参数:LEADING:28 SLIDINGWINDOW:4:20 MINLEN:19) 修剪读数2。将每个读数的条形码针对384²个预定条形码的列表进行匹配,并校正高达两个核苷酸错配的误差。丢弃具有与参考列表分离多于两个核苷酸的条形码的读数。然后将读取分割到条形码特异性文件中用于映射和UMI过滤。

[0311] 使用Bowtie (版本0.12.0,参数:-n 1-l 15-e 300-m 200-best-strata-a) 将修剪的读数与小鼠转录组进行比对。数据集还用不同的bowtie参数集重新处理,而不改变分析的定性结果。使用来自UCSC mm10基因组装配的所有注释转录物 (用125bp多聚T尾延伸) 构建参考转录组。使用定制的Python和PySAM脚本将映射的读数处理成每个基因的UMI过滤的转录物的计数。来自bowtie的比对按以下方式过滤: (1) 对于每个读数,通过选择最接近转

录物末端的比对,保留所有同工型之间每个基因至多一个比对。(2) 如果读数与多个基因比对,我们排除距离转录物末端超过400bp的任何比对;这是由CEL-SEQ方法的强3' 偏倚驱动的。(3) 排除映射至超过10个基因的读数,以及(4) 进行下面段落中描述的UMI过滤步骤。最后,(5) 如果读数在UMI过滤后仍然与多于2个基因对齐,则完全排除读数。在报告计数时,对于每个基因,还报告了在至少一个读数中不能从其中区分的任何其它基因;这允许排除由映射模糊性导致的在我们的下游分析中的假相关。通过以每个读数最多1-4个比对重新处理数据来确认流水线对该最终步骤的稳健性。在对每个样品分别进行步骤(1-5)后,将得到的基因表达表连锁起来并加载到MATLAB中用于分析。

[0312] UMI过滤(上述步骤4)如下进行。每个不同的UMI通过携带UMI的读数集与一组基因相关联。对于每个UMI,鉴定了可以解释具有该UMI的全组读数的最小基因组。这个问题被称为“命中集合问题(Hitting Set Problem)”(或“集合覆盖问题(Set Cover Problem)”)。应用贪婪算法获得每个UMI的最节省的基因集。仅保留一个读数/基因/UMI。使用这种方法,基因的一些子集可能仍然是彼此不可区分的,因为它们由相同的一组不明确比对的读数支持。因此,上一段中的步骤(5)用于消除超出预定阈值的模糊读数。为了举例说明UMI过滤步骤,考虑具有两个读数的UMI,第一个与基因A和B比对,第二个与基因B和C比对。虽然两个读数都不明确地比对,但是基因B可以解释两个读数的存在,因此丢弃与基因A和C的比对,并且保留基因B的两个读数中的仅一个。

[0313] UMI过滤计数归一化。在归一化之前,每个样品条形码的总UMI-过滤映射(UMIFM)的计数的变异为21%至55%(变异系数),参见表1。在分化期间CV似乎增长,表明总UMIFM计数中的一些变异产生于细胞大小的差异而不是RT效率的变化。所有计数通过总计数归一化来归一化,即细胞*i*中基因*j*的归一化计数以未归一化计数 $m_{i,j}$ 的方式给出为 $\hat{m}_{i,j} = m_{i,j} \bar{M} / M_i$,其中 $M_i = \sum_j m_{i,j}$ 和 \bar{M} 是所有细胞中 M_i 的平均值。使用子样本归一化也获得类似的结果。

[0314] 预测方法灵敏度。此部分导出图14E中的灵敏度曲线(实线曲线)的形成,该曲线是对于其中对检测的唯一限制是捕获效率 β 或beta(其被假定为在所有基因转录物中是无偏和均匀的)的情况预测的。所有其它偏倚,例如序列特异性或长度特异性偏倚,均被假定为可忽略。良好的拟合强化了这些假设。假设*n*为给定液滴中给定基因的转录物的数量。检测该液滴中的基因的零转录物的概率是 $p_0(n) = e^{-\beta n}$ 。然后通过将 $p_0(n)$ 在*n*(在纯RNA样品的情况下,其是关于平均值 \bar{n} 泊松分布的)的分布上边缘化来获得灵敏度*S*。获得 $S(\bar{n}) = 1 - \sum_{n=0}^{\infty} p_0(n) \text{Poiss}[n, \bar{n}]$,从而给出 $S(\bar{n}) = e^{-\bar{n}(1-e^{-\beta})}$,其是在图14E中绘制的曲线,其中 β (beta)的值从图14D测量。该曲线还可以被鉴定为在 β (beta)处评估的泊松分布的矩生成函数(MGF)。拟合的质量要求液滴之间的 β (beta)的变化小,这与对照液滴之间的总计数的低CV一致。对于非对照样品,每个基因的输入分布不再是泊松分布,并且检测频率 $S(\bar{n})$ 对于每个基因而言是不同的,并且在这里给出的假设下,等于在 β (beta)处评估的基础表达分布的MGF。

[0315] 用于PCA和tSNE分析主基因集的选择和过滤。由于每个基因携带与其它基因不相关的固有采样噪声,预期对于全转录组数据,在所有基因中观察到的大部分变异性将不会由顶部主成分解释。由于相同的原因,如果在评价细胞间相关性中包括大量“旁观者”基因(其在群体之间变化很小),则细胞亚群之间的差异可能看起来很弱。为了克服这些采样限制,使用仅选择来反映已知ES细胞生物学的基因子集来分析ES细胞群体结构,同时还报告

每个时间点的最可变基因。选择合适的基因集的一般策略如下：(1) 对于每个时间点，包括前200个最可变的基因，如通过 v -得分确定的(表2)，其与基因Fano因素密切相关；这些基因用参与ES细胞生物学的基因的精选列表补充。(2) 为了减少基因集，使用初始基因集对细胞群体进行初步主成分分析(PCA)，并使用结果来仅选择“主基因”，即促成非随机主成分(PC)的基因，如图150E中的矩阵随机化所确定的。主基因是对于每个非随机PC具有最高负载系数的基因，其中为每个PC动态地选择阈值设置以反映负载系数分布的结构。(3) 对于集合中的每个基因 g ，随后将集合再扩增以包括多达两个与 g 最强相关的另外的基因。该最终步骤允许包括不存在于初始集合中但与高度可变的基因集强烈相关的基因。在步骤(3)结束时获得的最终基因集用于在每个时间点的后续PCA和tSNE分析。

[0316] 网络邻域分析。使用距离度量 $d = (1 - (\text{皮尔逊相关性}))$ 来定义两个基因之间的距离，其中对所有细胞获得相关性。如下构建未加权的定向网络：对于感兴趣的给定基因 G_0 ，定义了其 N 个最近毗邻基因 G_1 (即与 G_0 具有最高相关性的基因)的有向边。将 N 个附加有向边从集合 G_1 的每个成员添加到其 N 个最近毗邻物，一起形成集合 G_2 。所得初步网络总共具有 $(N+1)*N$ 个有向边，并且具有代表 G_0 、 G_1 和 G_2 的多达 $1+(N+1)*N$ 个顶点。然后通过移除具有少于 X 个进入边缘的任何顶点来迭代地修剪网络。最终网络是“基因 G_0 的 X -连接的邻域”。如果它不是空集，则它具有：基因 G_0 ； G_1 的一些成员，其也是 G_1 的至少 $X-1$ 个其它成员的最近毗邻物；和 G_2 的一些成员，其是 G_1 的至少 X 个成员的最近毗邻物。对于图16中绘制的网络，使用参数 $N=50$ ， $X=3$ 。

[0317] 表1. 测序运行统计信息

[0318]

样 品 (LIF 后 天 数)	乳 液 体 积 (uL)	平 台	总 文 库 读 数 (未 过 滤)	细 胞 数 量/ 条 形 码	平 均 过 滤 读 数/ 细 胞	UMI 过滤的 映 射 (UMIFM) 计数/细胞	
						平 均	变 异 系 数 (CV)
Pure RNA 对照	16	HiSe q 2500	166,03 1,332	953	89,11 6	24,19 1	21%
mES LIF+	40	Next Seq	413,13 8,104	935	199,1 93	29,23 9	36%
mES 第 2 天 , (早 些 时 候)	6	HiSe q 2500	119,85 9,024	145	119,3 86	20,52 4	35%
mES 第 2 天 , (晚 些 时 候)	26	MiSe q	17,660 ,550	303	38,78 8	8,441	36%
mES 第 4 天	40	MiSe q	11,557 ,428	683	10,23 7	4,661	43%

[0319]

mES 第 7 天	8	HiSe q 2500	92,805 ,168	169	153,0 35	27,06 5	38%
mES 第 7 天	40	Next Seq	250,18 7,951	799	208,2 31	26,21 6	55%
mES 第 2 天, 早 些 时 候	95	HiSe q 2500	33,751 ,186	2,168	4,987	2,608	42%

[0320] 表3.跨时间点的mES细胞亚群标记物

[0321]

数据集	细胞组描述	细胞组大小 (数量/总数)	聚类索引 (-1<x<1)	高表达基因
mES 第 0 天	原始内胚层样	6/935	0.76	<i>Gsn</i> , <i>Col4a1/2</i> , <i>Serpinh1</i> , <i>Lama1/b1/c1</i> , <i>Sparc</i> , <i>Srgn</i> , <i>P4ha2</i> , <i>Lrpap1</i> , <i>Podxl</i> , <i>Ctsl</i> , <i>S100a10</i> , <i>Pgk1</i> , <i>Slc2a3</i> , <i>Tfpi</i> , <i>Amn</i> , <i>Fbp2</i> , <i>Gpx3</i> , <i>Man2c1os</i> , <i>Lpar3</i> , <i>Cd63</i>
	外胚层样	40/935 (全部) 6/935 (<i>Krt8</i> -高)	0.20 (全部) 0.55 (<i>Krt8</i> -高)	<i>Actg1</i> , <i>Anxa2</i> , <i>Krt8/18/19</i> , <i>Plaur</i> , <i>Cnn1</i> , <i>Tagln</i> , <i>Plin2</i> , <i>Flnc</i> , <i>Tinagl1</i> , <i>Slc2a1</i> , <i>Fam160b2</i> , <i>Mmab</i> , <i>Sfn</i> , <i>Plec</i> , <i>S100a6</i> , <i>Flnb</i> ,

[0322]

				<i>Ngfrap1</i>
	<i>Hsp90</i> -高	10/935	0.47	<i>Atf5, Calr, Hsp90b1, Hspa5, Manf, Pdia6, Creld2, Hyou1, Derl3, Prph, Chchd10</i>
	<i>Prdm1</i> -高	13/935	0.47	<i>Prdm1, Baat, Nsun6, Parp4, Srgn, Ssh1</i>
	<i>Trim71</i> -高	12/935	0.31	<i>Trim71, Cd44, Med10, Myo15, Bcl2/Cep83, Kdm1b, Sbk1, Csf1r, Psg18/20, Prss44</i>
LIF 撤 除 后 4、7 天	<i>H19/Rhox6/9+</i>	14/683, 10/899	0.36, 0.20	<i>H19, Igf2, Rhox6/9, Fabp3, Igfbp2, Sct, Vgf, Pmp22, Rhox5, Itm2a, Rhox5, 1700001F09Rik, Peg10</i>
	多能性-高	21/683, 31/899	0.16, 0.15	<i>Trim28, Tex19.1, Tdh, Tdgfl, Spry4, Sox2, Psors1c2, Pou5f1, Phc1, Ogfod3, Mylpf, Mt1/2, Mkrn1, Mkm1, L1td1, Kcnj14, Gad1, G3bp2, Dnmt3l, Cdh16, Nlrp1a,</i>

[0323]

				4930526L06Rik, 3110021A11Rik
	Zscan4-hi	4/683, 0/899	0.45, N/A	Zscan4a/c/d/f, Fbxo15, Tctst1/3, Dazl, Calcoco2, Mylpf, Dcdc2c, Lmx1a, Ddit4l, Aqr, Clp1, Tmem92, Usp17la, 2310039L15Rik, B020031M17Rik, Gm4027, Gm20767, Gm7102, Gm8994
	原始内胚层样	7/683, 4/899	0.55, 0.65	Gata6, Amn, Cd63, Ctstl, Col4a1/2, Lama1/b1/c1, Upp1, Sparc, P4ha2, Serpinh1, Fst, Lrpap1, P4hb, Ctsh, Clu, Epas1, Pga5

[0324] 表4.DNA寡核苷酸列表

[0325]

1. BHM 合成:	
掺入水凝胶的 DNA 引物	5’-/5Acryd/iSpPC/CGATGACG TAATACGACTCACTATAGGG ATACCACCATGG CTCTTTCCCTACACGACGCTCTTC-3’ (SEQ ID NO:7)

[0326]

条形码 1 (W1*-bc1-PE1*)	5'-AAGGCGTCACAAGCAATCACTC 10987654321 AGATCGGAAGAGCGTCGTGTAGGGAAAGA G-3' (SEQ ID NO:8)
条形码 2/UMI (T19V*-UMI-bc2- W1*):	5'-BAAAAAAAAAAAAAAAAAAAAA NNNNNN 87654321 AAGGCGTCACAAGCAATCACTC-3' (SEQ ID NO:9)
FAM-PE1*	/56-FAM/AGATCGGAAGAGCGTCGTGTAGG GAAAGAG (SEQ ID NO:10)
FAM-W1*	/56-FAM/AAGGCGTCACAAGCAATCACTC (SEQ ID NO:11)
FAM-A20	/56-FAM/AAAAAAAAAAAAAAAAAAAAA (SEQ ID NO:12)
完全装配的 DNA 引物:	CGATGACG TAATACGACTCACTATAGGG ATACCACCATGG CTCTTTCCCTACACGACGCTCTTCCGATCT 1234567890 GAGTGATTGCTTGTGACGCCTT 12345678 NNNNNN TTTTTTTTTTTTTTTTTTTTTV (SEQ ID NO:13)
2. 文库质保:	
RNA 连接:	/5Phos/AGATCGGAAGAGCGGTTCAGCAGGA ATGCC/3SpC3/ (SEQ ID NO:14)
第二 RT 引物	GTCTCGGCATTCCTGCTGAAC (SEQ ID NO:15)
PCR 富集引物:	AATGATACGGCGACCACCGAGATCTACACT

[0327]

CTTTCCTA CACGA (SEQ ID NO:16) CAAGCAGAAGACGGCATACGAGATCGGTC TCGGCATTCT GCTGAAC (SEQ ID NO:17)
--

[0328] 实施例16

[0329] 本实施例举例说明了将细胞包封至液滴中的方法。在本实施例中,使用液滴条形码化测序平台将细胞包封在具有裂解缓冲液、逆转录(RT)试剂和条形码化的寡核苷酸引物的液滴中。从每个裂解细胞释放的mRNA仍保持在相同的液滴中,并在互补DNA(cDNA)的合成期间被条形码化。条形码化后,来自所有细胞的材料通过破坏液滴而组合,并处理cDNA文库用于测序(图27)。

[0330] 在本实施例中,合成了与细胞共包封的条形码化的水凝胶微球(BHM)文库(图27)。BHM携带编码 384^2 (即147,456)个预定义条形码之一的共价偶联的、可光释放的引物。该库大小允许以99%的独特标记随机标记3,000个细胞,并且可以被条形码化的细胞的数量通过使用文库条形码标记每个约3k个细胞的收集管而变得远更大。如果需要在单个文库中进行大规模细胞捕获,则可以以直接的方式扩展该方法。

[0331] 图27显示用于DNA条形码化数千个细胞的液滴微流体平台。单细胞液滴条形码化的示意图。细胞与裂解缓冲液、逆转录(RT)混合物和携带条形码化RT引物的水凝胶微球共包封;包封后,从水凝胶中释放引物,并且在反转录过程中用DNA条形码标记液滴中的cDNA产物。然后将液滴破坏并在测序前线性扩增来自所有细胞的材料。UMI=随机六聚体独特分子标识符。

[0332] 虽然本文已经描述和示出了本发明的若干实施方案,但是本领域普通技术人员将容易想到用于执行本文描述的功能和/或获得本文描述的结果和/或一个或多个优点的各种其它工具和/或结构,并且这些变化和/或修改中的每一个被认为在本发明的范围内。更一般地,本领域技术人员将容易地理解,本文描述的所有参数、尺寸、材料和构造意在是示例性的,并且实际参数、尺寸、材料和/或构造将取决于使用了本发明的教导的具体应用。本领域技术人员将通过使用不超过常规实验而认识到或者能够确定本文所述的本发明的具体实施方案的许多等同物。因此,应当理解,前述实施方案仅以举例的方式给出,并且在所附权利要求及其等同物的范围内,本发明可以以不同于具体描述和请求保护的方式实施。本发明涉及本文所述的每个单独的特征、系统、制品、材料、试剂盒和/或方法。此外,如果这些特征、系统、制品、材料、试剂盒和/或方法不相互矛盾,则两个或更多个此类特征、系统、制品、材料、试剂盒和/或方法的任何组合包括在本发明的范围。

[0333] 本文定义和使用的所有定义应理解为优先于字典定义、通过引用并入的文献中的定义和/或所定义术语的普通含义。

[0334] 除非清楚地相反指示,否则本文在说明书和权利要求书中使用的不定冠词“一个”和“一种”应当被理解为是指“至少一个/一种”。

[0335] 短语“和/或”,如在本文中在说明书和权利要求中所用,应当被理解为意指所连接的元素的“任一个或两者”,即元素在一些情况下结合地存在以及在其它情况下分离地存

在。利用“和/或”列出的多个元素应当以相同的方式来解释,即所连接的元素的“一个或多个”。除通过“和/或”从句明确确定的元素外,还可任选地存在其它元素,无论与明确确定的那些元素相关还是无关。因此,作为非限定性实例,对“A和/或B”的提及,当与开放性措辞诸如“包含/包括”结合使用时,在一个实施方案中可仅指A(任选地包括除B外的元素);在另一个实施方案中可仅指B(任选地包括除A外的元素);在另一个实施方案中,可指A和B(任选地包括其它元素);依此类推。

[0336] 如本文中在说明书和权利要求中所用,“或”应当被理解为具有与如上定义的“和/或”相同的含义。例如,当在列表中分开各项时,“或”或“和/或”应当被理解为是包含性的,即包含至少一个,但也包括许多个元素或一系列元素中的多于一个,以及任选地,包括另外未列出的项。只有明确地指明相反的术语,例如“……中的仅一个”或“……中的恰好一个”或当用于权利要求中时的“由……组成”将指包含多个元素或一系列元素中的正好一个元素。一般地,如本文中所用的术语“或”应当仅在冠有排他性的术语诸如“任一”、“……之一”、“……中的仅一个”或“……中的恰好一个”时被解释为表示排他性选择(即“一个或另一个但非两个”)。“基本上由……组成”,当用于权利要求中时,应当具有其在专利法领域中使用的普通含义。

[0337] 如本文中在说明书和权利要求中所用,关于一系列一个或多个元素的短语“至少一个”应当被理解为意指选自一系列元素中的任何一个或多个元素的至少一个元素,但不一定包括元素列表内明确列出的每一个元素的至少一个并且不排除元素列表中的元素的任何组合。该定义还允许可任选地存在除短语“至少一个”所指的元素列表内明确确定的元素外的元素,无论与那些明确确定的元素相关还是不相干。因此,作为非限定性实例,“A和B的至少一个”(或,等同地,“A或B的至少一个”或,等同地“A和/或B的至少一个”)可在一个实施方案中指至少一个(任选地包括不止一个)A而无B存在(且任选地包括除B外的元素);在另一个实施方案中指至少一个(任选地包括不止一个)B而无A存在(且任选地包括除A外的元素);在另一个实施方案中指至少一个(任选地包括不止一个)A和至少一个(任选地包括不止一个)B(且任选地包括其它元素);依此类推。

[0338] 当本文中使用的单词“约”来指称数值时,应理解本发明的另一实施方式包括了不以单词“约”修饰的该数值。

[0339] 还应当理解,除非明确地指明与之相反,否则在包括不止一个步骤或行为的本文请求保护的任何方法中,所述方法的步骤或行为的顺序不必须地限定于其中叙述所述方法的步骤或行为的顺序。

[0340] 在权利要求中,以及在上述说明书中,所有过渡短语诸如“包含”、“包括”、“携带”、“具有”、“含有”、“拥有”、“牵涉”、“持有”、“由……组成”等被理解为开放性的,即意指包括但不限于。只有过渡短语“由……组成”和“基本上由……组成”应当分别是封闭性的或半封闭性的过渡短语,如美国专利局专利审查程序手册第2111.03节中所示的。

- [0001] 序列表
- [0002] <110> PRESIDENT AND FELLOWS OF HARVARD COLLEGE
- [0003] VILNIUS UNIVERSITY
- [0004] <120> 用于条形码化核酸的系统和方法
- [0005] <130> H0498.70512W000
- [0006] <140> Not Yet Assigned
- [0007] <141> 2015-04-17
- [0008] <150> US 62/072,944
- [0009] <151> 2014-10-30
- [0010] <150> US 62/066,188
- [0011] <151> 2014-10-20
- [0012] <150> US 62/065,348
- [0013] <151> 2014-10-17
- [0014] <150> US 61/982,001
- [0015] <151> 2014-04-21
- [0016] <160> 17
- [0017] <170> PatentIn version 3.5
- [0018] <210> 1
- [0019] <211> 64
- [0020] <212> DNA
- [0021] <213> 人工序列
- [0022] <220>
- [0023] <223> 合成的多核苷酸
- [0024] <400> 1
- [0025] cgatgacgtt atacgactca ctatagggat accacatgg ctctttccct acacgacgct 60
- [0026] cttc 64
- [0027] <210> 2
- [0028] <211> 20
- [0029] <212> DNA
- [0030] <213> 人工序列
- [0031] <220>
- [0032] <223> 合成的多核苷酸
- [0033] <400> 2
- [0034] taatacgact cactataggg 20
- [0035] <210> 3
- [0036] <211> 24
- [0037] <212> DNA
- [0038] <213> 人工序列

[0039]	<220>
[0040]	<223> 合成的多核苷酸
[0041]	<400> 3
[0042]	ctcttttcct acacgacgct cttc 24
[0043]	<210> 4
[0044]	<211> 22
[0045]	<212> DNA
[0046]	<213> 人工序列
[0047]	<220>
[0048]	<223> 合成的多核苷酸
[0049]	<400> 4
[0050]	aaggcgtcac aagcaatcac tc 22
[0051]	<210> 5
[0052]	<211> 20
[0053]	<212> DNA
[0054]	<213> 人工序列
[0055]	<220>
[0056]	<223> 合成的多核苷酸
[0057]	<400> 5
[0058]	tttttttttt tttttttttv 20
[0059]	<210> 6
[0060]	<211> 134
[0061]	<212> DNA
[0062]	<213> 人工序列
[0063]	<220>
[0064]	<223> 合成的多核苷酸
[0065]	<220>
[0066]	<221> misc_feature
[0067]	<222> (71) .. (78)
[0068]	<223> n是条形码1
[0069]	<220>
[0070]	<221> misc_feature
[0071]	<222> (101) .. (114)
[0072]	<223> n是条形码2
[0073]	<400> 6
[0074]	cgatgacgta atacgactca ctatagggat accaccatgg ctcttttcct acacgacgct 60
[0075]	cttcgatct nnnnnnnnaa ggcgtcaca gcaatcactc nnnnnnnnnn nnnntttttt 120
[0076]	tttttttttt tttv 134
[0077]	<210> 7

[0078]	<211> 64
[0079]	<212> DNA
[0080]	<213> 人工序列
[0081]	<220>
[0082]	<223> 合成的多核苷酸
[0083]	<400> 7
[0084]	cgatgacgta atacgactca ctatagggat accaccatgg ctctttccct acacgacgct 60
[0085]	cttc 64
[0086]	<210> 8
[0087]	<211> 63
[0088]	<212> DNA
[0089]	<213> 人工序列
[0090]	<220>
[0091]	<223> 合成的多核苷酸
[0092]	<220>
[0093]	<221> misc_feature
[0094]	<222> (23) .. (33)
[0095]	<223> n是条形码1
[0096]	<400> 8
[0097]	aaggcgtcac aagcaatcac tcnnnnnnnn nnnagatcgg aagagcgtcg tgtagggaaa 60
[0098]	gag 63
[0099]	<210> 9
[0100]	<211> 56
[0101]	<212> DNA
[0102]	<213> 人工序列
[0103]	<220>
[0104]	<223> 合成的多核苷酸
[0105]	<220>
[0106]	<221> misc_feature
[0107]	<222> (21) .. (34)
[0108]	<223> n是条形码2
[0109]	<400> 9
[0110]	baaaaaaaaa aaaaaaaaaa nnnnnnnnnn nnnnaaggcg tcacaagcaa tcactc 56
[0111]	<210> 10
[0112]	<211> 30
[0113]	<212> DNA
[0114]	<213> 人工序列
[0115]	<220>
[0116]	<223> 合成的多核苷酸

[0117]	<400> 10
[0118]	agatcggag agcgtcgtgt agggaaagag 30
[0119]	<210> 11
[0120]	<211> 22
[0121]	<212> DNA
[0122]	<213> 人工序列
[0123]	<220>
[0124]	<223> 合成的多核苷酸
[0125]	<400> 11
[0126]	aaggcgtcac aagcaatcac tc 22
[0127]	<210> 12
[0128]	<211> 20
[0129]	<212> DNA
[0130]	<213> 人工序列
[0131]	<220>
[0132]	<223> 合成的多核苷酸
[0133]	<400> 12
[0134]	aaaaaaaaa aaaaaaaaaa 20
[0135]	<210> 13
[0136]	<211> 136
[0137]	<212> DNA
[0138]	<213> 人工序列
[0139]	<220>
[0140]	<223> 合成的多核苷酸
[0141]	<220>
[0142]	<221> misc_feature
[0143]	<222> (71) .. (80)
[0144]	<223> n是条形码1
[0145]	<220>
[0146]	<221> misc_feature
[0147]	<222> (103) .. (110)
[0148]	<223> n是条形码2
[0149]	<220>
[0150]	<221> misc_feature
[0151]	<222> (111) .. (116)
[0152]	<223> n是a, c, g, 或t
[0153]	<400> 13
[0154]	cgatgacgta atacgactca ctatagggat accaccatgg ctctttccct acacgacgct 60
[0155]	cttccgatct nnnnnnnnnn gaggatgtgc ttgtgacgcc ttnnnnnnnn nnnnnntttt 120

[0156] tttttttttt tttttv 136
[0157] <210> 14
[0158] <211> 30
[0159] <212> DNA
[0160] <213> 人工序列
[0161] <220>
[0162] <223> 合成的多核苷酸
[0163] <400> 14
[0164] agatcggaag agcggttcag caggaatgcc 30
[0165] <210> 15
[0166] <211> 21
[0167] <212> DNA
[0168] <213> 人工序列
[0169] <220>
[0170] <223> 合成的多核苷酸
[0171] <400> 15
[0172] gtctcgcat tcctgctgaa c 21
[0173] <210> 16
[0174] <211> 44
[0175] <212> DNA
[0176] <213> 人工序列
[0177] <220>
[0178] <223> 合成的多核苷酸
[0179] <400> 16
[0180] aatgatacgg cgaccaccga gatctacact ctttcctac acga 44
[0181] <210> 17
[0182] <211> 47
[0183] <212> DNA
[0184] <213> 人工序列
[0185] <220>
[0186] <223> 合成的多核苷酸
[0187] <400> 17
[0188] caagcagaag acggcatacg agatcggtct cggcattcct gctgaac 47

