



(12) 发明专利申请

(10) 申请公布号 CN 119156454 A

(43) 申请公布日 2024.12.17

(21) 申请号 202380038228.7

(22) 申请日 2023.05.30

(30) 优先权数据

63/347,111 2022.05.31 US

(85) PCT国际申请进入国家阶段日

2024.11.04

(86) PCT国际申请的申请数据

PCT/EP2023/064443 2023.05.30

(87) PCT国际申请的公布数据

W02023/232811 EN 2023.12.07

(71) 申请人 雀巢产品有限公司

地址 瑞士

(72) 发明人 L·特沙里 S·赫尔齐格

A·佐林格 潘元龙 P·谷特

P·施泰纳 L·卢卡扬科

(74) 专利代理机构 北京市中咨律师事务所

11247

专利代理师 魏子翔 于静

(51) Int.Cl.

C12Q 1/6876 (2006.01)

C12Q 1/6883 (2006.01)

权利要求书5页 说明书38页

序列表(电子公布) 附图8页

(54) 发明名称

用于改善狗的健康寿命的方法

(57) 摘要

本发明提供了用于确定狗的死亡风险和/或健康寿命的概率的方法;所述方法包括a)提供来自该狗获得的样本的DNA甲基化图谱;以及b)使用该DNA甲基化图谱来确定该狗的死亡风险和/或健康寿命的概率。

1. 用于确定狗的死亡风险和/或健康寿命的概率的方法;所述方法包括:
 - a) 提供来自从所述狗获得的样本的DNA甲基化图谱;以及
 - b) 使用所述DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率。
2. 用于确定狗的生物学年龄的方法;所述方法包括:
 - a) 提供来自从所述狗获得的样本的DNA甲基化图谱;以及
 - b) 使用所述DNA甲基化图谱来确定所述狗的生物学年龄,其中所述DNA甲基化图谱与所述狗的死亡风险和/或健康寿命的概率相关。
3. 用于为狗选择生活方式方案、膳食方案或治疗干预的方法,所述方法包括:
 - a) 提供来自从所述狗获得的样本的DNA甲基化图谱;
 - b) 使用所述DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及
 - c) 基于步骤b)中确定的所述死亡风险和/或健康寿命的概率来为所述狗选择合适的生活方式方案、膳食方案或治疗干预。
4. 用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险的功效的方法,所述方法包括:
 - a) 将生活方式方案、膳食方案或治疗干预应用于所述狗,其中已经根据权利要求3所述的方法选择所述生活方式方案、膳食方案或治疗干预;
 - b) 在将所述生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;
 - c) 在遵循所述生活方式方案、膳食方案或治疗干预的所述时间段后,确定所述狗的所述死亡风险和/或健康寿命的概率是否存在变化。
5. 用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率的功效的方法,所述方法包括:
 - a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;
 - b) 将基于步骤a)中确定的所述死亡风险和/或健康寿命的概率选择的生活方式方案、膳食方案或治疗干预应用于所述狗;
 - c) 在将生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;
 - d) 确定所述狗的所述死亡风险和/或健康寿命的概率在步骤a)和步骤c)之间是否存在变化。
6. 用于开发抗衰老生活方式方案、膳食方案或治疗干预的方法;所述方法包括:
 - a) 使用来自从狗获得的样本的DNA甲基化图谱来确定所述狗的第一死亡风险和/或健康寿命的概率;
 - b) 将生活方式方案、膳食方案或治疗干预应用于所述狗;
 - c) 在将所述生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的第二样本的DNA甲基化图谱来确定所述狗的第二死亡风险和/或健康寿命的概率;
 - d) 在遵循所述生活方式方案、膳食方案或治疗干预的所述时间段后,确定所述狗的所述第一死亡风险和所述第二死亡风险和/或健康寿命的概率是否存在变化;

其中如果所述生活方式方案、膳食方案或治疗干预降低所述死亡风险和/或增加健康寿命的所述概率;以及/或者降低所述狗的死亡风险的增加率和/或增加所述狗的健康寿命的所述概率的减少率,则确定其是抗衰老的。

7. 用于预防或降低狗发展疾病的风险的方法;所述方法包括:

a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;其中对所述狗确定的所述死亡风险和/或健康寿命的概率与发展所述疾病的增加的可能性相关联;以及

b) 基于步骤a)中确定的所述死亡风险和/或健康寿命的概率来为所述狗选择生活方式方案、膳食方案或治疗干预;

其中所述生活方式方案、膳食方案或治疗干预预防或降低所述狗发展所述疾病的风险;优选地其中所述疾病是年龄相关疾病。

8. 用于选择适合接受抗衰老生活方式方案、膳食方案或治疗干预的狗的方法;所述方法包括:

a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;

b) 如果与狗的实足年龄相比,所述狗具有增加的死亡风险和/或降低的健康寿命的概率,则将所述狗选择为适合接受抗衰老生活方式方案、膳食方案或治疗干预。

9. 根据任一前述权利要求所述的方法,其中用于确定所述狗的所述死亡风险和/或健康寿命的概率的所述方法还包括将所述DNA甲基化图谱与所述狗的所述实足年龄、品种和/或性别中的一者或多者组合。

10. 根据权利要求3至9中任一项所述的方法,其中基于与所述狗的实足年龄相比所述狗具有增加的死亡风险和/或降低的健康寿命的概率的确定来选择生活方式方案、膳食方案或治疗干预。

11. 根据权利要求3至10中任一项所述的方法,其中所述生活方式方案、膳食方案或治疗干预是膳食干预。

12. 根据权利要求11所述的方法,其中所述膳食干预是热量限制饮食、老年饮食或低蛋白饮食。

13. 根据任一前述权利要求所述的方法,其中所述样本是血液样本。

14. 根据权利要求1至3或5至8中任一项所述的方法,其中步骤a)还包括确定来自从所述狗获得的样本的DNA甲基化图谱。

15. 根据任一前述权利要求所述的方法,其中所述DNA甲基化图谱包含至少一个如表3中所列的甲基化位点。

16. 根据权利要求15所述的方法,其中所述DNA甲基化图谱包含至少3个、至少5个、至少10个、至少20个、至少50个、至少100个、至少150个、至少200个或如表3中所列的甲基化位点中的每个甲基化位点。

17. 根据权利要求15所述的方法,其中所述DNA甲基化图谱包含如表4中所列的甲基化位点。

18. 根据权利要求15所述的方法,其中所述DNA甲基化图谱包含如表5中所列的甲基化位点。

19. 根据权利要求15所述的方法, 其中所述DNA甲基化图谱包含如表6中所列的甲基化位点。

20. 根据权利要求15所述的方法, 其中所述DNA甲基化图谱包含如表7中所列的甲基化位点。

21. 根据权利要求14至20中任一项所述的方法, 其中使用包括以下步骤中的一者或多者的方法来确定DNA甲基化: (a) 用APOBEC处理样本DNA以使胞嘧啶脱氨基; (b) 基于捕获的富集; 和/或 (c) 高通量测序。

22. 根据任一前述权利要求所述的方法, 其中所述DNA甲基化图谱与以下项的发病率或预测的发病率相关联: (i) 组织; (ii) 器官; 或 (iii) 生理系统, 诸如免疫系统、胃肠系统、泌尿系统、肌肉系统、心血管系统和/或神经系统。

23. 根据权利要求22所述的方法, 所述方法还包括对所述狗应用适用于改善根据权利要求22所述的组织; 器官; 或生理系统的所述发病率或预测的发病率的生活方式方案、膳食方案或治疗干预。

24. 用于降低狗的死亡风险和/或增加狗的健康寿命的概率的膳食干预或治疗方式, 其中将所述膳食干预施用于狗, 其中死亡风险通过根据权利要求1、7至10或13至23中任一项所述的方法确定。

25. 膳食干预治疗方式在降低狗的死亡风险和/或增加狗的健康寿命的概率中的用途, 其中将所述膳食干预施用于狗, 其中死亡风险通过根据权利要求1、7至20或13至23中任一项所述的方法确定。

26. 计算机可读介质, 所述计算机可读介质包括指令, 所述指令在被执行时使一个或多个处理器执行根据权利要求1、2、3、7至16或23中任一项所述的方法。

27. 用于确定狗的死亡风险和/或健康寿命的概率的计算机系统; 所述计算机系统被编程为使用所述狗的DNA甲基化图谱来确定所述狗的死亡风险。

28. 用于为狗选择合适的生活方式方案、膳食方案或治疗干预的计算机系统, 所述计算机系统被编程为执行以下步骤中的一者或多者:

a) 使用来自所述狗的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率; 以及

b) 基于步骤a) 中确定的所述死亡风险和/或健康寿命的概率来为所述狗选择合适的生活方式方案、膳食方案或治疗干预。

29. 用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率的功效的计算机系统, 所述计算机系统被编程为执行以下步骤中的一者或多者:

a) 使用来自在所述生活方式方案、膳食方案或治疗干预之前从所述狗获得的样本以及在所述生活方式方案、膳食方案或治疗干预之后从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率; 以及

b) 在已经应用所述生活方式方案、膳食方案或治疗干预之前和之后确定所述狗的所述死亡风险和/或健康寿命的概率在从所述狗获得的所述样本之间是否存在变化。

30. 用于确定狗将受益于抗衰老生活方式方案、膳食方案或治疗干预的可能性的计算机系统; 所述计算机系统被编程为执行以下步骤中的一者或多者:

a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健

康寿命的概率;

b) 如果与狗的实足年龄相比,所述狗具有增加的死亡风险和/或降低的健康寿命的概率,则将所述狗标识为可能对抗衰老生活方式方案、膳食方案或治疗干预有反应。

31. 计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机使用狗的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率。

32. 计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机使用来自狗的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及基于使用DNA甲基化图谱来确定的所述死亡风险和/或健康寿命的概率来为所述狗选择合适的生活方式方案、膳食方案或治疗干预。

33. 计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机:a) 使用来自在生活方式方案、膳食方案或治疗干预之前从狗获得的样本以及在所述生活方式方案、膳食方案或治疗干预之后从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及b) 在已经应用所述生活方式方案、膳食方案或治疗干预之前和之后确定所述狗的所述死亡风险和/或健康寿命的概率在从所述狗获得的所述样本之间是否存在变化。

34. 计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机:a) 使用来自从狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及b) 如果与狗的实足年龄相比,所述狗具有增加的死亡风险和/或降低的健康寿命的概率,则将所述狗标识为可能对抗衰老生活方式方案、膳食方案或治疗干预有反应。

35. 用于创建表观遗传时钟以确定狗的死亡风险和/或健康寿命的概率的方法;所述方法包括:a) 提供来自从所述狗获得的样本的DNA甲基化图谱;b) 计算两个gompertz或weibull函数,(i) 第一个gompertz或weibull函数,其将存活率建模为所述甲基化图谱、实足年龄和任选的品种类别和/或性别的函数;和(ii) 第二个gompertz或weibull函数,其将存活率建模为实足年龄和任选的仅品种类别和/或性别的函数;其中基于时间变量来确定所述死亡风险和/或健康寿命的概率,在所述时间变量处由所述第二个gompertz或weibull函数给出的所述狗的存活概率等于由所述第一个gompertz或weibull函数给出的在它们的实足年龄处的存活概率。

36. 根据权利要求27至35中任一项所述的计算机系统或计算机程序产品;其中所述计算机系统或计算机程序准备并共享详述步骤(b)的结果的报告。

37. 用于确定狗的死亡风险和/或健康寿命的概率的试剂盒,所述试剂盒包括用于收集和任选地稳定来自受试者的样本的装置和试剂;以及用于收集和任选地稳定所述样本的说明书,并且其中所述样本用于邮寄以用于后续DNA提取,用于测量从所述样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱。

38. 用于使用狗的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率的方法;所述方法包括(a) 提供来自所述狗的样本,任选地使用根据权利要求37所述的试剂盒;(b) 提供用于后续DNA提取的所述样本以用于测量从所述样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱;(c) 使用根据权利要求27所述的计算机系统或根据权利要求31所述

的计算机程序产品来确定所述狗的死亡风险和/或健康寿命的概率;其中所述计算机系统准备并共享详述步骤(c)的结果的报告。

39. 用于使用来自狗的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及基于使用DNA甲基化图谱来确定的所述死亡风险和/或健康寿命的概率来为所述狗选择合适的生活方式方案、膳食方案或治疗干预的方法;所述方法包括(a)提供来自所述狗的样本,任选地使用根据权利要求37所述的试剂盒;(b)提供用于后续DNA提取的所述样本以用于测量从所述样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱;(c)使用根据权利要求28所述的计算机系统或根据权利要求32所述的计算机程序产品来确定所述狗的死亡风险和/或健康寿命的概率以及为所述狗选择合适的生活方式方案、膳食方案或治疗干预;其中所述计算机系统准备并共享详述步骤(c)的结果的报告。

40. 用于使用DNA甲基化图谱来确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率的功效的方法,所述方法包括:(a)提供在所述生活方式方案、膳食方案或治疗干预之前从所述狗获得的样本以及在所述生活方式方案、膳食方案或治疗干预之后从所述狗获得的样本,任选地使用根据权利要求37所述的试剂盒;(b)提供用于后续DNA提取的所述样本以用于测量从所述样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱;(c)使用根据权利要求29所述的计算机系统或根据权利要求33所述的计算机程序产品在已经应用所述生活方式方案、膳食方案或治疗干预之前和之后确定所述狗的所述死亡风险和/或健康寿命的概率在从所述狗获得的所述样本之间是否存在变化;其中所述计算机系统准备并共享详述步骤(c)的结果的报告。

41. 使用DNA甲基化图谱将狗标识为可能对抗衰老生活方式方案、膳食方案或治疗干预有反应的方法;所述方法包括:(a)提供从所述狗获得的样本;任选地使用根据权利要求37所述的试剂盒;(b)提供用于后续DNA提取的所述样本以用于测量从所述样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱;(c)使用根据权利要求30所述的计算机系统或根据权利要求34所述的计算机程序产品,如果与狗的实足年龄相比,所述狗具有增加的死亡风险和/或降低的健康寿命的概率,则将所述狗标识为可能对抗衰老生活方式或膳食有反应;其中所述计算机系统准备并共享详述步骤(c)的结果的报告。

用于改善狗的健康寿命的方法

技术领域

[0001] 本发明涉及用于使用DNA甲基化图谱来确定狗的健康状态的方法。特别地,本发明涉及基于由DNA甲基化图谱确定的健康状态而为狗选择生活方式方案、膳食方案或治疗干预,或确定生活方式方案、膳食方案或治疗干预的功效的方法。

背景技术

[0002] 确定关于狗的健康的信息的能力是告知狗的总体健康和福祉所期望的。

[0003] 众所周知,实足年龄是总体健康状态的主要指标,其中实足年龄的增加与健康状况的降低相关联。然而,取决于遗传、营养和生活方式的不同,个体衰老的速度可能比实足年龄慢或快。因此,实足年龄可能并不总是反映个体的衰老速率或健康状况降低的风险。另一方面,某一个体的生物学年龄(基于例如临床生物化学和细胞生物学量度),与相同实足年龄的其他个体相比,可能会有所不同。用于确定生物学年龄的方法可能有助于比基于个体的实足年龄的预期更早地标识出处于年龄相关病症风险中的个体(参见例如W02019/046725)。