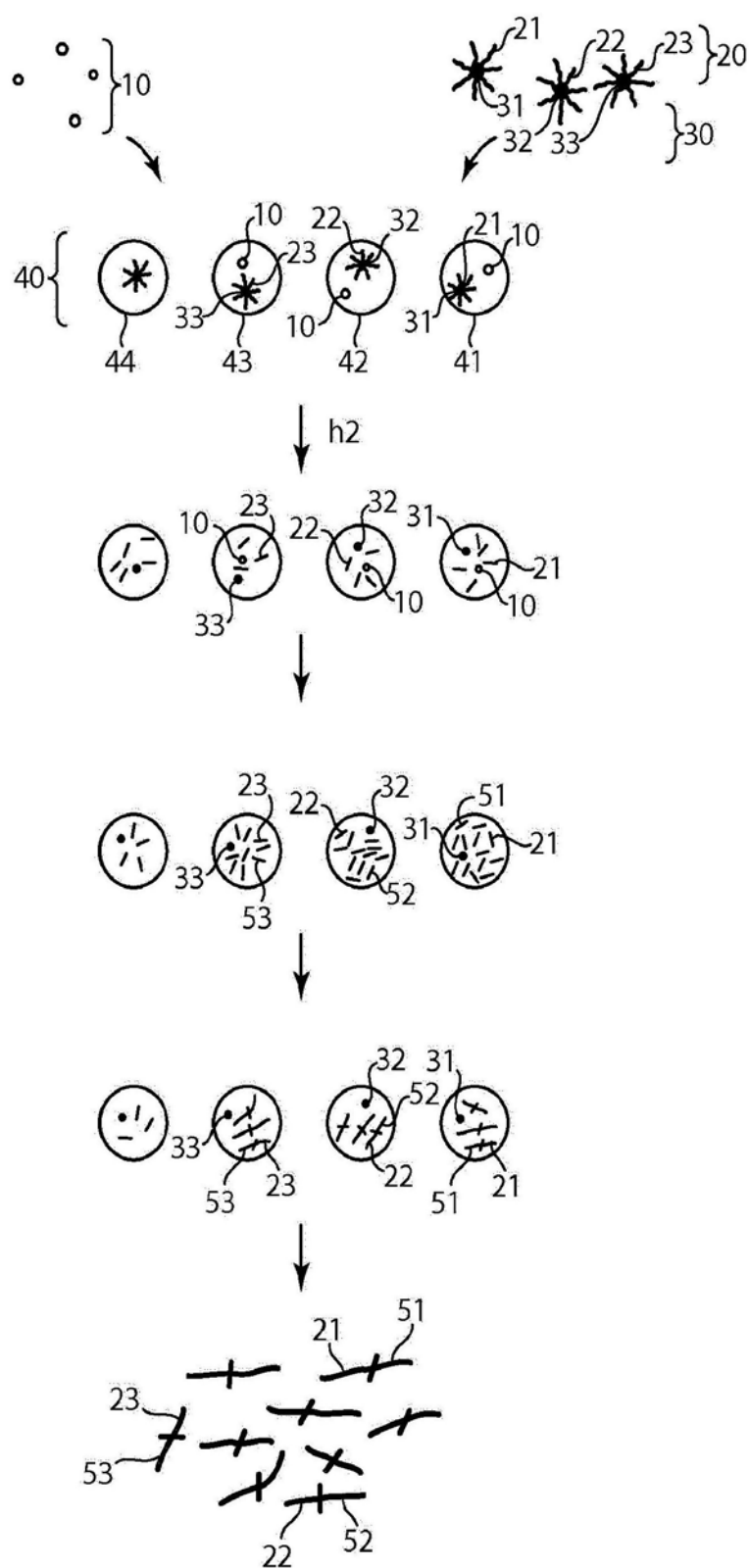


图1

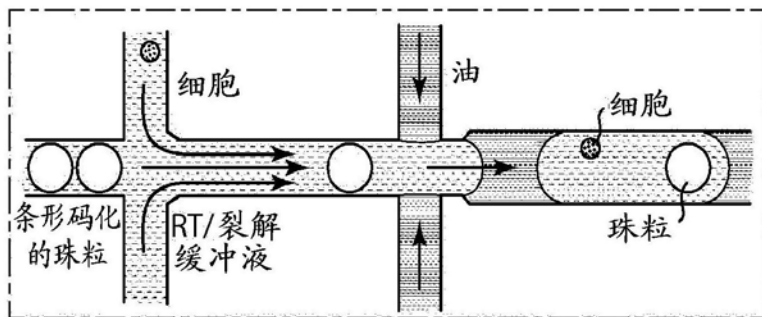


图2A

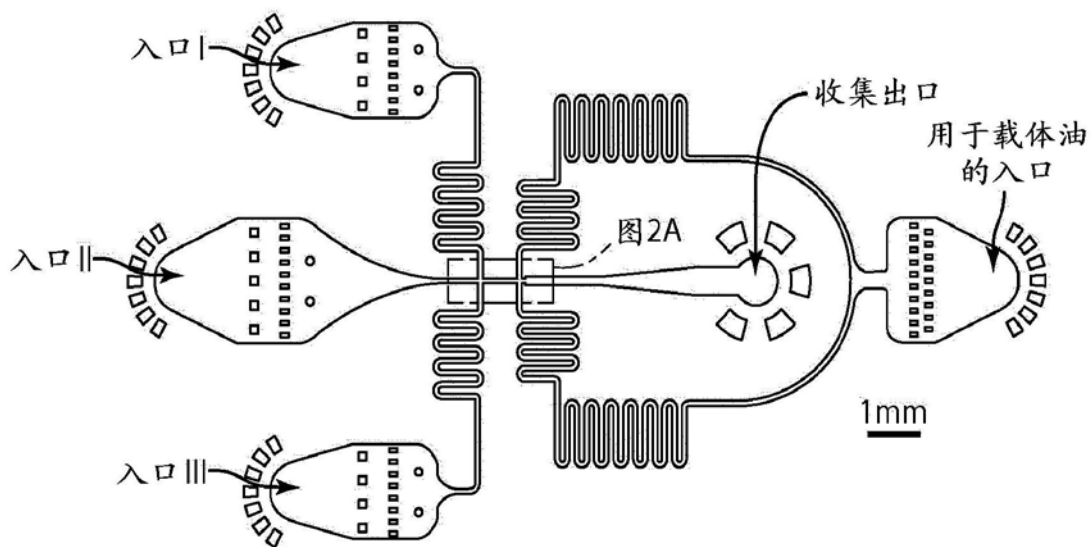


图2B

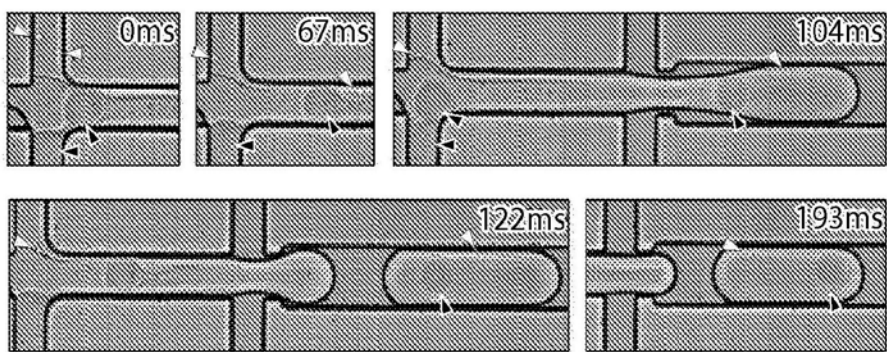


图3

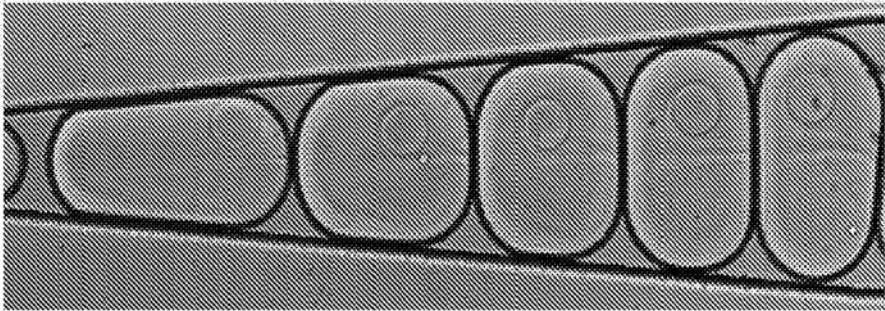


图4

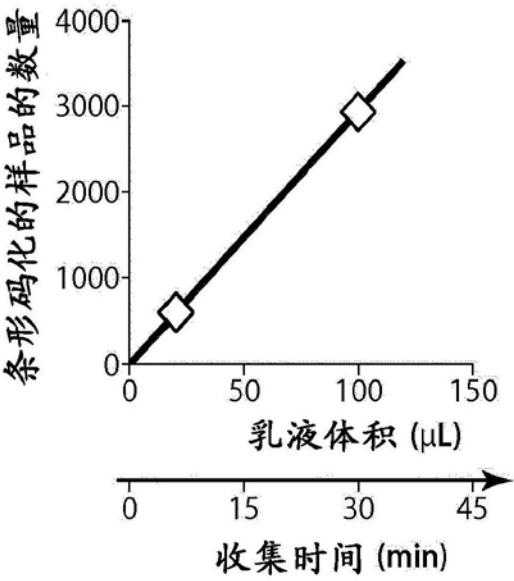


图5

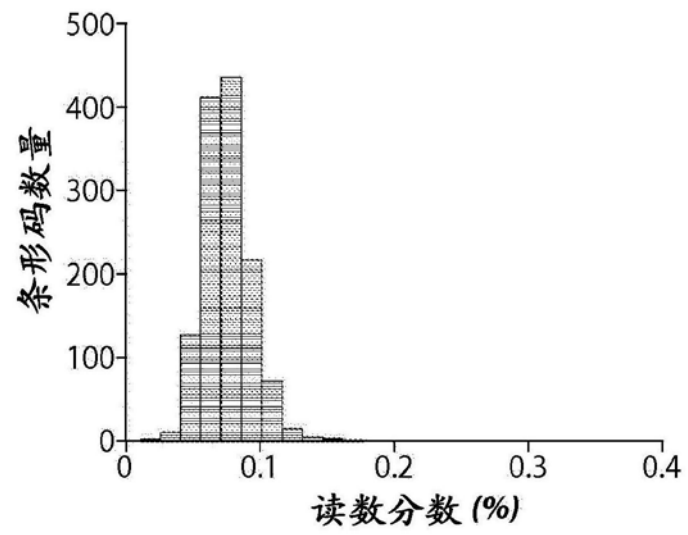


图6

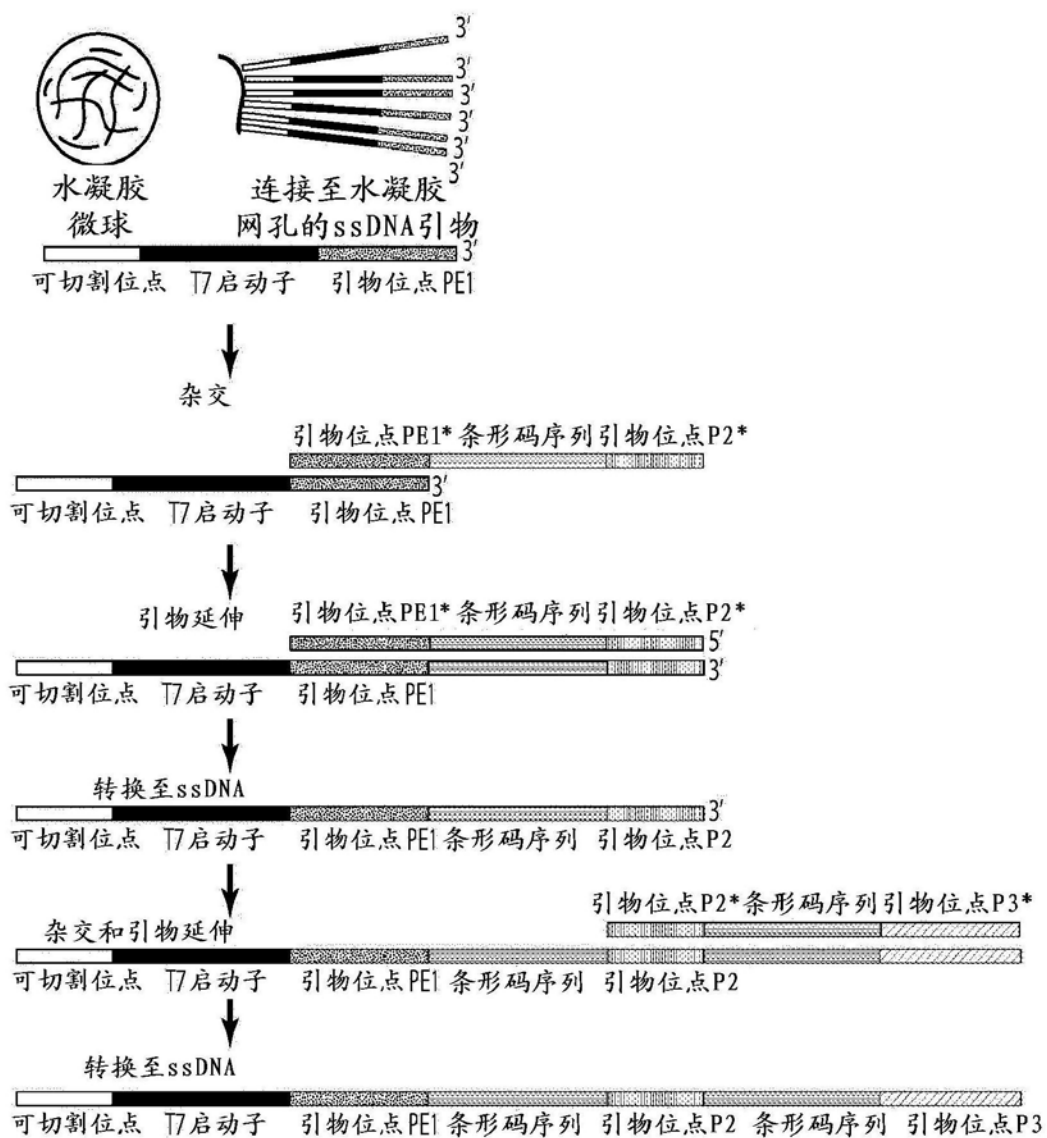


图7

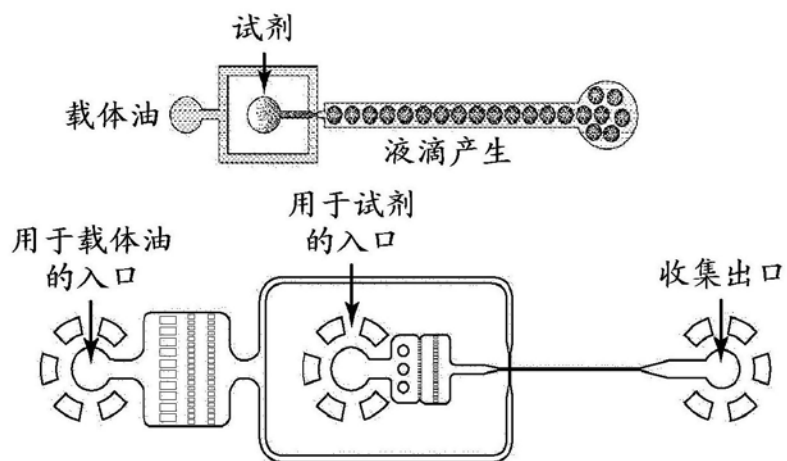


图8

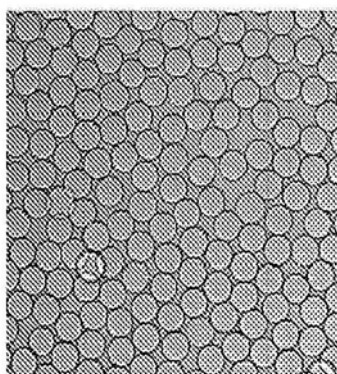


图9A

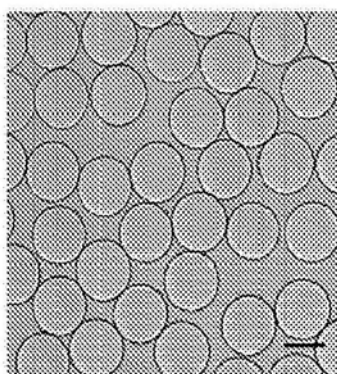


图9B

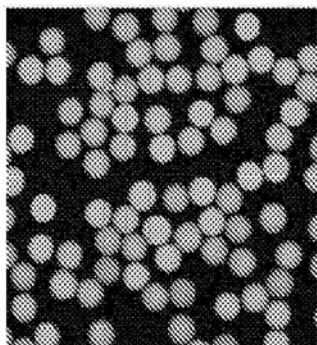


图10A

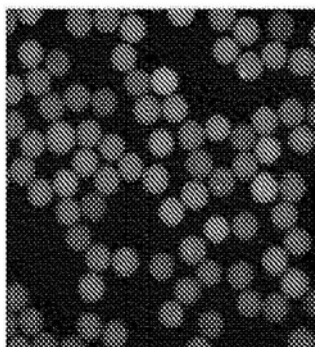


图10B

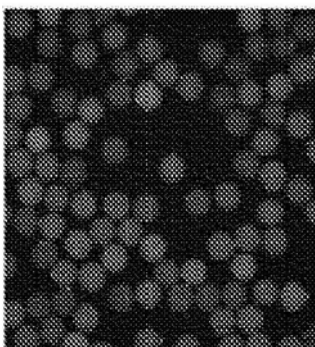


图10C

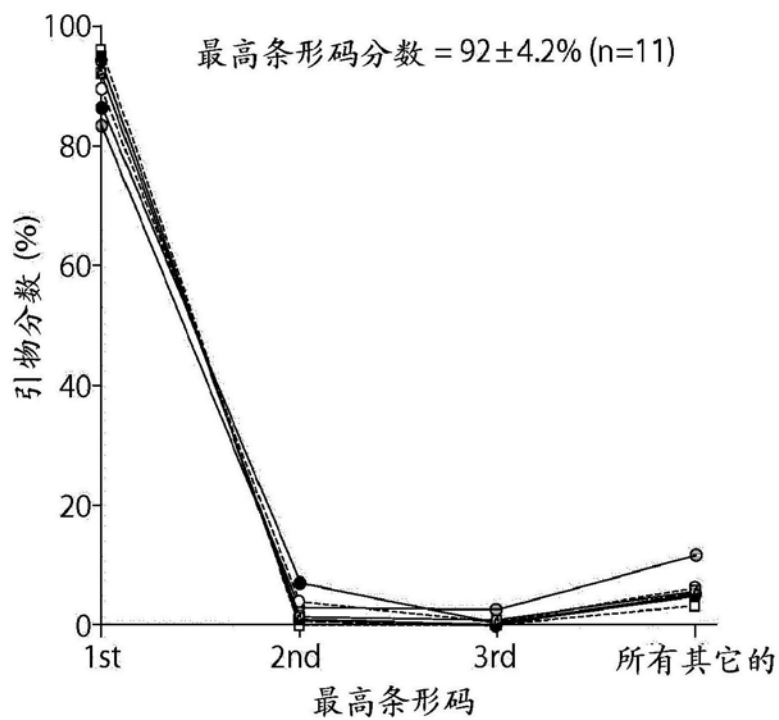


图11

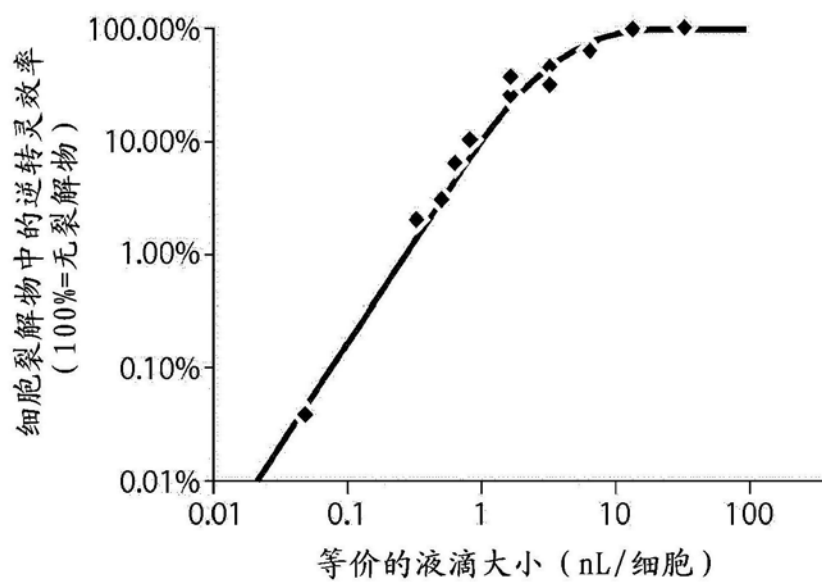


图12

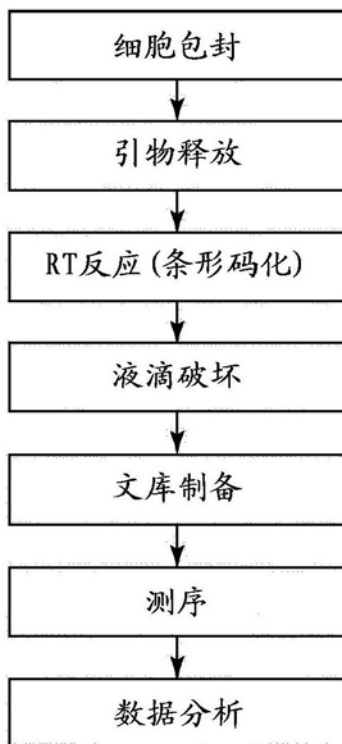


图13A

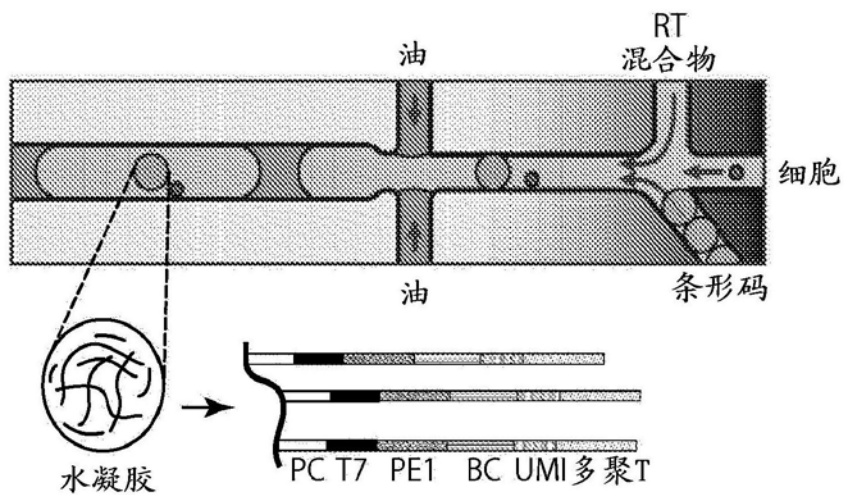


图13B

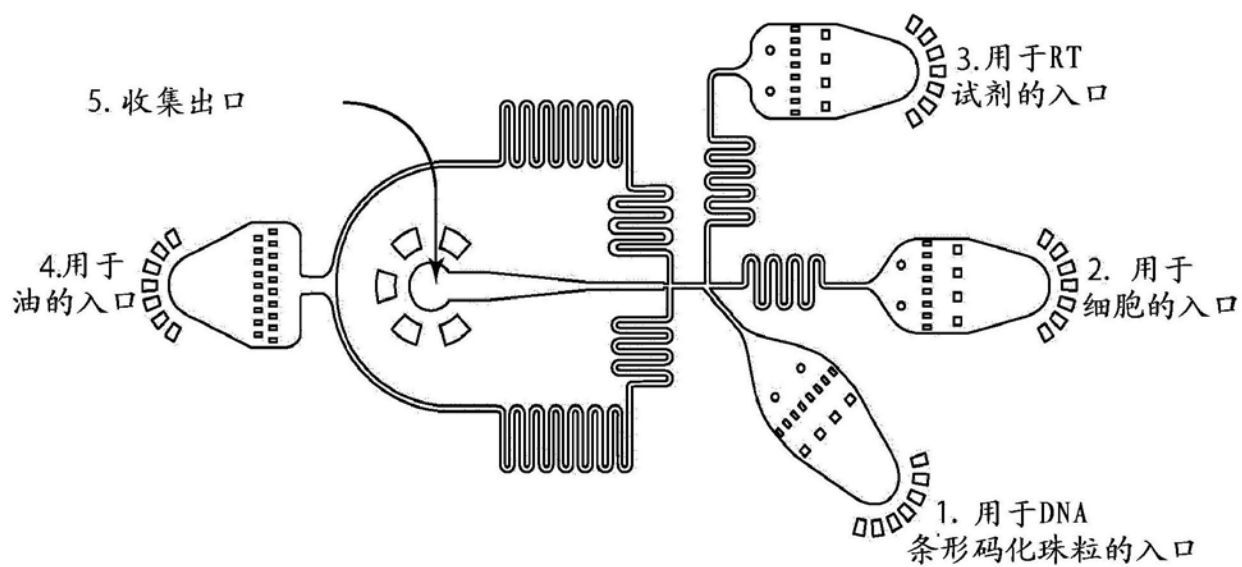


图13C

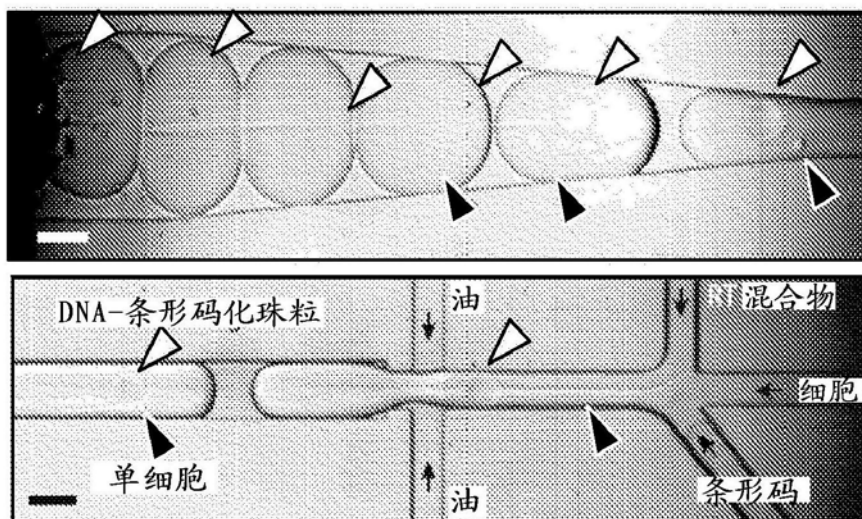


图13D

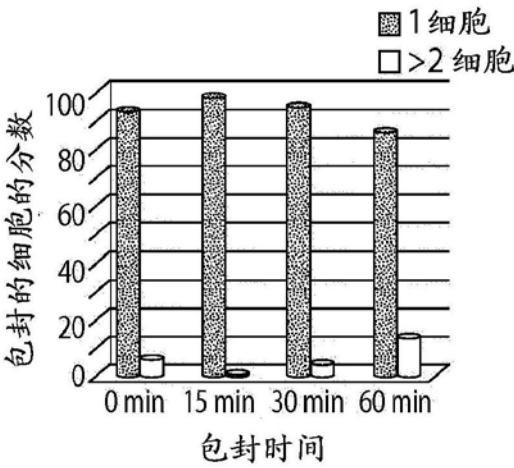


图13E

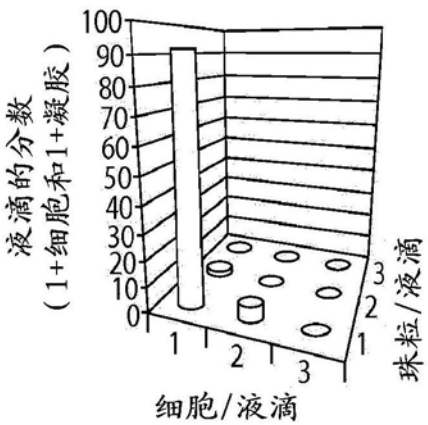


图13F

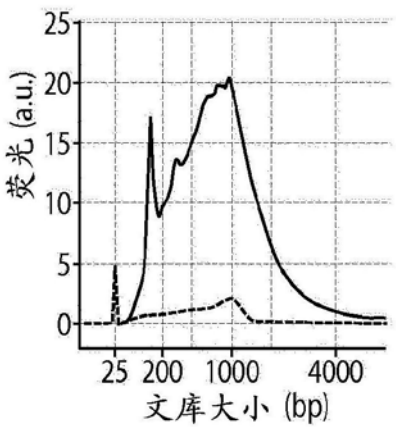


图13G

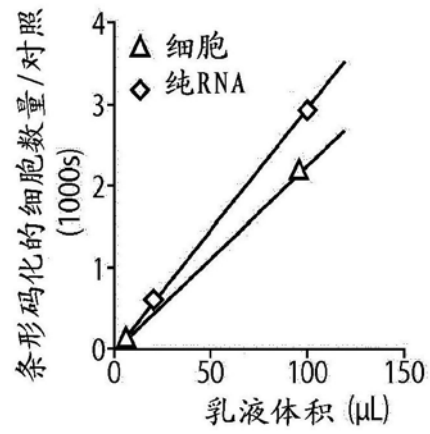


图13H

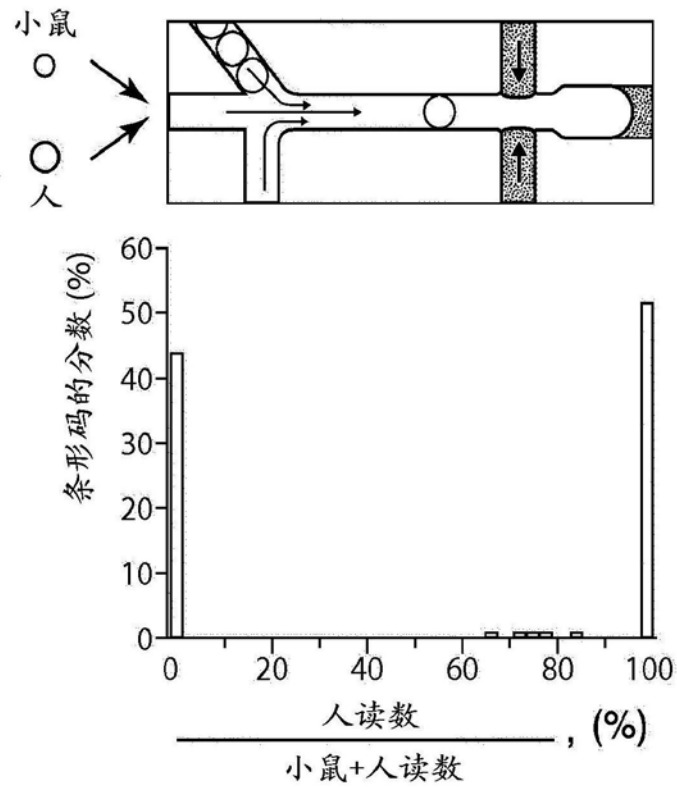


图14A

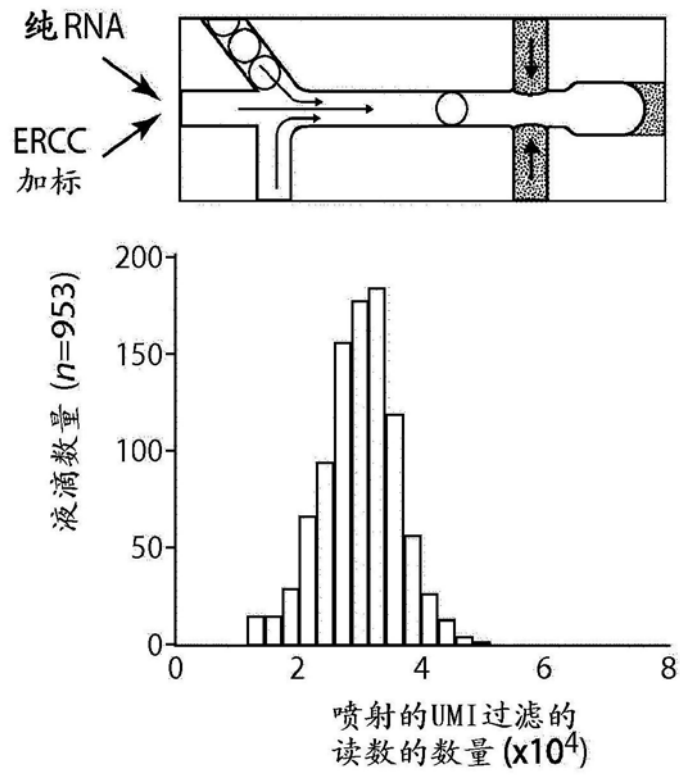


图14B

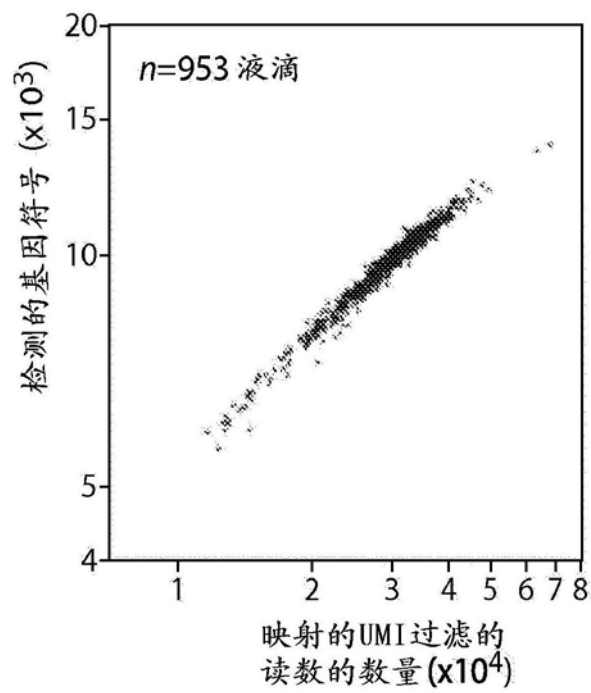


图14C

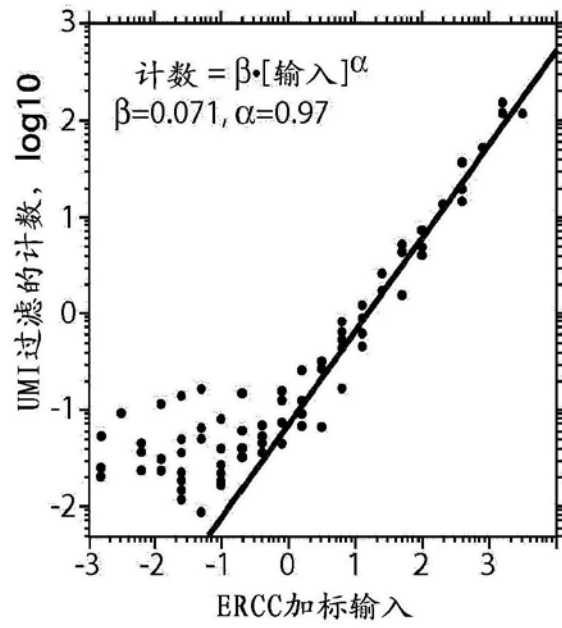


图14D

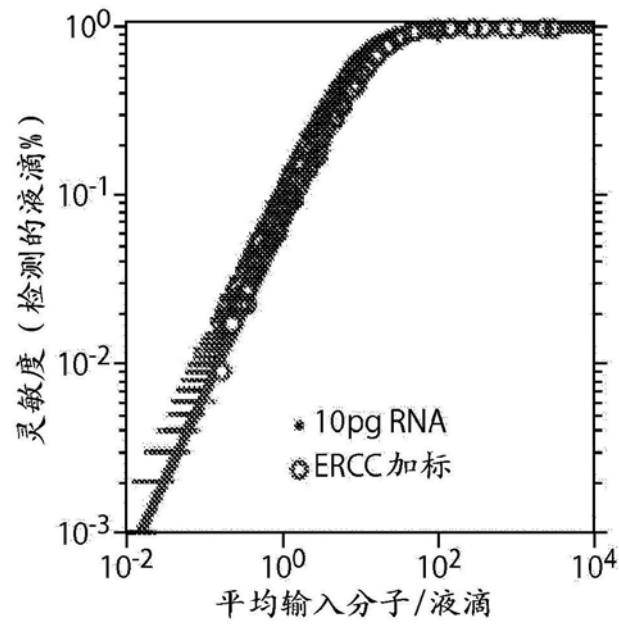


图14E

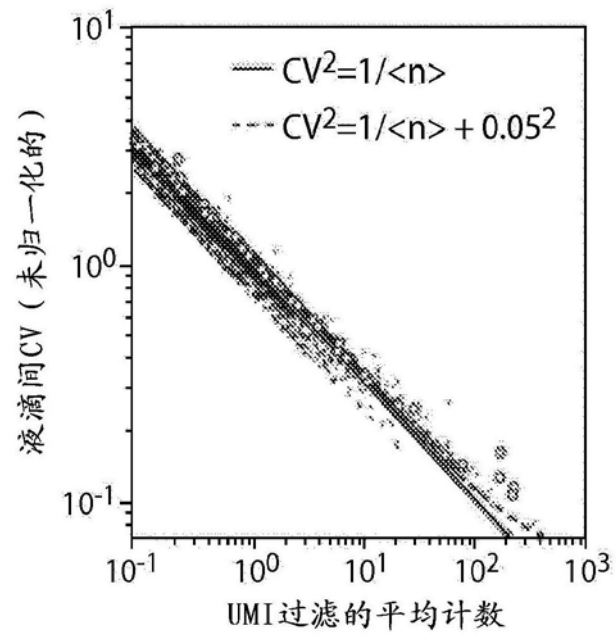


图14F

$$\begin{array}{l}
 \text{1) 泊松噪声以上:} \\
 \underbrace{CV^2 - \frac{1}{\mu}}_{\text{观察到的}} = \underbrace{\left(CV_{\text{bio}}^2 - \frac{1}{\mu_{\text{bio}}} \right)}_{\text{生物学变异性}} \underbrace{(1 + CV_{\text{方法}}^2)}_{\text{噪声膨胀}} + \underbrace{CV_{\text{方法}}^2}_{\text{背景}} \\
 \text{2) } F = \underbrace{\beta F_{\text{bio}} + (1 - \beta)}_{\text{采样减少}} + \underbrace{CV_{\text{方法}}^2}_{\text{背景}} \\
 \text{3) 基因相关性:} \\
 \underbrace{\text{corr}_{i,j}}_{\text{观察到的}} \simeq \underbrace{\text{corr}_{i,j}^{(\text{bio})}}_{\text{生物学的}} \sqrt{\underbrace{(1 - F_i^{-1})(1 - F_j^{-1})}_{\text{噪声膨胀}}}
 \end{array}$$

图14G

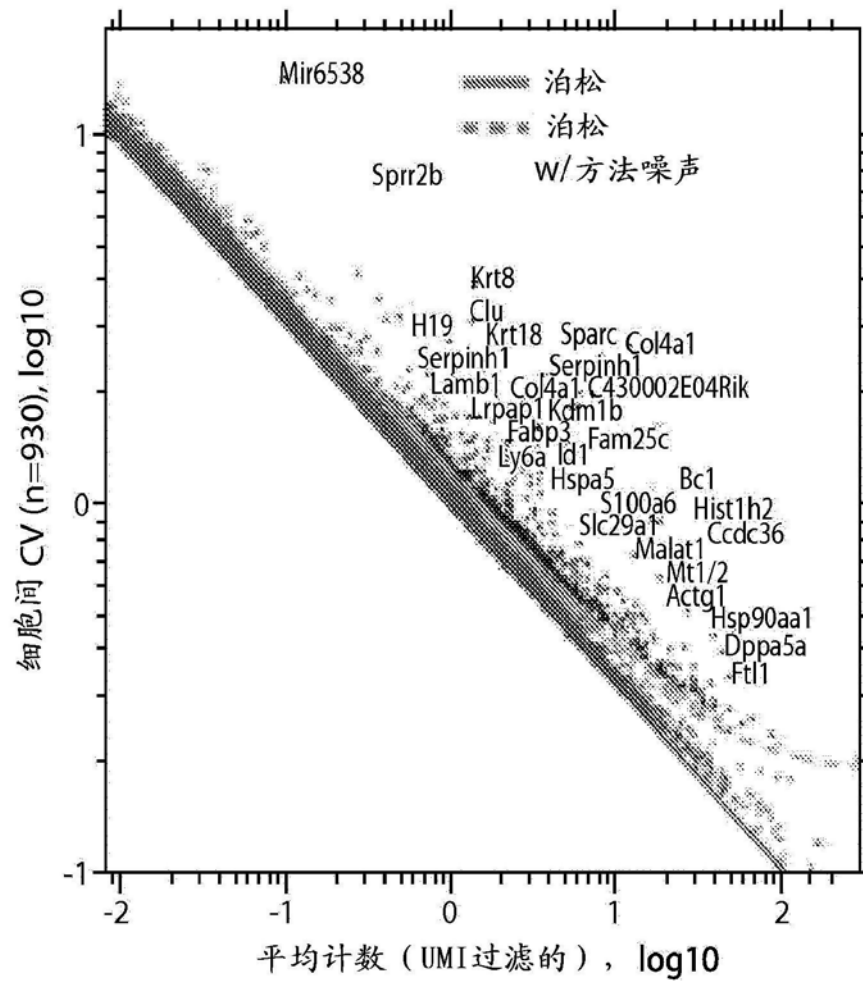


图15A

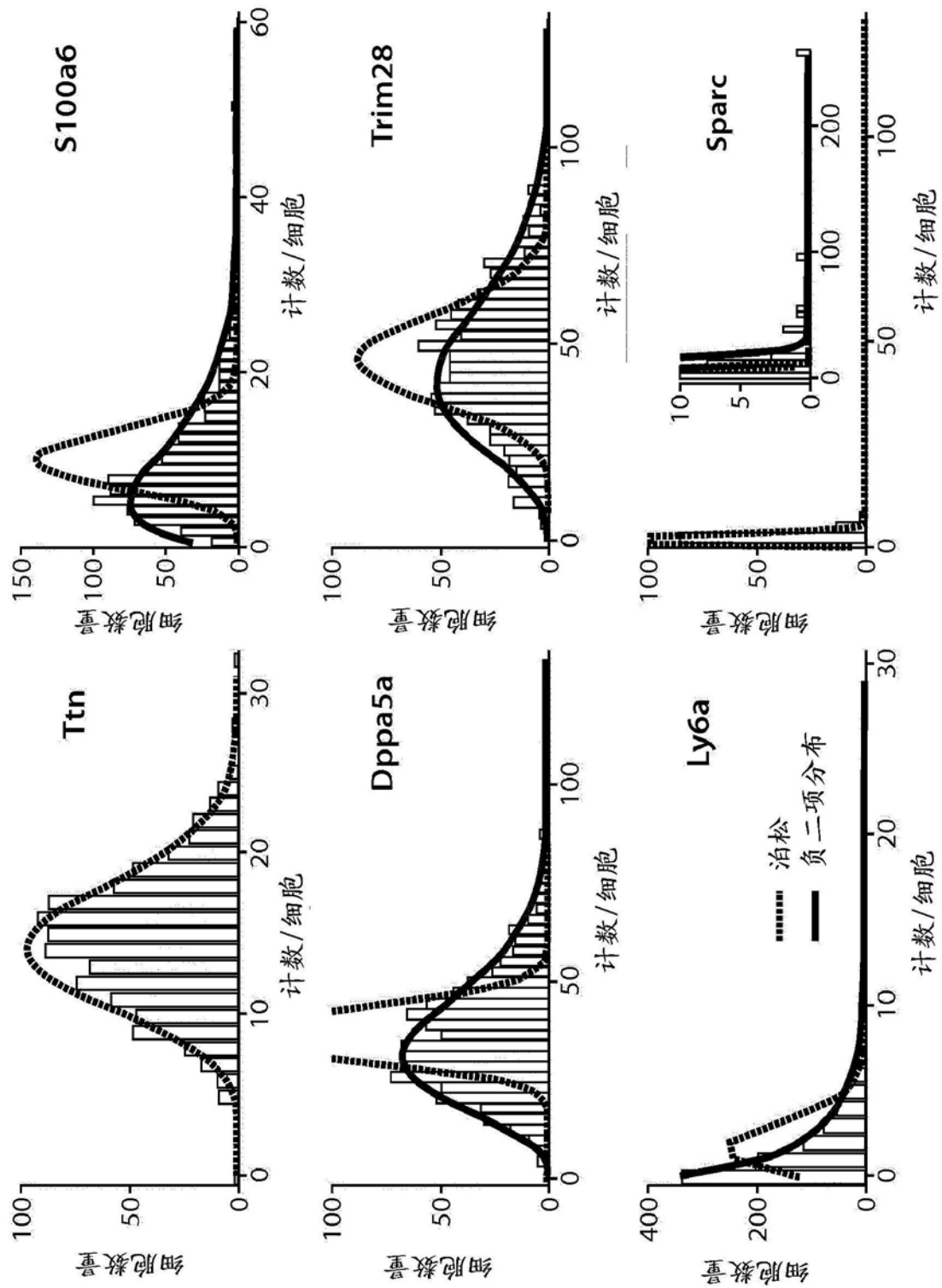


图15B

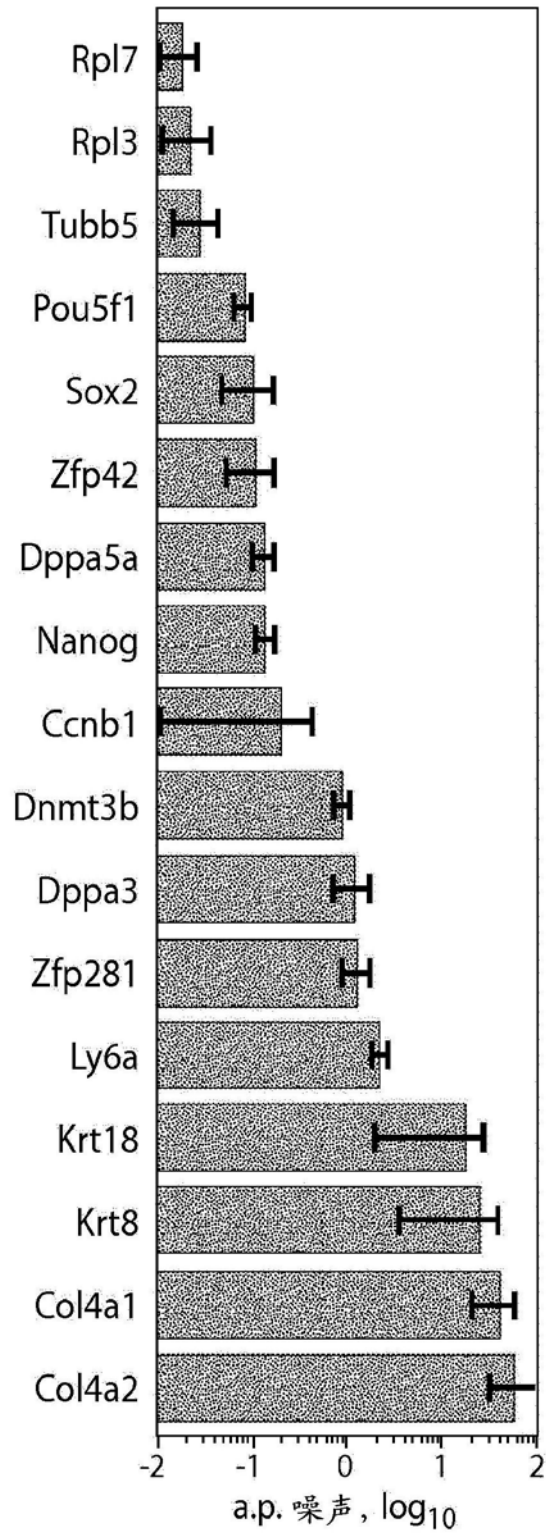


图15C

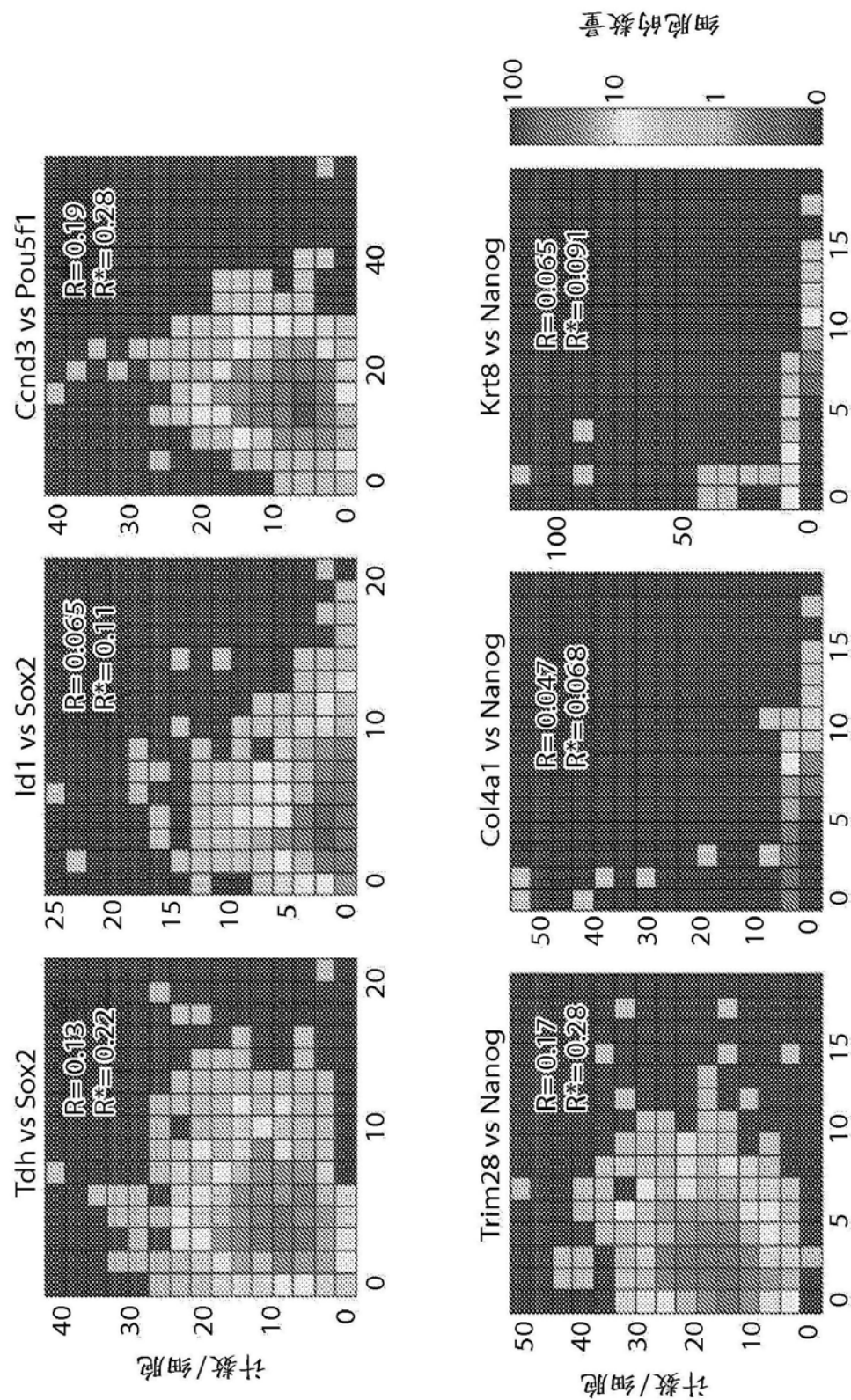


图15D

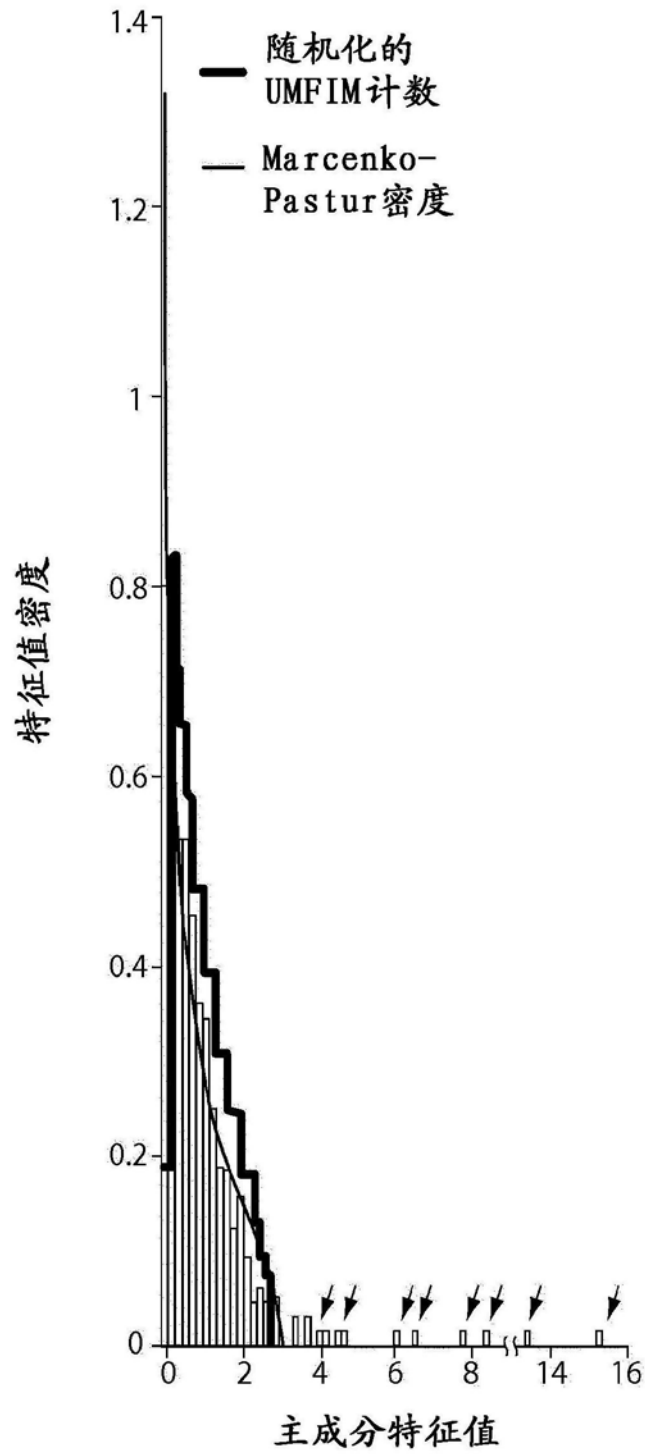


图15E

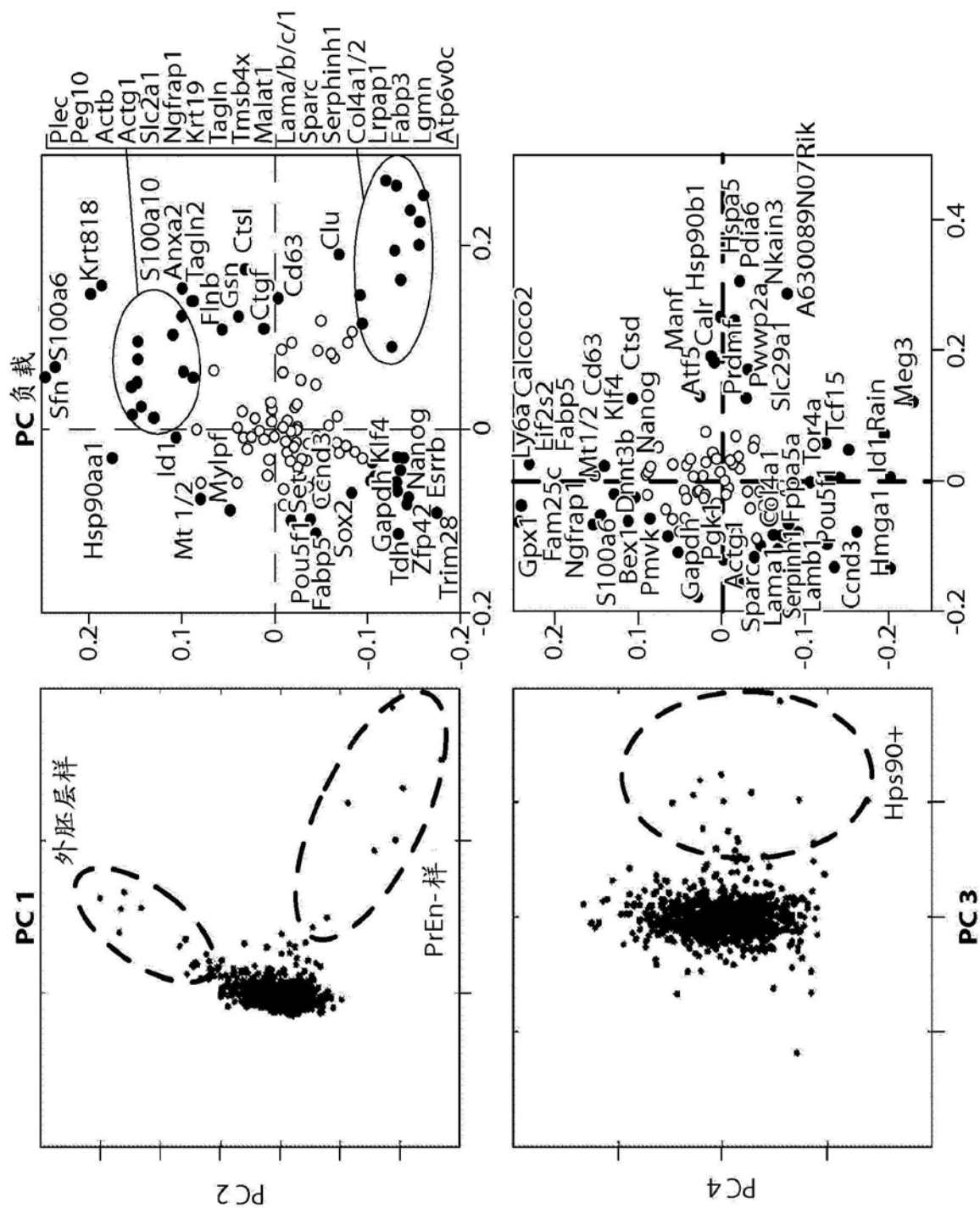


图15F

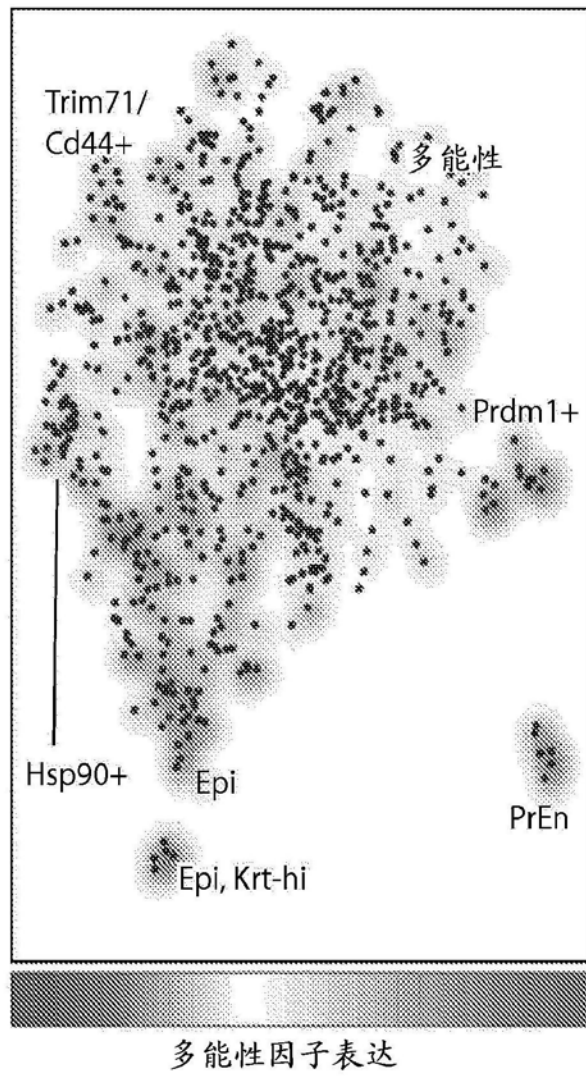


图15G

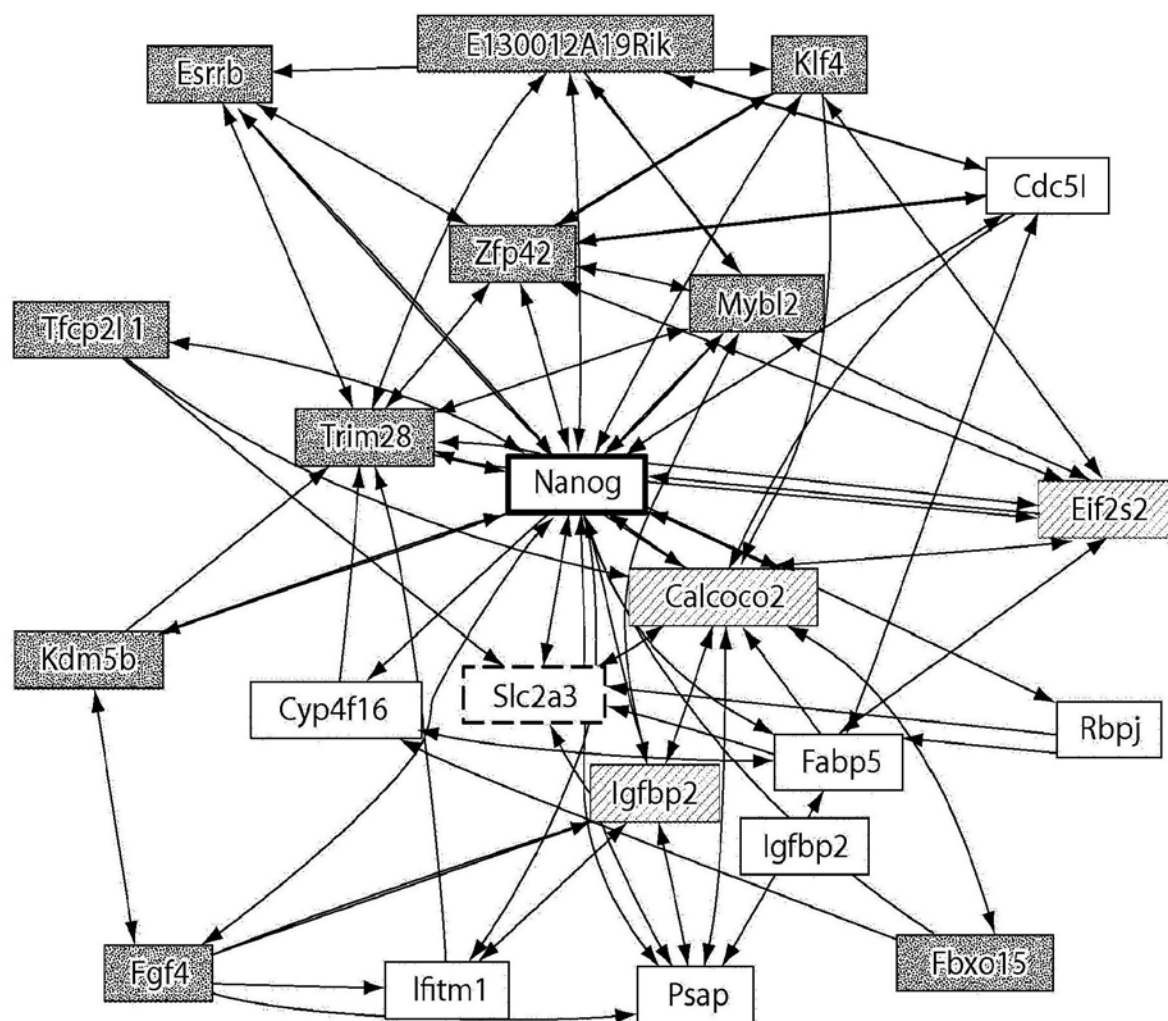


图16A

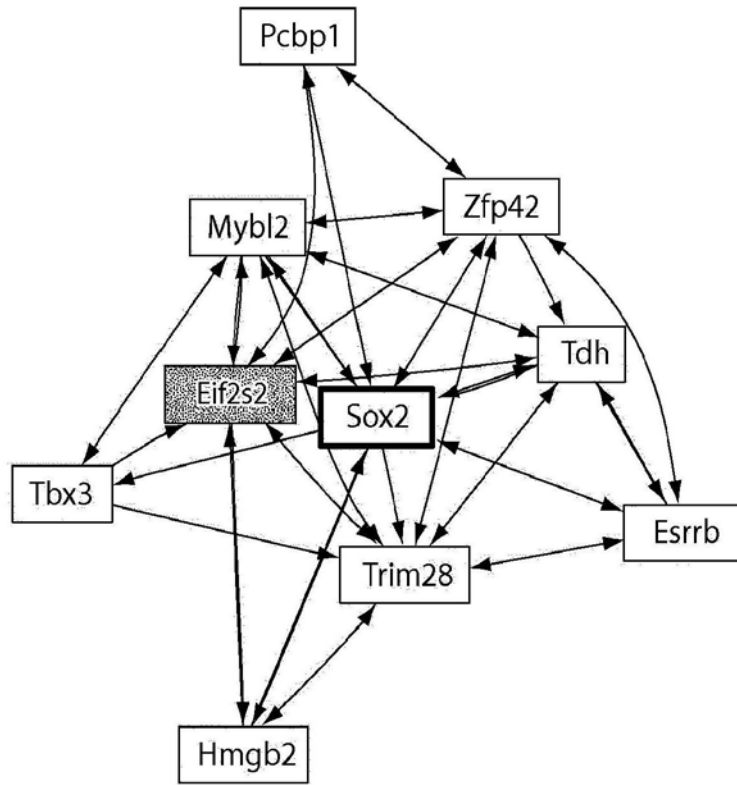


图16B

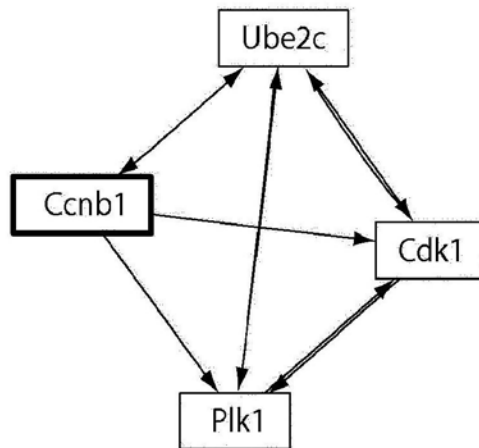


图16C

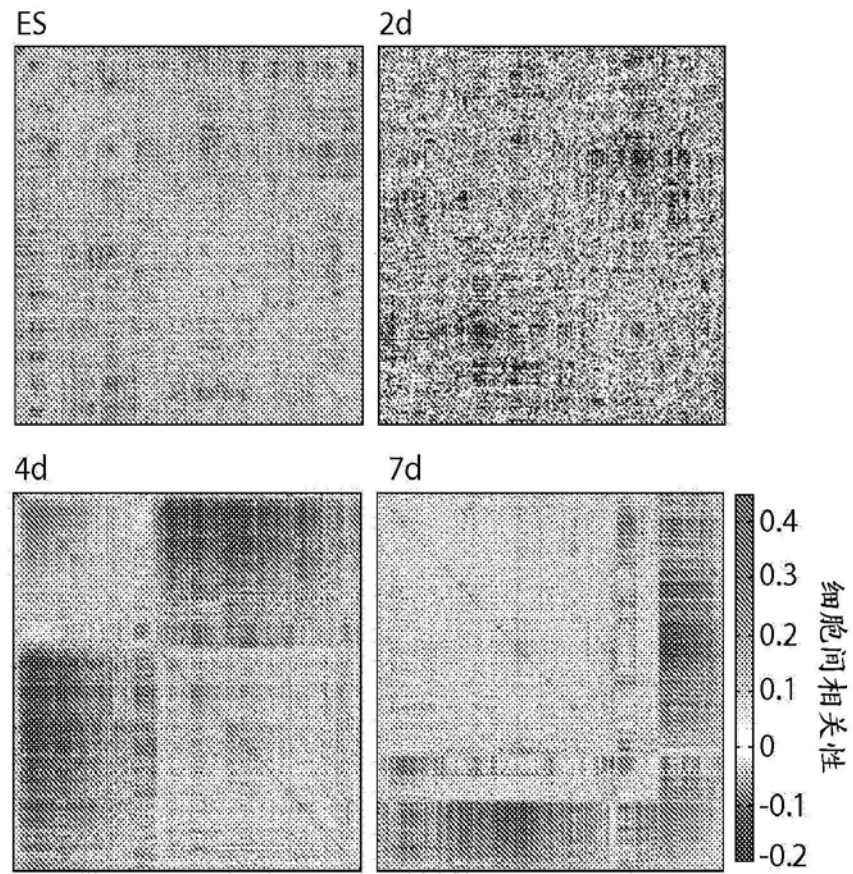


图17A

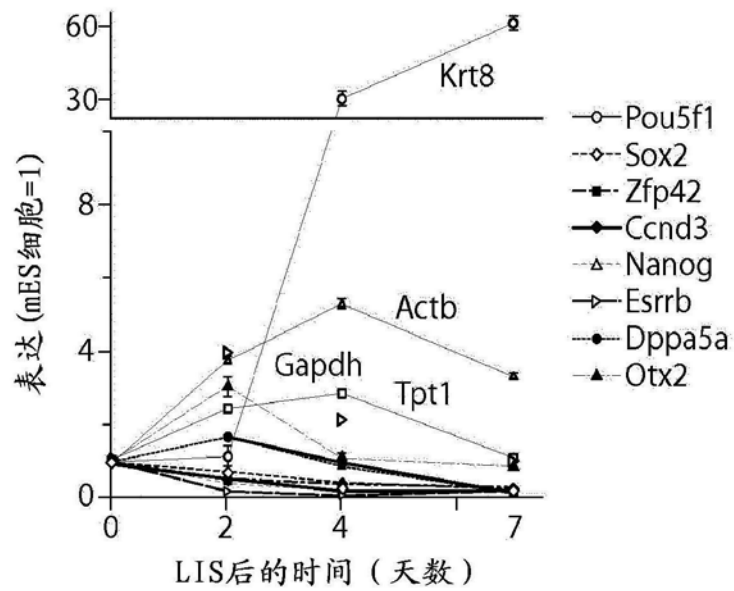


图17B

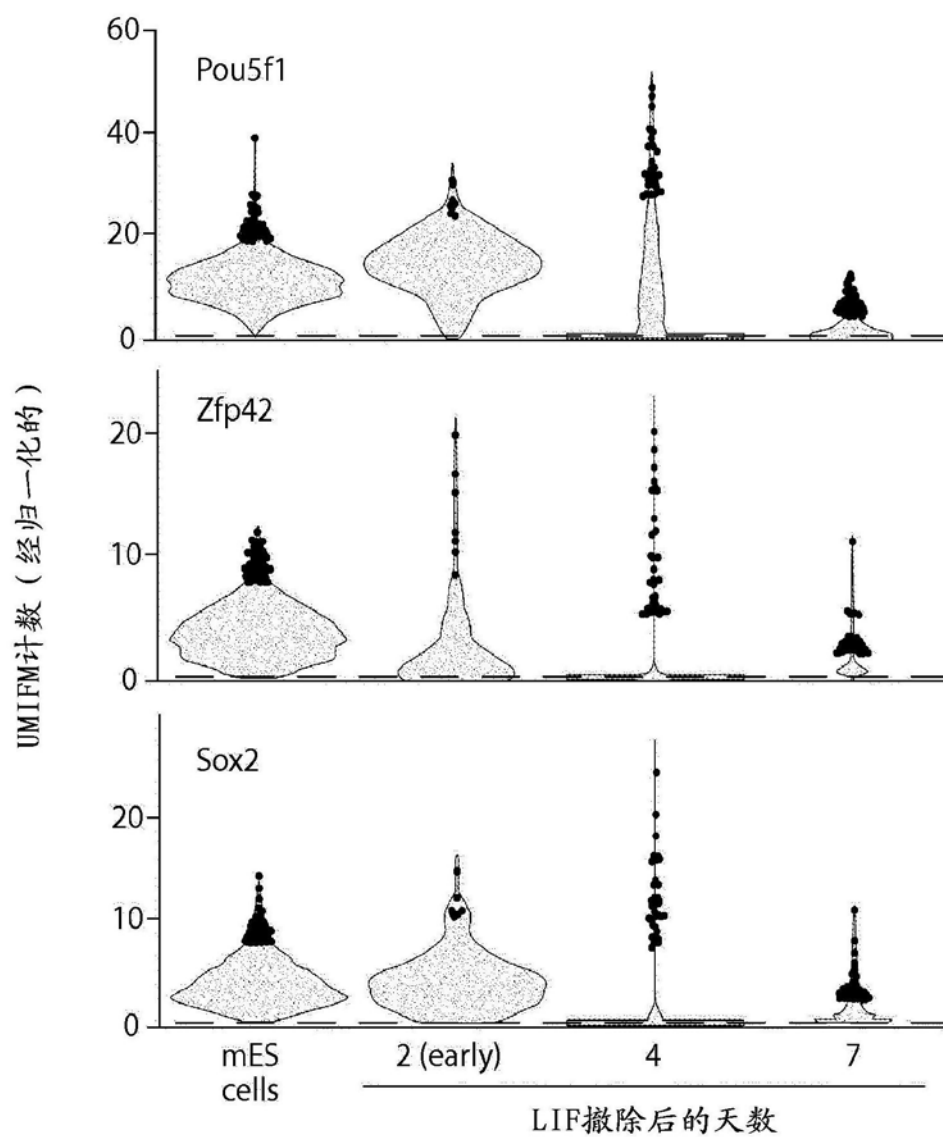


图17C-1

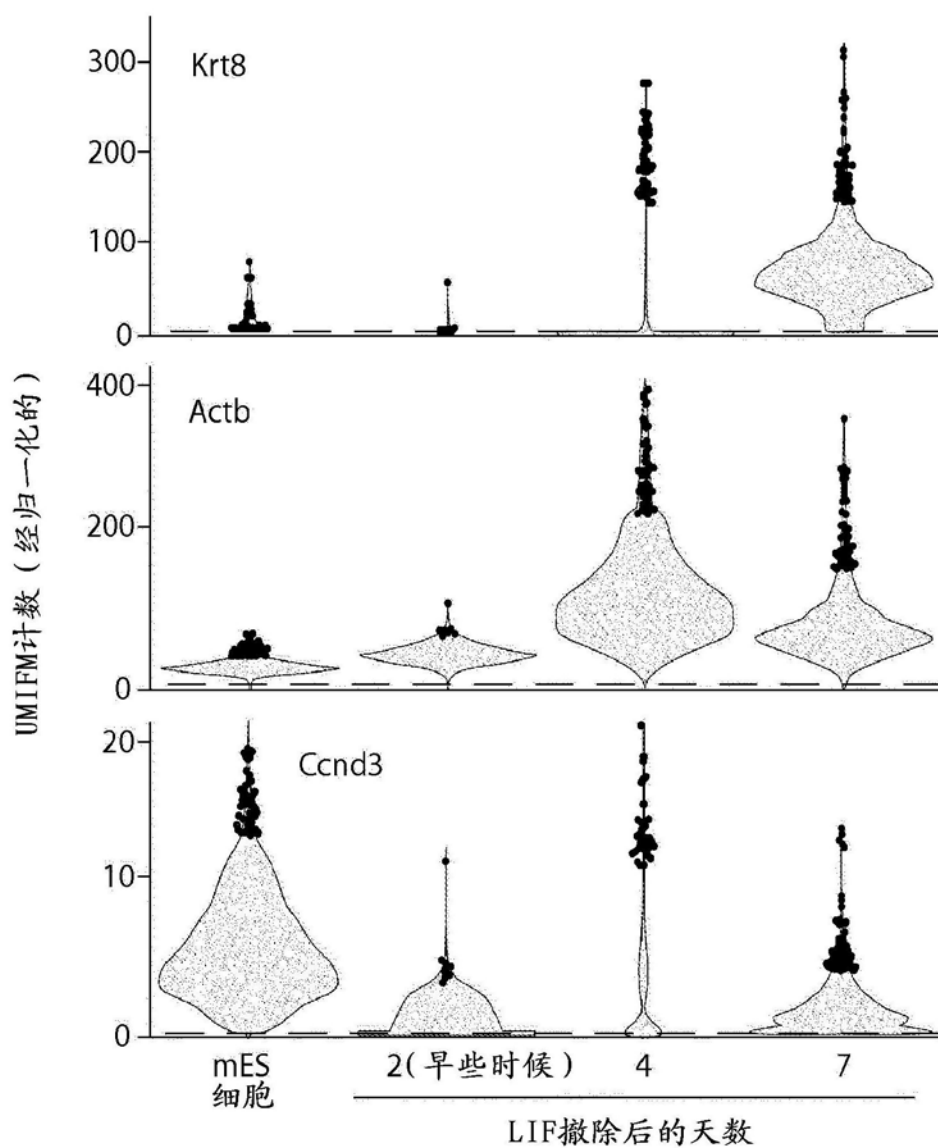


图17C-2