[0004] 然而,需要确定狗的生物学年龄并利用生物学年龄的测量来改善狗的健康结果的其他方法。

发明内容

[0005] 本发明涉及用于基于DNA甲基化图谱来量化狗的健康状态的方法。该方法能够通过评估来自狗的DNA甲基化图谱来确定狗的死亡风险和/或健康寿命的概率。

[0006] 评估狗的健康状态的现有方法基于DNA甲基化和实足年龄之间的相关性来确定生物学年龄。计算动物的生物学年龄可包括确定与给定实足年龄的预期DNA甲基化图谱相比的DNA甲基化图谱。因此,这种方法基于使用实足年龄作为总体健康的主要指标。

[0007] 相比之下,本发明考虑了DNA甲基化图谱对死亡风险和/或健康寿命的概率的直接预测值。举例来说,给定的DNA甲基化标志物可能不与实足年龄直接相关,但是可指示特定的病理状况,从而指示死亡风险增加和/或健康寿命减少的概率。因此,本方法可描述为识别狗的死亡风险和/或健康寿命的概率。因此,本发明的DNA甲基化标志物和DNA甲基化图谱不一定与实足年龄相关,但与狗的表型年龄和实足年龄之间的差异有关。

[0008] 在第一方面,本发明提供了用于确定狗的死亡风险的方法;所述方法包括:a)提供来自从所述狗获得的样本的DNA甲基化图谱;以及b)使用所述DNA甲基化图谱来确定所述狗的死亡风险。

[0009] 确定死亡风险可以是指确定狗与例如相同实足年龄、性别和品种的同等狗相比活的时间更长或更短的可能性。因此,本方法可确定狗相较于例如相同实足年龄、性别和品种的同等狗的寿命、健康跨度和/或长寿的概率。此外,用于改善狗的死亡风险和/或健康寿命的概率的方法可改善狗的可能寿命、健康跨度和/或长寿。

[0010] 如本文所用,“寿命”可以指受试者活的时间长度(例如数年)。“健康跨度”可以指

没有疾病的生命的时间长度(例如数年)。“长寿”可以指受试者活了超过其预期寿命的时间长度(例如数年)。

[0011] 合适地,死亡风险可等同于狗的健康寿命的概率;其中死亡风险的降低等同于狗的更长健康寿命的概率增加,或者死亡风险的增加等同于狗的更长健康寿命的概率降低。死亡风险可表示为狗的确定年龄(即生物学年龄)和实足年龄之间的差异。例如,通过本方法确定的生物学年龄之间的差异相较于实足年龄的增加可指示狗的死亡风险增加。通过本方法确定的生物学年龄之间的差异相较于实足年龄的降低可指示狗的死亡风险降低。合适地,死亡风险和/或健康寿命的概率可描述为狗的生物学年龄。合适地,死亡风险和/或健康寿命的概率可描述为狗的表观遗传年龄。合适地,本发明的生物时钟可称为表观遗传时钟。

[0012] 合适地,确定狗的生物学年龄大于其实足年龄指示较高的死亡风险。合适地,确定狗的生物学年龄小于其实足年龄指示死亡风险的降低。合适地,确定狗的生物学年龄大于其实足年龄指示更长健康寿命的概率降低。合适地,确定狗的生物学年龄小于其实足年龄指示更长健康寿命的概率增加。

[0013] 合适地,本方法可用于基于狗的死亡风险和/或健康寿命的概率来确定狗的生物学年龄。

[0014] 因此,在另一个方面,本发明提供了用于确定狗的生物学年龄的方法;所述方法包括:a)提供来自从所述狗获得的样本的DNA甲基化图谱;以及b)使用所述DNA甲基化图谱来确定所述狗的生物学年龄,其中所述DNA甲基化图谱与所述狗的死亡风险和/或健康寿命的概率相关。

[0015] 在所有本方法中,确定或改善狗的死亡风险和/或健康寿命的概率也适用于确定或改善狗的生物学年龄;其中使用与狗的死亡风险和/或健康寿命的概率相关的DNA甲基化图谱来确定狗的生物学年龄。

[0016] 在另一个方面,本发明提供了用于为狗选择生活方式方案、膳食方案或治疗干预的方法,所述方法包括:a)提供来自从所述狗获得的样本的DNA甲基化图谱;b)使用所述DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;以及c)基于步骤b)中确定的所述死亡风险而为所述狗选择合适的生活方式方案、膳食方案或治疗干预。

[0017] 如本文所用,“为狗选择合适的生活方式方案、膳食方案或治疗干预”还可涵盖“为狗推荐生活方式方案、膳食方案或治疗干预”或“为狗提供推荐的生活方式方案、膳食方案或治疗干预”。

[0018] 在另一方面,本发明提供了用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率的功效的方法,所述方法包括:a)将生活方式方案、膳食方案或治疗干预应用于所述狗,其中已经根据本发明的前述方面选择所述生活方式方案、膳食方案或治疗干预;b)在将所述生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;c)在遵循所述生活方式方案、膳食方案或治疗干预的所述时间段后,确定所述狗的所述死亡风险和/或健康寿命的概率是否存在变化。

[0019] 在另一个方面,本发明提供了用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康生命的概率的功效的方法,所述方法包括:a)使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;b)将基于步骤a)中确定的所述死

亡风险选择的生活方式方案、膳食方案或治疗干预应用于所述狗;c) 在将生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;d) 确定所述狗的所述死亡风险在步骤a) 和步骤c) 之间是否存在变化。

[0020] 合适地,改善狗的死亡风险和/或健康寿命的概率可以指狗的生物年龄和实足年龄之间的差异的减少,其中狗的生物年龄大于其实足年龄。此外,改善狗的死亡风险和/或健康寿命的概率可以指保持或进一步增加狗的生物年龄和实足年龄之间的差异,其中狗的生物年龄小于其实足年龄。另选地,狗的死亡风险和/或健康寿命的概率的恶化可以指狗的生物年龄和实足年龄之间的差异的增加,其中狗的生物年龄大于其实足年龄。狗的死亡风险和/或健康寿命的概率的恶化还可以指狗的生物年龄和实足年龄之间的差异的减少,其中狗的生物年龄小于其实足年龄。

[0021] 合适地,改善狗的死亡风险和/或健康寿命的概率可以指狗的生物年龄和实足年龄之间的变化率的降低,其中狗的生物年龄大于其实足年龄。例如,每增加1岁的实足年龄,狗的生物年龄可能已经增加1.5岁。在生活方式和膳食方案干预后,变化率的降低使得每增加1岁的实足年龄狗的生物年龄随后增加1.25岁,这可提供狗的死亡风险和/或健康寿命的概率的改善。

[0022] 改善死亡风险和/或健康寿命的概率还可以指保持或增加狗的生物年龄和实足年龄之间的变化率,其中狗的生物年龄小于其实足年龄。例如,每增加1岁的实足年龄,狗的生物年龄可能已经增加小于1岁(例如0.9岁)。在生活方式、膳食方案或治疗干预后,变化率可改变,使得每增加1岁的实足年龄狗的生物年龄随后增加例如0.8岁或更少,这可提供狗的生物年龄的改善。

[0023] 用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率的功效的本方法可有利地允许对生活方式方案、膳食方案或治疗干预对改善或保持狗的健康的有效性进行持续监测。这种方法的使用可有利地允许识别特别有效的生活方式方案、膳食方案或治疗干预。相比之下,如果基于狗的死亡风险和/或健康寿命的概率确定生活方式方案、膳食方案或治疗干预是无效的;则可实施替代的生活方式方案、膳食方案或治疗干预。

[0024] 因此,本方法使得能够基于如从DNA甲基化图谱所确定的狗的死亡风险和/或健康寿命的概率来为狗选择合适的生活方式方案、膳食方案或治疗干预。例如,通常基于狗的实足年龄推荐高度可消化和高质量的蛋白饮食。例如,可能会推荐,让狗在7、8岁左右改用老年饮食。然而,在本发明的上下文中,与狗的实足年龄相比,确定狗的死亡风险增加和/或健康寿命的概率降低(即生物年龄增加)可允许确定让狗在更早的年龄改用老年饮食。相比之下,与狗的实足年龄相比,死亡风险降低和/或健康寿命的概率增加(即生物年龄减少)的狗可能能够更长时间地继续享用成年饮食。

[0025] 合适地,本方法可包括在确定与狗的实足年龄相比狗具有增加的死亡风险和/或降低的健康寿命的概率后,对狗选择和/或应用生活方式方案、膳食方案或治疗干预。

[0026] 在另一个方面,本发明提供了用于预防或降低狗发展疾病的风险的方法;所述方法包括:

[0027] a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/

或健康寿命的概率;其中对所述狗确定的所述死亡风险和/或健康寿命的概率与发展所述疾病的增加的可能性相关联;以及

[0028] b) 基于步骤a)中确定的所述死亡风险和/或健康寿命的概率来为所述狗选择生活方式方案、膳食方案或治疗干预;

[0029] 其中所述生活方式方案、膳食方案或治疗干预预防或降低所述狗发展所述疾病的风险。

[0030] 合适地,所述疾病是年龄相关疾病。例如,年龄相关疾病骨关节炎、痴呆、认知功能障碍、糖尿病前状况、糖尿病、癌症、心脏病、肥胖症、胃肠道病症、失禁、肾病、少肌症、视力丧失、听力丧失、骨质疏松、白内障、脑血管疾病和/或肝病。

[0031] 该方法还可任选地包括将生活方式方案、膳食方案或治疗干预施用于狗。合适地,生活方式方案可以是膳食干预或治疗方式。

[0032] 在另一方面,本发明提供了用于选择适合接受抗衰老生活方式方案、膳食方案或治疗干预的狗的方法;所述方法包括:a)使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;b)如果与狗的实足年龄相比,所述狗具有增加的死亡风险和/或降低的健康寿命的概率,则将所述狗选择为适合接受抗衰老生活方式方案、膳食方案或治疗干预。

[0033] 合适地,虽然抗衰老生活方式方案、膳食方案或治疗干预对于基于实足年龄的狗可能是有效的,但当应用于与其实足年龄相比具有增加的死亡风险和/或降低的健康寿命的概率的狗时,其可能是特别有效的。因此,本方法可有利地使得能够选择对抗衰老生活方式方案、膳食方案或治疗干预具有增加的应答可能性或改善的应答量的狗。

[0034] 生活方式方案、膳食方案或治疗干预可基于与狗的实足年龄相比狗具有增加的死亡风险和/或降低的健康寿命的概率(即增加的生物年龄)的确定来选择。

[0035] 生活方式方案、膳食方案或治疗干预可以是膳食干预。膳食干预可以是热量限制饮食、老年饮食或低蛋白饮食。

[0036] DNA甲基化图谱可能与以下项的生物年龄增加相关联:(i)组织;(ii)器官;或(iii)生理系统,诸如免疫系统、胃肠系统、泌尿系统、肌肉系统、心血管系统和/或神经系统。

[0037] 本发明还提供了用于降低狗的死亡风险和/或增加狗的健康寿命的概率的膳食干预,其中将膳食干预施用于狗,其中死亡风险和/或健康寿命的概率通过本发明的方法确定。

[0038] 本发明进一步涉及膳食干预在降低狗的死亡风险和/或增加狗的健康寿命的概率中的用途,其中将膳食干预施用于狗,其中死亡风险和/或健康寿命的概率通过本发明的方法确定。

[0039] 在另一方面,本发明提供了计算机可读介质,该计算机可读介质包括指令,这些指令在被执行时使一个或多个处理器执行本发明的方法。

[0040] 在另一方面,本发明提供了用于确定狗的死亡风险的计算机系统;所述计算机系统被编程为使用所述狗的DNA甲基化图谱来确定所述狗的死亡风险。

[0041] 在另一方面,本发明提供了用于为狗选择合适的生活方式方案、膳食方案或治疗干预的计算机系统,所述计算机系统被编程为执行以下步骤中的一者或多者:a)使用来自

所述狗的DNA甲基化图谱来确定所述狗的死亡风险;以及b) 基于步骤a) 中确定的所述死亡风险而为所述狗选择合适的生活方式方案、膳食方案或治疗干预。

[0042] 在另一方面中,本发明提供了用于确定生活方式方案、膳食方案或治疗干预对改善狗的死亡风险的功效益的计算机系统,所述计算机系统被编程为执行以下步骤中的一者或多者:a) 使用来自在所述生活方式方案、膳食方案或治疗干预之前从所述狗获得的样本以及在所述生活方式方案、膳食方案或治疗干预之后从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;以及b) 在已经应用所述生活方式方案、膳食方案或治疗干预之前和之后确定所述狗的所述死亡风险在从所述狗获得的所述样本之间是否存在变化。

[0043] 在另一方面,本发明提供了用于确定狗将受益于抗衰老生活方式方案、膳食方案或治疗干预的可能性的计算机系统;所述计算机系统被编程为执行以下步骤中的一者或多者:a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;b) 如果与狗的实足年龄相比,所述狗具有增加的死亡风险,则将所述狗标识为可能对抗衰老生活方式方案、膳食方案或治疗干预有反应。

[0044] 在另一方面,本发明提供了计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机使用狗的DNA甲基化图谱来确定所述狗的死亡风险。

[0045] 在另一方面,本发明提供了计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机使用来自狗的DNA甲基化图谱来确定所述狗的死亡风险;以及基于使用DNA甲基化图谱来确定的所述死亡风险而为所述狗选择合适的生活方式方案、膳食方案或治疗干预。

[0046] 在另一方面,本发明提供了计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机:a) 使用来自在生活方式方案、膳食方案或治疗干预之前从狗获得的样本以及在所述生活方式方案、膳食方案或治疗干预之后从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;以及b) 在已经应用所述生活方式方案、膳食方案或治疗干预之前和之后确定所述狗的所述死亡风险在从所述狗获得的所述样本之间是否存在变化。

[0047] 在另一方面,本发明提供了计算机程序产品,所述计算机程序产品包括计算机可实施指令,所述计算机可实施指令用于使可编程计算机:a) 使用来自从狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险;以及b) 如果与狗的实足年龄相比,所述狗具有增加的死亡风险,则将所述狗标识为可能对抗衰老生活方式方案、膳食方案或治疗干预有反应。

[0048] 有利地,本发明可允许基于多个器官系统和功能的标志物来确定死亡风险和/或健康寿命的概率。因此,本方法可有利地涵盖了一系列潜在的器官功能障碍。

[0049] 评价狗的死亡风险和/或健康寿命的概率允许人们测试动物福祉的几个方面。首先,它可预测该动物是否更需要膳食或基于补充剂的干预。它还可用来测试膳食或基于补充剂的干预对衰老的功效益。

附图说明

[0050] 图1——预测死亡风险的血液生物标志物的标识。cox比例危险模型适于所评估的28种生物标志物中的每一种生物标志物,包括性别和品种类别(小型或中型)。针对每个参

数的p值来调整值,以说明多重比较(按错误发现率(fdr))。参数显示是调整后fdr低于0.05的参数。

[0051] 图2——生物标志物有助于用于确定表型年龄的多参数模型的预测能力的证明。

[0052] 图3——在热量限制研究中验证本发明的例示性时钟。变化量对应于实足年龄相对于预测的phenoDNAmAge的回归模型的残差。与控制饮食的狗相比,热量限制的狗具有显著更低的生物年龄(更低的变化量)。