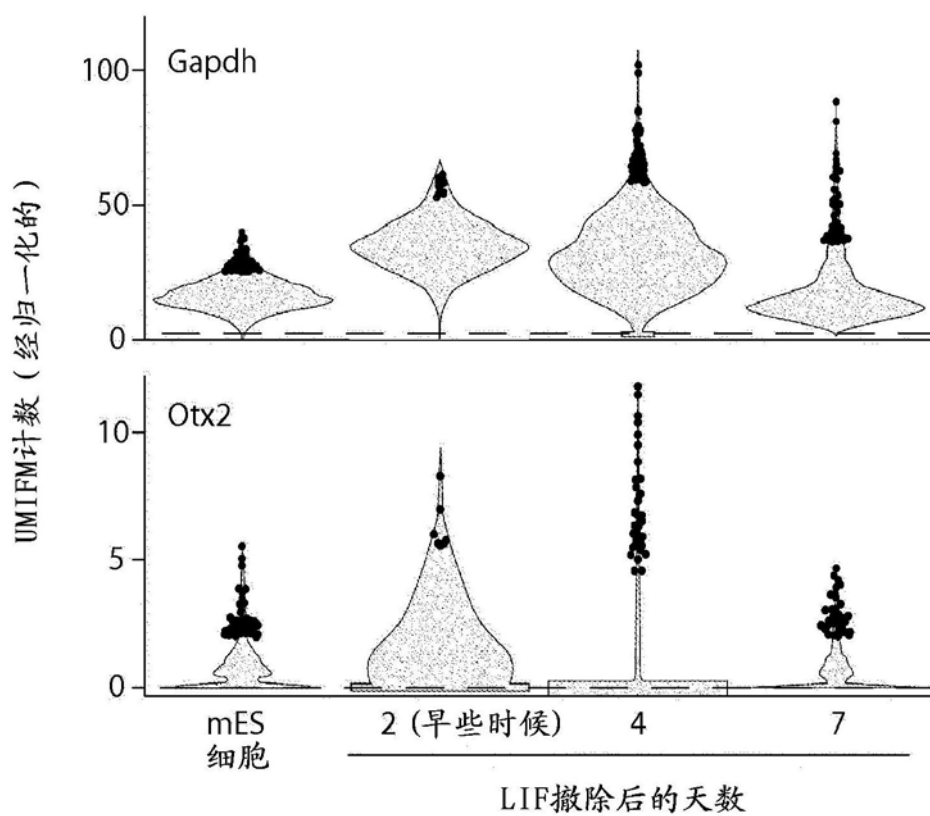


图17C-3

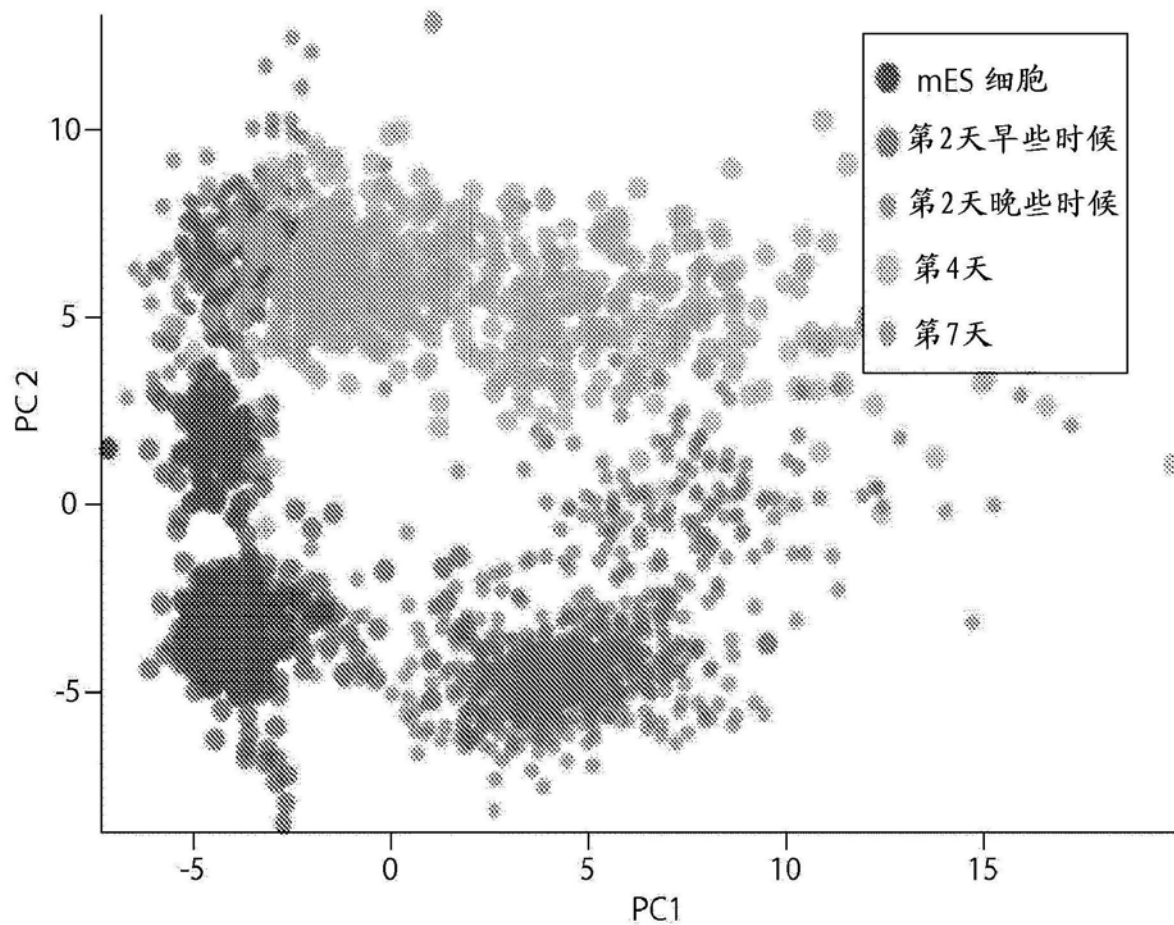


图17D

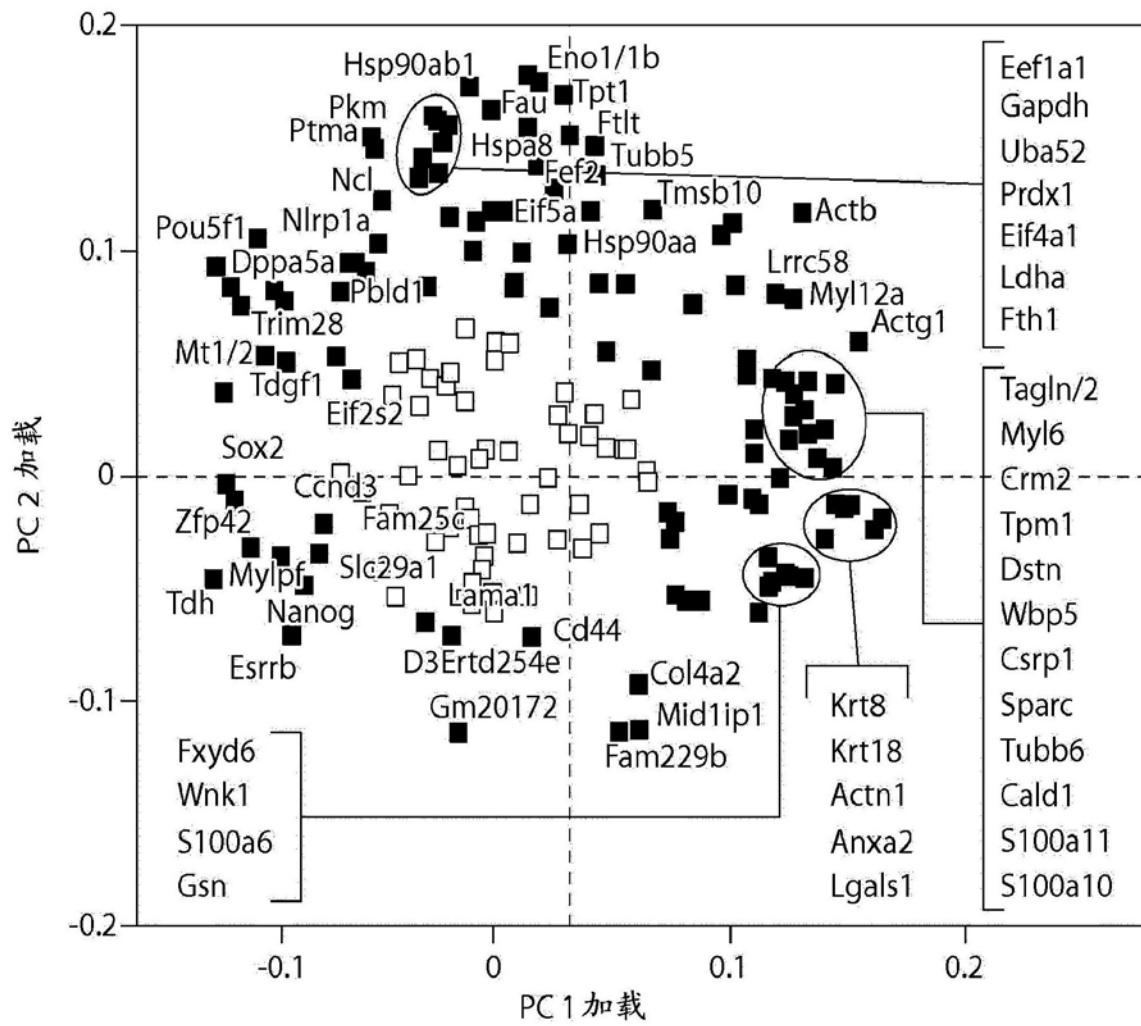


图17E

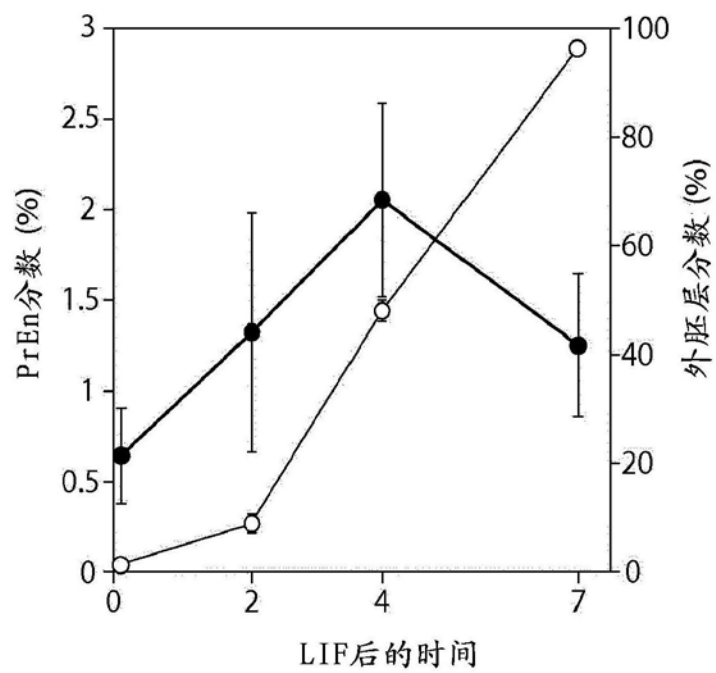


图17F

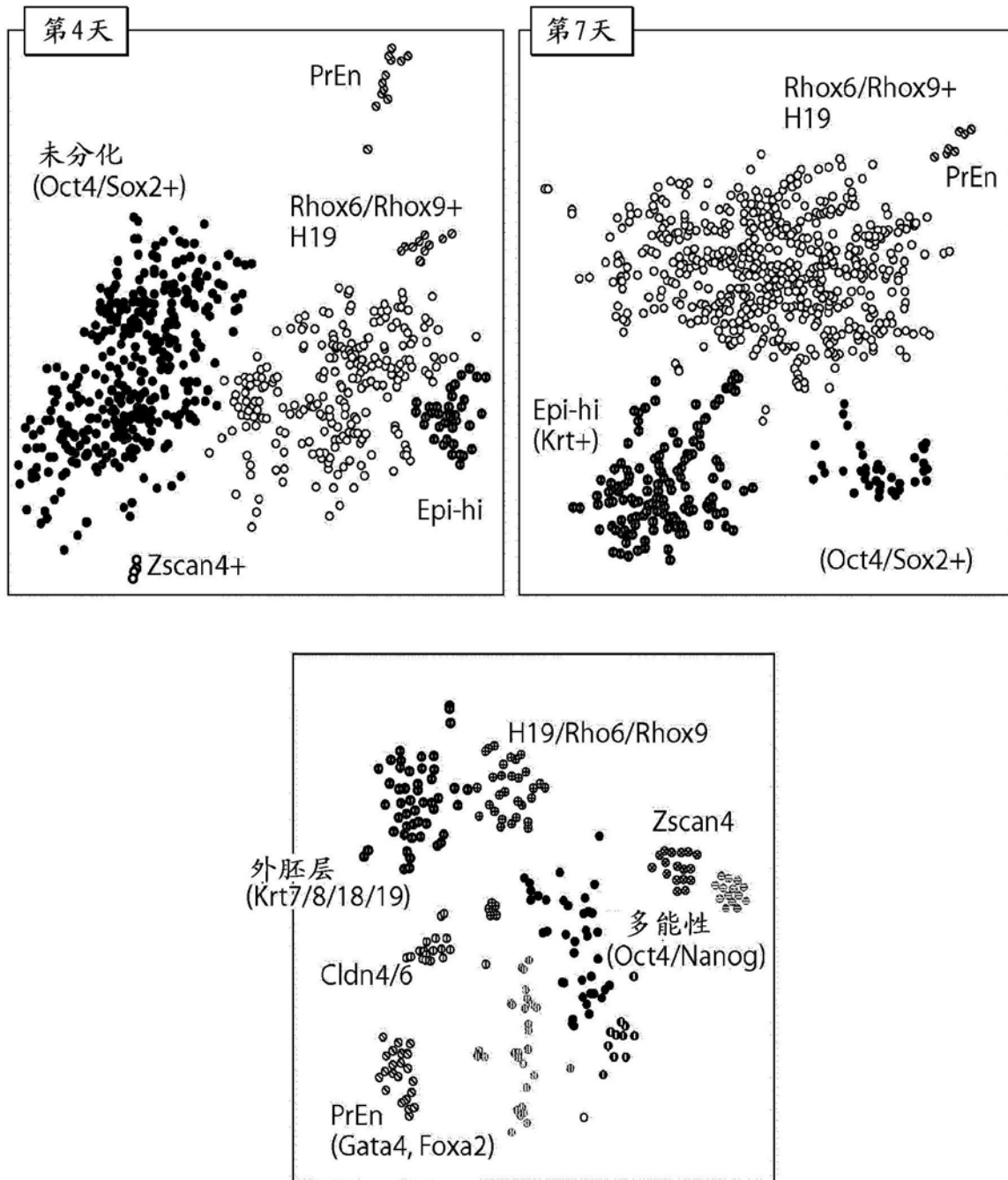


图17G

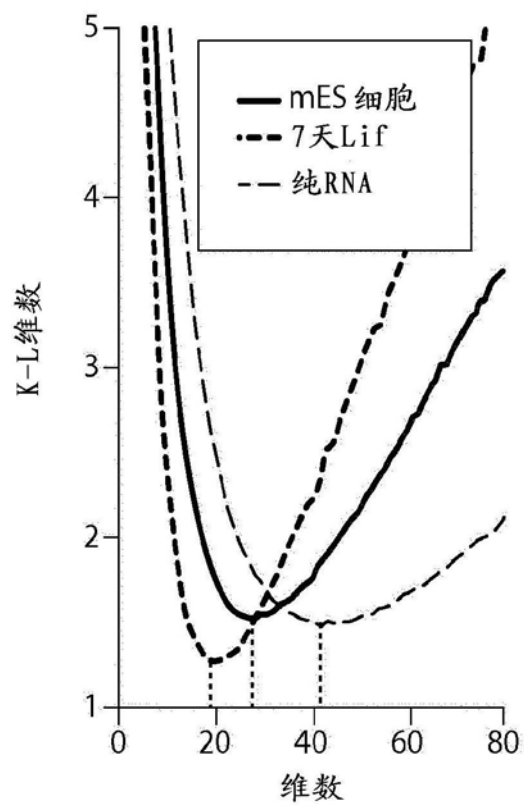


图17H

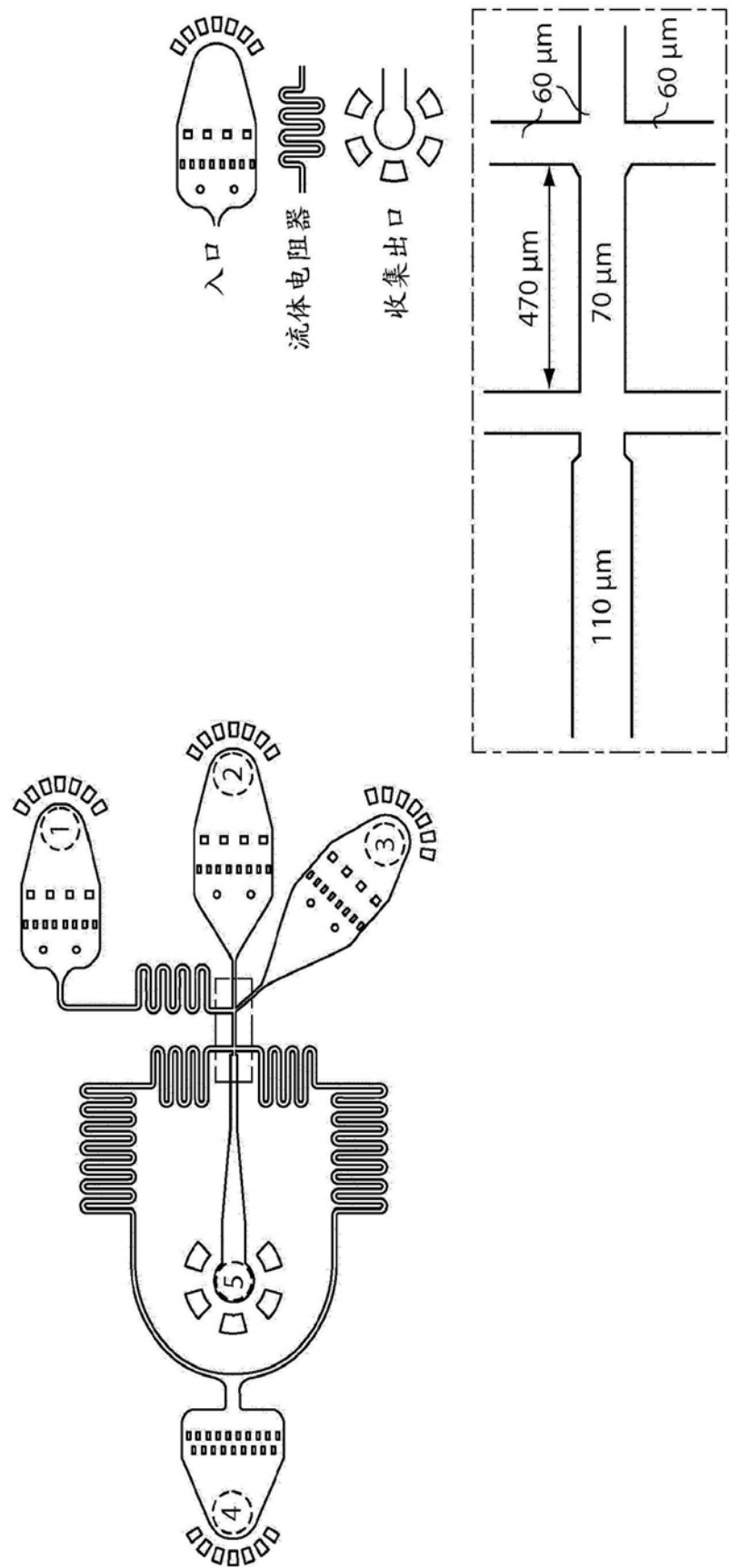


图18

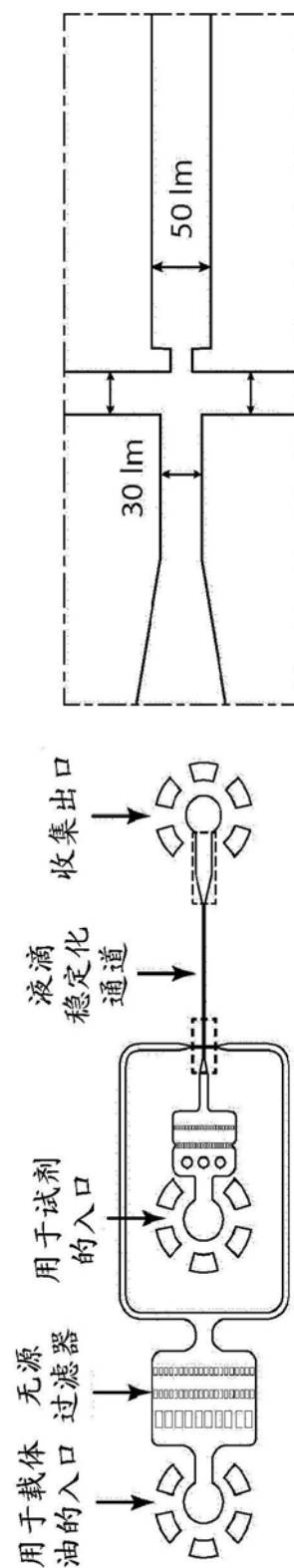


图19A

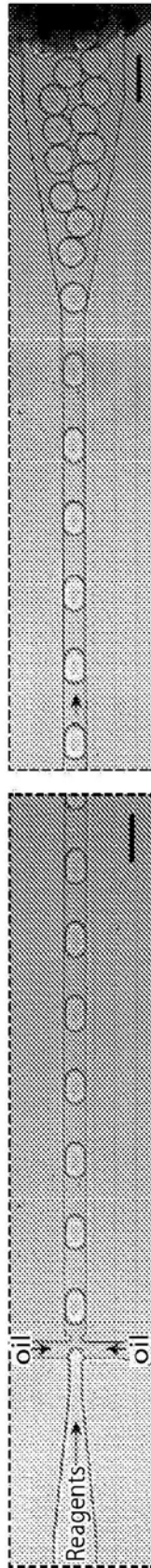


图19B

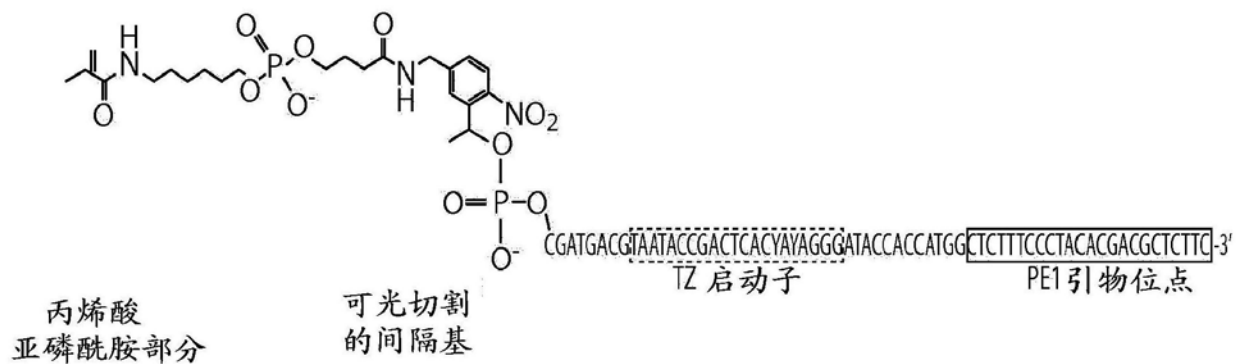


图20A

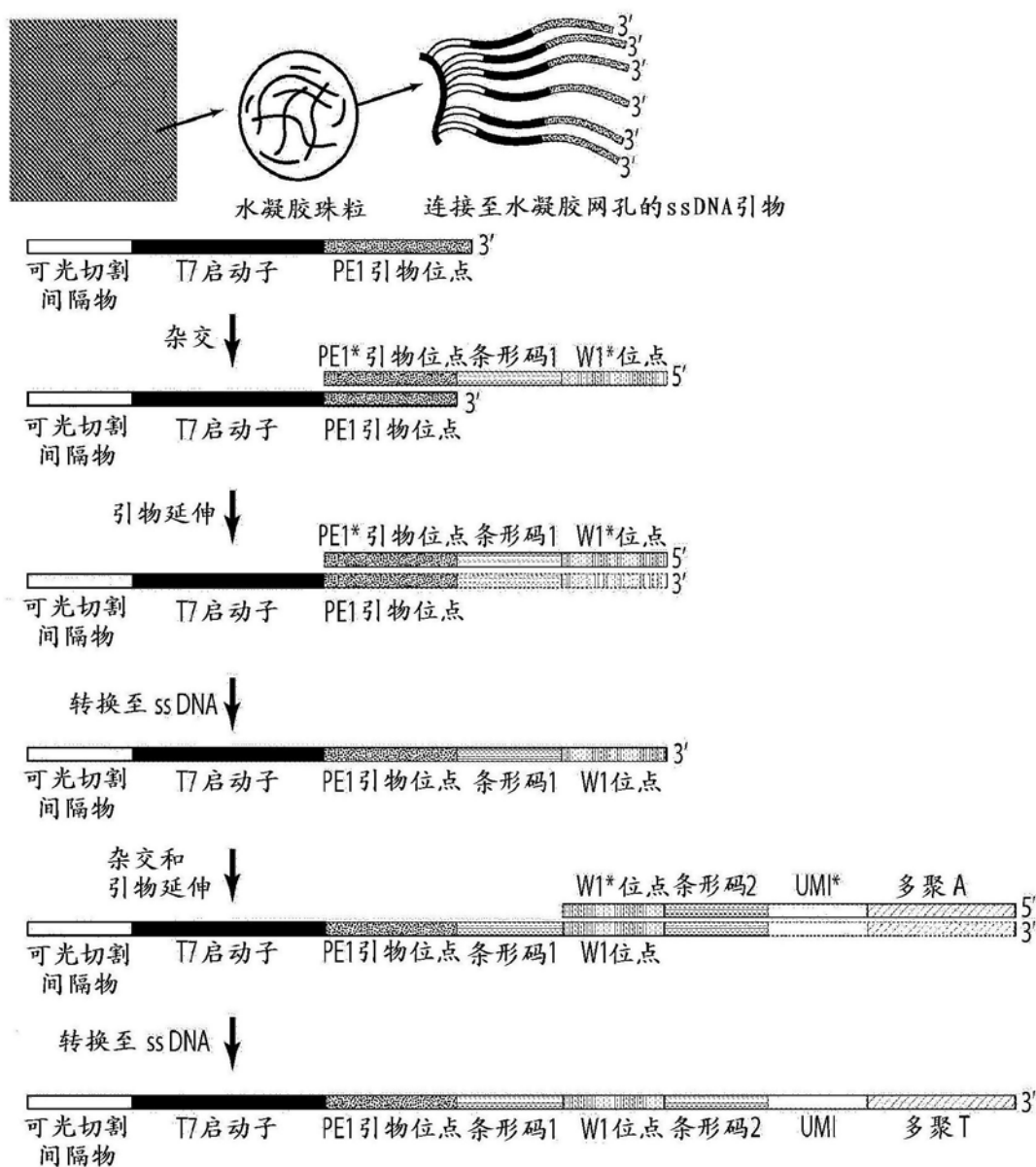


图20B

```
5'-[5Acrd//iSpPC/CGATGACG TAATACGACTCACTATAGGG ATACCACCATGG CTCTTCCCTACACGACGCTCTTC  
CGATCT[barcode1]AAGGCGTCACAAGCAATCACTC[barcode2]NNNNNNTTTTTTTTTTTTTTTTTTT]-3'
```