[0053] 图4——在6岁时的热量限制研究中验证本发明的例示性时钟。变化量对应于实足年龄相对于预测的phenoDNAmAge的回归模型的残差。与控制饮食的狗相比,热量限制的狗具有更低的生物年龄(更低的变化量)。

[0054] 图5——调整线性混合效应模型(DogID作为随机效应)。与控制饮食的狗相比,热量限制的狗具有显著更低的生物年龄。

[0055] 图6——通过本发明的例示性时钟所确定的生物学上年轻的狗和生物学上年长的狗之间的存活率

[0056] 图7——Cox比例风险模型使用性别和变化量phenoDNAmAge(来自实足年龄相对于预测的phenoDNAmAge的残差)进行拟合,在训练集上按品种类别(小型或中型)分层。变化量的增加与更高的死亡风险相关联

[0057] 图8——技术重复之间的再现性评估

[0058] 图9——包含来自完整生物时钟的前3个甲基化位点的例示性生物时钟与实足年龄相关

[0059] 图10——包含来自完整生物时钟的前5个甲基化位点的例示性生物时钟与实足年龄相关

[0060] 图11——包含来自完整生物时钟的前10个甲基化位点的例示性生物时钟与实足年龄相关

[0061] 图12——包含来自完整生物时钟的前20个甲基化位点的例示性生物时钟与实足年龄相关

[0062] 图13——包含来自完整生物时钟的前50个甲基化位点的例示性生物时钟与实足年龄相关

具体实施方式

[0063] 现将通过非限制性实施例来描述本发明的各优选特征和实施方案。技术人员将理解,在不脱离所公开的本发明范围的前提下,他们可以组合本文所公开的本发明的所有特征。

[0064] 必须指出的是,如本文和所附权利要求中所用的,单数形式“一个”、“一种”和“该”包括复数指代物,除非上下文另外明确规定。

[0065] 如本文所用,术语“包含(”和“由……构成”与“包括”或“含有”同义,并且包括端值在内或是开放式的,并且不排除附加的未列举的成员、要素或方法步骤。术语“包含”和“由……构成”也包括术语“由……组成”。

[0066] 数字范围包括限定该范围的数字。

[0067] 本文所讨论的出版物仅是对于其在本专利申请的提交日期之前的公开内容而提

供的。本文中的任何内容均不应理解为承认此类出版物构成本文所附权利要求的现有技术。

[0068] 本文公开的方法和系统可由兽医、医护专业人员、实验室技术人员、宠物护理提供者等使用。

[0069] 受试者

[0070] 本方法针对犬受试者。因此，本发明的受试者是狗。

[0071] 在一个另选的方面，受试者可以是猫科动物受试者。因此，在本发明的另选方面，受试者是猫。除非另有说明，否则本文所有公开内容同样适用于猫。

[0072] 品种

[0073] 本方法可利用关于狗的品种的信息。例如，狗可分类为玩赏型、小型、中型、大型或巨型品种。合适地，可基于狗的重量对狗品种进行分类。合适地，可基于给定品种的狗的平均重量对狗品种进行分类。

[0074] 合适地，狗可被分类为小型或中型品种。合适地，分类由该品种的成年狗的平均重量确定。合适地，平均重量低于10kg的品种被分类为小型品种和/或平均重量高于10kg的品种被分类为中型品种。

[0075] 在受试者是猫的另选方面，猫可以是家猫。合适地，猫可以是家养短毛猫。

[0076] 性别

[0077] 合适地，狗的性别可被分类为雄性或雌性。

[0078] 实足年龄

[0079] 实足年龄可定义为从受试者出生到给定日期所经过的时间量。实足年龄可用年、月、日等来表示。

[0080] 合适地，本方法可应用于任何实足年龄的狗。在某些实施方案中，狗可为至少约2岁大。合适地，狗可为至少约2岁大、至少约3岁大、至少约4岁大、至少约5岁大、至少约6岁大、至少约7岁大、至少约8岁大、至少约9岁大或至少约10岁大。

[0081] 合适地，狗可为至少约7岁大。

[0082] 样本

[0083] 本发明包括提供或确定来自从受试者获得的一个或多个样本的DNA甲基化图谱的步骤。

[0084] 合适地，样本是血液、毛囊、口腔拭子、唾液或组织样本。

[0085] 合适地，样本是毛囊、口腔拭子或唾液样本。如果例如在兽医环境之外提供样本——例如使用根据本发明的试剂盒，则此类样本类型是特别适用的。

[0086] 合适地，样本源自血液。样本可包含血液成分或可以是全血。样本优选包括全血。样本可包括外周血单核细胞(PBMC)或淋巴细胞样本。从受试者收集样本并从样本中提取DNA(例如基因组DNA)的技术是本领域熟知的。

[0087] 可对从受试者获得的一个或多个样本执行本方法。例如，可使用在给定时间点获得的第一样本和在获得第一样本后的时间间隔之后获得的第二样本来执行该方法。可对在一段时间内从同一只狗获得的样本执行该方法不止一次。例如，可每月一次、每年一次或每两年一次重复获得样本。合适地，样本可大约每年获得一次(例如，在年度兽医健康检查期间)。这可能有助于确定特定治疗或生活方式改变，诸如膳食干预或运动方案改变的效果。

[0088] 在一个实施方案中,该方法可应用于在生活方式改变(例如膳食产品干预或运动方案改变)之前从受试者获得的样本。在另一个实施方案中,该方法可应用于在例如膳食产品干预或运动方案改变之前和之后从受试者获得的样本。该方法还可应用于在例如整个膳食产品干预或运动方案改变的过程中的预定时间获得的样本。这些预定时间在例如整个膳食产品干预或运动方案改变的过程中可以是周期性的,例如每天或每三天,或者可取决于所测试的受试者。

[0089] DNA甲基化

[0090] DNA甲基化是将甲基基团(CH_3)共价添加到作为DNA分子一部分的胞嘧啶碱基上的过程。在体内,该过程由DNA甲基转移酶(Dnmt)家族催化,其通过从S-腺苷甲硫氨酸(SAM)转移甲基基团而产生修饰的胞嘧啶。胞嘧啶在第5个碳原子上被修饰,并且修饰的残基称为5-甲基胞嘧啶(5mC)。DNA甲基化也可包括5-羟甲基胞嘧啶(5hmC)。

[0091] DNA甲基化是表观遗传机制的一个示例,即它能够在不修饰潜在DNA序列的情况下修饰基因表达。DNA甲基化可例如通过充当抑制因子的募集信号或通过直接阻断转录因子募集来抑制基因的表达。DNA甲基化主要发生在哺乳动物体细胞的基因组中形成二核苷酸(CpG)的相邻胞嘧啶和鸟嘌呤的位点处。虽然在胚胎发育中观察到非CpG甲基化,但在成人中这些修饰在大多数细胞类型中大大减少。CpG岛是具有高CpG密度但通常未甲基化的DNA片段。这些区域与启动子区域,特别是管家基因的启动子区域相关联,并且被认为保持在允许基因表达的允许状态。

[0092] 已经发现DNA甲基化在人类和其他动物中随年龄而变化。老化的哺乳动物组织显示总体DNA低甲基化,这被认为是由于DNMT1甲基转移酶活性的逐渐丧失或错误靶向,但CpG岛的局部超甲基化。局部超甲基化可导致某些基因的抑制,并且这可促成年龄相关疾病。DNA甲基化的表观遗传变化与年龄之间的联系允许使用“DNA甲基化时钟”来估计“生物学年龄”。通常,已经使用监督机器学习方法针对实足年龄来训练这些时钟,并且“时钟年龄”与个体的实际实足年龄的偏差被认为是“生物学”年龄的指标。这与个体的实足年龄相关,但与相关性的偏差可指示个体中年龄相关疾病或疾患的潜在风险。

[0093] 特异性甲基化的DNA的检测可通过多种方法完成(参见例如Zuo等人,2009; Epigenomics.1(2):331-345)和Rauluseviciute等人;Clinical Epigenetics;2019;11(193))。许多方法可用于检测样本(诸如血液、尿液、粪便或唾液)中特定基因座处的差异甲基化的DNA。这些方法能够将5-甲基胞嘧啶或甲基化的DNA与未非甲基化的DNA区分开,并随后对特定基因组位点的甲基化和未甲基化的DNA的比例进行量化。

[0094] 本方法可包括使用任何合适的方法来确定狗的DNA甲基化图谱。合适的方法包括但不限于下文所述的那些。

[0095] 酶促甲基化测序法(EM-seq)

[0096] 合适地,酶促方法用于检测5mC和5hmC。举例来说,可使用酶促甲基化测序法(EM-seq)。

[0097] 通常在EM-seq中,在第一个酶促步骤中,通过Tet甲基胞嘧啶双加氧酶2(TET2)的活性将5mC氧化为5hmC,然后氧化为5fC,最后氧化为5caC。此外,T4-BGT酶的使用使预先存在的5hmC和由TET2活性产生的5hmC两者糖基化。在第二个酶促步骤中,在双链DNA变性后,使用酶载脂蛋白B mRNA编辑酶催化多肽样3A(APOBEC3A)来使胞嘧啶脱氨基,但不能使5mC

和5hmC的氧化或糖基化形式脱氨基。只有未甲基化的胞嘧啶被脱氨基形成尿嘧啶碱基。在第一个酶促步骤之前,DNA片段可由机械剪切和末端修复、A-加尾以及连接至测序衔接子来产生,这可使用例如**NEBNext**[®] DNA Ultra II试剂(NEB)来进行。在第二个酶促步骤后,可通过PCR反应,使用可扩增含有尿嘧啶的模板的聚合酶(诸如**NEBNext**[®] Q5U[™])来扩增脱氨基的单链DNA,并且可以与通过亚硫酸氢盐测序产生的DNA样本相同的方式对所得文库进行测序或分析。EM-seq的输出通常与全基因组亚硫酸氢盐测序相同,但使用较少的DNA损伤试剂,这因此减少样本损失,并且在胞嘧啶甲基化调用的覆盖率、灵敏度和准确度方面可优于亚硫酸氢盐转化制备的样本。例示性EM-seq方法由Vaisvila等人(Genome Research; 2021;31:1-10)所述。

[0098] 基于亚硫酸氢盐转化的方法

[0099] 当用亚硫酸氢钠处理时,亚硫酸氢盐转化利用未甲基化的胞嘧啶向尿嘧啶的选择性转化。变性的DNA用亚硫酸氢钠进行处理,亚硫酸氢钠将所有未修饰的胞嘧啶转化为尿嘧啶,并且随后的PCR扩增将这些残基转化为胸腺嘧啶。分析产生的DNA序列可经由许多不同的方法进行,该方法的示例包括但不限于:变性凝胶电泳、单链构象多态性、解链曲线、荧光实时PCR(MethyLight)、MALDI质谱、阵列杂交和测序(例如全基因组亚硫酸氢盐测序WGBS)。最近开发的技术(诸如SeqCap Epi)在测序之前富集所关注序列,这使得能够在更集中的区域进行更深入的覆盖。亚硫酸氢盐转化的样本中的序列丰度与未处理的对照的序列丰度的比较允许分析靶位点处的甲基化,其中转化的序列的比例指示靶位点处的甲基化水平。

[0100] 亚硫酸氢盐转化方法的其他变型是可用的,其能够将5mC与氧化形式5-羟甲基胞嘧啶(5hmC)区分开,后者在标准亚硫酸氢盐转化下表现与5mC相同,并且能够检测进一步的修饰5-甲酰基胞嘧啶(5fC)。这些方法,诸如oxBS-Seq和redBS-Seq,利用这些标志物的氧化和还原来改变每种物质对亚硫酸氢盐转化的敏感性,并且通过比较分析来量化靶基因座处的每种修饰的量。

[0101] 选择性限制性核酸内切酶消化方法

[0102] 存在的分析DNA甲基化模式的方法可涉及限制性酶的使用。这些方法包括例如限制性界标基因组扫描(RLGS)(Costello等人,2000;Nat Genet.;24(2):132-8)、甲基化敏感性代表性差异分析(MS-RDA)(Ushijima等人,Proc Natl Acad Sci U S A.1997年3月18日;94(6):2284-9)和差异甲基化杂交(DMH)(Huang等人,Cancer Res.1997年3月15日;57(6):1030-4)。限制性核酸内切酶的消化活性可以是甲基化依赖性的。这种特异性可用于区分甲基化的序列和未甲基化的序列。某些限制性酶(例如BstUI、HpaII和NotI)对甲基化的识别序列敏感。其他限制性酶(诸如McrBC)对甲基化的序列具有特异性。

[0103] 例如,差异甲基化杂交(DMH)(Huang等人,如上所述)需要用大量基因组限制性酶(诸如MseI)对基因组进行初始片段化,该酶将基因组片段化为小于200bp的长度。在该步骤后,使用甲基化敏感的限制性核酸内切酶(MRE)或在该技术的一些版本中使用MRE的混合物来对基因组片段进行消化以改善覆盖率。根据所使用的一种或多种酶的特异性,甲基化的序列或未甲基化的序列将被降解。消化的序列在随后的PCR步骤中不会被扩增。所得PCR产物适用于通过测序或微阵列杂交与荧光染料组合进行进一步处理和分析。

[0104] 合适地,本方法利用通过包括使用一种或多种MRE的方法产生的DNA甲基化图谱。

[0105] 合适的比较物可用于研究条件之间的甲基化状态。可将来自健康受试者的DNA与老年或患病受试者进行比较以检测甲基化状态的变化 (Huang等人, Hum Mol Genet. 1999年3月; 8(3): 459-70)。另选地, 甲基化不敏感形式的二级消化酶 (诸如HpaII同裂酶MspI) 可用于产生对照样本, 使得可进行基因组内或基因组间DNA甲基化比较 (Khulan等人, Genome Res. 2006年8月; 16(8): 1046-55)。

[0106] 在一些实施方案中, 用于检测甲基化的方法包括对基因组DNA进行随机剪切或随机片段化, 用甲基化依赖性或甲基化敏感的限制性酶切割DNA, 随后选择性鉴定和/或分析切割或未切割的DNA。选择性鉴定可包括例如分离切割和未切割的DNA (例如, 通过大小) 并对切割的或另选地未切割的所关注序列进行量化。另选地, 该方法可涵盖在限制性酶消化后扩增完整的DNA, 从而仅扩增在扩增区域中未被限制性酶切割的DNA。在一些实施方案中, 可使用基因特异性的引物进行扩增。另选地, 可将衔接子添加到随机片段化的DNA的末端, 可用甲基化依赖性或甲基化敏感的限制性酶对DNA进行消化, 可使用与衔接子序列杂交的引物来扩增完整的DNA。在这种情况下, 可进行第二步以确定DNA扩增池中的特定基因的存在、不存在或量。在一些实施方案中, 使用实时定量PCR来扩增DNA。

[0107] 合适地, 通过探针或引物与未消化的核酸的选择性杂交来检测核酸的消化。另选地, 探针选择性地与消化的核酸和未消化的核酸两者杂交, 但例如通过电泳促进两种形式之间的区分。用于实现与杂交探针的选择性杂交的合适的检测方法包括例如Southern或其他核酸杂交。

[0108] 合适的杂交条件可基于包含探针的核酸双链体的解链温度 (T_m) 来确定。本领域技术人员将意识到, 对于每种探针, 最佳杂交反应条件应根据经验确定, 但可应用一些一般性原则。优选地, 以低至中等严格性进行采用短寡核苷酸探针的杂交。在富含GC的探针或引物或更长的探针或引物的情况下, 高严格性杂交和/或洗涤是优选的。高严格性在本文中定义为在约 $0.1 \times$ SSC 缓冲液和/或约 0.1% (w/v) SDS 或更低盐浓度中和/或在至少 65°C 的温度或等效条件处进行的杂交和/或洗涤。本文提及的特定严格性水平涵盖使用除本领域技术人员已知的SSC以外的洗涤/杂交溶液的等效条件。

[0109] 基于亲和力富集的方法

[0110] 通过使用含有甲基-CpG-结合结构域 (MBD) 的抗体 (诸如抗5mC和/或甲基化的CpG结合蛋白) 可实现甲基化的DNA与未甲基化的DNA的区分。MBD-结构域蛋白的抗体能够相对于未甲基化的DNA特异性分离甲基化的DNA。利用抗体的方法通常称为MeDIP, 而利用甲基化的CpG结合蛋白的方法通常称为MBD或MIRA方法。

[0111] 这些方法需要基因组的初始片段化, 这可通过用频繁切割的酶 (诸如MseI) 进行大量基因组消化, 随后进行甲基化的片段的亲和纯化来进行。可通过微阵列杂交或测序将输入DNA与纯化的甲基化的DNA进行比较, 以获得特定位点处的甲基化水平的比较分析。

[0112] 基于亲和力富集的方法的其他变型是可用的, 诸如MethylCap-Seq或MBD-Seq。这些方法通过使用盐梯度以甲基-CpG-丰度依赖性方式洗脱甲基化的DNA片段, 将CpG岛和其他高度甲基化的基因座与CpG密度较低的基因座分离来降低样本复杂性。然后可分别对这些部分进行测序, 从而提高序列覆盖率。

[0113] 基于单分子测序的方法

[0114] 当代测序方法能够直接对单个分子进行测序。单分子实时 (SMRT) DNA测序是可用

的,例如来自Pacific Biosciences的Sequel系统,并且已经显示能够基于聚合酶动力学鉴定修饰的碱基(诸如甲基化的胞嘧啶)。能够对长链DNA进行单独测序的纳米孔测序装置(诸如来自Oxford Nanopore Technologies的MinION纳米孔测序仪)也能够检测碱基修饰,包括甲基化。

[0115] DNA甲基化位点

[0116] 合适地,DNA甲基化位点可以指在单个胞嘧啶,合适地单个CpG二核苷酸处存在或不存在5mC。

[0117] 合适地,DNA甲基化位点可以指跨DNA区域内的多个CpG位点的甲基化的存在或不存在(即5mC的数目或5mC的百分比)。合适地,DNA位点甲基化位点可以指跨DNA区域内的多个CpG位点的甲基化水平(即5mC的数目或5mC的百分比)。“DNA区域”可以指基因组DNA的特定部分。这些DNA区域可通过参考基因名称或一组染色体坐标来指定。基因名称和染色体坐标两者是本领域技术人员熟知和理解的。

[0118] 合适地,基因名称和/或坐标可基于“Tasha”狗参考基因组(https://www.ncbi.nlm.nih.gov/assembly/GCF_000002285.5;Jagannathan等人;Genes (Bsacl); 2021;12(6);847)。

[0119] 例如,DNA区域可限定基因的启动子附近的DNA部分。已知启动子区域富含CpG。举例来说,DNA区域可以指启动子上游约3kb至下游约3kb;上游约2kb至下游约2kb;上游约2kb至下游约1kb;上游约2kb至下游约0.5kb;上游约1kb至下游约0.5kb;上游约0.5kb至下游约0.5kb。合适地,DNA区域可以指启动子上游约1kb至下游约0.5kb。

[0120] 合适地,DNA区域可包含相隔小于约5000个、小于约4000个、小于约3000个、小于约2000个、小于约1000个、小于约500个或小于约200个碱基的CpG位点或由其组成。

[0121] 合适地,DNA区域可包含相隔约200至约5000个、约200至约4000个、约200至约3000个、约200至约2000个或约200至约1000个碱基的CpG位点或由其组成。

[0122] 合适地,DNA区域可包含一个或多个CpG岛。合适地,DNA区域可由CpG岛组成。

[0123] “CpG岛”可以指包含至少200bp、GC百分比大于50%和观察到的与预期的CpG比率大于60%的DNA区域。

[0124] 合适地,DNA甲基化位点不包含X和/或Y染色体CpG。

[0125] 合适地,DNA甲基化位点不包含已知在CpG处包含SNP的CpG。

[0126] 对上文详述的每个基因/DNA区域的引用应理解为对这些分子的所有形式及其片段或变体的引用。如本领域技术人员所理解的,已知一些基因在个体之间表现出等位基因变异或单核苷酸多态性。变体包括来自共享至少90%、95%、98%、99%序列同一性的相同区域的核酸序列,即相对于本文所述的DNA区域具有一个或多个缺失、添加、取代、反向序列等。因此,本发明应被理解为延伸到此类变体,就本申请而言,尽管个体之间可能存在实际核酸序列之间的微小遗传变异,但这些变体仍能实现相同的结果。因此,本发明应被理解为延伸到由任何其他突变、多态性或等位基因变异产生的所有形式的DNA。

[0127] 就筛选这些基因区域的甲基化而言,应当理解,该测定可设计成筛选特定DNA。基于染色体坐标选择分析哪条链并靶向该链完全在本领域技术人员的技能范围内。在一些情况下,可建立测定来筛选两条链。

[0128] “甲基化状态”可理解为是指在DNA区域内的一个或多个特定核苷酸处甲基化的存

在、不存在和/或量。特定DNA序列(例如如本文所述的DNA区域)的甲基化状态可指示序列中每个碱基的甲基化状态或者可指示序列内碱基对的子集的甲基化状态(例如胞嘧啶的甲基化状态或一种或多种特定限制性酶识别序列的甲基化状态),或者可指示关于序列内的区域性甲基化密度的信息而不提供序列中何处发生甲基化的确切信息。甲基化状态可任选地由“甲基化值”表示或指示。

[0129] 合适地,可使用EM-Seq策略来确定DNA甲基化。在这种方法中,可将甲基化水平确定为在酶和APOBEC3A转化处理后在靶CpG位点“i”处的“C”+“U”总碱基中的“C”碱基的分数。在其他实施方案中,可将甲基化水平确定为在酶和APOBEC3A转化处理以及随后的核苷酸扩增后在位点“i”处的“C”+“T”总碱基中的“C”碱基的分数。然后可评价每个位点处的平均甲基化水平以确定是否满足一个或多个阈值。

[0130] 在一些实施方案中,特别是当使用亚硫酸氢盐转化和测序方法时,可将甲基化水平确定为在亚硫酸氢盐处理后在靶CpG位点“i”处的“C”+“U”总碱基中的“C”碱基的分数。在其他实施方案中,可将甲基化水平确定为在亚硫酸氢盐处理和随后的核苷酸扩增后在位点“i”处的“C”+“T”总碱基中的“C”碱基的分数。然后可评价每个位点处的平均甲基化水平以确定是否满足一个或多个阈值。

[0131] 另选地,可例如通过对在用甲基化依赖性限制性酶进行限制性消化后存在的完整DNA的量进行定量来产生甲基化值。在该示例中,如果使用定量PCR对DNA中的特定序列进行定量,则约等于模拟处理的对照的模板DNA的量表明该序列不是高度甲基化的,而基本上小于模拟处理的样本中出现的模板的量表明该序列处存在甲基化的DNA。因此,例如来自上述示例的值(即甲基化值)表示甲基化状态,因此可用作甲基化状态的定量指标。当需要将样本中序列的甲基化状态与阈值进行比较时,这是特别有用的。

[0132] 本发明不受被认为指示生物学年齡的甲基化残基的确切数目的限制,因为样本之间将发生一些变化。本发明也不必受甲基化残基(例如,特定甲基化位点)的定位限制。

[0133] 在一个实施方案中,可采用筛选方法,该方法特异性地针对评估一个或多个特定胞嘧啶残基或相对DNA链上位置n+1处的对应胞嘧啶的甲基化状态。

[0134] 富集和检测方法

[0135] 确定DNA甲基化图谱可包括富集DNA样本中所选DNA区域的步骤。例如,该方法可包括富集DNA样本中包含DNA甲基化位点的DNA区域的步骤,该DNA甲基化位点包含DNA甲基化图谱。

[0136] 合适的富集方法是本领域已知的并且包括例如基于扩增或杂交的方法。扩增富集通常是指例如使用针对待富集的DNA区域的引物的基于PCR的富集。可使用任何合适的扩增形式,例如聚合酶链反应(PCR)、滚环扩增(RCA)、反向聚合酶链反应(iPCR)、原位PCR、链置换扩增或循环探针技术。

[0137] 杂交富集或基于捕获的富集通常是指使用与待富集的DNA区域杂交的杂交探针(或捕获探针)。

[0138] 杂交探针可直接附接到固体载体,或者可包含一个部分,例如生物素,以允许与适用于捕获生物素部分的固体载体(例如包被有链霉亲和素的珠子)结合。在任何情况下,可捕获包含与探针互补的序列的DNA,从而允许将包含所关注DNA区域的DNA与不包含所关注DNA区域的DNA分开。因此,此类捕获步骤允许富集所关注DNA区域。例如,DNA区域可以是接

近基因启动子的DNA区域。

[0139] 本文所用的阵列可根据探针组成和阵列的期望用途而变化。例如,在阵列中检测到的核酸(或CpG位点)可以是至少10个、100个、1,000个、10,000个、10万个、100万个、1000万个、1亿个或更多。另选地或另外地,检测到的核酸(或CpG位点)可选择为不超过1亿个、1000万个、100万个、10万个、10,000个、1,000个、100个或更少。使用核酸测序方法,诸如本领域已知的那些;例如下一代或大规模平行测序,可获得类似的范围。

[0140] 合适地,富集步骤可在分离或区别处理甲基化和未甲基化的DNA的步骤之前或之后进行。

[0141] 如本文所用,术语“富集”或者“DNA”或“DNA区域”的“富集”意指与起始材料中包含期望序列的DNA的量和/或比例相比,包含期望序列的DNA的(绝对)量和/或比例增加的过程。就这一点而言,通过扩增进行的富集增加了期望序列的量和比例。通过基于捕获的富集进行的富集增加了包含期望序列的DNA的比例。

[0142] 在处理DNA以区分甲基化和未甲基化的位点后,本方法还可包括鉴定甲基化或未甲基化的位点(即在原始样本中)的步骤。

[0143] 鉴定步骤可包括本领域已知的任何合适的方法,例如阵列检测或测序(例如下一代测序)。

[0144] 测序鉴定步骤优选包括下一代测序(大规模平行或高通量测序)。下一代测序方法是本领域熟知的,并且原则上,任何方法都可考虑用于本发明。下一代测序技术可根据制造商的说明书(例如由Roche、Illumina或Applied Biosystems提供)进行。

[0145] 在一个优选的实施方案中,通过使用酶促反应转化DNA甲基化、进行全基因组文库制备以及由测序(EM-Seq)测量甲基化图谱来处理样本。

[0146] 在一个特别优选的实施方案中,通过使用酶促反应转化DNA甲基化、进行全基因组文库制备、将全基因组转化的文库制备物与捕获探针(优选能够捕获基因启动子附近的DNA区域的捕获探针)杂交;以及由测序(EM-Seq)测量甲基化图谱来处理样本。

[0147] DNA甲基化图谱

[0148] “DNA甲基化图谱”或“甲基化图谱”可以指在一个或多个DNA甲基化位点处5mC的存在、不存在、量或水平。优选地,“甲基化图谱”是指多个DNA甲基化位点处5mC的存在、不存在、量或水平。因此,可评估多个位点内每个单独的DNA甲基化位点处5mC的存在、不存在、量或水平,并且有助于确定狗的死亡风险和/或健康寿命的概率。因此可通过组合来自多种DNA甲基化标志物的值来改善该方法的质量和/或功效。

[0149] 合适地,本发明的生物时钟包含来自多个甲基化位点的甲基化图谱。

[0150] 合适地,来自至少3个、至少5个、至少10个、至少20个、至少50个、至少100个、至少200个、至少500个、至少1000个、至少2000个、至少5000个、至少10000个、至少50000个、至少100000个、至少250000个或至少500000个DNA甲基化位点的5mC的存在或不存在可用于确定狗的死亡风险和/或健康寿命(即生物年龄)的概率。

[0151] 合适地,甲基化图谱可以指来自至少100个、至少200个、至少500个、至少1000个或至少2000个DNA甲基化位点的5mC的存在或不存在。

[0152] 合适地,甲基化图谱可以指来自约100个、约200个、约500个、约1000个或约2000个DNA甲基化位点的5mC的存在或不存在。

[0153] 为了产生用于确定死亡风险和/或健康寿命的概率的生物时钟,可处理或简化初始甲基化图谱以产生限制性甲基化图谱,然后将其用于产生生物时钟。

[0154] 举例来说,可通过例如使用DNA区域而不是单个胞嘧啶,通过选择与特定生理或生化途径相关联的甲基化位点的子集,进行相关性分析并保留每个簇的一个或多个代表性DNA甲基化位点,或进行差异分析以预先选择DNA甲基化位点或保留在年轻狗和年长狗之间变化更大的DNA甲基化位点来处理或简化初始甲基化图谱,

[0155] 例如,DNA区域可以是如本文所定义的任何DNA区域。

[0156] 合适地,甲基化图谱可以指与特定生理或生化途径相关联的基因的DNA甲基化位点。因此,甲基化图谱可以使得能够确定特定组织、器官或生理系统的生物年龄。确定特定组织、器官或生理系统的生物年龄可有利地允许以集中于该组织、器官或生理系统的病理和疾病的方式使用该方法。例如,如果已知特定品种的狗与肌肉或心血管疾病相关联,则确定该生理系统的生物年龄可能是有利的。

[0157] 合适地,生理系统可以是炎症系统、肌肉系统、心血管系统和/或神经系统。

[0158] 特定组织、器官或生理系统的生物年龄可使用DNA甲基化图谱来确定,该DNA甲基化图谱包含来自该组织、器官或生理系统优先或特异性表达的基因的甲基化位点或由其组成。通过特定组织、器官或生理系统对基因的分类可在例如Gene Ontology (<http://geneontology.org/>)、KEGG途径数据库 (<https://www.genome.jp/kegg/>) 或MSigDB (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) 处公开获得。

[0159] 在一些实施方案中,阈值选择对于表观遗传年龄预测因子具有最高排序的平均甲基化值的那些位点。例如,阈值可以是具有平均甲基化水平的那些位点,该平均甲基化水平是针对预测因子(例如生物时钟)测试的所有位点“i”的平均甲基化水平的前50%、前40%、前30%、前20%、前10%、前5%、前4%、前3%、前2%或前1%。