图20C

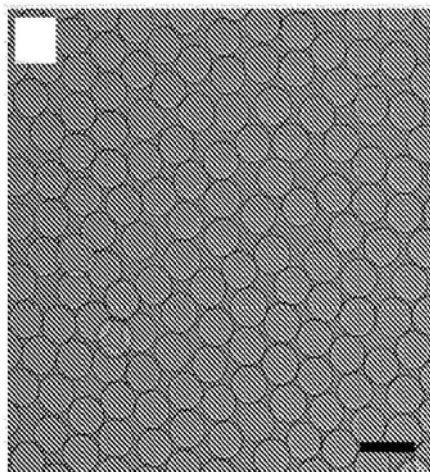


图21A

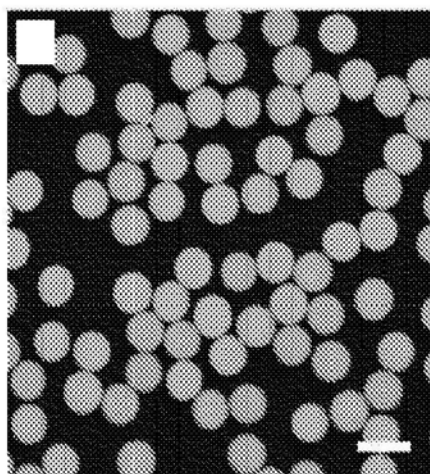


图21B

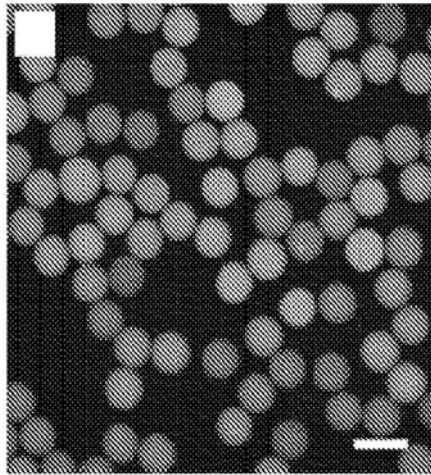


图21C

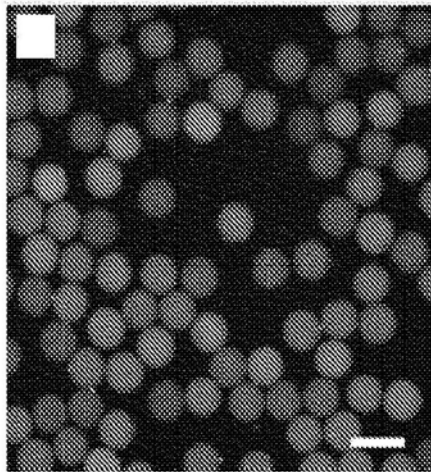


图21D

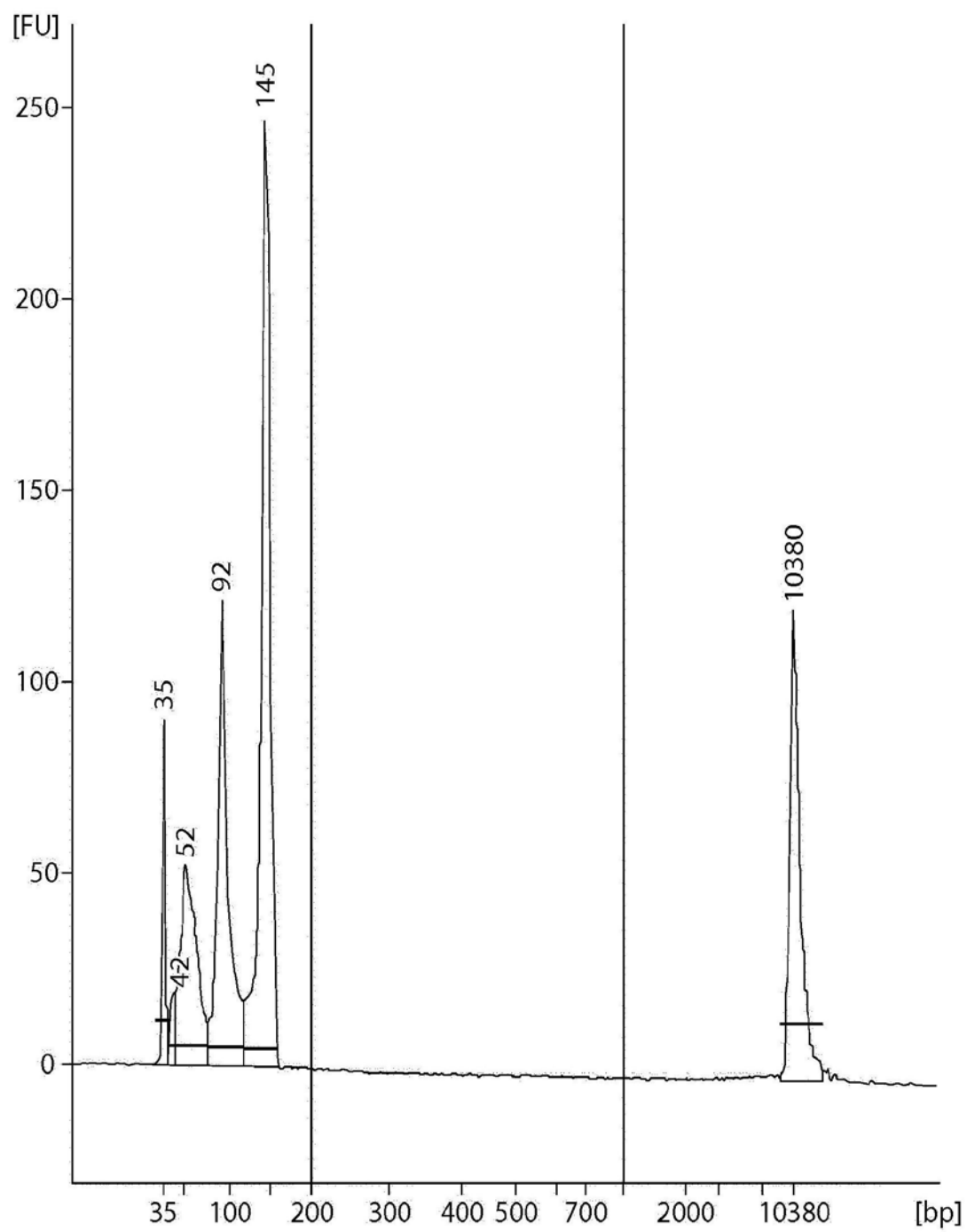


图21E

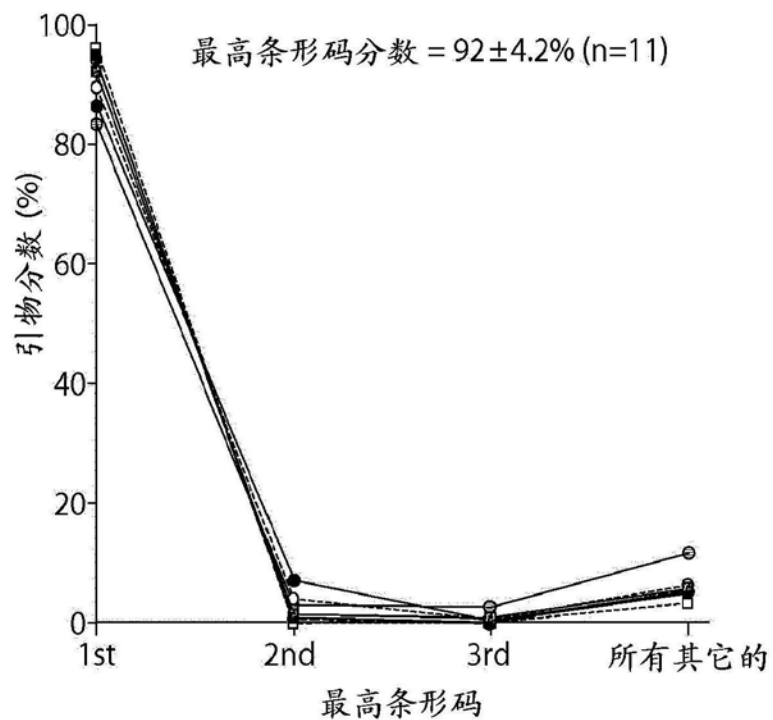


图21F

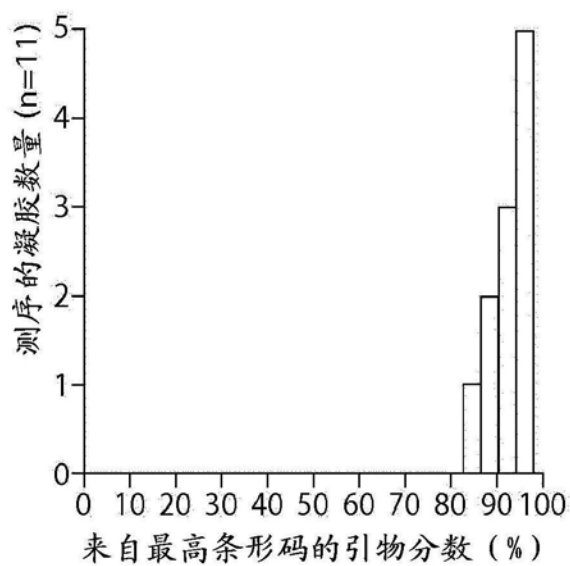


图21G

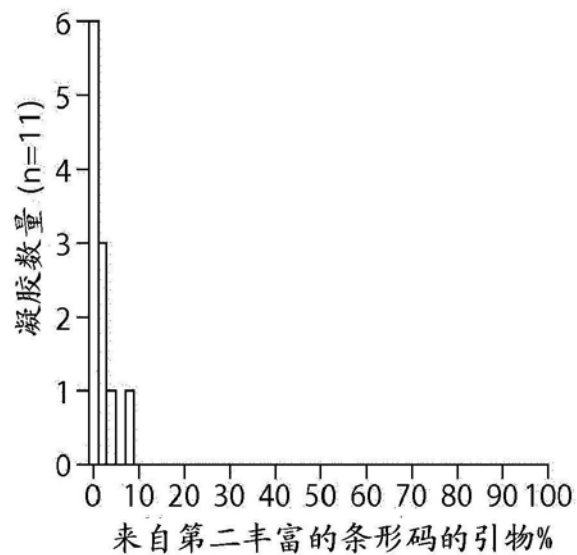


图21H

条形码化的细胞数量/对照

条形码化的的细胞数量/对照

随机 384² 条形码化p值

140	1/0.6	1/0.9	0/0.9	1/0.9	3/2.1	3/2.3	2/2.8
0.88	608	2/3.8	4/3.9	7/3.9	9/8.9	17/10	14/12.1
0.83	0.95	921	7/5.9	5/5.9	22/13.5	26/15.1	20/18.3
0.81	0.91	0.92	945	8/6.1	12/13.9	15/15.5	24/18.8
0.81	0.89	0.91	0.82	951	11/14	17/15.6	23/18.9
0.68	0.92	0.98	0.95	0.86	2,159	36/35.6	47/43.1
0.81	0.92	0.89	0.97	0.99	0.99	2,420	70/48.1
0.94	0.87	0.93	0.98	0.95	0.93	0.93	2,930