[0160] 另选地,阈值可以是平均甲基化水平处于大于或等于50、60、70、80、90、95、96、97、98或99的百分等级的那些位点。在其他实施方案中,阈值可基于平均甲基化水平的绝对值。例如,阈值可以是平均甲基化水平大于99%、大于98%、大于97%、大于96%、大于95%、大于90%、大于80%、大于70%、大于60%、大于50%、大于40%、大于30%、大于20%、大于10%、大于9%、大于8%、大于7%、大于6%、大于5%、大于4%、大于3%或大于2%的那些位点。相对阈值和绝对阈值可单独或组合应用于每个位点“i”处的平均甲基化水平。作为组合阈值应用的说明,可选择在通过平均甲基化水平测试的所有位点的前3%中并且还具有大于6%的绝对平均甲基化水平的位点的子集。该选择过程的结果是特定超甲基化位点(例如,CpG位点)的DNA甲基化图谱,其被认为对死亡风险和/或健康寿命确定的概率最具信息性。

[0161] 合适地,DNA甲基化图谱可包含至少一个如表3中所列的甲基化位点。

[0162] 合适地,甲基化位点可定义为存在于SEQ ID NO:1-261中的任一者或多者中的甲基化标志物。SEQ ID NO:1-261显示了“Tasha”狗参考基因组中甲基化标志物的任一侧的序列 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000002285.5; Jagannathan等人; Genes (Bsa1); 2021; 12 (6); 847)。“CG”甲基化标志物是序列中的第26个和第27个核苷酸(即在甲基化标志物之前有25个核苷酸,在甲基化标志物之后有25个核苷酸)。

[0163] 合适地,甲基化位点可定义为表3中标记为“位点”的列中的插入位置。例如,对于

位点chr1:3844418-3844420,甲基化标志物是chr1:3844419。合适地,DNA甲基化图谱可包含至少3个、至少5个、至少10个、至少20个、至少50个、至少100个、至少150个、至少200个或优选地如表3中所列的甲基化位点中的每个甲基化位点。

[0164] 合适地,DNA甲基化图谱可包含甲基化位点chr2.32494387.32494389;chr22.46374563.46374565;和chr6.45773846.45773848。这些位点在表4中示出。

[0165] 合适地,DNA甲基化图谱可包含甲基化位点chr2.32494387.32494389;chr22.46374563.46374565;chr6.45773846.45773848;chr5.61645225.61645227;和chr3.70831746.70831748。这些位点在表5中示出。

[0166] 合适地,DNA甲基化图谱可包含甲基化位点chr2.32494387.32494389;chr22.46374563.46374565;chr6.45773846.45773848;chr5.61645225.61645227;chr3.70831746.70831748;chr15.10856498.10856500;chr16.8886545.8886547;chr3.46843750.46843752;chr20.44942213.44942215;和chr20.57347921.57347923。这些位点在表6中示出。

[0167] 合适地,DNA甲基化图谱可包含甲基化位点chr2.32494387.32494389;chr22.46374563.46374565;chr6.45773846.45773848;chr5.61645225.61645227;chr3.70831746.70831748;chr15.10856498.10856500;chr16.8886545.8886547;chr3.46843750.46843752;chr20.44942213.44942215;chr20.57347921.57347923;chr24.21051024.21051026;chr33.26512711.26512713;chr23.37987526.37987528;chr5.32347978.32347980;chr6.60675998.60676000;chr19.44679975.44679977;chr10.7411293.7411295;chr33.26512692.26512694;chr1.121182146.121182148;和chr17.62156690.62156692。这些位点在表7中示出。

[0168] 合适地,DNA甲基化图谱可包含甲基化位点chr2.32494387.32494389;chr22.46374563.46374565;chr6.45773846.45773848;chr5.61645225.61645227;chr3.70831746.70831748;chr15.10856498.10856500;chr16.8886545.8886547;chr3.46843750.46843752;chr20.44942213.44942215;chr20.57347921.57347923;chr24.21051024.21051026;chr33.26512711.26512713;chr23.37987526.37987528;chr5.32347978.32347980;chr6.60675998.60676000;chr19.44679975.44679977;chr10.7411293.7411295;chr33.26512692.26512694;chr1.121182146.121182148;chr17.62156690.62156692;chr35.23853863.23853865;chr17.49222883.49222885;chr14.5949013.5949015;chr5.67427228.67427230;chr20.43366718.43366720;chr24.12518910.12518912;chr29.19479611.19479613;chr4.36180013.36180015;chr35.23852932.23852934;chr1.58324050.58324052;chr9.12522577.12522579;chr5.32711703.32711705;chr33.23473772.23473774;chr2.51831832.51831834;chr20.46098568.46098570;chr11.24476652.24476654;chr1.100289525.100289527;chr17.33851462.33851464;chr6.12827818.12827820;chr20.54301453.54301455;chr1.19541665.19541667;chr5.55848493.55848495;chr10.44524681.44524683;chr9.50976682.50976684;chr33.26512695.26512697;chr24.24968751.24968753;chr10.18237423.18237425;chr35.23853763.23853765;chr9.1046331.1046333;和chr20.49940551.49940553。

[0169] 指示死亡风险和/或健康寿命的概率的DNA甲基化位点/DNA甲基化图谱的确定

[0170] 本发明包括利用DNA甲基化图谱来确定狗的死亡风险和/或健康寿命的概率。因此,本发明包括利用DNA甲基化图谱产生与死亡风险和/或健康寿命的概率相关联的生物时钟。本发明的生物时钟也可称为“表观遗传时钟”。

[0171] 例如,可通过训练数据集和机器学习方法来实现指示死亡风险和/或健康寿命的概率的DNA甲基化位点或DNA甲基化图谱的提供。合适地,机器学习方法可以是监督机器学习方法。

[0172] 举例来说,DNA甲基化位点或DNA甲基化图谱可针对包括具有已知死亡结果(存活或死亡)和实足年龄的狗的数据集进行训练。合适地,DNA甲基化位点或DNA甲基化图谱可针对包括具有已知死亡结果和实足年龄与已知品种和/或性别的组合的狗的数据集进行训练。

[0173] 例如,可通过使用机器学习框架针对具有已知死亡结果(存活或死亡)和实足年龄的狗的训练数据集来训练多个DNA甲基化位点处的甲基化状态数据集,并且针对保留的队列进行测试以验证模型的准确性,来提供指示死亡风险和/或健康寿命的概率的DNA甲基化位点或DNA甲基化图谱的模型。

[0174] 机器学习框架可包括例如使用glmnet R包将惩罚模型拟合到具有已知死亡结果(存活或死亡)和实足年龄(以及任选的品种和/或性别)的狗的训练数据集。

[0175] 合适地,惩罚模型可以是例如惩罚Cox回归、解的最小角回归路径(LARS)Cox回归或惩罚生存模型。

[0176] 机器学习框架可包括例如使用glmnet R包将惩罚Cox回归拟合到具有已知死亡结果(存活或死亡)和实足年龄(以及任选的品种和/或性别)的狗的训练数据集。

[0177] 合适地,机器学习框架可包括拟合由DNA甲基化图谱和实足年龄(以及任选的品种和/或性别)解释的已知死亡结果(存活或死亡)/存活率的惩罚模型,优选惩罚Cox回归。

[0178] 合适地,机器学习框架可包括拟合由DNA甲基化图谱、实足年龄、品种和性别解释的已知死亡结果(存活或死亡)/存活率的惩罚模型,优选惩罚Cox回归。

[0179] 如本文所用,“已知死亡结果(存活或死亡)”也可称为“存活率”。

[0180] 合适地,机器学习框架可用于确定包括指示死亡风险和/或健康寿命的概率的一组DNA甲基化位点或DNA甲基化图谱的模型。

[0181] 模型可包括多个DNA甲基化位点处的甲基化状态;其中在模型中通过乘以系数值来考虑每个位点处的甲基化状态。

[0182] 合适地,性别可被编码为数值,其中0代表雌性,而1代表雄性。

[0183] 合适地,品种可被编码为数值,其中0代表小型品种,而1代表中型品种。

[0184] 狗的生物学年龄可用年、月、日等来表示。

[0185] 每个参数的系数值通常取决于模型中的所有变量的测量单位。如技术人员将理解,每个系数值的值因此将取决于例如模型中使用的不同参数的数量和性质以及所提供的训练数据的性质。因此,可将常规统计方法应用于训练数据集,以便得到系数值。这种方法包括例如在训练集上计算两个gompertz或weibull函数(例如,在已知狗的状态(存活或死亡)的情况下),一个函数将存活率建模为甲基化图谱、实足年龄、品种类别(小型或中型狗)和性别的函数(模型1),而第二个函数仅考虑实足年龄、品种类别和性别(模型2)。这些模型

可在R软件环境中使用flexsurv包(v2.1)拟合。

[0186] 生物学年龄可被定义为时间变量(“实足年龄”),在该时间变量下,模型2给出的动物的存活概率等于模型1给出的在它们的实足年龄下的存活概率。

[0187] 可通过针对在DNA样本收集年龄预测的表型年龄来训练多个DNA甲基化位点处的甲基化状态的数据集,并且针对保留的队列进行测试以验证模型的准确性,来提供指示死亡风险和/或健康寿命的概率的DNA甲基化位点或DNA甲基化图谱的模型。

[0188] 用于确定狗或猫的表型年龄的方法分别描述在PCT/EP2023/061058和PCT/EP2023/061059中。表型年龄的计算考虑了血液生物标志物对死亡风险和/或健康寿命的概率的直接预测值。举例来说,给定的生物标志物可能不与实足年龄直接相关,但是可指示特定的病理状况,从而指示死亡风险增加和/或健康寿命减少的概率。

[0189] 确定狗的表型年龄可包括确定从狗获得的一个或多个样本中的一种或多种生物标志物的水平,其中该一种或多种生物标志物选自白细胞计数、血清白蛋白、血清碱性磷酸酶、血清肌酸激酶、血红蛋白、血细胞比容、平均红细胞血红蛋白、血清葡萄糖、平均红细胞体积、血清球蛋白、血清钙、血小板计数和/或红细胞计数。

[0190] 合适地,狗的表型年龄可由以下项提供:

[0191] a. 确定以下生物标志物的水平;从所述狗获得的一个或多个样本中的白细胞计数、血清白蛋白、血清碱性磷酸酶、血清肌酸激酶、血红蛋白、血细胞比容、平均红细胞血红蛋白、血清葡萄糖、平均红细胞体积和血清球蛋白;以及

[0192] b. 使用公式(1)确定狗的表型年龄(Phenoage):

$$[0193] \quad \text{Phenoage} = \ln \left(\frac{\gamma_{\text{品种}} * e^{xb * \{e^{\gamma * \text{年龄}} - 1\}}}{e^{\left\{ \text{品种} * \beta_{\text{品种}2} \right\} + \left\{ \text{性别} * \beta_{\text{性别}2} \right\} + \beta_{02} * \gamma}} + 1 \right) * \frac{1}{\gamma_{\text{品种}}}$$

[0194] 其中xb是根据公式(2)的每种生物标志物、性别和品种的值乘以它们的相应系数的总和:

$$[0195] \quad xb = \sum_{u=1}^p x_u \beta_u + \beta_0$$

[0196] 其中性别被编码为数值,其中0代表雌性,而1代表雄性,

[0197] 其中品种被编码为数值,其中0代表小型品种,而1代表中型品种,

[0198] 并且其中使用表型年龄来确定狗的死亡风险和/或健康寿命的概率。

[0199] 每个参数的系数值通常取决于模型中的所有变量的测量单位。如技术人员将理解,每个系数值的值因此将取决于例如模型中使用的不同参数的数量和性质以及所提供的训练数据的性质。因此,可将常规统计方法应用于训练数据集,以便得到用于上述公式的系数值。这种方法包括例如在训练集上计算两个gompertz或weibull函数(例如,在已知狗的状态(存活或死亡)的情况下),一个函数将存活率建模为所选生物标志物、实足年龄、品种类别(小型或中型狗)和性别的函数(模型1),而第二个函数仅考虑实足年龄、品种类别和性别(模型2)。这些模型可在R软件环境中使用flexsurv包(v 2.1)拟合。

[0200] 合适地,给定生物标志物的负系数意味着生物标志物的较高水平对降低死亡风险

具有正面影响,和/或生物标志物的较低水平对降低死亡风险具有负面影响。合适地,给定生物标志物的正系数意味着生物标志物的较高水平对降低死亡风险具有负面影响,和/或生物标志物的较低水平对降低死亡风险具有正面影响。

[0201] 例示性系数和 γ 和 $\gamma_{\text{品种}}$ 值在下表中提供。

[0202]		系数
	γ	0.491790219
	β_0	-6.036261473
	β 白细胞计数	0.091862564
	β 血红蛋白	-0.009131623
	β 平均红细胞体积	-0.007486146
	β 血细胞比容	-0.018418391
	β 平均红细胞血红蛋白	-0.128195615
	β 血清葡萄糖	0.009169677
	β 血清球蛋白	0.132755858
	β 血清肌酸激酶	0.332818902
	β 血清白蛋白	-0.744060565
	β 血清碱性磷酸酶	0.262594338
[0203]	β 品种	1.138018960
	β 性别	0.151826455
	$\gamma_{\text{品种}}$	0.5668399
	β_{02}	-9.5204440
	$\beta_{\text{品种}2}$	1.2299804
	$\beta_{\text{性别}2}$	0.2678798

[0204] 表型年龄可被定义为时间变量(“实足年龄”),在该时间变量下,模型2给出的动物的存活概率等于模型1给出的在它们的实足年龄下的存活概率。

[0205] 狗的表型年龄可用年、月、日等来表示。

[0206] 用于确定表型年龄的生物标志物可使用本领域的标准方法来确定,并且通常作为标准血液测试的一部分来测量,以确定动物的疾病状态。例如,生物标志物通常被确定为标准临床全血细胞计数(cbc)和标准临床血液化学分析的一部分。

[0207] 合适地,指示针对表型年龄训练的死亡风险和/或健康寿命的概率的DNA甲基化位点或DNA甲基化图谱的模型可在两步过程中提供。

[0208] 在第一步中,机器学习框架可包括例如使用glmnet R包拟合由如本文所述的一种或多种血液生物标志物和实足年龄(以及任选的性别和/或品种)解释的表型年龄(PhenoAge)的惩罚模型。优选地,机器学习框架可包括拟合由如本文所述的一种或多种血液生物标志物、实足年龄、性别和品种解释的表型年龄(PhenoAge)的惩罚模型。

[0209] 合适地,惩罚模型可以是例如惩罚Cox回归、解的最小角回归路径(LARS)Cox回归或惩罚生存模型。

[0210] 机器学习框架可包括拟合由如本文所述的一种或多种血液生物标志物、实足年龄、性别和品种解释的表型年龄(PhenoAge)的惩罚Cox回归。

[0211] 在第二步中,机器学习框架可包括拟合由DNA甲基化解释的表型年龄的惩罚回归。合适地,机器学习框架可包括拟合由DNA甲基化图谱解释的表型年龄的惩罚回归。

[0212] 惩罚回归可以是弹性网络回归。

[0213] 如本文所用,术语“一种或多种生物标志物”可包括至少一种、至少两种、至少三种、至少四种、至少五种、至少六种、至少七种、至少八种、至少九种、至少十种、至少十一种、至少十二种或至少十三种生物标志物。

[0214] 如本文所用,术语“一种或多种生物标志物”可包括一种、两种、三种、四种、五种、六种、七种、八种、九种、十种、十一种、十二种或十三种生物标志物。

[0215] 合适地,DNA甲基化位点或DNA甲基化图谱可与本文所述的一种或多种血液生物标志物的水平组合以便产生指示死亡风险和/或健康寿命的概率的模型。例如,可通过针对具有已知死亡结果(存活或死亡)和实足年龄的狗的训练数据集来训练多个DNA甲基化位点处的甲基化状态和一种或多种血液生物标志物的水平的数据集,并且针对保留的队列进行测试以验证模型的准确性,来提供包括本文所述的DNA甲基化图谱和一种或多种血液生物标志物的水平的组合的模型。

[0216] 机器学习框架可包括例如使用glmnet R包将惩罚回归拟合到具有已知死亡结果(存活或死亡)和实足年龄(以及任选的品种和/或性别)的狗的训练数据集。

[0217] 机器学习框架可包括例如使用glmnet R包将惩罚Cox回归拟合到具有已知死亡结果(存活或死亡)和实足年龄(以及任选的品种和/或性别)的狗的训练数据集。

[0218] 合适地,机器学习平台可包括一个或多个深度神经网络。神经网络是在非循环图中连接的神经元(也称为单元)的集合。神经网络模型通常被组织成不同的神经元层。对于大多数神经网络,最常见的层类型是全连接层,其中两个相邻层之间的神经元完全成对连接,但单层内的神经元不共享连接。深度神经网络的主要特征之一是神经元由非线性激活函数控制。这种非线性与深度架构组合,使得输入特征的更复杂组合成为可能,最终导致对它们之间关系的更广泛理解,并因此导致更可靠的最终输出。深度神经网络已经应用于范围介于结构数据到化学描述符或转录组学数据的许多类型的数据。

[0219] 合适地,机器学习平台包括一个或多个生成式对抗网络。合适地,机器学习平台包括对抗自动编码器架构。合适地,机器学习平台包括特征重要性分析以用于通过DNA甲基化位点在生物年龄确定中的重要性来对它们进行排序。

[0220] 狗的生物年龄可用年、月、日等来表示。

[0221] 优选地,死亡风险和/或健康寿命的概率被表示为狗的生物年龄和实足年龄之间的差异。

[0222] 与参考或对照进行比较

[0223] 本方法还可包括将测试样本中一个或多个位点处的DNA甲基化差异与一个或多个参考或对照进行比较的步骤。参考或对照中一个或多个位点处的DNA甲基化的存在或不存在可能与预先定义的死亡风险和/或健康寿命(即生物年龄)的概率相关联。在一些实施

方案中,参考值是之前针对具有已知死亡风险和/或健康寿命(即生物学年龄)的概率的受试者或受试者群组获得的值。参考值可基于来自具有已知死亡状态(存活或死亡)、实足年龄、品种和/或性别的受试者群组的一个或多个位点处的已知DNA甲基化状态,例如平均或中值水平。

[0224] 将生物标志物水平与另外的量度和/或特性相组合

[0225] 合适地,本方法还包括将DNA甲基化图谱与狗的实足年龄、品种和/或性别中的一者或多者相组合。通过将该信息相组合,可确定与死亡风险和/或健康寿命的概率相关联的生物学年龄。

[0226] 受试者分类

[0227] 还可将通过本发明的方法确定的生物学年龄与一个或多个预定的阈值(即与实足年龄的差异)进行比较。使用这种阈值,可将受试者分层到指示确定的风险(例如,低、中或高确定风险)的各个类别中。与阈值的发散程度可用于确定哪些个体将最受益于某些干预。这样,可优化膳食干预和生活方式改变。

[0228] 用于选择/监测受试者的生活方式方案、膳食方案或治疗干预的方法

[0229] 在另一个方面,本发明提供了用于为受试者选择生活方式方案、膳食方案或治疗干预的方法。生活方式的改变可以是本文所述的任何改变,例如膳食干预和/或运动方案的改变。生活方式的改变可以是治疗方式的施用。

[0230] 可在任何合适的时间段内将生活方式方案、膳食方案或治疗干预应用于狗。在所述一段时间后,可使用本方法再次确定狗的死亡风险和/或健康寿命的概率,以便确定生活方式方案、膳食方案或治疗干预对降低狗的死亡风险和/或增加狗的健康寿命的概率的功效。举例来说,可应用生活方式方案、膳食方案或治疗干预至少2周、至少4周、至少8周、至少16周、至少32周或至少64周。可应用生活方式方案、膳食方案或治疗干预至少3个月、至少6个月、至少12个月、至少24个月、至少36个月、至少48个月或至少60个月。

[0231] 生活方式方案、膳食方案或治疗干预可称为抗衰老生活方式方案、膳食方案或治疗干预。

[0232] 优选地,该改变是如本文所述的膳食干预。术语“膳食干预”是指应用于受试者并引起受试者的饮食变化的外部因素。更优选地,膳食干预包括至少一种膳食产品或膳食方案或营养补充剂的施用。

[0233] 膳食干预可以是膳食、膳食方案、补充剂或补充剂方案或膳食和补充剂的组合,或膳食和多种补充剂的组合。

[0234] 本文所述的膳食干预或膳食产品可以是任何合适的膳食方案,例如热量限制饮食、老年饮食、低蛋白饮食、磷饮食、低蛋白饮食、钾补充剂饮食、多不饱和脂肪酸(PUFA)补充剂饮食、抗氧化剂补充剂饮食、维生素B补充剂饮食、液体饮食、硒补充剂饮食、 ω 3-6比例饮食、或补充有肉碱的饮食、支链氨基酸或衍生物、核苷酸、烟酰胺前体(诸如烟酰胺单核苷酸(MNM)或烟酰胺核苷(NR))或上述的任何组合。

[0235] 合适地,膳食干预或膳食产品可以是热量限制饮食、老年饮食或低蛋白饮食。合适地,膳食干预或膳食产品可以是热量限制饮食。合适地,膳食干预或膳食产品可以是低蛋白饮食。

[0236] 可基于狗的基线维持能量需求(MER)来确定膳食干预。合适地,MER可以是稳定狗

体重(三周内变化小于5%)的食物量。

[0237] 举例来说,通常认为年轻的正在成长的狗受益于高能量/高蛋白饮食;然而,年长狗可能具有较低的能量需求,因此饮食可以适当调整。特别是,许多制造商生产“老年”系列狗粮,这种狗粮热量较低,纤维含量较高,但蛋白质和脂肪水平适合于年长狗。

[0238] 合适地,热量限制饮食可占狗MER的约50%、约55%、约60%、约65%、约75%、约80%、约85%或约90%。合适地,热量限制饮食可占狗MER的约60%或约75%。

[0239] 合适地,低蛋白饮食可包含少于20%的蛋白质(干物质%)。例如,低蛋白饮食可包含少于19%的蛋白质(干物质%)。

[0240] 这些饮食通常是基于狗的实足年龄而推荐的。例如,可能会推荐,让狗在7、8岁左右改用老年饮食。然而,在本发明的上下文中,相比于在给定实足年龄的情况下所预期而确定狗的死亡风险增加可允许确定让狗在更早的年龄改用老年饮食。相比之下,与狗的实足年龄相比,死亡风险降低的狗可能能够更长时间地继续享用成年饮食。

[0241] 膳食干预可包括食物、补充物和/或饮料,其包含模拟热量限制(CR)的益处而不限制每日热量摄入的营养物和/或生物活性剂。例如,食品、补充剂和/或饮料可包含具有类似CR的益处的功能性成分。合适地,食物、补充物和/或饮料可包含自噬诱导剂。合适地,食物、补充剂和/或饮料可包含水果和/或坚果(或它们的提取物)。合适的示例包括但不限于石榴、草莓、黑莓、卡姆果、核桃、栗子、开心果、碧根果。合适地,食品、补充剂和/或饮料可包含益生菌,含或不含水果提取物或坚果提取物。

[0242] 改变受试者的生活方式还包括指示受试者需要改变生活方式,例如规定更多的运动。类似于膳食干预,相比于在给定实足年龄的情况下所预期而确定狗的死亡风险增加可允许确定让狗改用适当的运动方案。

[0243] 改变受试者的生活方式还包括选择或推荐治疗方式或方案。治疗方式或方案可以是用于治疗 and/或预防例如关节炎、牙科疾病、内分泌病症、心脏病、糖尿病、肝病、肾病、前列腺病症、癌症以及行为或认知障碍的方式。合适地,可对由于死亡风险增加和/或基于已知与疾病相关途径相关联的特定生物标志物而被标识为处于这种病症风险中的狗施用预防性治疗。在其他实施方案中,可更规律地监测被确定为处于某些病状的风险中的狗(由于死亡风险增加和/或基于已知与疾病相关途径相关联的特定生物标志物),以便可尽可能早地开始诊断和治疗。

[0244] 本发明还涉及监测和/或确定抗衰老疗法的功效或开发抗衰老疗法。抗衰老疗法可包括例如“复壮”干预。复壮干预旨在引起受试者的表观遗传或生物学年龄的降低。合适地,复壮干预可将表观遗传年龄重新编程为非常年轻的狗的表观遗传年龄。此类复壮干预的示例包括但不限于合适地将表观遗传年龄重新编程为非常年轻的狗的表观遗传年龄的基因疗法。监测和/或确定生活方式方案、膳食方案或治疗干预的功效或开发生活方式方案、膳食方案或治疗干预以减少生物学年龄的本方法特别适用于该方面。

[0245] 因此,本发明可有利地实现对预期对给定干预(例如生活方式方案、膳食方案或治疗干预)反应特别好的狗的标识。因此,干预可以更有针对性的方式应用于预期有反应的狗。

[0246] 在一个方面,本发明提供了用于确定生活方式方案、膳食方案或治疗干预对降低狗的死亡风险和/或增加狗的健康寿命的概率的功效的方法,所述方法包括:a)将生活方式

方案、膳食方案或治疗干预应用于所述狗,其中已经根据本发明的方法选择所述生活方式方案、膳食方案或治疗干预;b) 在将所述生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;c) 在遵循所述生活方式方案、膳食方案或治疗干预的所述时间段后,确定所述狗的所述死亡风险是否存在变化。

[0247] 在另一个方面,本发明提供了用于确定生活方式方案、膳食方案或治疗干预对降低狗的死亡风险和/或增加狗的健康寿命的概率的功效的方法,所述方法包括:a) 使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;b) 将基于步骤a) 中确定的所述死亡风险和/或健康寿命的概率选择的生活方式方案、膳食方案或治疗干预应用于所述狗;c) 在将生活方式方案、膳食方案或治疗干预应用于所述狗一段时间后;使用来自从所述狗获得的样本的DNA甲基化图谱来确定所述狗的死亡风险和/或健康寿命的概率;d) 确定所述狗的所述死亡风险和/或健康寿命的概率在步骤a) 和步骤c) 之间是否存在变化。

[0248] 合适地,在确定第一死亡风险和/或健康寿命的概率之前,可能已将生活方式方案、膳食方案或治疗干预应用于狗一段时间;然而,仍然可通过在生活方式方案、膳食方案或治疗干预的应用期间的两个或更多个时间确定死亡风险和/或健康寿命的概率来监测生活方式方案、膳食方案或治疗干预对改善狗的死亡风险和/或健康寿命的概率(即降低死亡风险和/或增加健康寿命的概率)的有效性。

[0249] 合适地,本方法可包括“生态系统”;特别是数字生态系统。合适地,本方法可包括提供从狗获得的样本,任选地使用根据本发明的试剂盒;以及(b) 提供用于后续DNA提取的样本(例如通过邮寄)以用于测量从样本提取的DNA中的DNA甲基化,从而获得DNA甲基化图谱。

[0250] 然后可根据本方法中的任一种方法使用DNA甲基化图谱;优选地使用根据本发明的计算机系统或计算机程序产品。

[0251] 然后,计算机系统或计算机程序可以例如为狗选择或推荐合适的生活方式方案、膳食方案或治疗干预的形式准备并共享详述分析/方法的结果或本方法的任何其他结果的报告。

[0252] 合适地,样本可以是可由狗主人在家获得的样本(例如不需要兽医或医护专业人员)。合适地,样本可以是毛囊、口腔拭子或唾液样本。

[0253] 膳食干预的用途

[0254] 在一个方面,本发明提供了用于降低狗的死亡风险和/或增加狗的健康寿命的概率的膳食干预,其中将膳食干预施用于狗,其中死亡风险和/或健康寿命的概率通过本方法确定。

[0255] 在另一方面,本发明提供了膳食干预在降低狗的死亡风险和/或增加狗的健康寿命的概率中的用途,其中将膳食干预施用于狗,其中死亡风险和/或健康寿命的概率通过本方法确定。

[0256] 如本文所述,膳食干预可以是膳食产品或膳食方案或营养补充剂。

[0257] 计算机程序产品

[0258] 本方法可使用计算机来执行。因此,本方法可在计算机上执行。

[0259] 合适地,计算机可准备并共享详述本方法的结果的报告。

[0260] 本文所述的方法可实施为在通用硬件诸如一个或多个计算机处理器上运行的计算机程序。在一些实施方案中,本文所述的功能可通过诸如智能手机、平板终端或个人计算机的装置实施。

[0261] 在一个方面中,本发明提供了一种计算机程序产品,该计算机程序产品包括计算机可实施指令,这些计算机可实施指令用于使可编程计算机如本文所述地确定狗的死亡风险和/或健康寿命的概率。

[0262] 在一个实施方案中,使用者向装置输入如本文定义的DNA甲基化标志物中的一种或多种DNA甲基化标志物的水平,任选地连同实足年龄、品种和性别一起输入。装置然后处理该信息并提供狗的生物学年龄的确定。另选地,装置然后处理该信息并基于生物学年龄为狗提供合适的生活方式方案、膳食方案或治疗干预的确定。

[0263] 该装置通常可以为网络上的服务器。然而,可使用任何装置,只要其可使用处理器、中央处理单元(CPU)等处理生物标志物数据和/或附加参数或特性即可。例如,装置可以是智能电话、平板终端或个人计算机,并且输出指示所确定的狗的生物学年龄或基于生物学年龄的狗的合适的生活方式方案、膳食方案或治疗干预的确定的信息。

[0264] 本领域的技术人员将理解,在不脱离本文所公开的本发明范围的前提下,他们可以自由地组合本文所述的本发明的所有特征。

[0265] 实施例

[0266] 现将通过实施例进一步描述本发明,这些实施例意在帮助本领域的技术人员实施本发明,而不以任何方式限制本发明的范围。

[0267] 实施例1-用于产生表观遗传生物时钟的例示性方法

[0268] DNA甲基化位点的鉴定

[0269] 通过进行DNA提取、通过使用酶促反应转化DNA甲基化、进行全基因组文库制备、将全基因组转化的文库制备物与针对基因启动子的捕获探针杂交以及由测序(EM-Seq)测量甲基化图谱来分析来自犬科动物队列的全血样本。

[0270] 捕获探针针对约40,000个靶标(启动子区域——启动子上游约1kb至下游0.5kb)。这些靶区域包含潜在的所关注的甲基化位点(可被甲基化的单个胞嘧啶残基)。

[0271] 在测序之后和进一步分析之前进行以下生物信息学步骤:

[0272] 使用fastQC进行fastq质量检查——<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

[0273] 使用trimGalore进行衔接子修剪——https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

[0274] 使用bwa-meth(<https://github.com/brentp/bwa-meth>)或Bismark与狗基因组进行比对——<https://www.bioinformatics.babraham.ac.uk/projects/bismark/>