图22A

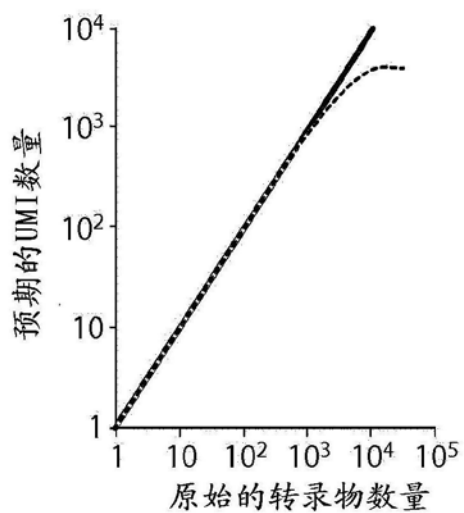


图22B

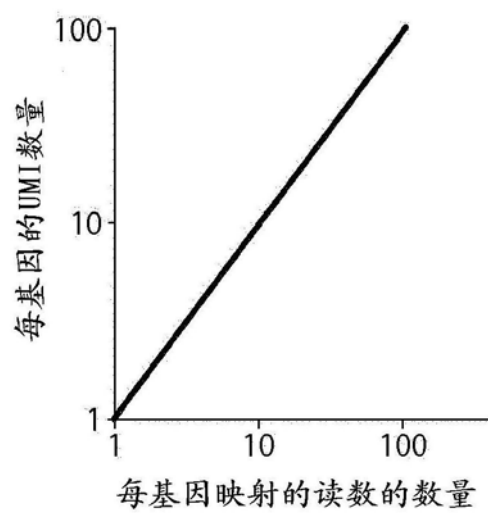


图22C

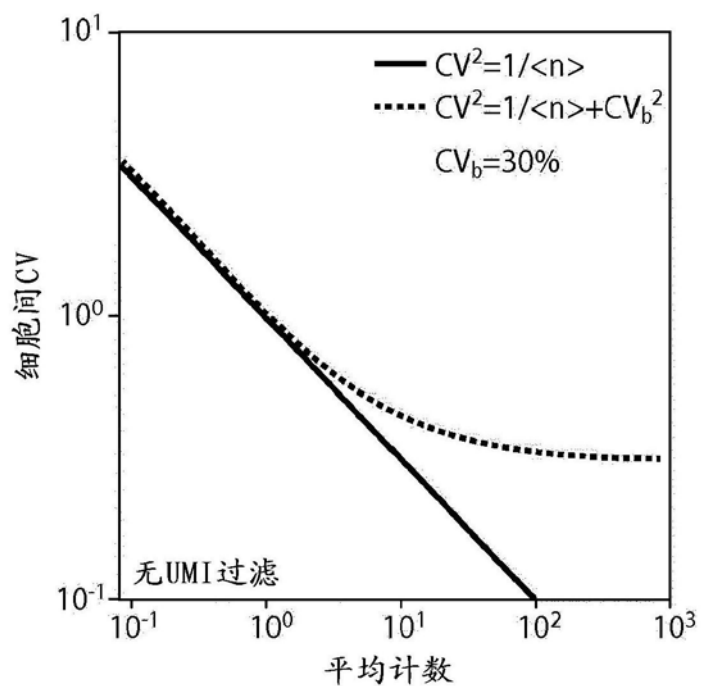


图22D

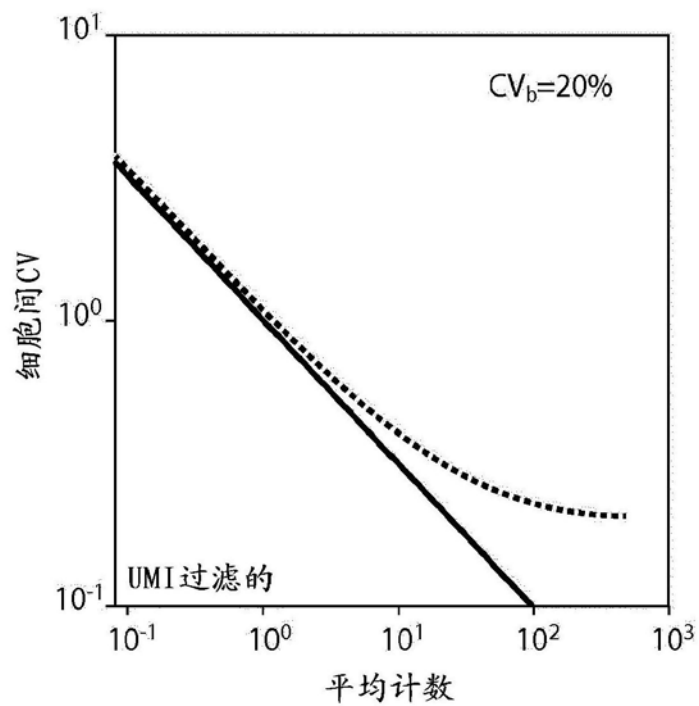


图22E

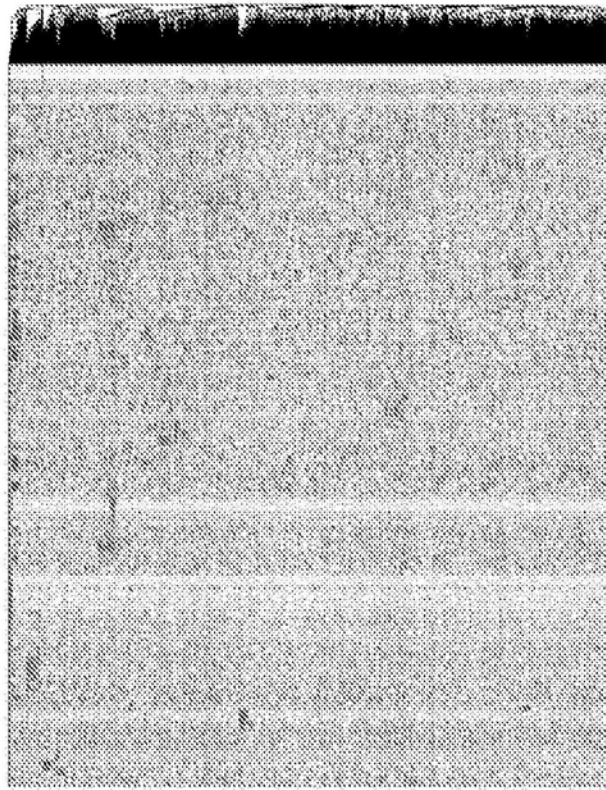


图23A

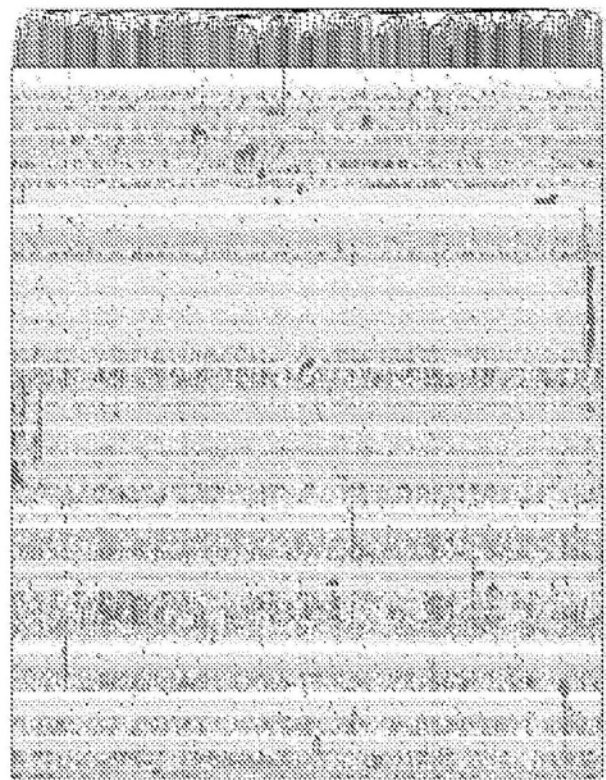


图23B

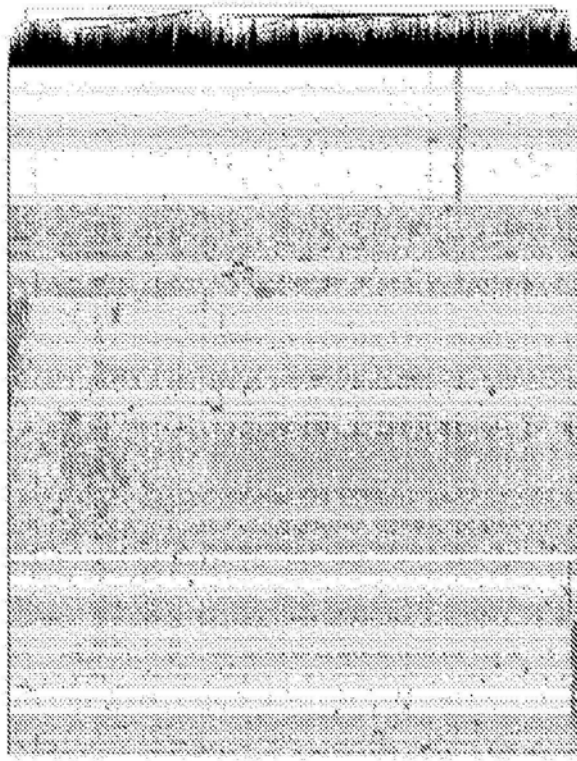


图23C

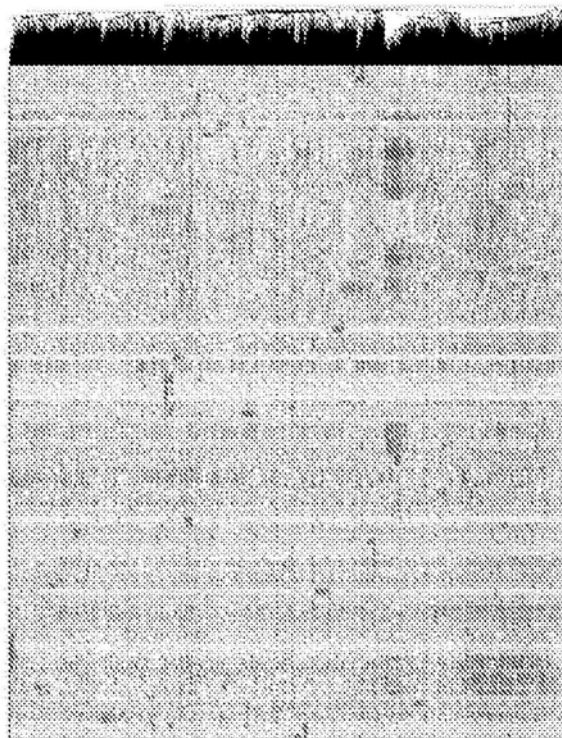


图23D

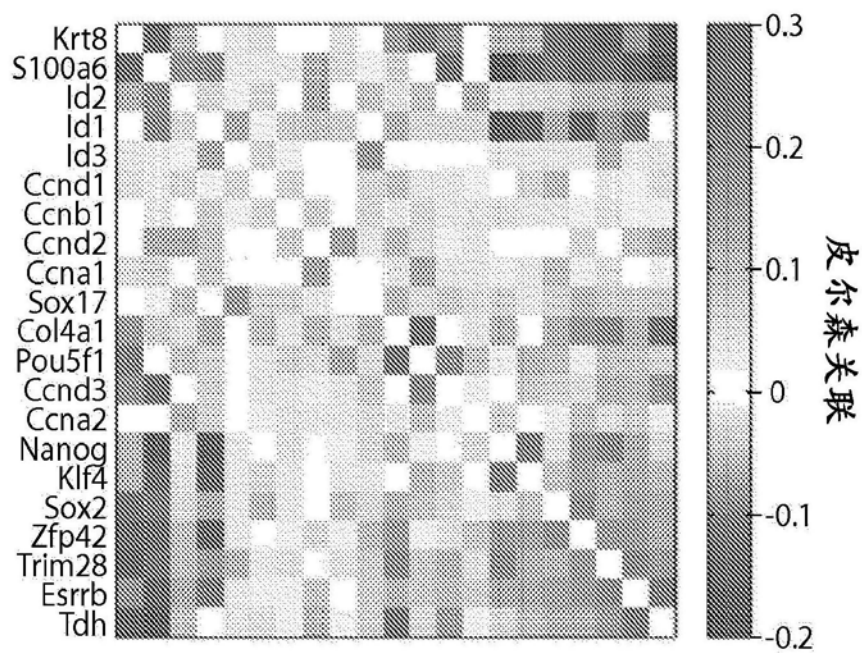


图24A

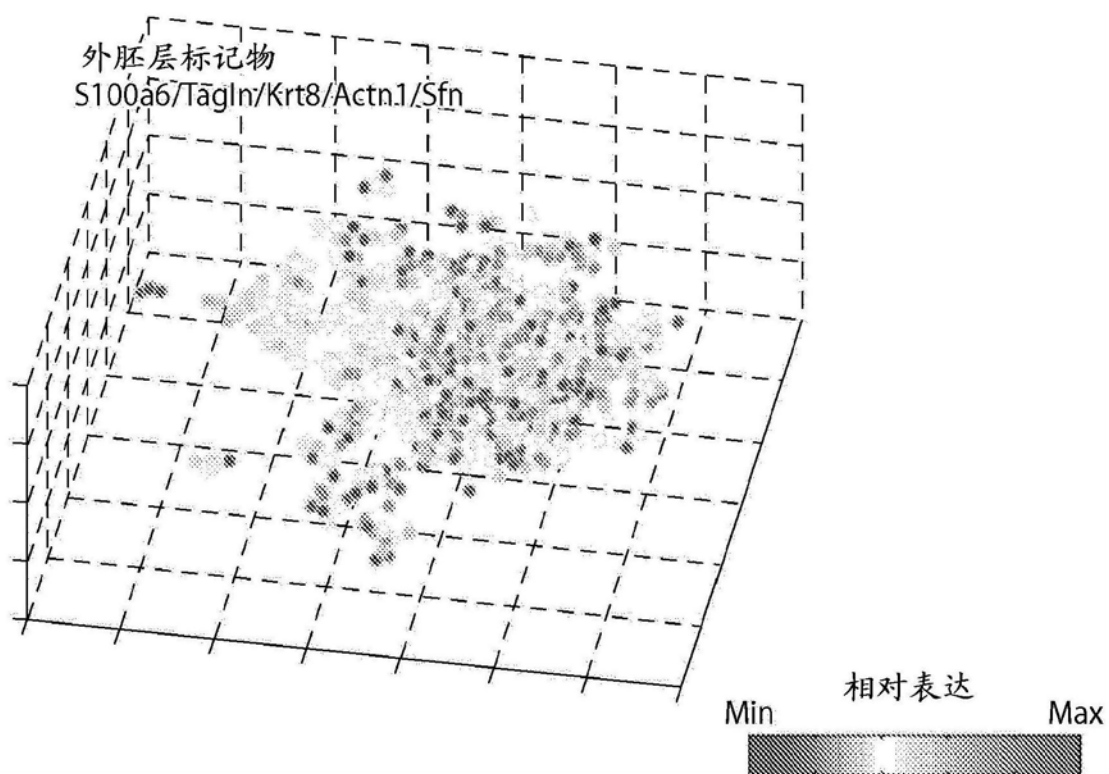


图24B

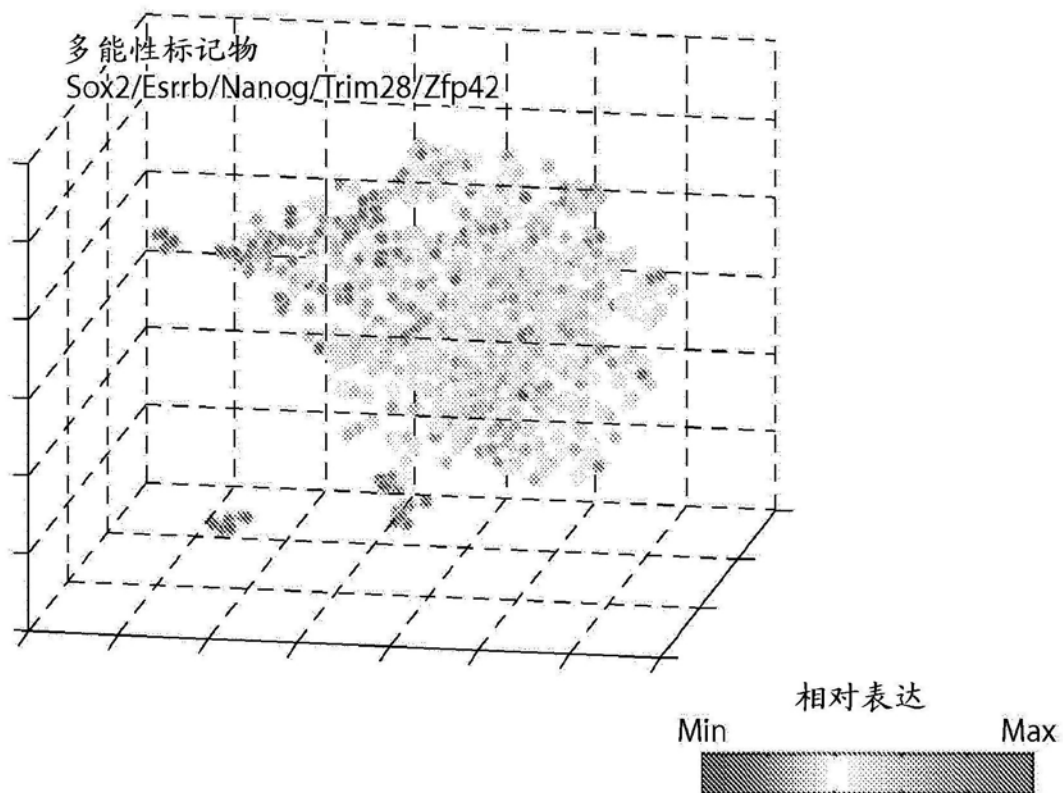


图24C

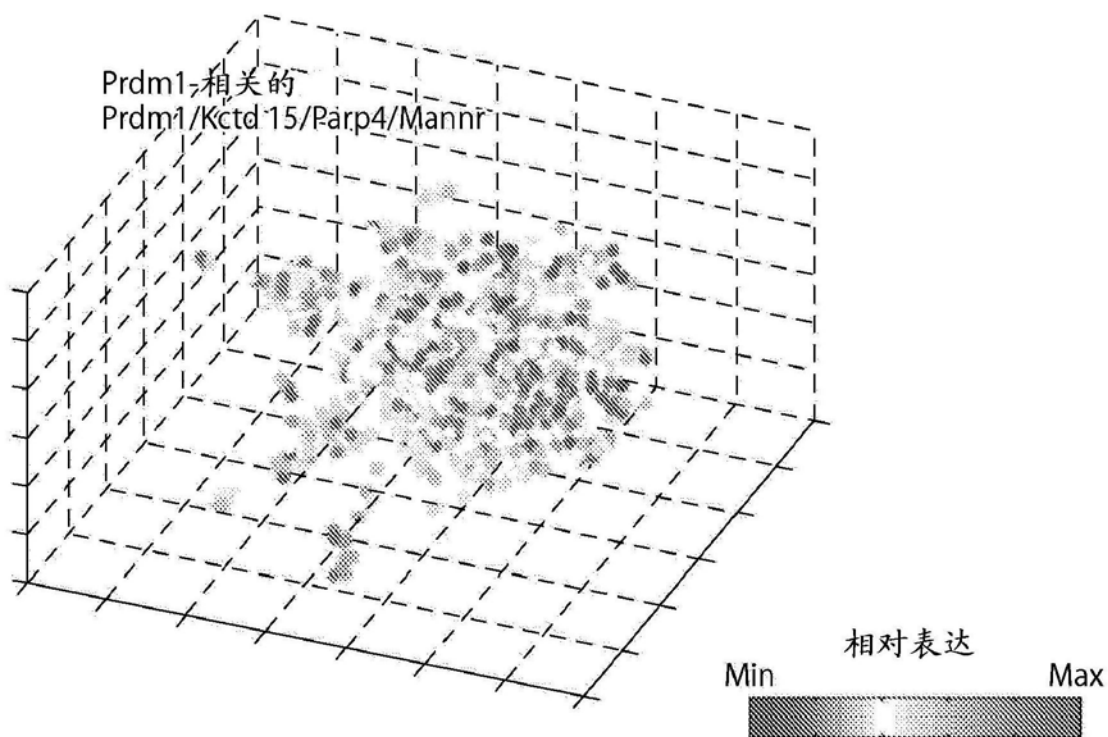


图24D

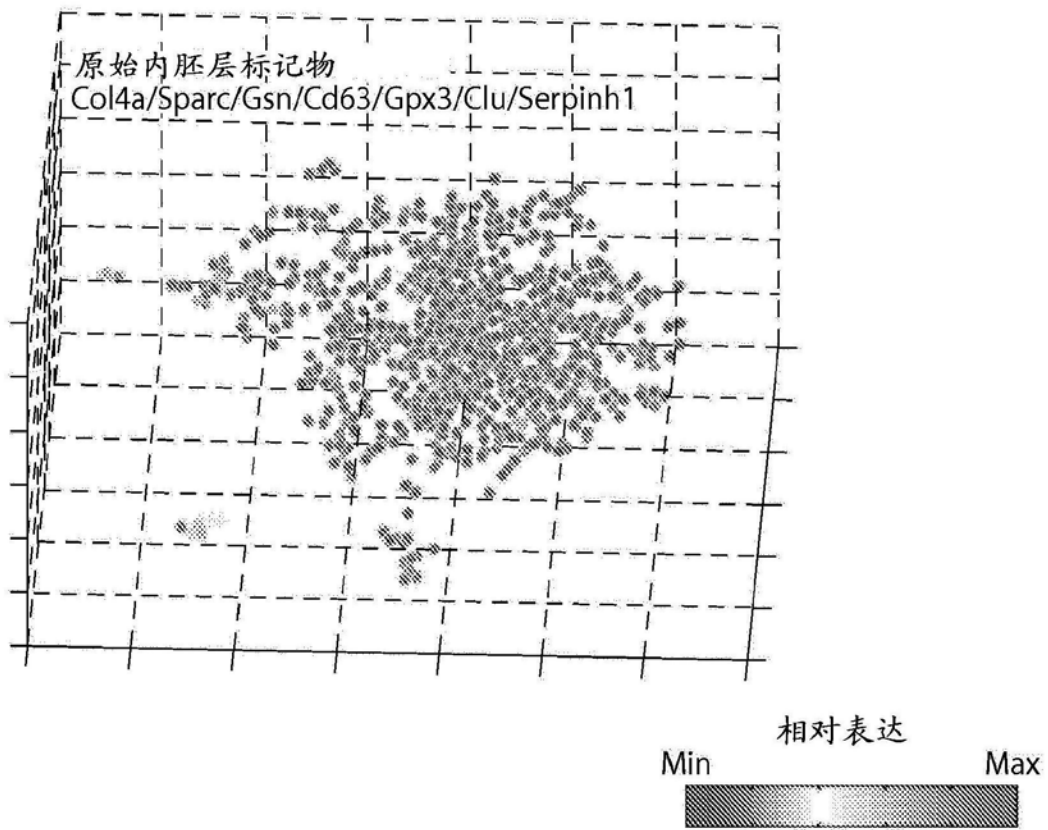


图24E

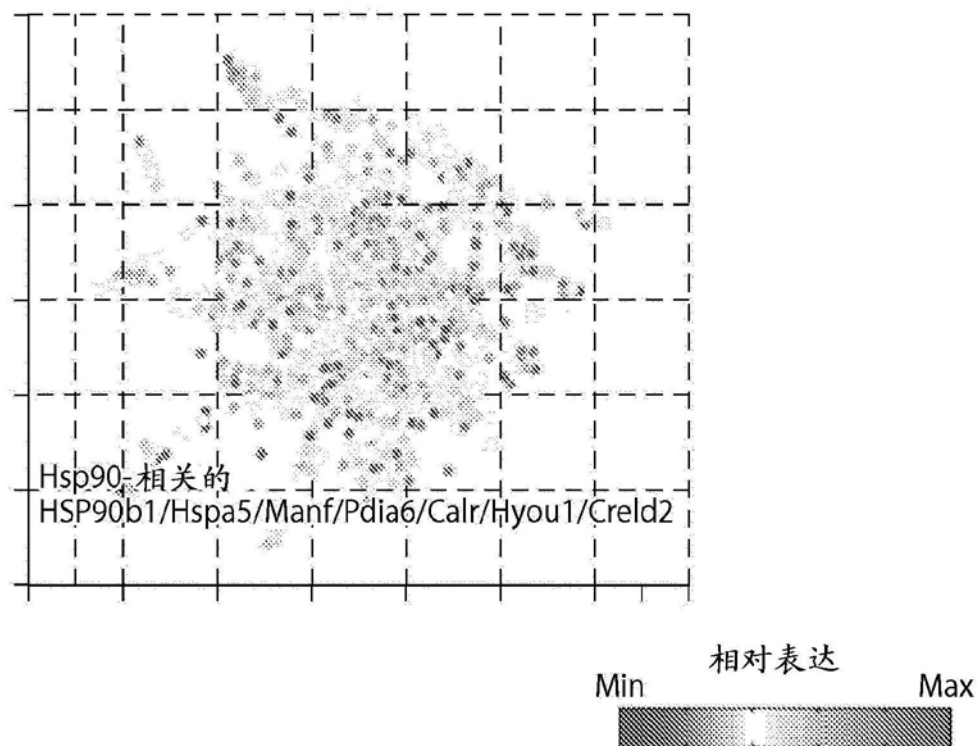


图24F

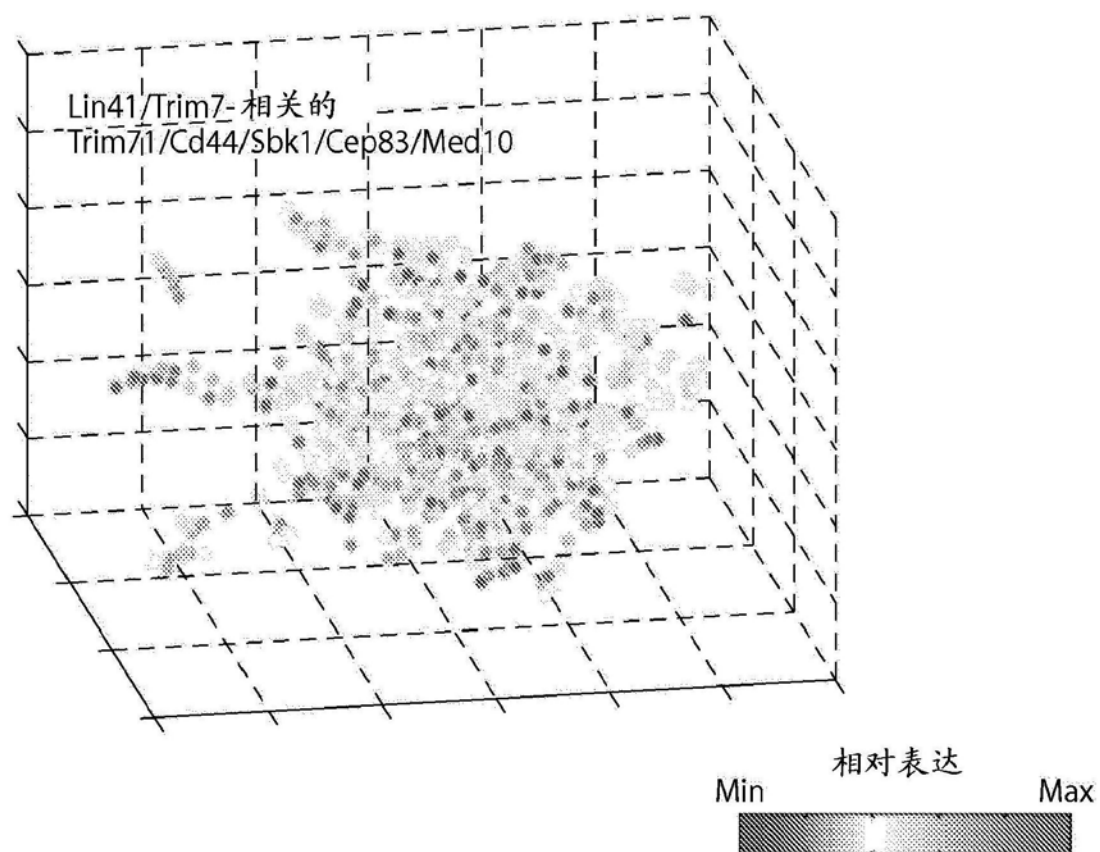


图24G

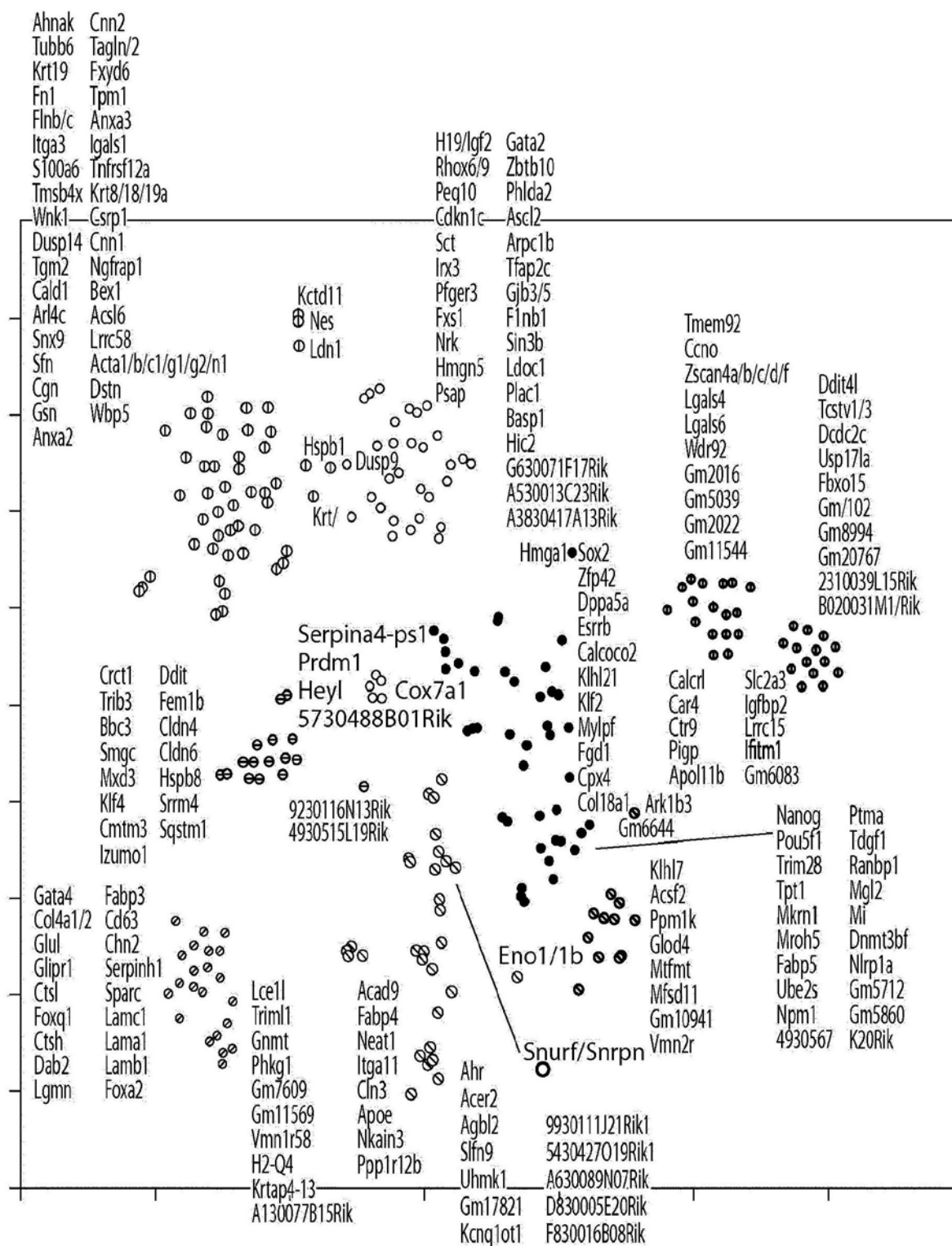


图25

(表2)

基因符号	v-得分	Q值 (BH- FDR校正 的p值)	最大丰度 (归一化的 技术, 经 四舍五入的)	基因符号	1.967	0.024794	22	Elf4ebp1	1.577	0.075054	34
Spr2b	21.646	0	94	Lapm5	1.954	0.024794	35	Ptcd2	1.574	0.075054	14
Krt8	19.012	0	115	Dnmt3l	1.939	0.024794	11	Marcksl1	1.573	0.075054	19
Sparc	14.051	0	259	Krt19	1.936	0.024794	65	Actb	1.570	0.075054	94
Bcl1	12.630	0	269	Wtap	1.924	0.024794	24	Cep83	1.569	0.075054	16
Clu	12.060	0	146	Dusp9	1.923	0.024794	53	Tceb2	1.569	0.075054	29
Krt18	9.637	0	82	Gpx4	1.912	0.024794	41	2410006H16Rik	1.567	0.075054	19
H19	7.345	0	59	Tdh	1.909	0.024794	32	Pkp2	1.566	0.075054	17
Ccdc36	7.124	0	248	P4hb	1.898	0.024794	23	Blvrb	1.562	0.075054	17
Serpinh1	6.580	0	67	Tspan17	1.897	0.024794	34	Ube2l6	1.561	0.075054	15
Hist1h2ao	6.577	0	104	S100a10	1.895	0.032493	40	Pcolce	1.557	0.075054	14
Col4a1	6.194	0	53	Slc2a3	1.891	0.032493	35	Tcl1	1.553	0.075054	11
Fam25c	5.457	0	92	Ldha	1.871	0.038635	36	Klf2	1.553	0.075054	18
Lamb1	4.808	0	55	Peg10	1.863	0.038635	15	Hmces	1.550	0.075054	17
Gm7102	4.557	0	59	Rhox13	1.860	0.038635	24	D11Wsu47e	1.548	0.075054	12
S100a6	4.253	0	60	Igfbp2	1.858	0.038635	31	Mfge8	1.545	0.075054	16
Sfn	4.209	0	44	Tdgf1	1.858	0.038635	15	Rpl39l	1.544	0.075054	12
Malat1	3.667	0	83	Vim	1.849	0.038635	53	Pdia6	1.544	0.075054	23
Lama1	3.666	0	44	Tor4a	1.840	0.053653	36	Creld2	1.543	0.075054	17
Mt1	3.441	0	123	Lamc1	1.818	0.060827	14	Utf1	1.541	0.075054	28
Hspa5	3.373	0	56	Ldhb	1.809	0.067887	28	Surf4	1.539	0.075054	21
Fabp3	3.283	0	52	Lgmn	1.806	0.071607	19	Gng5	1.539	0.075054	26
Tagln2	3.200	0	55	Klf5	1.805	0.071607	17	Stmn2	1.538	0.075054	15
Hist1h2ah	3.005	0	38	Upp1	1.805	0.071607	45	Trib3	1.533	0.075054	16
Acaa1b	2.982	0	34	A630089N07Rik	1.800	0.071607	37	Hist1h4i	1.532	0.075054	9
Ly6a	2.982	0	30	Ngfrap1	1.791	0.071607	14	Hist1h3b	1.528	0.075054	6
Sohlh2	2.966	0	47	Dusp1	1.790	0.071607	30	Exoc3l	1.528	0.075054	12
				Slc2a1	1.786	0.071607	30	Set	1.527	0.075054	67
				BC018473	1.783	0.071607	18	Tpd52	1.527	0.075054	18
				Pmvk	1.780	0.071607	26	Gabarrap12	1.527	0.075054	14
				Calr							

图26-1

Id1	2.965	0	24	Pbld1	1.777	0.071607	54	H2afx	1.526	0.075054	15
Fgd1	2.927	0	73	Fabp5	1.777	0.071607	16	Nxn	1.526	0.075054	13
Lrpap1	2.873	0	55	Slc25a5	1.764	0.071607	21	Cks2	1.522	0.075054	18
Gm6878	2.862	0	28	Elf2s2	1.741	0.071607	52	Nme6	1.521	0.075054	23
Cd63	2.840	0	51	Asb13	1.738	0.07173	28	Cd9	1.519	0.075054	30
Bscl2	2.821	0	44	Kif15	1.738	0.07173	22	2810403A07Rik	1.519	0.075054	25
Mt2	2.690	0	197	Slc3a2	1.737	0.07173	39	Arl6ip1	1.518	0.075054	17
Gm20594	2.633	0	35	Gm7325	1.726	0.07173	25	Pim3	1.518	0.075054	23
Papd5	2.629	0	51	Atf5	1.724	0.07173	15	Psap	1.518	0.075054	18
Slc29a1	2.622	0.016109	36	Id3	1.703	0.075054	15	Sssca1	1.517	0.075054	21
Med10	2.605	0.016109	31	Hspb1	1.684	0.075054	52	Stmn1	1.516	0.075054	15
Actg1	2.560	0.016109	162	Ifitm1	1.681	0.075054	14	Zfp296	1.514	0.075054	11
Dcdc2c	2.515	0.016109	38	Erdr1	1.675	0.075054	40	Pdlim7	1.512	0.075054	17
Lgals1	2.451	0.01882	34	Pgk1	1.669	0.075054	59	Actn1	1.506	0.075054	24
Anxa2	2.419	0.01882	26	Htra1	1.660	0.075054	20	Glxr5	1.506	0.075054	14
Rhox5	2.381	0.01882	15	Kpna2	1.657	0.075054	24	Epha2	1.505	0.075054	23
Mylpf	2.380	0.01882	27	Morf4l2	1.656	0.075054	26	Exosc2	1.505	0.075054	29
Hsp90b1	2.365	0.01882	43	Ctsd	1.652	0.075054	29	Psmd7	1.504	0.075054	21
Meg3	2.362	0.01882	31	Cdc20	1.651	0.075054	20	Gm6083	1.503	0.075054	30
2700046A07Rik	2.317	0.01882	32	Gpx1	1.649	0.075054	40	Pkig	1.503	0.075054	10
Slc25a4	2.315	0.01882	32	Id2	1.649	0.075054	14	Bloc1s1	1.502	0.075054	20
Ctgef	2.309	0.01882	34	Madd	1.648	0.075054	38	Supt4a	1.502	0.075054	19
Tmsb4x	2.279	0.01882	61	Rasgrp2	1.647	0.075054	12	Dnmt3b	1.501	0.075054	21
Gsn	2.264	0.01882	30	Zfp286	1.644	0.075054	16	Apoe	1.500	0.075054	14
Bhmt	2.254	0.01882	20	Pfdn2	1.637	0.075054	18	Clic4	1.496	0.075054	24
Ccnd3	2.245	0.01882	39	Aes	1.632	0.075054	25	Bex1	1.496	0.075054	35
Sept12	2.210	0.01882	17	Ccne1	1.631	0.075054	19	Cenpn	1.495	0.075054	12
Manf	2.206	0.01882	32	Sort1	1.629	0.075054	28	Cox10	1.492	0.075054	15
Trim71	2.184	0.01882	29	Kctd15	1.629	0.075054	24	Mdh2	1.492	0.075054	23
Tagln	2.138	0.01882	20	Elp5	1.628	0.075054	23	Nkain3	1.491	0.075054	66
Col4a2	2.135	0.01882	35	Bcas2	1.617	0.075054	41	Srm	1.490	0.075054	24
Hsp90aa1	2.1215	0.01882	190	Plec	1.611	0.075054	21	Tsc22d1	1.490	0.075054	27

图26-2

Ahcy12	2.105	0.01882	30	Map4k5	1.604	0.075054	35	Qsox1	1.490	0.075054	26
Best2	2.094	0.01882	19	Flnb	1.600	0.075054	28	Ulk1	1.489	0.075054	17
Asns	2.091	0.01882	28	Fth1	1.600	0.075054	65	Gpx3	1.488	0.075054	17
Calco2	2.064	0.01882	25	Glud1	1.599	0.075054	48	Srgn	1.488	0.075054	12
Ctsl	2.063	0.01882	28	Akr1b3	1.598	0.075054	23	Xrc1	1.488	0.075054	12
Dppa5a	2.036	0.01882	133	Mycn	1.596	0.075054	19	Rpl10l	1.487	0.075054	13
Fam229b	2.024	0.01882	38	Nanog	1.596	0.075054	18	Ung	1.486	0.075054	17
Chchd10	2.018	0.01882	29	Acot13	1.595	0.075054	21	Itpk1	1.486	0.075054	14
Sbk1	1.987	0.024794	21	Ikbp	1.594	0.075054	23	Mthfd2	1.485	0.075054	12
E130012A19Rik	1.975	0.024794	31	Ckb	1.594	0.075054	20	Srrd	1.485	0.075054	16
Parvb	1.481	0.075054	19	Echdc2	1.589	0.075054	12	Hspa12b	1.483	0.075054	26
Imp3	1.479	0.075054	14	Tex19.1	1.426	0.075054	18	Qdpr	1.396	0.075054	12
Thoc3	1.479	0.075054	21	Ube3b	1.425	0.075054	14	Midn	1.396	0.075054	20
Hn1	1.478	0.075054	16	Col18a1	1.425	0.075054	15	Mars	1.395	0.075054	16
1110008F13Rik	1.478	0.075054	18	Lias	1.425	0.075054	20	Dap	1.395	0.075054	10
2310036O22Rik	1.477	0.075054	20	Fscn1	1.423	0.075054	19	2200002D01Rik	1.394	0.075054	11
Crip1	1.476	0.075054	16	Ltb	1.423	0.075054	14	Zdhc7	1.392	0.075054	10
Esrrb	1.475	0.075054	20	Psm2	1.422	0.075054	27	Hap1	1.392	0.075054	12
Gtf3a	1.475	0.075054	13	Pmm1	1.422	0.075054	18	Cdv3	1.392	0.075054	24
Aurka	1.474	0.075054	10	Arpc1b	1.422	0.075054	16	Polr2e	1.392	0.075054	22
Ass1	1.473	0.075054	15	Mrpl34	1.422	0.075054	17	Ctnnal1	1.390	0.075054	15
Rnaseh2c	1.473	0.075054	14	Grrn	1.422	0.075054	15	Nlrp1a	1.390	0.075054	41
Tbc1d16	1.472	0.075054	23	Cct5	1.422	0.075054	24	Snx6	1.390	0.075054	10
Ube2s	1.471	0.075054	31	Cfdp1	1.421	0.075054	20	Xpnpep1	1.390	0.075054	18
Cdk2ap1	1.470	0.075054	15	Ltbp4	1.421	0.075054	21	Gm10845	1.390	0.075054	19
Tmem230	1.469	0.075054	21	Tubb2b	1.421	0.075054	14	Trim25	1.390	0.075054	18
Pdgfa	1.468	0.075054	10	Lima1	1.421	0.075054	15	Gpx2	1.390	0.075054	8
Rrp7a	1.468	0.075054	13	B3gnt7	1.420	0.075054	14	Esd	1.389	0.075054	14
Tyw1	1.468	0.075054	36	Tubb3	1.419	0.075054	12	Sgk1	1.389	0.075054	13
Erf	1.464	0.075054	22	Fkbp3	1.419	0.075054	17	Mrpl13	1.389	0.075054	12
Timm8b	1.464	0.075054	12	Mybl2	1.419	0.075054	25	Stmn3	1.389	0.075054	12
				Txnrd1	1.417	0.075054	22	Mrpl53	1.389	0.075054	9

图26-3

D230025D16Rik	1.464	0.075054	14	Trmt1	1.417	0.075054	15	Rwdd1	1.388	0.075054	16
Scd2	1.463	0.075054	19	Fkbp1a	1.417	0.075054	32	Fkbp4	1.388	0.075054	19
Med13l	1.463	0.075054	19	Uchl1	1.417	0.075054	16	Txn1	1.388	0.075054	20
Syce2	1.462	0.075054	13	Psmc11	1.416	0.075054	12	Asb6	1.387	0.075054	11
Arf1	1.462	0.075054	34	Mif4gd	1.416	0.075054	13	Chchd4	1.387	0.075054	13
Fcf1	1.461	0.075054	10	Pfdn6	1.416	0.075054	20	Sqstm1	1.387	0.075054	19
Etv4	1.459	0.075054	17	1700007K13Rik	1.416	0.075054	8	Guca1a	1.386	0.075054	7
Eid2	1.458	0.075054	8	Rabac1	1.416	0.075054	11	Glpr2	1.386	0.075054	13
Polr2h	1.456	0.075054	18	Snrpa1	1.415	0.075054	16	Prpsap1	1.386	0.075054	17
Rad51	1.456	0.075054	13	Dtd1	1.414	0.075054	11	Rbmxl1	1.386	0.075054	16
Dppa4	1.455	0.075054	17	Ddx56	1.413	0.075054	20	Elf3c	1.385	0.075054	18
Sertad1	1.454	0.075054	12	Srpk1	1.413	0.075054	15	Tspan4	1.385	0.075054	11
Lincenc1	1.453	0.075054	11	Krtcap2	1.413	0.075054	14	Dmkn	1.384	0.075054	10
Cox7a2l	1.451	0.075054	19	Gorasp2	1.412	0.075054	14	Rian	1.384	0.075054	11
Sf3b5	1.451	0.075054	13	Malsu1	1.412	0.075054	12	Copb2	1.384	0.075054	13
Tra2b	1.451	0.075054	19	Wbp5	1.411	0.075054	18	Ahsa1	1.384	0.075054	19
Tinag1l	1.450	0.075054	14	Mark3	1.409	0.075054	24	Ndufb6	1.384	0.075054	12
Tubb6	1.450	0.075054	14	Tsta3	1.408	0.075054	12	Atp6v0e	1.383	0.075054	13
Chchd3	1.450	0.075054	13	Tpx2	1.408	0.075054	16	Ccdc22	1.382	0.075054	11
Fam195a	1.449	0.075054	9	Sec11c	1.407	0.075054	15	Rbm42	1.381	0.075054	12
Abcd4	1.449	0.075054	27	H2afz	1.407	0.075054	34	Cyc1	1.380	0.075054	21
F830016B08Rik	1.448	0.075054	26	Nfyb	1.407	0.075054	13	Lmna	1.380	0.075054	15
U2af1	1.446	0.075054	21	Atg101	1.407	0.075054	14	Clic1	1.380	0.075054	10
Rbfa	1.444	0.075054	10	Wars	1.407	0.075054	17	Klf4	1.379	0.075054	14
Fn1	1.444	0.075054	23	Atp6v0d1	1.406	0.075054	12	Slc30a2	1.379	0.075054	11
Cope	1.444	0.075054	11	Coq7	1.406	0.075054	10	Psmc6	1.379	0.075054	16
Dym	1.443	0.075054	11	Ctsh	1.406	0.075054	16	Scamp3	1.378	0.075054	16
Calu	1.443	0.075054	18	281042815Rik	1.406	0.075054	15	Calm1	1.378	0.075054	24
Crif2	1.442	0.075054	11	Fam96b	1.406	0.075054	14	Qars	1.376	0.075054	15
Top2a	1.441	0.075054	14	Hbegf	1.406	0.075054	15	Mirpl43	1.376	0.075054	38
Cstb	1.441	0.075054	10	Samm50	1.405	0.075054	15	Adamts13	1.376	0.075054	18
Ppa1	1.440	0.075054	17	Zc3hc1	1.404	0.075054	19	Ppie	1.376	0.075054	11

图26-4

Arl6ip4	1.440	0.075054	10	Rpap1	1.404	0.075054	17	Mea1	1.375	0.075054	11
Tuba4a	1.439	0.075054	24	Tpi1	1.403	0.075054	11	Prpf3	1.375	0.075054	16
Bzw2	1.438	0.075054	19	Elof1	1.403	0.075054	12	Atf4	1.375	0.075054	13
Churc1	1.438	0.075054	34	Tcea3	1.403	0.075054	13	Spata33	1.375	0.075054	11
Bud31	1.438	0.075054	20	1500012F01Rik	1.402	0.075054	14	Atp5j2	1.374	0.075054	19
Aifm1	1.437	0.075054	16	Cenpa	1.401	0.075054	14	Lrc59	1.374	0.075054	23
Ndufa13	1.436	0.075054	12	Pecr	1.400	0.075054	13	Fam129b	1.374	0.075054	14
Dstn	1.435	0.075054	18	Myod1	1.400	0.075054	10	Nhp2	1.373	0.075054	22
Ocel1	1.435	0.075054	15	Degs1	1.400	0.075054	14	Aprt	1.372	0.075054	22
Galk1	1.435	0.075054	11	Ndr4	1.399	0.075054	16	Cdc5l	1.371	0.075054	19
Hmox1	1.433	0.075054	12	Lonp1	1.399	0.075054	16	Gna11	1.371	0.075054	11
Ubal2	1.433	0.075054	15	Ppa2	1.399	0.075054	15	Spns1	1.370	0.075054	19
Romo1	1.432	0.075054	14	Apitd1	1.399	0.075054	10	Ntan1	1.370	0.075054	20
Phc1	1.432	0.075054	23	Dcun1d5	1.399	0.075054	11	Actl6a	1.369	0.075054	13
Emc10	1.432	0.075054	13	Ndufa4	1.398	0.075054	19	Ywhaq	1.369	0.075054	24
Ptov1	1.431	0.075054	17	Stx5a	1.398	0.075054	9	Chmp4b	1.369	0.075054	17
Snf8	1.430	0.075054	21	Rheb	1.398	0.075054	21	Fto	1.368	0.075054	24
Calm2	1.429	0.075054	21	Ddt	1.397	0.075054	14	Sdhc	1.368	0.075054	17
Mrpl14	1.429	0.075054	13	Trappc3	1.397	0.075054	12	Plin2	1.368	0.075054	11
Pih1d1	1.427	0.075054	14	Cbx3	1.397	0.075054	19	Eif3g	1.368	0.075054	19
Dhx16	1.427	0.075054	19	Slirp	1.396	0.075054	21	Mcm7	1.368	0.075054	16
Siva1	1.427	0.075054	18	Mrpl12	1.396	0.075054	14	Crip2	1.368	0.075054	11
Pop5	1.426	0.075054	11	Tcf15	1.396	0.075054	12	Slc6a8	1.367	0.075054	20
Saal1	1.367	0.075054	14	Gins4	1.349	0.075054	19	Rpl39	1.334	0.075054	53
Mrpl23	1.367	0.075054	17	Slc15a4	1.349	0.075054	14	Srp14	1.333	0.075054	15
Cdkn1a	1.367	0.075054	9	Agpat2	1.349	0.075054	9	Umps	1.333	0.075054	14
Emg1	1.366	0.075054	14	Cdca5	1.349	0.075054	11	Use1	1.333	0.075054	10
Ccdc124	1.366	0.075054	19	A430005L14Rik	1.348	0.075054	12	Nxt1	1.332	0.075054	11
Brl3	1.366	0.075054	14	Mrpl30	1.348	0.075054	15	Ehd1	1.332	0.075054	27
Csrp1	1.366	0.075054	17	Rcl1	1.348	0.075054	15	Timm10	1.332	0.075054	10
Glul	1.366	0.075054	20	Tmem216	1.347	0.075054	8	Srsf11	1.331	0.075054	15
Socs2	1.366	0.075054	10	Sall4	1.347	0.075054	20	Cyb5a	1.331	0.075054	13

图26-5

Ftl1	1.365	0.075054	185	Bex4	1.347	0.075054	14	Cox7a1	1.331	0.075054	7
Sox2	1.365	0.075054	21	Pttg1ip	1.347	0.075054	14	Taf11	1.331	0.075054	8
Cdkn2a1p	1.364	0.075054	15	Immt	1.347	0.075054	21	Pdhh	1.331	0.075054	16
Hs1bp3	1.364	0.075054	13	Wdr74	1.347	0.075054	15	Eif2b2	1.331	0.075054	11
Fdft1	1.364	0.075054	13	Pttg1	1.347	0.075054	10	Lcmt1	1.330	0.075054	14
Exosc8	1.364	0.075054	13	Bub3	1.346	0.075054	10	Fzr1	1.330	0.075054	9
Guk1	1.364	0.075054	13	Nodal	1.346	0.075054	9	Cggbp1	1.330	0.075054	22
Deb1	1.363	0.075054	8	Pdp1f	1.346	0.075054	10	Prdx6	1.330	0.075054	18
Ndufb10	1.362	0.075054	12	Atp5g1	1.346	0.075054	24	Fbxo6	1.330	0.075054	8
Rnf7	1.362	0.075054	19	Eftud2	1.345	0.075054	16	Tomm70a	1.330	0.075054	22
Ndufb11	1.362	0.075054	13	Nop10	1.345	0.075054	15	Serpinb6a	1.330	0.075054	12
Dynl1	1.362	0.075054	22	Rbpms2	1.344	0.075054	13	Rmrp	1.330	0.075054	5
Gstm1	1.361	0.075054	9	Impact	1.344	0.075054	19	Psmal	1.330	0.075054	18
Smn1	1.361	0.075054	14	Vbp1	1.344	0.075054	14	Ppm1a	1.329	0.075054	11
Vdac3	1.361	0.075054	14	Plk1	1.344	0.075054	13	Chtf18	1.329	0.075054	13
Cdkn2a1pnl	1.360	0.075054	14	Pebp1	1.344	0.075054	26	Sec13	1.329	0.075054	11
Safb2	1.360	0.075054	15	Skap2	1.343	0.075054	11	Ndufa10	1.329	0.075054	18
Dctn6	1.360	0.075054	9	Atp6v0b	1.343	0.075054	16	Slc7a3	1.328	0.075054	15
Gnb2l1	1.360	0.075054	137	Maged1	1.343	0.075054	12	Dapk3	1.328	0.075054	10
Psma5	1.360	0.075054	17	Trmt112	1.343	0.075054	14	Uqcc2	1.328	0.075054	17
Dazap2	1.360	0.075054	9	Tfgr	1.343	0.075054	11	Rarg	1.327	0.075054	12
Isyna1	1.360	0.075054	11	Tytw3	1.342	0.075054	15	Polr2k	1.327	0.075054	13
Odc1	1.360	0.075054	35	Ubqln1	1.342	0.075054	17	Mfsd3	1.327	0.075054	12
Psmb4	1.360	0.075054	25	Ddost	1.341	0.075054	28	Gm9855	1.326	0.075054	12
Timm23	1.359	0.075054	23	Ccnb1	1.341	0.075054	13	Scand1	1.326	0.075054	7
1110038F14Rik	1.359	0.075054	9	Slc25a20	1.341	0.075054	10	Polr2i	1.326	0.075054	14
Tc2n	1.359	0.075054	16	Apoc1	1.340	0.075054	9	Atp5e	1.326	0.075054	18
Srtt	1.359	0.075054	22	Dscr3	1.340	0.075054	17	Ppp1r8	1.326	0.075054	15
Defa6	1.359	0.075054	5	Ssr2	1.340	0.075054	14	Rilpl1	1.326	0.075054	8
Hax1	1.359	0.075054	11	Fgf4	1.340	0.075054	12	Podxl	1.326	0.075054	14
Epas1	1.358	0.075054	13	Myl12a	1.340	0.075054	26	Rpph1	1.325	0.075054	4
Socs3	1.358	0.075054	12	Mcm6	1.340	0.075054	22	Lsr	1.325	0.075054	10

图26-6

Rtcb	1.358	0.075054	21	Mrps12	1.339	0.075054	14	Psmc1	1.325	0.075054	18
Psmc2	1.358	0.075054	26	Ubl5	1.339	0.075054	14	Lin7b	1.325	0.075054	6
Llg12	1.358	0.075054	12	Trap1	1.339	0.075054	11	Gabarapl1	1.325	0.075054	13
Mcl1	1.358	0.075054	16	Atp5o	1.339	0.075054	21	Zfp42	1.325	0.075054	17
Cct4	1.357	0.075054	21	Gtf2f1	1.339	0.075054	17	Snrf	1.325	0.075054	24
Sucg1	1.357	0.075054	13	Pacsin2	1.339	0.075054	26	Bcl3	1.324	0.075054	10
1110058L19Rik	1.357	0.075054	9	Dcaf7	1.339	0.075054	15	Atic	1.324	0.075054	17
Ybx3	1.356	0.075054	24	2410004N09Rik	1.338	0.075054	11	Ddx39b	1.324	0.075054	17
P4ha2	1.356	0.075054	11	Cisd3	1.338	0.075054	10	Ppp2r5a	1.324	0.075054	15
Fkbp2	1.356	0.075054	12	Asnsd1	1.338	0.075054	13	Rangap1	1.323	0.075054	15
Fblim1	1.355	0.075054	15	Pomgnt1	1.337	0.075054	9	Abhd17a	1.323	0.075054	10
Sf3b1	1.355	0.075054	31	Rpa1	1.337	0.075054	17	Smyd5	1.323	0.075054	11
Psmc3	1.355	0.075054	34	Mrpl4	1.337	0.075054	12	Dpp7	1.323	0.075054	8
Acer2	1.355	0.075054	26	Adprh	1.337	0.075054	10	Dnpep	1.323	0.075054	11
Cacng7	1.354	0.075054	14	Tomn40l	1.336	0.075054	13	Bms1	1.323	0.075054	16
Mrps36	1.354	0.075054	10	Ndufa6	1.336	0.075054	21	Ptp4a2	1.323	0.075054	15
Phf23	1.354	0.075054	12	B4galt3	1.336	0.075054	12	Lage3	1.323	0.075054	14
Eif2b5	1.353	0.075054	14	Mvd	1.336	0.075054	12	Cdca8	1.323	0.075054	16
Ndufs7	1.352	0.075054	13	Rbmx12	1.336	0.075054	13	Ankrd46	1.323	0.075054	8
Tubg1	1.352	0.075054	9	Hsph1	1.336	0.075054	22	Ddx23	1.323	0.075054	13
Dnajc2	1.352	0.075054	17	Lin28a	1.336	0.075054	12	Rfc2	1.323	0.075054	14
Capns1	1.351	0.075054	22	Got1	1.336	0.075054	15	Uqcrb	1.323	0.075054	13
Utp3	1.351	0.075054	11	Utp14a	1.336	0.075054	9	Txn1	1.322	0.075054	27
Ctnna1	1.351	0.075054	17	Efcc1	1.336	0.075054	21	Snhg5	1.322	0.075054	11
Prr13	1.351	0.075054	11	Ndufs3	1.336	0.075054	15	Bcat1	1.322	0.075054	31
Psph	1.351	0.075054	10	Fam64a	1.335	0.075054	9	Yrdc	1.322	0.075054	16
Prkcsb	1.351	0.075054	21	Txnip	1.335	0.075054	13	Tuba1b	1.322	0.075054	43
Mrpl54	1.351	0.075054	13	AU022252	1.335	0.075054	12	Vipas39	1.321	0.075054	8
Slmo2	1.351	0.075054	9	Tbrg1	1.335	0.075054	15	Cops8	1.321	0.075054	15
Mrpl46	1.350	0.075054	10	Blmh	1.335	0.075054	21	Pcna	1.321	0.075054	17
Tprgl	1.350	0.075054	13	Uqcr10	1.335	0.075054	25	Ydjic	1.320	0.075054	12
Trim28	1.350	0.075054	50	Hexa	1.334	0.075054	9	Flna	1.320	0.075054	19
Ctc1	1.350	0.075054	11	Mrpl44	1.334	0.075054	10	Ampd2	1.320	0.075054	12
Ube2c	1.349	0.075054	26	Add1	1.334	0.075054	11	Gpn1	1.320	0.075054	11
Txn2	1.349	0.075054	15	Mrps24	1.334	0.075054	16	Sars	1.320	0.075054	11

图26-7

Pf3dn4	1.319	0.075054	13	Tnk2	1.308	0.075054	12	Coro1a	1.297	0.075054	10
Mfsd5	1.319	0.075054	11	Nutf2	1.308	0.075054	12	H1fx	1.297	0.075054	8
Igf2bp1	1.319	0.075054	45	Ddx47	1.307	0.075054	10	Psm6	1.297	0.075054	21
Pgp	1.318	0.075054	13	Gpatch4	1.307	0.075054	16	Stra13	1.297	0.075054	13
Ear10	1.318	0.075054	6	Rrp12	1.307	0.075054	11	N6amt2	1.297	0.075054	10
Cldn7	1.318	0.075054	14	Tma7	1.307	0.075054	23	Dus1l	1.297	0.075054	11
Cpsf4l	1.318	0.075054	10	Cuta	1.307	0.075054	11	Smarrcb1	1.296	0.075054	13
Mif	1.317	0.075054	29	Cox14	1.307	0.075054	10	Tax1bp1	1.296	0.075054	12
Nsmce4a	1.317	0.075054	14	Mdh1	1.307	0.075054	16	Mtdh	1.296	0.075054	21
Pold1	1.317	0.075054	16	Tufm	1.307	0.075054	16	Rlim	1.296	0.075054	23
Pnp	1.316	0.075054	15	Tomm5	1.306	0.075054	11	S100a11	1.296	0.075054	8
Cib1	1.316	0.075054	9	Cog8	1.306	0.075054	13	Gde1	1.295	0.075054	8
Tipin	1.316	0.075054	12	Mbd3	1.306	0.075054	24	Plrg1	1.295	0.075054	10
Vdac1	1.316	0.075054	26	Uqcr11	1.306	0.075054	21	Gabarap	1.295	0.075054	12
4833439L19Rik	1.316	0.075054	12	Ap4m1	1.306	0.075054	11	Ilk	1.295	0.075054	8
L1td1	1.315	0.075054	24	Rbpj	1.306	0.075054	18	Fads1	1.294	0.075054	12
Slc25a1	1.315	0.075054	9	Nedd8	1.306	0.075054	18	Dnajc21	1.294	0.075054	10
Ints3	1.315	0.075054	13	Gtsf1l	1.305	0.075054	11	Tpst2	1.294	0.075054	10
Cdc16	1.315	0.075054	11	2500004C02Rik	1.305	0.075054	21	Lamtor4	1.294	0.075054	13
Sdhb	1.315	0.075054	10	Uqcrh	1.305	0.075054	20	Skp1a	1.294	0.075054	19
Sfr1	1.315	0.075054	13	Myl9	1.305	0.075054	14	Msn	1.294	0.075054	17
Phlda3	1.315	0.075054	12	Zfp771	1.305	0.075054	8	Grhpr	1.294	0.075054	9
Ndufs2	1.314	0.075054	12	Ifi30	1.305	0.075054	10	Rsl1d1	1.293	0.075054	17
Psrc1	1.314	0.075054	10	Lsm4	1.304	0.075054	19	Nrp	1.293	0.075054	8
Prdx2	1.314	0.075054	18	Ywhab	1.304	0.075054	20	Psmc2	1.293	0.075054	15
Fam168b	1.314	0.075054	13	Dnajb1	1.304	0.075054	10	Me1	1.293	0.075054	12
Ebp	1.314	0.075054	10	Polr2g	1.304	0.075054	9	Mirps35	1.293	0.075054	7
Ss18	1.314	0.075054	10	Oaz2	1.304	0.075054	14	Gsta4	1.292	0.075054	18
Sumo2	1.314	0.075054	22	Pycrl	1.304	0.075054	8	Ap1s1	1.292	0.075054	13
Hdgf	1.314	0.075054	41	Dpy30	1.304	0.075054	19	Zdhhc4	1.292	0.075054	8
Ik	1.314	0.075054	16	Parp1	1.303	0.075054	16	Phf10	1.292	0.075054	9
Farsa	1.314	0.075054	10	Ugp2	1.303	0.075054	12	Tfpi	1.292	0.075054	21

图26-8

Aimp2	1.314	0.075054	12	Wdr5	1.303	0.075054	17	Fdx1	1.292	0.075054	10
Pgls	1.314	0.075054	10	Pold2	1.303	0.075054	11	Ccng1	1.292	0.075054	12
B9d1	1.314	0.075054	9	Ergic3	1.303	0.075054	10	Rab2a	1.291	0.075054	11
Chmp6	1.314	0.075054	8	Rhoc	1.303	0.075054	12	Bax	1.291	0.075054	27
Aup1	1.314	0.075054	11	Bub1b	1.303	0.075054	14	Ecsit	1.291	0.075054	10
Aldh2	1.314	0.075054	14	Pbk	1.303	0.075054	20	Mrps30	1.290	0.075054	10
Dnajc9	1.314	0.075054	15	Phgdh	1.302	0.075054	11	Etv5	1.290	0.075054	15
Tmed10	1.313	0.075054	13	Cnn2	1.302	0.075054	12	Ino80e	1.290	0.075054	12
Sap30	1.313	0.075054	10	Slc50a1	1.302	0.075054	13	2810008D09Rik	1.290	0.075054	5
Maf1	1.313	0.075054	11	Ak4	1.302	0.075054	15	Ubxn1	1.290	0.075054	21
Drg1	1.313	0.075054	17	Ndufb4	1.302	0.075054	10	Hspa9	1.290	0.075054	24
Snd1	1.312	0.075054	18	Ptrf	1.302	0.075054	10	Nudt1	1.289	0.075054	12
Acaa1a	1.312	0.075054	12	Scyl1	1.302	0.075054	12	2700094K13Rik	1.289	0.075054	18
Polr2j	1.312	0.075054	16	Cyth2	1.302	0.075054	10	Clstn1	1.289	0.075054	10
Pes1	1.312	0.075054	17	Xrcc5	1.301	0.075054	14	Psmc12	1.289	0.075054	14
N4bp3	1.312	0.075054	9	Cox6a1	1.301	0.075054	26	Cmtm7	1.289	0.075054	10
Gars	1.312	0.075054	17	Zdhc18	1.301	0.075054	14	Kat2a	1.289	0.075054	15
Ccz1	1.312	0.075054	9	Bloc1s2	1.301	0.075054	8	Phkg1	1.289	0.075054	20
Tuba1c	1.312	0.075054	37	Mast2	1.301	0.075054	10	Arhgap8	1.288	0.075054	12
Nfkbia	1.312	0.075054	9	1110038B12Rik	1.301	0.075054	18	Atp5k	1.288	0.075054	31
Nono	1.312	0.075054	29	Aspsr1	1.300	0.075054	15	Tdrp	1.288	0.075054	12
Mtch1	1.312	0.075054	21	Cops3	1.300	0.075054	13	Nutrf2-ps1	1.288	0.075054	16
Ndufb8	1.312	0.075054	16	Nenf	1.300	0.075054	11	Mrpl47	1.288	0.075054	10
Sec61b	1.311	0.075054	17	Ndufs6	1.299	0.075054	18	Pdia3	1.288	0.075054	13
Ppm1g	1.311	0.075054	18	1190005I06Rik	1.299	0.075054	6	Iscu	1.288	0.075054	12
Mrps2	1.311	0.075054	11	Hsbp1	1.299	0.075054	12	Jmjd6	1.288	0.075054	10
Gnai2	1.311	0.075054	14	Tmed9	1.299	0.075054	10	Ndufs1	1.288	0.075054	17
Zw10	1.311	0.075054	9	Gm3258	1.299	0.075054	10	Kif22	1.287	0.075054	12
Sfxn1	1.311	0.075054	12	Rpl36al	1.299	0.075054	43	Scamp4	1.287	0.075054	11
Atp9a	1.310	0.075054	15	Jam2	1.299	0.075054	21	Drosha	1.287	0.075054	12
Ccdc93	1.309	0.075054	22	Eci2	1.299	0.075054	10	Snrnp25	1.287	0.075054	9
2010107E04Rik	1.309	0.075054	18	Gm16062	1.299	0.075054	14	Rrp15	1.287	0.075054	12

图26-9

Rpp25	1.309	0.075054	11	Bola1	1.299	0.075054	8	Ruvbl1	1.287	0.075054	16
Cd99	1.309	0.075054	10	Mirpl20	1.298	0.075054	15	Cfta	1.287	0.075054	19
Pfn1	1.309	0.075054	10	Ubezq1	1.298	0.075054	9	Pdlim1	1.287	0.075054	11
Dcaf15	1.309	0.075054	11	Csrp2	1.298	0.075054	15	Eras	1.287	0.075054	9
Dus2	1.309	0.075054	18	Map1lc3b	1.298	0.075054	16	Prdx4	1.287	0.075054	11
Ezh2	1.308	0.075054	13	Hnrnp1	1.298	0.075054	17	Rab1b	1.287	0.075054	13
Cnn3	1.308	0.075054	22	Ccdc12	1.298	0.075054	8	Idh3g	1.287	0.075054	10
Sumo1	1.308	0.075054	15	Ddx1	1.298	0.075054	13	Rcor2	1.287	0.075054	14
Lsm10	1.308	0.075054	9	Sar1b	1.298	0.075054	8	1110008P14Rik	1.287	0.075054	10
Pa2g4	1.308	0.075054	30	Elp3	1.298	0.075054	17	Csnk1e	1.287	0.075054	24
Tomm20	1.308	0.075054	41	Mafg	1.297	0.075054	12	Grccl0	1.287	0.075054	10
Tmem147	1.308	0.075054	9	Irf3	1.297	0.075054	22	Hras	1.287	0.075054	15
Nudt21	1.287	0.075054	11	Ier3ip1	1.278	0.075054	17	Cap1	1.268	0.075698	14
Gjb3	1.286	0.075054	8	Stau1	1.278	0.075054	21	Dcald	1.268	0.075698	10
Acaa2	1.286	0.075054	11	Tmem208	1.278	0.075054	9	Pabpc4	1.268	0.075698	16
Ppp2r2d	1.286	0.075054	13	Arrb2	1.278	0.075054	16	Gak	1.268	0.075698	13
Nthl1	1.286	0.075054	9	Aldh7a1	1.277	0.07553	10	Mpdu1	1.267	0.075698	15
Chmp2a	1.286	0.075054	11	Elf6	1.277	0.07553	15	Pomp	1.267	0.075698	20
Gatsl3	1.286	0.075054	8	Vmp1	1.277	0.07553	10	Idh1	1.267	0.075698	13
Erp29	1.285	0.075054	14	Trpt1	1.277	0.07553	11	Poldip2	1.267	0.075698	12
Hsd17b10	1.285	0.075054	10	Ddx21	1.277	0.075698	17	Rpusd1	1.267	0.075698	10
Plaur	1.285	0.075054	10	Rps19bp1	1.277	0.075698	12	Agl	1.267	0.075698	43
Capza1	1.285	0.075054	15	Gdl2	1.277	0.075698	18	Sept1	1.267	0.075698	8
Mtch2	1.285	0.075054	10	Sephs2	1.276	0.075698	15	Lsm7	1.267	0.075698	13
Bcap31	1.284	0.075054	12	Gbx2	1.276	0.075698	7	Nubp2	1.267	0.075698	8
Eed	1.284	0.075054	15	Sdha	1.276	0.075698	15	Atp5l	1.267	0.075698	20
Ccnb2	1.284	0.075054	11	Nop2	1.276	0.075698	12	Fxr1	1.267	0.075698	11
Ddit4	1.284	0.075054	10	H2afy	1.275	0.075698	15	Pam16	1.267	0.075698	15
Vps4b	1.284	0.075054	16	Pld3	1.275	0.075698	13	Epcam	1.267	0.075698	13
Apex1	1.284	0.075054	11	Gss	1.275	0.075698	11	Edf1	1.267	0.075698	19
Thap4	1.284	0.075054	11	Map7d2	1.275	0.075698	16	Parl	1.267	0.075698	10
Imp4	1.284	0.075054	12	Cox7a2	1.275	0.075698	21	Rpn1	1.266	0.075698	14

图26-10

Atp5g3	1.284	0.075054	29	Med24	1.275	0.075698	11	Gcat	1.266	0.075698	13
3110001122Rik	1.284	0.075054	17	Ccm4l	1.275	0.075698	10	Slc9a3r1	1.266	0.075698	11
Herpud1	1.283	0.075054	8	Stmn1-rs1	1.275	0.075698	12	Sf3b3	1.266	0.075698	18
Cirbp	1.283	0.075054	9	Mgea5	1.274	0.075698	15	Gon4l	1.266	0.075698	14
Rrm2	1.283	0.075054	14	Bnip3l	1.274	0.075698	16	Uba2	1.266	0.075698	14
Pef1	1.283	0.075054	11	Hyou1	1.274	0.075698	13	Stoml1	1.266	0.075698	12
Ctps	1.283	0.075054	13	Chrac1	1.274	0.075698	8	Sod2	1.266	0.075698	11
Ly6e	1.283	0.075054	13	Apol7a	1.274	0.075698	13	Rangrf	1.265	0.075698	13
Gldc	1.283	0.075054	11	Trp53	1.274	0.075698	18	Slc20a2	1.265	0.075698	13
Gsr	1.283	0.075054	12	Adck5	1.274	0.075698	8	Usmg5	1.265	0.075698	20
Spccs2	1.283	0.075054	15	Nid2	1.274	0.075698	14	Gm56l	1.265	0.075698	9
Elf3b	1.283	0.075054	34	Zfp451	1.273	0.075698	14	Polr1d	1.265	0.075698	19
Usp10	1.283	0.075054	13	Tnfr1	1.273	0.075698	10	Uhrf1	1.265	0.075698	15
Rpa2	1.283	0.075054	19	Mgat4b	1.273	0.075698	10	Atp6v1g1	1.265	0.075698	13
Mrpl55	1.283	0.075054	13	Snrpd2	1.273	0.075698	30	Rps27l	1.265	0.075698	29
Mrps15	1.282	0.075054	13	Timm13	1.273	0.075698	18	Igf2	1.265	0.075698	12
Rab35	1.282	0.075054	9	Fbxo15	1.273	0.075698	12	Phb2	1.265	0.075698	19
Ubc	1.282	0.075054	16	Syngn2	1.273	0.075698	10	Arcpc2	1.265	0.075698	14
0610007P14Rik	1.282	0.075054	8	Cox7b	1.273	0.075698	24	Plvap	1.265	0.075698	11
Cpsf3	1.282	0.075054	11	Atp6v1e1	1.273	0.075698	10	Cd81	1.265	0.075698	15
Tceb1	1.282	0.075054	13	Card10	1.273	0.075698	8	Mapkapk2	1.264	0.075698	11
Wnt7b	1.282	0.075054	10	Mpzl1	1.273	0.075698	11	Nme4	1.264	0.075698	8
G3bp2	1.282	0.075054	24	Slc29a4	1.272	0.075698	7	Lgals3	1.264	0.075698	11
Ubqln4	1.282	0.075054	13	Mfsd11	1.272	0.075698	22	Lrrc14	1.264	0.075698	11
Unc45a	1.281	0.075054	9	Mrps9	1.272	0.075698	12	Ccdc3	1.264	0.075698	9
Wdr61	1.281	0.075054	9	Lpo4	1.272	0.075698	11	Bet1l	1.264	0.075698	9
Zfp11	1.281	0.075054	10	6430706D22Rik	1.272	0.075698	11	Hist1h1b	1.264	0.075698	8
Nup62	1.281	0.075054	17	Ppp1r14b	1.271	0.075698	14	Hnrnp3	1.264	0.075698	11
Cox6c	1.281	0.075054	16	Msh2	1.271	0.075698	20	Tbc1d13	1.264	0.075698	15
Nars	1.281	0.075054	12	Acta1	1.271	0.075698	14	Lamtor2	1.264	0.075698	15
Cipc	1.281	0.075054	13	Sec23b	1.271	0.075698	10	Aph1a	1.263	0.075698	14
Psmd13	1.281	0.075054	22	Ninj1	1.271	0.075698	8	Nme1	1.263	0.075698	16

图26-11

Slc25a3	1.281	0.075054	28	Gps1	1.270	0.075698	13	Hnrnpdl	1.263	0.075698	16
Sin3b	1.281	0.075054	12	Cryz	1.270	0.075698	10	Dynlrb1	1.263	0.075698	10
Pde6d	1.281	0.075054	12	Elf3m	1.270	0.075698	14	Prmt7	1.263	0.075698	13
Ndufaf2	1.281	0.075054	11	F3	1.270	0.075698	16	Pwp2	1.263	0.075698	12
Klf9	1.280	0.075054	17	Mrpl2	1.270	0.075698	9	Lamp1	1.262	0.075698	24
Znrd1	1.280	0.075054	12	Lta4h	1.270	0.075698	16	Ndufa9	1.262	0.075698	10
Cox6b1	1.280	0.075054	13	Cnpy2	1.270	0.075698	12	Zfp511	1.262	0.075698	10
Ece2	1.280	0.075054	11	Nsmce1	1.270	0.075698	11	Rexo2	1.262	0.075698	16
Sh3bgrl3	1.280	0.075054	12	Rfc5	1.270	0.075698	14	Nol7	1.262	0.075698	13
Icam1	1.280	0.075054	9	Tomm22	1.270	0.075698	22	Dlst	1.262	0.075698	18
Mageb16	1.280	0.075054	11	Ssrp1	1.269	0.075698	20	Atox1	1.261	0.076367	20
Myo9b	1.279	0.075054	12	Psmc4	1.269	0.075698	21	Ccm2	1.261	0.076367	11
Sra1	1.279	0.075054	10	Timm50	1.269	0.075698	7	Poli	1.261	0.076367	13
Ddx49	1.279	0.075054	12	Elov6	1.269	0.075698	15	Gtf2h4	1.261	0.076367	12
Ptms	1.279	0.075054	15	Snhg1	1.269	0.075698	13	Copz1	1.260	0.076367	16
Psmc4	1.279	0.075054	18	Polr1c	1.269	0.075698	12	Ift43	1.260	0.076367	10
Bsg	1.279	0.075054	20	Zmynd19	1.269	0.075698	11	Fsd1	1.260	0.076367	8
Ak1	1.279	0.075054	16	Lphn1	1.269	0.075698	21	Hdac1	1.260	0.076367	14
Dohh	1.279	0.075054	12	Atp1b1	1.269	0.075698	12	Prss50	1.260	0.076367	14
Wdr75	1.279	0.075054	11	Ssbp1	1.268	0.075698	18	Arl6ip5	1.260	0.076367	10
Snai3	1.279	0.075054	8	Dusp12	1.268	0.075698	11	Cyb5r3	1.260	0.076367	14
Aen	1.279	0.075054	13	Spr	1.268	0.075698	13	Myl6	1.260	0.076367	32
Tbc1d10b	1.279	0.075054	12	Cir1	1.268	0.075698	15	Hs2st1	1.259	0.077442	19
Riok2	1.278	0.075054	12	Ywhah	1.268	0.075698	13	Nup85	1.259	0.077442	15
Tfdp1	1.259	0.077442	10	Noc4l	1.252	0.077442	13	Rapgef1	1.245	0.079253	20
Rheb1	1.259	0.077442	13	Zswim7	1.252	0.077442	8	3110040N11Rik	1.245	0.079253	8
Ndufb7	1.259	0.077442	14	Coro1b	1.252	0.077442	11	Akt1	1.245	0.079253	11
Ltv1	1.259	0.077442	13	Ier2	1.252	0.077442	9	Hirip3	1.245	0.079508	11
Mtfr1l	1.259	0.077442	10	Timm10b	1.251	0.077442	11	Samd1	1.245	0.079508	11
Tram1	1.258	0.077442	9	Arhgap21	1.251	0.077442	12	AKO10878	1.245	0.079508	21
Ssr4	1.258	0.077442	9	Pxn	1.251	0.077442	10	Mrpl16	1.245	0.079508	9
Arl3	1.258	0.077442	8	Mrpl52	1.251	0.077442	9	Rgl2	1.245	0.079508	8

图26-12

Snrpc	1.258	0.077442	16	Dynl12	1.251	0.077442	14	Pasma4	1.245	0.079508	26
Park7	1.258	0.077442	27	Fmr1nb	1.251	0.077442	11	Ogfod3	1.244	0.079508	9
Elf4ebp3	1.258	0.077442	8	Bag1	1.251	0.077442	13	Mtx1	1.244	0.080016	12
Drap1	1.258	0.077442	19	Tmem11	1.251	0.077442	9	Ahcy	1.244	0.080453	17
Rab33b	1.258	0.077442	24	Cdt1	1.251	0.077442	15	Nelfe	1.244	0.080453	11
Flot1	1.258	0.077442	9	Me2	1.251	0.077442	11	Tmem60	1.244	0.080453	10
Fntb	1.258	0.077442	11	Coa3	1.251	0.077442	8	Phax	1.244	0.080453	15
Wdr45b	1.257	0.077442	11	Nipsnap3b	1.251	0.077442	9	Rab14	1.244	0.080453	12
Odf2	1.257	0.077442	11	Bag6	1.250	0.077442	12	Fam89b	1.244	0.080453	9
Slc10a1	1.257	0.077442	17	Mcm5	1.250	0.077442	16	2410015M20Rik	1.244	0.080453	13
Cenpt	1.257	0.077442	11	Zfp13	1.250	0.077442	15	Acadvl	1.244	0.080453	10
Nrde2	1.257	0.077442	14	Rad21	1.250	0.077442	15	Pmf1	1.243	0.080453	10
Ndufab1	1.257	0.077442	12	Laptm4a	1.250	0.077442	14	Psmb2	1.243	0.080453	15
Gcsh	1.257	0.077442	10	Hat1	1.250	0.077442	24	Farsb	1.243	0.080453	16
Msantd2	1.257	0.077442	7	Fbl	1.250	0.077442	21	Ube2a	1.243	0.080453	11
Haus4	1.257	0.077442	9	Ppil3	1.250	0.077442	12	Pced11b	1.243	0.080453	12
Itm2b	1.257	0.077442	14	Mest	1.250	0.077442	13	Hspb8	1.243	0.080453	7
Ict1	1.256	0.077442	9	Dtymk	1.250	0.077442	16	Ywhae	1.243	0.080453	26
Ndufa3	1.256	0.077442	11	Slbp	1.249	0.07782	11	BC031181	1.243	0.080453	10
Dnajc8	1.256	0.077442	13	Llph	1.249	0.07782	20	Zcchc17	1.243	0.080453	8
Akap8l	1.256	0.077442	12	Perp	1.249	0.07782	9	Fam162a	1.242	0.080453	9
Dnajc19	1.256	0.077442	18	Pagr1a	1.249	0.07782	11	Otud6b	1.242	0.080453	14
1810022K09Rik	1.256	0.077442	10	Erc3	1.249	0.07782	12	Sap30l	1.242	0.080453	9
Smim11	1.256	0.077442	8	1110004F10Rik	1.249	0.078195	15	Nop56	1.242	0.080453	29
Dhx30	1.256	0.077442	10	Rp9	1.249	0.078195	10	Anxa7	1.242	0.080515	18
Ccnd1	1.256	0.077442	9	Pop4	1.249	0.078195	9	Rbm12b2	1.242	0.080515	9
Mgl2	1.256	0.077442	23	Arg2	1.249	0.078195	11	Hiat1	1.242	0.080515	9
Atp6v1f	1.256	0.077442	12	Elf4b	1.249	0.078195	24	Dclk2	1.241	0.080515	12
Rpl30	1.255	0.077442	49	Pitrm1	1.249	0.078255	13	Akr1e1	1.241	0.080515	10
Ubap2	1.255	0.077442	12	Jagn1	1.249	0.078255	9	Blvra	1.241	0.080515	6
Rhoa	1.255	0.077442	26	Utp11l	1.249	0.078255	8	Pgam1	1.241	0.080515	25
Map2k2	1.255	0.077442	12	Snupn	1.249	0.078255	10	Irf2bpl	1.241	0.080515	16

图26-13

Mrps34	1.255	0.077442	6	Drg2	1.249	0.078255	10	Wbp1	1.241	0.080515	10
Zdhc12	1.255	0.077442	10	Pecam1	1.248	0.078255	8	Serf2	1.241	0.080515	24
Wbp11	1.255	0.077442	16	Psat1	1.248	0.078255	16	2810004N23Rik	1.241	0.080515	10
Birc5	1.255	0.077442	18	Nucks1	1.248	0.078255	27	Msh6	1.241	0.080515	13
Shmt1	1.255	0.077442	15	Evl	1.248	0.078255	8	Bap1	1.241	0.080515	11
Cst3	1.255	0.077442	14	Uhrf2	1.248	0.078255	9	Morc1	1.240	0.080515	12
Maea	1.254	0.077442	15	Ap2m1	1.248	0.078686	17	Ccdc163	1.240	0.080515	9
Gadd45gip1	1.254	0.077442	13	Tra2a	1.248	0.078686	11	Tmem97	1.240	0.080515	9
Ppilb	1.254	0.077442	16	Tfcp2l1	1.248	0.078686	22	Trappc1	1.240	0.080515	9
Gpr108	1.254	0.077442	8	Tex264	1.248	0.078686	9	Bcat2	1.240	0.080515	8
Chchd6	1.254	0.077442	7	Csnk1g2	1.248	0.078686	10	Cenpm	1.240	0.080515	9
Elf2d	1.254	0.077442	13	Myeov2	1.247	0.078686	16	Cotl1	1.240	0.080515	9
Zmiz2	1.254	0.077442	16	Ints9	1.247	0.078686	14	Mpv17l2	1.240	0.080515	7
Clpp	1.254	0.077442	22	Myadm	1.247	0.078686	16	Nob1	1.240	0.080515	13
Skiv2l2	1.253	0.077442	10	Tns3	1.247	0.078686	12	1810037l17Rik	1.240	0.080515	9
Anp32e	1.253	0.077442	15	Tecr	1.247	0.078686	11	Pno1	1.240	0.080515	9
Wdr46	1.253	0.077442	17	Pycr2	1.247	0.078686	9	Klf4	1.240	0.080515	15
Rpn2	1.253	0.077442	12	Arcp5l	1.247	0.078686	10	Rrp9	1.240	0.080515	14
Bcam	1.253	0.077442	8	Cox5b	1.247	0.078686	26	Ostc	1.239	0.080515	10
Abce1	1.253	0.077442	13	Adap1	1.247	0.078686	11	Nit2	1.239	0.080515	12
Slc25a12	1.253	0.077442	14	Lypla2	1.247	0.078686	10	Ftsj2	1.239	0.080515	12
Anxa5	1.253	0.077442	14	Hk2	1.247	0.078686	17	Uchl3	1.239	0.080515	9
Zc3h10	1.253	0.077442	11	Mirpl40	1.246	0.078686	11	Yeats4	1.239	0.080515	12
Dnttip2	1.253	0.077442	16	Msmo1	1.246	0.078686	6	Dppa3	1.239	0.080515	15
Plscr1	1.252	0.077442	10	Ndufa8	1.246	0.078686	10	1110001J03Rik	1.239	0.080515	7
Cdk2ap2	1.252	0.077442	9	Nifk	1.246	0.078686	15	2310009B15Rik	1.239	0.080515	7
Gm12669	1.252	0.077442	12	St6galnac2	1.246	0.078686	11	Atp5c1	1.239	0.080515	22
Smpd1	1.252	0.077442	8	Nradd	1.246	0.078686	13	Ndufb9	1.239	0.080515	16
Mff	1.252	0.077442	11	Mcts2	1.246	0.078792	8	Pet112	1.239	0.080515	11
Mogs	1.252	0.077442	7	Cramp1l	1.246	0.078792	19	Egfl7	1.238	0.080515	14
Atp1a1	1.252	0.077442	15	Nacc1	1.246	0.078792	20	Taf9	1.238	0.080515	8
Rsrp1	1.252	0.077442	9	Magoh	1.246	0.078792	13	2610507B11Rik	1.238	0.080515	16
Atp1f1	1.252	0.077442	13	Tspo	1.246	0.078792	11	Arl2	1.238	0.080515	11
Srp9	1.252	0.077442	14	Prss43	1.246	0.078792	26	Rabggfb	1.238	0.080515	14
Wbp11	1.252	0.077442	16	Ing1	1.245	0.078792	11	Mat2a	1.238	0.080515	19
Rnf220	1.252	0.077442	11	Sept11	1.245	0.078792	13	Avpi1	1.238	0.080515	10