[0275] 使用Picard标记重复项——https://gatk.broadinstitute.org/hc/en-us/articles/360037052812-Mar_kDuplicates-Picard-

[0276] 使用Methyldackel调用甲基化——<https://github.com/dpryan79/Methyldackel>

[0277] 用于过滤位点以产生用于训练生物时钟的限制性DNA甲基化图谱的方法

[0278] 在确定每个样本中的甲基化位点的甲基化状态后,可过滤/处理初始甲基化图谱以便产生包含较少离散甲基化位点的限制性甲基化图谱。这种过滤的目的是提供包含例如50000至500000个可用于训练生物时钟的甲基化位点的限制性甲基化图谱。

[0279] 用于过滤初始甲基化图谱的方法包括:

[0280] 1) 除去所有样本中(未)甲基化的位点

[0281] 2) 除去在至少90%的样本中不具有至少5个计数的位点

[0282] 3) 除去染色体X和/或Y位点

[0283] 减少离散甲基化位点数目的其他潜在过滤步骤包括:

[0284] 1) 将甲基化位点定义为DNA区域(CpG岛,相隔小于例如1000bp的位点被认为是相同的)

[0285] 2) 将靶标限制到例如炎症位点,例如与来自GO/KEGG途径的炎症相关联的基因,并选择这些基因上的位点

[0286] 3) 相关分析和每个簇的代表物的保留(例如,加权相关网络分析(WGCNA)(Langfelder&Horvath;BMC Bioinformatics;9(559);2008或EBModules(Zollinger等人;Biostatistics,19(2),153-168;(2018))。

[0287] 4) 差异甲基化分析以预先选择DNA甲基化位点(例如使用逻辑回归)或鉴定在年轻狗和年长狗之间变化更大的位点(例如使用方差分析)。如果进行差异甲基化分析(使用R中的函数preprocessQuantile),则在过滤步骤之前进行归一化(quantile)。

[0288] 产生生物时钟

[0289] 包括关于狗队列的死亡状态(存活或死亡)、实足年龄、品种和性别的信息的数据集被分成训练集和测试集(例如,2/3的数据用于训练,1/3的数据用于测试,从而确保关于元数据的良好划分;例如在训练集和测试集中每个品种/每个性别的类似比例)。

[0290] 使用甲基化位点、品种/品种类别、性别和实足年龄作为预测因子,将惩罚Cox模型拟合到存活率。使用R中的glmnet包来估计模型的参数以及惩罚参数。

[0291] 进行两个gompertz或Weibull函数的计算,一个函数将存活率建模为甲基化图谱、实足年龄、品种类别(小型或中型狗)和性别的函数(模型1),而第二个函数仅考虑实足年龄、品种类别和性别(模型2)。该模型在R软件环境中使用flexsurv包(v 2.1)拟合。

[0292] 死亡风险和/或健康寿命年龄的概率可被定义为时间变量(“实足年龄”),在该时间变量下,模型2给出的动物的存活概率等于模型1给出的在它们的实足年龄下的存活概率。

[0293] 选择惩罚参数以提供合理数目的DNA甲基化位点(例如最多1000个),其具有对于死亡风险和/或健康寿命的概率的良好模型拟合。

[0294] 然后在测试集上评估基于DNA甲基化位点的生物年龄模型。

[0295] 在本实施例中,进行以下步骤:

[0296] 1) 将具有少于15个计数覆盖的所有甲基化位点设定为“缺失”;

[0297] 2) 使用“Boostme”算法进行插补(Zou等人;BMC Genomics 19,390(2018))并且将训练集、测试集和验证集大小设定为200,000个随机选择的位点;

[0298] 3) 除去ChrX位点;

[0299] 4) 使用EBmodule进行位点过滤(Zollinger等人;Biostatistics,19(2),153-168;

(2018));通过对每个靶标(TSS周围1500bp)和每个染色体的位点进行分组,将数据集分成约5000个位点的块。然后计算应用EB模块的相关矩阵——如Zollinger等人所述。然后由medoid位点表示每个模块。一些位点不属于任何模块并且称为分散位点(如Zollinger等人所述那样定义);

[0300] 5) 除去存在于哺乳动物CpG阵列上的位点(Arneson等人;Nature Communications;13;783;2022);

[0301] 6) 使用phenoAge_pred (DNA收集年龄时表型年龄的预测值)作为响应变量在500K位点上调整弹性网络回归,这选择261个位点作为形成生物时钟的位点(参见表3)。

[0302] 使用一系列评估来验证时钟。图3示出了热量限制中时钟的验证,其中与控制饮食的狗相比,热量限制的狗具有更低的生物年龄(更低的变化量)。图3中的变化量对应于实足年龄相对于预测的phenoDNAmAge的回归模型的残差。

[0303] 进一步的验证在图4中示出,其中两种饮食之间的变化量(phenoDNAmAge)存在差异,限制饮食的狗在6岁时在生物学上更年轻。

[0304] 图5示出了当调整线性混合效应模型(DogID作为随机效应)时,两种饮食之间存在显著差异。

[0305] 图6示出了生物学上更年轻的狗和生物学上更年长的狗之间存活率的显著差异。

[0306] 图7示出了当Cox比例风险模型使用性别和变化量phenoDNAmAge(来自实足年龄相对于预测的phenoDNAmAge的残差)进行拟合,在训练集上按品种类别(小型或中型)分层时,变化量phenoDNAmAge的增加与死亡风险的增加显著相关。

[0307] 图8示出了生物时钟在技术重复之间是高度可重复的。

[0308] 进行单变量分析以确定261个位点中的每个位点单独在与生物年龄相关方面是否具有统计学意义(参见表3)。

[0309] 仅使用表3中示出的完整位点列表中的前3个、前5个、前10个、前20个和前50个位点还产生了进一步的生物时钟,并且每个位点显示与生物年龄相关(参见图9-图13)。这些时钟是通过基于全时钟的系数的绝对值(按降序,首先取大系数)选择前n个位点而产生的。使用前n个位点作为预测因子拟合分别解释实足年龄的线性模型。前3个、前5个、前10个和前20个时钟的细节在表4-表7中示出。

[0310] 实施例2——与狗的死亡风险相关联的血液生物标志物的确定

[0311] 从由标准临床全血细胞计数(cbc)和标准临床血液化学分析组成的生物标志物组确定预测性血液生物标志物。过夜禁食后采集血清样本,并使用标准兽医临床实践进行测量。

[0312] 表1——临床全血细胞计数(cbc)和临床血液化学分析

[0313]	参数名称	量度单位
	血细胞比容	%
	血红蛋白	g/dL
	平均红细胞血红蛋白	pg

[0314]

平均红细胞血红蛋白浓度	g/dL
平均红细胞体积	fL
血小板	$10^3/\mu\text{L}$
红细胞	$10^3/\mu\text{L}$
白细胞	$10^3/\mu\text{L}$
血清白蛋白+	g/dL
血清碱性磷酸酶*	U/L
血清 ALT *	U/L
血清 AST *	U/L
血清钙	mg/dL
血清氯化物	mmol/L
血清胆固醇	mg/dL
血清肌酸激酶*	IU/L
血清肌酸酐, Jaffe 法*	mg/dL
血清 GGT *	g/dL
血清球蛋白	g/dL
血清葡萄糖	mg/dL
血清镁	mg/dL
血清磷	mg/dL
血清钾	mmol/L
血清钠	mmol/L
血清总胆红素*	mg/dL
血清总蛋白质	g/dL
血清甘油三酯*	mg/dL
血清尿素氮*	mg/dL

[0315] *数值使用自然对数进行对数变换。

[0316] 我们对狗进行了纵向研究,对这些参数以及关于狗的状态(存活或死亡)、性别和品种进行了重复测量。我们首先基于该品种的成年犬的平均重量(分别低于10kg或高于10kg)将品种分类为小型或中型。然后我们使用R编程语言组织数据。对于每只狗,我们使用左边开放和右边关闭的时间间隔(即(tstart,tstop))将生物标志物记录为时间依赖性协变量,其中生物标志物信息对应于间隔的开始,并且事件(存活或死亡)被记录为最后的tstop值。为此,我们在R(v.3.2-13)中使用了存活包的tmerge函数。然后,我们针对28种生物标志物中的每种生物标志物,将cox比例危险模型拟合到该数据,包括性别和品种类别(小型或中型)。然后,我们调整每个参数的p值以说明多重比较(按错误发现率(fdr)),并且选择调整后的fdr低于0.05的特征(图1)。

[0317] 使用该方法,我们标识处13种生物标志物,它们可单独预测狗的存活概率:

[0318] 白细胞计数($10^3/\mu\text{L}$)

[0319] 血清白蛋白(g/dL)

[0320] 血清碱性磷酸酶(U/L, $\ln$ 变换)

[0321] 血清肌酸激酶(IU/L, $\ln$ 变换)

[0322] 血红蛋白(g/dL)

[0323] 血细胞比容(%)

[0324] 平均红细胞血红蛋白(pg)

[0325] 血清钠(mmol/L)

[0326] 平均红细胞体积(fL)

[0327] 血清球蛋白(g/dL)

[0328] 血清钙(mg/dL)

[0329] 血清血小板计数($10^3/\mu\text{L}$)

[0330] 红细胞计数($10^3/\mu\text{L}$)

[0331] 实施例3——用于预测死亡风险的多参数模型

[0332] 接下来,我们构建了同时考虑多个参数的最佳模型,因为这更有可能涵盖随着年龄增长而发生的各种器官功能障碍。然而,选择几个可能相互关联的特征会有偏差。为了避免这个问题,我们使用了使用glmnet包(v4.1-3)的惩罚回归方法。我们在数据上拟合了LASSO惩罚cox比例危险模型,并且使用20倍交叉验证来比较惩罚参数 λ 的不同值。这种方法选出了10个最能预测存活率的血液生物标志物,按重要性排序如下:

[0333] • 白细胞计数($10^3/\mu\text{L}$)

[0334] • 血清白蛋白(g/dL)

[0335] • 血清碱性磷酸酶(U/L, $\ln$ 变换)

[0336] • 血清肌酸激酶(IU/L, $\ln$ 变换)

[0337] • 血红蛋白(g/dL)

[0338] • 血细胞比容(%)

[0339] • 平均红细胞血红蛋白(pg)

[0340] • 血清葡萄糖(mg/dL)

[0341] • 平均红细胞体积(fL)

[0342] • 血清球蛋白(g/dL)

[0343] 我们还发现,该列表中的前3种生物标志物是最具预测性的,并且可通过结合接下来的7种生物标志物中的每一种生物标志物来提高性能。

[0344] 为了提取动物的表型年龄,我们在我们的训练集上计算了两个不同的gompertz函数,一个函数将存活率建模为所选生物标志物、年龄、品种类别(小型或中型狗)和性别的函数(模型1),而第二个函数仅考虑年龄、品种类别和性别(模型2)。这些模型使用flexsurv包(v 2.1)进行拟合。表型年龄被定义为时间变量(“年龄”),在该时间变量下,模型2给出的动物的存活概率等于模型1给出的在它们的实足年龄下的存活概率。这导致了将血液生物标志物与表型年龄联系起来的数学函数,并且由以下公式给出:

$$[0345] \quad \text{Phenoage} = \ln \left(\frac{\gamma_{\text{品种}} * e^{xb * \{e^{\gamma * \text{年龄}} - 1\}}}{e^{\left\{ \text{品种} * \beta_{\text{品种}2} \right\} + \left\{ \text{性别} * \beta_{\text{性别}2} \right\} + \beta_{02} * \gamma}} + 1 \right) * \frac{1}{\gamma_{\text{品种}}}$$

[0346] 其中xb是每种生物标志物、性别和品种的值乘以它们的相应系数的总和。性别和品种被编码为数值,对于性别,0代表雌性,而1代表雄性;对于品种,0代表小型品种,而1代表中型品种。系数由在我们的训练集上训练的两个gompertz函数给出。

[0347] 作为示例,系数以及 γ 和 $\gamma_{\text{品种}}$ 值已经针对生物标志物的完整列表从我们的训练集进行了测量,并且在表2中给出。

$$[0348] \quad xb = \sum_{u=1}^p x_u \beta_u + \beta_0$$

[0349] 表2——系数和 γ 和 $\gamma_{\text{品种}}$ 值已经从训练集进行了测量

[0350]

	系数
γ	0.491790219
β_0	-6.036261473
$\beta_{\text{白细胞计数}}$	0.091862564
$\beta_{\text{血红蛋白}}$	-0.009131623
$\beta_{\text{平均红细胞体积}}$	-0.007486146
$\beta_{\text{血细胞比容}}$	-0.018418391
$\beta_{\text{平均红细胞血红蛋白}}$	-0.128195615
$\beta_{\text{血清葡萄糖}}$	0.009169677
$\beta_{\text{血清球蛋白}}$	0.132755858
$\beta_{\text{血清肌酸激酶}}$	0.332818902
$\beta_{\text{血清白蛋白}}$	-0.744060565
$\beta_{\text{血清碱性磷酸酶}}$	0.262594338
$\beta_{\text{品种}}$	1.138018960
$\beta_{\text{性别}}$	0.151826455
$\gamma_{\text{品种}}$	0.5668399
β_{02}	-9.5204440
$\beta_{\text{品种}2}$	1.2299804
$\beta_{\text{性别}2}$	0.2678798

[0351] 此外,通过从列表顶部开始系统地去掉一种生物标志物来约简该10种生物标志物的集合,我们观察到存活率预测的强度的降低(p值)。第一个参数的下降最明显,证实了它们的最大贡献,但我们观察到每次集合约简时预测质量的变化,表明每个参数都对整体预测均有贡献(图2)。

[0352] 在上述说明书中提到的所有出版物均以引用方式并入本文。在不脱离本发明的范围和实质的情况下,本发明所公开的方法、组合物和用途的各种修改和变型对技术人员将是显而易见的。虽然已结合具体优选的实施方案对本发明进行了公开,但是应当理解,受权利要求书保护的本发明不应不当地受限于此类具体实施方案。实际上,对技术人员显而易