图26-14

Dctpp1	1.238	0.080515	20	Cdk4	1.231	0.08545	13	Mt3	1.224	0.089235	10
Dctn3	1.238	0.080999	12	Mtg2	1.231	0.08545	7	Chmp1a	1.224	0.089235	11
Mrps18a	1.238	0.081183	9	Inpp5d	1.231	0.08545	11	Msc	1.224	0.089235	10
Slc35a4	1.238	0.081183	11	Cct2	1.231	0.08545	27	Mcm4	1.224	0.089235	17
Metap1	1.238	0.081183	11	Actr3	1.231	0.08545	17	Ube2d3	1.224	0.089235	17
Mrps10	1.237	0.081183	11	Top3b	1.230	0.08545	10	Nhp2l1	1.224	0.089235	19
Lmb1	1.237	0.081183	20	Ube2g2	1.230	0.08545	9	Ranbp1	1.224	0.089235	24
Rnf167	1.237	0.081183	10	Ergic1	1.230	0.08545	12	Ifrd1	1.224	0.089235	7
Ywhag	1.237	0.081515	23	Ndufa11	1.230	0.08545	18	Ndufa2	1.224	0.089235	21
Zfp706	1.237	0.081515	24	Rer1	1.230	0.08545	10	Pou5f1	1.224	0.089235	55
Plod1	1.237	0.081515	7	Aco2	1.230	0.085902	14	Timm17a	1.224	0.089235	10
Gtf2f2	1.237	0.081515	12	Pspc1	1.230	0.087351	9	Cks1b	1.224	0.089235	16
Farp1	1.237	0.081646	10	Gltscr2	1.230	0.087351	15	Gmppa	1.224	0.089235	9
D8Ertd738e	1.237	0.081646	11	1700037H04Rik	1.230	0.087351	14	Fbxo45	1.224	0.089235	9
Ndufa12	1.237	0.081646	13	Mphosph8	1.229	0.087439	11	Dppa2	1.224	0.089235	12
Jarid2	1.237	0.081646	20	Urah	1.229	0.087439	10	Orail	1.224	0.089235	7
Ntmt1	1.237	0.081646	9	Stk11	1.229	0.087439	10	Prr5	1.224	0.089235	7
Naca	1.237	0.081646	38	Rpl7l1	1.229	0.087439	48	Nrbp1	1.224	0.089235	10
Aldoa	1.237	0.081646	19	Aqp3	1.229	0.087439	6	Uqcrrf1	1.223	0.089235	10
Zswim1	1.236	0.081646	9	Ascc1	1.229	0.087439	10	Gm10094	1.223	0.089235	10
Pgs1	1.236	0.081646	10	Polr2f	1.229	0.087439	16	Josd2	1.223	0.089235	8
Phldb1	1.236	0.081973	11	Kif11	1.229	0.087439	16	Ndufb3	1.223	0.089235	8
Cers5	1.236	0.081973	8	Mapre1	1.229	0.087439	24	Coq2	1.223	0.089235	9
Cul7	1.236	0.081973	11	Nr4a1	1.229	0.087439	9	Upf3a	1.223	0.089235	8
Tspan3	1.236	0.081973	9	Ints5	1.228	0.087439	11	Pxk	1.223	0.089235	13
F2rl1	1.236	0.081973	12	Usp1	1.228	0.087439	11	Pgd	1.222	0.089235	17
Lsm3	1.236	0.082051	15	Usp22	1.228	0.087439	10	Yars	1.222	0.089235	12
Elp2	1.236	0.082051	10	Rad23b	1.228	0.087439	17	Cdc45	1.222	0.089235	9
Ift20	1.236	0.082051	12	Elf3l	1.228	0.087439	10	Dnajc11	1.222	0.089235	11
Rev1	1.236	0.082051	13	Dut	1.228	0.087439	15	Gtf2e2	1.222	0.089235	11
Tbcb	1.236	0.082051	10	Cep89	1.228	0.087439	8	Tax1bp3	1.222	0.089235	8
Ankrd54	1.236	0.0820515	11	Tmbim6	1.228	0.087439	17	Tarbp2	1.222	0.089235	14

图16-15

Nkap1l	1.236	0.082051	11	Sugt1	1.228	0.088122	11	Cops5	1.222	0.089235	16
Kcmf1	1.236	0.082051	17	Fxr2	1.228	0.088122	15	Lsm6	1.222	0.089235	12
Mkm1	1.236	0.082374	25	Zglp1	1.228	0.088122	8	Rce1	1.222	0.089235	8
Arhgdia	1.236	0.082374	19	Slc35b1	1.228	0.088122	10	Gadd45b	1.222	0.089235	8
Gldnos	1.236	0.082374	19	H2afv	1.228	0.088122	17	Unc50	1.222	0.089235	9
Tbrg4	1.235	0.082374	18	Rhot2	1.228	0.088122	8	Akap8	1.222	0.089235	10
Tjap1	1.235	0.082374	7	Elf1	1.227	0.088122	35	Bckdha	1.221	0.089235	8
Chmp3	1.235	0.082426	10	Kdelr1	1.227	0.088486	12	Iscal	1.221	0.089235	9
Dtnbp1	1.235	0.082426	5	Unc119	1.227	0.088486	8	Derl1	1.221	0.089235	10
Nsf	1.235	0.082426	15	Lmb2	1.227	0.088486	10	Nusap1	1.221	0.089235	10
Grtp1	1.235	0.082426	8	Mphosph6	1.227	0.088486	11	Abtb1	1.221	0.089235	8
Supt5	1.235	0.082426	17	Hmgn2	1.227	0.088486	15	Evi5	1.221	0.089235	24
Morf4l1	1.235	0.082426	27	Hsdl1	1.227	0.088486	12	Lpar6	1.221	0.089235	9
Tcf3	1.235	0.082426	11	Dnmt1	1.227	0.088486	14	Rgs19	1.221	0.089235	12
Enoph1	1.235	0.082426	9	Rdh14	1.227	0.088486	13	Eprs	1.221	0.089235	16
Pole4	1.235	0.082426	14	Isc2	1.226	0.088486	7	Oser1	1.221	0.089235	8
Prelid1	1.234	0.082426	23	Tmem126a	1.226	0.088486	6	3110062M04Rik	1.221	0.089235	8
Atxn10	1.234	0.082426	14	Dhrs4	1.226	0.088486	11	2700060E02Rik	1.221	0.089235	20
Ppp2r2a	1.234	0.082426	18	Mak16	1.226	0.088486	10	Tagap1	1.221	0.089235	7
Slc39a7	1.234	0.082426	11	Ddx24	1.226	0.088486	11	Tmem192	1.221	0.089235	7
U2af2	1.234	0.082426	15	Pvrl2	1.226	0.088486	11	Trappc2l	1.220	0.089235	8
BC030867	1.234	0.082426	10	Man2c1	1.226	0.088486	8	Tbx3	1.220	0.089235	18
Ccndbp1	1.234	0.082426	7	Crelid1	1.226	0.088486	9	Adipor2	1.220	0.089235	16
Bod1	1.234	0.082426	9	Lgsn	1.226	0.088486	15	Dpm3	1.220	0.089235	10
Dctd	1.233	0.082426	9	Uck2	1.226	0.088486	8	Ddx41	1.220	0.089235	8
Tpm1	1.233	0.082426	15	Ldlrap1	1.226	0.088486	7	Tmem69	1.220	0.089235	12
Chchd1	1.233	0.082426	10	Ndufv3	1.226	0.088486	10	Hadh	1.220	0.089235	9
Preb	1.233	0.082938	10	Banf1	1.226	0.088486	18	Cstf2t	1.220	0.089235	21
Taldo1	1.233	0.083351	14	Phb	1.225	0.088486	18	Vps36	1.220	0.089235	7
Cox5a	1.233	0.083351	27	Prkar1a	1.225	0.088919	14	Acads	1.220	0.089235	10
Myo1c	1.233	0.083351	17	Timm8a1	1.225	0.088919	10	Cacybp	1.220	0.089235	17
Hs6st1	1.233	0.084271	14	Rps6ka1	1.225	0.089235	14	Tjp2	1.220	0.089235	13

图26-16

Polr2d	1.233	0.084271	9	Trmt6	1.225	0.089235	12	Mad2l2	1.220	0.089235	10
Mrpl33	1.232	0.084271	10	Dmwd	1.225	0.089235	9	Pdk1	1.220	0.089235	16
Gadd45a	1.232	0.084271	9	Ndufaf3	1.225	0.089235	7	Nop14	1.220	0.089235	13
Fdps	1.232	0.084581	12	Rassf7	1.225	0.089235	9	Asna1	1.220	0.089235	13
Orc2	1.232	0.084581	17	Gtf2i	1.225	0.089235	11	Taf13	1.220	0.089235	8
Fam207a	1.232	0.084581	14	Baz2b	1.225	0.089235	17	Ube2e3	1.219	0.089235	11
Smg7	1.232	0.084581	10	Scgb3a1	1.225	0.089235	8	Rnf25	1.219	0.089235	9
Haus1	1.231	0.08545	9	Egln2	1.225	0.089235	10	Angptl4	1.219	0.089235	7
Snrpg	1.231	0.08545	38	Naa38	1.225	0.089235	17	Csnkla1	1.219	0.089235	10
Ddx39	1.231	0.08545	15	Grb7	1.224	0.089235	10	Gns	1.219	0.089235	8
Glr3	1.231	0.08545	15	Ndufs8	1.224	0.089235	9	Snord64	1.219	0.089235	4
Tsen15	1.231	0.08545	8	Wdr83	1.224	0.089235	9	Pdcd5	1.219	0.089235	14
Stt3a	1.219	0.089235	12	Etf1	1.213	0.089235	16	Sqle	1.208	0.092108	8
Rab3gap1	1.219	0.089235	20	Dctn1	1.213	0.089235	12	Rab3ip	1.208	0.092108	10
Dync1i2	1.219	0.089235	11	Arlh2	1.213	0.089235	14	Slc25a28	1.207	0.092108	9
Champ1	1.219	0.089235	10	Nck1	1.213	0.089235	7	Srf	1.207	0.092108	9
Exosc5	1.219	0.089235	11	Bst2	1.213	0.089235	8	Pisd-ps3	1.207	0.092108	6
Ascc2	1.219	0.089235	8	Ppp1r10	1.213	0.089235	9	E430025E21Rik	1.207	0.092108	12
Cnot10	1.219	0.089235	17	Smyd2	1.213	0.089235	10	Mrpl15	1.207	0.092108	21
Mrps17	1.219	0.089235	8	Cpsf3l	1.212	0.089235	13	Echs1	1.207	0.092108	8
Tma16	1.218	0.089235	10	Smad4	1.212	0.089235	11	Fmnl3	1.207	0.092108	8
Gprc5a	1.218	0.089235	11	Klc3	1.212	0.089235	9	Mbtd1	1.207	0.092108	11
Dok2	1.218	0.089235	9	493044M15Rik	1.212	0.089235	9	Atp5d	1.207	0.092108	21
Snrpa	1.218	0.089235	18	Ccnt1	1.212	0.089235	7	Crls1	1.207	0.092108	11
Rnf115	1.218	0.089235	10	Atrx	1.212	0.089235	17	Exosc6	1.207	0.092108	10
Eml2	1.218	0.089235	9	Rbbp7	1.212	0.089235	12	Pcbp4	1.206	0.092108	8
Cox6b2	1.218	0.089235	8	Mbnl2	1.212	0.089235	12	Psme4	1.206	0.092108	20
Prdm15	1.218	0.089235	14	Phlda1	1.212	0.089235	11	Syf2	1.206	0.092108	11
Itga3	1.217	0.089235	11	Rbm17	1.212	0.089235	12	Nkain1	1.206	0.092108	11
Rnf126	1.217	0.089235	13	Tox4	1.212	0.089235	16	Psmd3	1.206	0.092108	19
Idh3b	1.217	0.089235	10	Srp72	1.212	0.089235	11	Srp68	1.206	0.092108	12
Cisd1	1.217	0.089235	10	Uqcrc	1.212	0.089235		Mrps21	1.206	0.092108	10

图26-17

Fam107b	1.217	0.089235	10	Mreg	1.212	0.089235	11	Lasp1	1.206	0.092108	15
Elf4e	1.217	0.089235	12	Epm2aip1	1.211	0.089235	22	Otx2	1.206	0.092108	8
Aars	1.217	0.089235	12	Zscan10	1.211	0.089235	10	Amn	1.206	0.092108	9
Nsf1c	1.217	0.089235	11	Lmo4	1.211	0.089235	15	2510039O18Rik	1.206	0.092108	9
Nup43	1.217	0.089235	10	Cenph	1.211	0.089235	9	Dbf4	1.206	0.092108	13
Mrps16	1.217	0.089235	15	Fgfr1	1.211	0.089235	11	Triml2	1.206	0.092108	8
Psmg4	1.216	0.089235	8	Apmap	1.211	0.089235	10	Hagh	1.206	0.092108	10
Zmat2	1.216	0.089235	13	E2f6	1.211	0.089235	10	Pigp	1.205	0.092108	27
Fxyd6	1.216	0.089235	11	Thap7	1.211	0.089235	10	LOC101056043	1.205	0.092108	10
Mcm3	1.216	0.089235	21	Itga6	1.211	0.089235	13	Glo1	1.205	0.092108	16
Rtn4	1.216	0.089235	14	Tppp3	1.211	0.089235	8	Tnfrsf12a	1.205	0.092384	8
Frat2	1.216	0.089235	10	Ube2m	1.211	0.089235	18	Taf12	1.205	0.092384	8
Adss	1.216	0.089235	14	Gm13051	1.211	0.089352	13	Tssc4	1.205	0.092384	10
Gemin2	1.216	0.089235	9	Nsun2	1.211	0.089352	15	Ptdss2	1.205	0.092384	9
Cct3	1.216	0.089235	29	Snora73b	1.211	0.089352	3	Clapin1	1.205	0.092583	21
Copa	1.216	0.089235	15	Gm5801	1.211	0.089352	13	Hint2	1.204	0.092583	8
Shfm1	1.216	0.089235	20	Tubb4b	1.210	0.089352	42	Znhit3	1.204	0.092583	11
Lamtor5	1.215	0.089235	10	Slc25a33	1.210	0.089352	9	Ola1	1.204	0.092583	14
Rcn1	1.215	0.089235	12	Strip1	1.210	0.089352	10	Lamtor3	1.204	0.092583	8
Rras2	1.215	0.089235	9	Chd4	1.210	0.090159	26	Cct7	1.204	0.092583	31
Khdrbs3	1.215	0.089235	9	Tuba1a	1.210	0.090584	43	Faf1	1.204	0.092583	11
Rab1	1.215	0.089235	12	Taf1c	1.210	0.09065	8	Emc4	1.204	0.092583	8
Psme1	1.215	0.089235	7	Ncapg	1.210	0.09065	11	Tyro3	1.204	0.092583	16
Ap1m1	1.215	0.089235	9	Dpp3	1.210	0.09065	9	Dpp9	1.204	0.092583	9
Vps28	1.215	0.089235	7	Myd88	1.210	0.09065	9	Phlda2	1.204	0.092583	6
Rtca	1.215	0.089235	8	Snmp200	1.210	0.09065	17	Ap4s1	1.204	0.092583	7
Fars2	1.215	0.089235	10	Folr1	1.210	0.09065	11	Ttc1	1.204	0.092583	8
Phf2	1.215	0.089235	12	Mrps28	1.209	0.09065	9	Ccnf	1.204	0.092583	10
Cd2bp2	1.215	0.089235	12	Lrch4	1.209	0.09065	11	Noc2l	1.203	0.092811	12
Nudt4	1.215	0.089235	12	Cd151	1.209	0.090849	8	Gadd45g	1.203	0.092811	9
Fam13b	1.215	0.089235	8	Mcts1	1.209	0.090849	8	Numa1	1.203	0.092811	15
Ap2b1	1.215	0.089235	13	Nudc	1.209	0.090849	22	Ssbp3	1.203	0.092811	13

图26-18

Raga	1.215	0.089235	9	Phyhd1	1.209	0.090849	12	Mrpl37	1.203	0.093266	10
Tmem41a	1.215	0.089235	8	Mtfmt	1.209	0.091736	20	Xrcc6	1.203	0.093266	10
Msantd3	1.215	0.089235	8	Hdgfrp2	1.209	0.091842	12	Ap2a2	1.203	0.093266	12
Zdhhc9	1.215	0.089235	8	Mapk13	1.209	0.091842	13	Mbtps1	1.203	0.093266	12
Aven	1.214	0.089235	9	Lman2	1.209	0.091842	13	Atp5j	1.203	0.093266	17
Ralbp1	1.214	0.089235	14	1500011K16Rik	1.209	0.091842	10	Akirin1	1.203	0.093266	11
Pdcd6	1.214	0.089235	8	Tmem134	1.209	0.091842	10	Rnps1	1.203	0.093266	15
Sptan1	1.214	0.089235	19	Wdr3	1.209	0.091842	11	Fam192a	1.203	0.093266	8
Gnb2	1.214	0.089235	18	Bloc1s4	1.209	0.091842	7	Jak3	1.203	0.093266	9
Ptpn6	1.214	0.089235	8	Ppp5c	1.209	0.091842	12	Tars	1.203	0.093623	15
Plekhj1	1.214	0.089235	9	Mdk	1.208	0.092006	12	Sept2	1.203	0.093623	18
Vmn2r15	1.214	0.089235	7	Cog4	1.208	0.092006	9	Cnih1	1.202	0.09376	10
Jam3	1.214	0.089235	9	2410003L11Rik	1.208	0.092006	6	Zic3	1.202	0.09376	17
Rbm19	1.214	0.089235	12	Ddx50	1.208	0.092006	13	Zfp428	1.202	0.09376	10
H1f0	1.214	0.089235	8	Selk	1.208	0.092006	10	Foxk2	1.202	0.09376	11
Ddx54	1.214	0.089235	11	Ddah1	1.208	0.092006	13	Txndc15	1.202	0.09376	7
Tmem115	1.214	0.089235	8	Fem1b	1.208	0.092006	15	Mcm10	1.202	0.094119	11
Klk8	1.214	0.089235	9	Ilf2	1.208	0.092006	10	Rmnd5a	1.202	0.094119	16
Ankfy1	1.214	0.089235	30	St14	1.208	0.092006	13	Arfgap1	1.202	0.094119	9
Rnh1	1.214	0.089235	8	Klhl13	1.208	0.092006	9	D2Wsu81e	1.202	0.094119	7
Sfswap	1.214	0.089235	9	Eif3f	1.208	0.092006	15	Rae1	1.202	0.094119	14
Scrib	1.213	0.089235	15	Sudla2	1.208	0.092006	9	Mrpl11	1.202	0.094119	13
Asrgl1	1.213	0.089235	7	Smdt1	1.208	0.092006	14	Timp1	1.202	0.094119	7
Ctbp2	1.213	0.089235	19	Nfatc2ip	1.208	0.092006	15	Ckap2l	1.202	0.094119	8
Ttc13	1.202	0.094119	8	Mrps33	1.197	0.095294	12	Nicn1	1.192	0.099753	15
Ccna2	1.202	0.094119	12	Timm22	1.197	0.095294	10	Nfu1	1.192	0.099753	7
Fam63a	1.202	0.094119	10	Lrwd1	1.197	0.095294	8	Rhox6	1.192	0.099753	6
Pafah1b3	1.202	0.094119	7	Mien1	1.197	0.095294	10	Vps4a	1.192	0.099753	9
Dpm2	1.201	0.094253	10	Golph3	1.197	0.095294	8	Zdhhc16	1.192	0.099753	9
Dhcr24	1.201	0.094253	13	Srsf7	1.197	0.095294	17	Plp2	1.191	0.099753	12
Uqcrc2	1.201	0.094253	15	Dqx1	1.197	0.095294	12	Cr1l	1.191	0.099753	9
Myo10	1.201	0.094253	11	Ncbp1	1.197	0.095294	9	Zfp330	1.191	0.099753	9

图26-19

Claspr	1.201	0.094253	10	Grpel1	1.196	0.095294	12	Elf3	1.191	0.099753	10
Pstk	1.201	0.094253	9	Mpc2	1.196	0.095294	9	Tomm34	1.191	0.099753	9
Rbm14	1.201	0.094399	8	Smc4	1.196	0.095294	11	Fam110a	1.191	0.099753	11
Brd2	1.201	0.094399	20	Rcc2	1.196	0.095294	29	Dguok	1.191	0.099753	8
Fastk	1.201	0.094399	9	BC005624	1.196	0.095294	11	Cdkal1	1.191	0.099753	9
Sec63	1.201	0.094399	13	Mirpl22	1.196	0.095294	7	1110007C09Rik	1.191	0.099753	7
Flot2	1.201	0.094399	9	Naa10	1.196	0.095294	11	Ncoa5	1.191	0.099753	10
Tmem259	1.201	0.094399	14	Mirps18b	1.196	0.095294	12	Anxa11	1.191	0.099753	7
Mob4	1.201	0.094399	12	Enah	1.196	0.095294	20	Sirt4	1.191	0.099753	8
Dnajc15	1.201	0.094399	7	Rbbp5	1.196	0.095459	9	Hmbs	1.191	0.099753	9
Trim35	1.201	0.094399	10	Fam96a	1.196	0.095459	8	Anxa3	1.191	0.099753	10
Pla2g1b	1.201	0.094399	7	Gm5860	1.196	0.095459	21	Ppp1cc	1.191	0.099753	18
Asun	1.200	0.094399	17	Arglu1	1.196	0.095459	11	Fbxo42	1.190	0.099753	13
Mtap	1.200	0.094399	10	Rpusd4	1.196	0.095459	7	Ufc1	1.190	0.099753	11
Znhit1	1.200	0.094399	9	Anp32b	1.196	0.095459	26	Nt5dc2	1.190	0.099753	8
Pepd	1.200	0.094399	10	Tmx2	1.196	0.095459	12	Fh1	1.190	0.099753	8
Ngdn	1.200	0.094399	8	Mettl9	1.196	0.095459	8	Polr2c	1.190	0.099753	12
Coq6	1.200	0.094399	7	Med28	1.196	0.095459	15	Vdac2	1.190	0.099753	23
Grwd1	1.200	0.094399	10	Mcm2	1.196	0.095459	20	Atxn7l3b	1.190	0.099753	15
Fig4	1.200	0.094399	11	Get4	1.196	0.095459	15	Afg3l2	1.190	0.099753	9
Cox20	1.200	0.094399	9	Tcf15	1.195	0.095459	9	Tex13	1.190	0.099753	12
Abcg2	1.200	0.094399	9	Ubr7	1.195	0.095459	11	Mrto4	1.190	0.099753	10
Ntpcr	1.200	0.094399	7	Srsf3	1.195	0.095459	21	Tspyl2	1.190	0.099753	8
Canx	1.200	0.094399	20	Xbp1	1.195	0.096124	11	Otulin	1.190	0.099753	8
Ptbp1	1.200	0.094399	29	Phyh	1.195	0.096124	8	Mrps11	1.190	0.099753	8
Etaa1	1.200	0.094399	12	Ddah2	1.195	0.096372	8	Vkorc1	1.190	0.099753	9
Atad3a	1.200	0.094399	13	Rbm39	1.195	0.096372	16	Ppp1r15b	1.190	0.099753	13
Taf10	1.199	0.094399	9	Rap1b	1.195	0.096372	10	Ngn	1.190	0.099753	9
Klhl21	1.199	0.094399	14	Basp1	1.195	0.096372	12	H3f3a	1.189	0.099753	25
Zcrb1	1.199	0.094399	11	Alg3	1.195	0.096666	9	Cwc15	1.189	0.099753	10
Mrpl18	1.199	0.094399	13	Nr0b1	1.195	0.096666	7	Grb2	1.189	0.099753	16
Sdf2l1	1.199	0.094399	12	Polg	1.194	0.096948	9	Loxl2	1.189	0.099753	28

图26-20

Ath11	1.199	0.094399	12	Rpp25l	1.194	0.096948	9	Snx32	1.189	0.099753	8
Cmtm6	1.199	0.094399	12	Lin37	1.194	0.096948	5	Npdc1	1.189	0.099753	8
Emc3	1.199	0.094399	9	R3hdm4	1.194	0.097607	8	2310011J03Rik	1.189	0.099753	9
Metn	1.199	0.094399	7	Gemin4	1.194	0.097607	10	Zfp936	1.189	0.099753	8
Thoc7	1.199	0.094399	13	BC052040	1.194	0.097607	9	Pdap1	1.189	0.099753	26
Prc1	1.199	0.094399	9	Zfp598	1.194	0.097607	8	Rgs2	1.189	0.099753	13
Srrm1	1.199	0.094399	17	Prmt1	1.194	0.097935	20	Cdc34	1.189	0.099753	22
Glit25d1	1.199	0.094399	11	Rfc3	1.194	0.097935	11	Dgkd	1.189	0.099753	9
Hars	1.199	0.094399	13	Cops4	1.194	0.09805	12	Snord47	1.189	0.099753	5
Rexo4	1.199	0.094399	10	Aplp2	1.194	0.09805	12	Tagln3	1.189	0.099753	6
Lpar5	1.199	0.094399	9	Mirps18c	1.194	0.09805	8	Gm12504	1.189	0.099753	41
Rab40c	1.199	0.094616	12	Pnpo	1.194	0.09805	11	Lamtor1	1.189	0.099753	13
Cnot2	1.198	0.094616	15	Vat1	1.193	0.09805	9	Suds3	1.189	0.099753	11
Thop1	1.198	0.094616	9	Nkiras2	1.193	0.09805	9	Prdx5	1.189	0.099753	7
Rad17	1.198	0.094616	9	Etfb	1.193	0.09805	11	Gar1	1.189	0.099753	15
Rela	1.198	0.094616	9	Esyrt1	1.193	0.098419	9	1110065P20Rik	1.188	0.099753	7
Syce1	1.198	0.094616	8	Tatdn2	1.193	0.098985	9	Adck4	1.188	0.099753	8
Jkamp	1.198	0.094616	9	Zfp787	1.193	0.098985	10	Mrpl10	1.188	0.099753	10
Rbm38	1.198	0.094616	9	Sdhb	1.193	0.098985	11	Cpsf4	1.188	0.099753	9
Bckd1b	1.198	0.094616	7	Vars	1.193	0.098985	12	Pbdc1	1.188	0.099921	22
Mfsd10	1.198	0.094616	8	Abcf1	1.193	0.098985	14	Gclm	1.188	0.099921	12
Mthfd1	1.198	0.094616	13	Prkab1	1.193	0.099182	9	Fam136a	1.188	0.099921	10
Exosc4	1.198	0.094616	12	Khsrp	1.193	0.099182	17	Mlh1	1.188	0.099921	7
Uqcrc1	1.198	0.094618	15	Acvr2b	1.192	0.099182	7	Usp48	1.188	0.099921	11
Zc3h15	1.198	0.094618	13	Nr1h4	1.192	0.099182	9	Dctn2	1.188	0.099921	14
Ccdc181	1.198	0.094618	8	Ccdc166	1.192	0.09942	7	Timeless	1.188	0.099921	14
Fbxw11	1.198	0.094618	14	Ehd4	1.192	0.09942	10	U2af114	1.188	0.099921	8
Rrbp1	1.198	0.094618	16	Sec11a	1.192	0.09942	14	Snhg8	1.187	0.099921	12
Mzt2	1.198	0.094618	9	Cpsf1	1.192	0.09942	10	Cdc42	1.187	0.099921	22
Surf2	1.198	0.094618	7	Ecd	1.192	0.099753	13	Hint1	1.187	0.099921	26
Las1l	1.198	0.094618	15	Mirps6	1.192	0.099753	7	Arpc3	1.187	0.099921	10
Pla2g6	1.198	0.094618	12	Hsf1	1.192	0.099753	9	Mfap3	1.187	0.099921	18

图26-21

Mknk2	1.198	0.094618	8	Wdr45	1.192	0.099753	11	Nif31	1.187	0.099921	12
Ttc27	1.198	0.094958	11	Smarca4	1.192	0.099753	14	Pdia4	1.187	0.099921	9
Ppp2r5d	1.197	0.094958	13	Sar1a	1.192	0.099753	16	Tmem160	1.187	0.099921	8
Prpf38a	1.197	0.095294	8	Csnk2a1	1.192	0.099753	24	Hhex	1.187	0.099921	9
Abcf2	1.187	0.099921	15	Mbip	1.186	0.099921	7	Nt5c	1.186	0.099921	7
Manbal	1.187	0.099921	10	Trmt61a	1.186	0.099921	7				
Fis1	1.186	0.099921	10	Ppp1r18	1.186	0.099921	8				
Pgrmc1	1.186	0.099921	8	Gm6792	1.186	0.099921	6				

图26-22

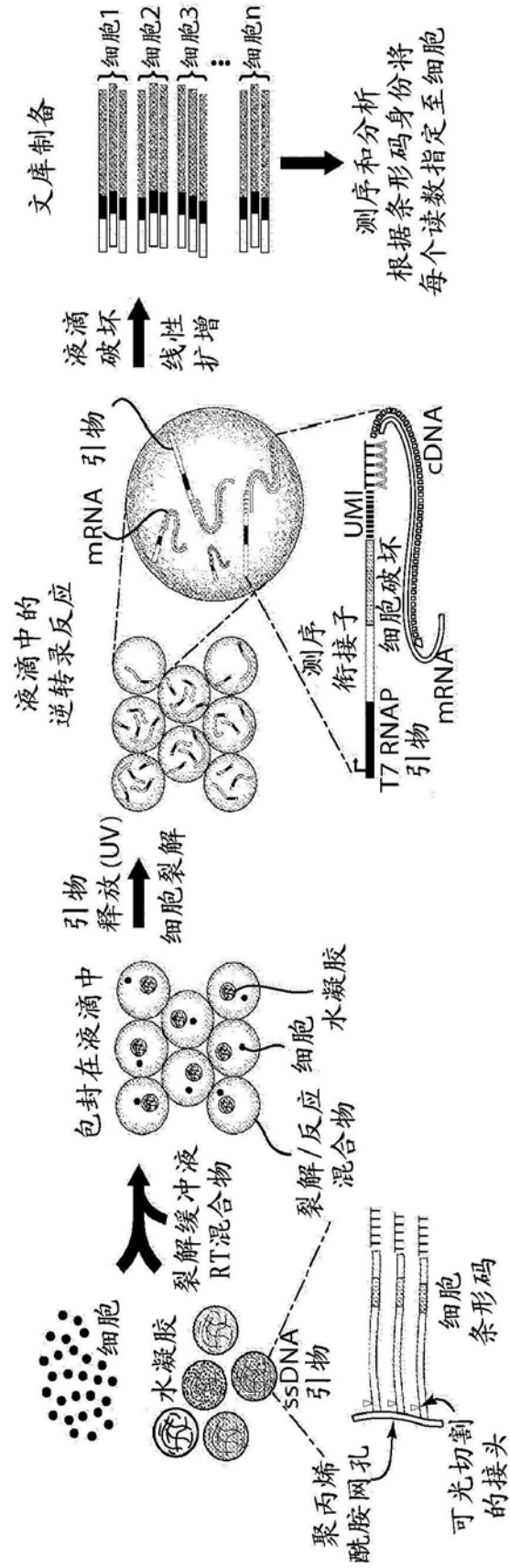


图27