见的对用于实践本发明所公开的模式的各种修改旨在落在以下权利要求书的范围内。

[0353] 表3

位点	系数	单变量分析 (p 值)	序列	SEQ ID NO.
截距	12.32478528	NA	NA	
chr1:3844418-3844420	-1 < x < 0	2.98E-26	GGCTGGGGGACGAGGACCGGGGCTCGGCCCTCCAGTTGGGCAATTTCTG	1
chr1:19541665-19541667	0 < x < 1	3.01E-24	GGTGCCCTGGCCACACGGGATCCCGGACTTGTCTGCAAAAGACCCCTTCGTTA	2
chr1:22213697-22213699	0 < x < 1	1.15E-65	TGACCCGAGAGGAGCCGACCCGCGCGCGGACGTTACCTCCGGGGGCGC	3
chr1:108982935-108982937	-1 < x < 0	5.96E-55	CTGAAGTCACTCCAGGCTAGAGGGGCGGGGCAATCCCCCTCCCTGCAATCC	4
chr1:112673186-112673188	-1 < x < 0	0.000323576	CCCCTCCTTTCACCCCTCATCTAACTCGAAGCAGCCGAGGGGGATCTTCTGG	5
chr1:121182146-121182148	> 1	2.29E-14	CTGGCCCGGAAGGCGCAGGGCATCTCGATGTCTCTCCCTCCCGGACTGTCA	6
chr1:58324050-58324052	0 < x < 1	1.82E-36	CTGCCGAGGGCCGGGAAGGCCCTCCGGACTCTCTCGAAAGCATCTCGACCG	7
chr1:97456109-97456111	0 < x < 1	3.39E-13	AAGTACAACCCGCTTCAGGAATGACTCGGTGACATCGGTGACGACCTTCTGC	8
chr1:98084469-98084471	-1 < x < 0	1.04E-37	ACTGTGTTTTTTGTCAACATCTCTCGATGATATGACATTTCCTTAGAAATCAC	9
chr1:98433022-98433024	0 < x < 1	7.27E-19	GGCGAAGAACGAGGCTGTGTCTGGCGGGCGGGCGGGCGGGCGGGCGGTCC	10
chr1:100289525-100289527	0 < x < 1	2.33E-47	TTCACAGACCCGACACTGAGGCTCGCGGGCTCGGACCTACTCAGGCTTCC	11
chr1:107216628-107216630	-1 < x < 0	1.54E-24	AACCTGATGAATGAGGCCAGCTGAGGCTCGCGGGGAGAGAGGGGAGCGGCTCTC	12
chr10:801511-801513	0 < x < 1	1.39E-49	GGCGCGCCAGGACGATGGAGTTGCCGGAGAGAGAGGGGAGCGGCTTCTCTGGT	13
chr10:7411293-7411295	> 1	1.96E-58	CAGCACCGGGCATCTGGACAGCAACCGGAGAGGAGCGGCTTCTCTCTGGT	14
chr10:18237423-18237425	0 < x < 1	1.50E-12	CCGGGGCGGTGGGAACCAAGCTCCGACAGAGCCCTCCAGGCCCTACACGCG	15
chr10:19851621-19851623	-1 < x < 0	4.61E-13	AGTGGGAGGCTTGAGACTGCGCAGCGCTTGTGTTCTCAGGGGGCCATGCCCC	16
chr10:44524681-44524683	-1 < x < 0	1.40E-16	CCAAGAGCGAGACGATCCCATTCAGCGCTTCTGGGTACTGTCTTGAGGAT	17
chr10:40787376-40787378	-1 < x < 0	4.45E-53	CATCTGTGTGACCAAGACTGACCCCTCGGAGCCTAAACCCCGAGCGGGGCC	18
chr10:49526790-49526792	0 < x < 1	1.95E-16	GCCTCATACTCTCTCCAGACAGCGCGCGGGCTCGGGCGGGGCGGGCTCCG	19
chr10:62719946-62719948	0 < x < 1	2.78E-62	CTACGAGGGTCTTTTACCCATTCCAGCCTGAGAAAGAGGCCATGACGATCCA	20
chr11:19910066-19910068	0 < x < 1	1.00E-44	CCCGCTCGGCCCCCGGAGCCCGGAGCCCGGAGCCCGGGGCTCCCGCCAGC	21
chr11:33523800-33523802	0 < x < 1	7.25E-40	TGTAAGAGGGCTTCGGAGTGTCTCTCGGCAAGCGCTACGACCCGCAAAA	22
chr11:24476652-24476654	-1 < x < 0	2.19E-11	ACTGTCCAGGGTGTACAAATCCACGCCAACATCCAAATATGCAATA	23
chr11:49947821-49947823	-1 < x < 0	2.35E-06	GTGGATTCAAGGATTATACCCCTACCCAGCCCAAGACTGTGAGCTCGGGAGA	24
chr11:49987374-49987376	-1 < x < 0	1.06E-11	CAGCAACCGACAGTGAAGTGTGAGCCCGGGTGAATCTCTGGGCAAGCCGAGT	25
chr11:50642620-50642622	-1 < x < 0	4.31E-36	TAGGAGAAAGCCCTGGGTGCTCCCGCCCTTCCCTCACCCTGCGCGCTGG	26
chr11:51974516-51974518	0 < x < 1	4.46E-20	CACGTTTCAATTGACAAACGCGCGGGCGAAGGCCTAGTGAGCGCGGGTGGG	27
chr11:53577274-53577276	0 < x < 1	9.73E-30	GGCGAGCCCGGAGCGGGGTCCCGGGGAGGCGCGGGGCGCGCGCGGGCG	28
chr11:53577353-53577355	0 < x < 1	5.21E-23	GGCGAAGCCGCCCTACAGTACATCGCGCTCATCGCCATGGCCATCGCGCA	29
chr11:53576970-53576972	0 < x < 1	3.53E-60	GTCCGGCAGCAGCGCGACCGGGCCCGCCCGCGCTCCGGACGCGAGCGG	30
chr11:66137549-66137551	-1 < x < 0	2.47E-29	AAGACATTTTGTGATAGATGTTCTCGGCAGTAACTAGCTGGGTTTTTGT	31
chr11:66814981-66814983	0 < x < 1	1.79E-16	CGGGTTCCCCCCCCACCGGCCTGGGCGCCCTCTCTGGCGCCTTGGGTGCTCT	32

[0354]

[0355]

chr12:10418644-10418646	-1<x<0	5.33E-36	CCCCAAATATCTCTCTCTGTGGTCCGGAAATCCCTCCCAACCCCTCTCCGAG	33
chr12:11605135-11605137	-1<x<0	9.20E-10	TTATCTATTACTTATGCTTGGTTACGATTGGTGTCTATCAGTTACTATTGAA	34
chr12:14294419-14294421	-1<x<0	1.16E-98	GTGAGCCAGCCTGTGTACAGGGCGGCGGAGGCCAAGGACCCGTGGTCCCTCA	35
chr12:48113683-48113685	-1<x<0	2.04E-20	ACATTCTACGTGAGACTTTCAGGTGCCCCACCACACTGCTGGTGTCTTAAGC	36
chr13:10113845-10113847	-1<x<0	3.57E-41	CGGGCACTCCCTGGACACAGGACGCGCTGGCCCCAACGTTAGGCGCAAT	37
chr13:27910113-27910115	0<x<1	4.23E-18	GGGAAGACTTTGGTGCCCCCGGTGCGTGGCGCTCTCCGGGGGGCCAGCG	38
chr13:36969021-36969023	0<x<1	2.13E-11	CAACAGCCAATGAGAAGAGGCCCTCGGCGGCGGCCCCAGGAGATGGGT	39
chr13:37474592-37474594	-1<x<0	3.38E-35	TTCCGGCTCCCGAACCCAGGGTGGCCAGCCAGGCTCTGCCCTCCGGGGGTC	40
chr13:37665667-37665669	-1<x<0	1.92E-25	CCCTGGTGGGGGCTTAGGCCATCCCGCCGAGCCAGCCAGGACGACGACGCG	41
chr13:37877786-37877788	0<x<1	2.89E-40	AGCTCGACCTGTGGCTCCCGCCGAGCCAGGCTAGCCGCAAGGACGACGCG	42
chr13:37877759-37877761	0<x<1	2.10E-28	TGACCGGGGCTCTCTGGGCGAGGGGAGCTCGCACCTGCTGGCTCCCGC	43
chr13:37906098-37906100	0<x<1	4.78E-48	TGACGCTCCCGGGCGGGCTCCCGCCGAGCCAGCCAGGCGCTCCGGCCCCA	44
chr14:345761-345763	-1<x<0	2.19E-05	GGGGGACCCCTGGACTGTGGCTGGGCTGCTGTGTAGGTGGCTGTGGGT	45
chr14:368735-368737	-1<x<0	2.43E-08	GTGAGACCCAGCAGCCCTGCCACCCGTTTGACCCCCACGCTGGGAGCTGC	46
chr14:5949013-5949015	>1	1.78E-30	AAGTTCTCGGACTCCCGCCGACACCGCGCTGGCGGAGGACGCGGGCGG	47
chr14:5949450-5949452	0<x<1	4.24E-65	CCGGCGTGGCAGCATGTCCGCCCGCTGGCTGGCTGGCTGCTGCTGCTG	48
chr14:26285429-26285431	0<x<1	2.51E-11	TCGGTGTGTGATTATCTGTAGCGGGCTCTCTGAATTACTTTAATGTTG	49
chr14:39846247-39846249	0<x<1	2.23E-34	TGTTCCAGGCTCCCGGGCCCGGAGCTTCACTCCCTTTGCGAATTAC	50
chr14:39735273-39735275	-1<x<0	1.58E-08	TGTTAGCTGGGGCCCCCGCTGGGCGGCTCTCCGCGCCCGGGGCGCTGT	51
chr14:58424689-58424691	-1<x<0	1.92E-61	TTTCTAGGACTACGCTGTGTGACCTGTGTGGCAATAAATATGCTAAC	52
chr15:2611795-2611797	0<x<1	3.97E-27	CCTTAGACCCCAACCTTTCGAGGTGGTCTCAGGTCCGGTCTCAGGTCTCG	53
chr15:7297280-7297282	0<x<1	2.78E-64	AATAACTGTCAGTCCCTGGGCTCCGAGGACGCTGCCCTCCCAACGCCC	54
chr15:10856498-10856500	>1	7.98E-14	CCCCAGCGGGCGGCTCGGTTTCGACGGAGGGCGGGCTGCGCGGCTC	55
chr15:10856620-10856622	0<x<1	1.84E-40	GGGACGCGGAGGCGTGGCAGCGCCGCGGAGGAGACCCGCGCCAGCAGGT	56
chr15:18190829-18190831	0<x<1	0.019356242	CACCGTGAGCTTCCCGCCCTGCGCAGCGCAGCGCTCTCGCGCTATA	57
chr15:20380665-20380667	-1<x<0	1.19E-06	AGTTCTTTTGGGAAGATGAACCCCGACGGAGGAGCTGCGCGGGGTGTGT	58
chr15:51319029-51319031	-1<x<0	1.01E-30	GAGTTCTGTGACGAGGTATCTCGGTCAAGTATTATCCACCCGCTCTC	59
chr15:51319050-51319052	-1<x<0	1.03E-19	TCCTCGGTCAAGTATTATCCACCCGCTCCCTCTCTCCATGCATCTTG	60
chr16:8886545-8886547	>1	8.35E-87	GGCGGCTCCCCGTGCGGACGCGCGGGCCGACCTCTCCGCGGGGCTCC	61
chr16:15770977-15770979	0<x<1	1.59E-51	GGTCCGTGGCCCCCGGGCCCCCGGCTGGGCCCCAGCCGCGGCGCCCT	62
chr17:33851462-33851464	0<x<1	5.12E-101	ACAGCAGGATGAAGCCGAAGCCCTCCGAGTCCGGTCCGAAGACTCTCGTT	63
chr17:34115428-34115430	-1<x<0	6.79E-40	TCCTGGGCCCCGCGGCGTGGCCCCCTCGGGCTCTCGTGAAGCCCCCTG	64
chr17:39172251-39172253	0<x<1	0.004509682	TCCTCAAGCGGCTCGGGCTGGGCGGCGGAGACTCCCTCGTCTGTTCCCTC	65
chr17:48274513-48274515	0<x<1	1.48E-38	ATGCTACTGACGGCACTGGCTGCTTCGGCGTGGCTCTGTGCGGTAGTCTC	66
chr17:48369176-48369178	-1<x<0	5.43E-12	AAAAAGCCAAATGCCATTACGTACGTTTGTGTTAATTACACTCCCATCG	67

[0356]

chr17:48538075-48538077	0 < x < 1	1.58E-35	AGATGCTCAGCCGAACGTCCCTCGCGCCCCCGGAGCCGGGGGCGGCCCC	68
chr17:49222883-49222885	>1	1.31E-26	GAAAGCAGGAGCCGCTCCCGAGGCGGTCCTCCGGGAGCAAGGGCGCCTCGG	69
chr17:58961294-58961296	0 < x < 1	2.10E-19	AAATATGACATTTCAGCGAGTGTACCGCTAATGAAGCCTTATCTGCGCATGA	70
chr17:60022289-60022291	-1 < x < 0	1.02E-57	ACAAATCCTTTTAAGGCAACGTGTCCCGCCTCTCCTAACCATATAAGTGCAAC	71
chr17:60438944-60438946	-1 < x < 0	2.77E-33	AACTTGTTCTCTAGTGAATGCAATGCGGTTTCCCTCCCTCTCCCATTTATTT	72
chr17:62156690-62156692	>1	0.086414775	CGGACAGGAGCTGGGCTTGGGTGCGGTGACCTTGGGGAAGTAGGCTTCC	73
chr18:5425112-5425114	0 < x < 1	5.52E-53	CAGACGCGAGCGAGGCCGCTGGAGGCGGGTGGCTGGCTGGACGGGACGGG	74
chr18:31999739-31999741	0 < x < 1	1.62E-17	GGTCCGTCGCGGAAGCGGGCCAGCCGGTTGGAGCTGCGCCCGCAGCCGCCG	75
chr18:38129770-38129772	-1 < x < 0	3.08E-63	GGGGAAGCTTGGTCTGATAAAAGCCGAGAGACCCAACTGTGTCTCCAGGG	76
chr18:44686312-44686314	-1 < x < 0	1.01E-30	GAAGGGCAGCAGCGCCGGCTGCCCCGGGAGCCAGCGAGGCCCCAGGCCCC	77
chr18:45473750-45473752	-1 < x < 0	0.008692376	GGGACCGGGTCGAGACACCTCTCCCGCTCCGCTTGGCCACTCCAGGCTTG	78
chr18:47074292-47074294	0 < x < 1	3.88E-20	GGAGCCGGAGGCGAGGGGGGGGACGCGTGTCTCTTTATTGCTTATAA	79
chr18:4690963-4690965	0 < x < 1	3.35E-38	GGCGGATCGGGGCGGCGCGGGCGGTCTACGAGACCTCGGGCGGCGCC	80
chr18:44679236-44679238	-1 < x < 0	7.48E-29	TCCCGGAAGTCTGTCTCACTAGCCCGGTGGCGGCCGGGTGTCACCCGCTG	81
chr18:44686394-44686396	-1 < x < 0	1.73E-66	GTGCCACCATCCCGGGTTGAAGACGGGTTCGAGTGGAGCCAGGAACACC	82
chr18:48681702-48681704	-1 < x < 0	9.82E-09	GGATCCGTCAATTGACCCCTTGGAGAACGGAATTTCTGGGGGCTCAAAGGAGT	83
chr18:50257730-50257732	-1 < x < 0	1.96E-27	CCTTCGAGCTGCACCCCTCTCCCGCTTTAGCCTGGGAAGGGGCCAG	84
chr18:50259206-50259208	-1 < x < 0	2.83E-46	CCAAATGACGCTCATCGGAGAGAACGGCCACAGCACCGACGCTCACGGG	85
chr18:51374197-51374199	0 < x < 1	6.02E-85	CCACAGCCCTCGCTCCCGGAGCGCGCACGCCCGCAGCGGTGCAAACTG	86
chr19:44679975-44679977	>1	2.92E-14	GCCCATCGCTGCGGCTCGGTCTCCGGGAGAGGAAAGTAACCTCCGAGC	87
chr2:25678434-25678436	-1 < x < 0	0.818276235	GGGAAGGTGGGTGGAGTTGCCAGAGCGCGGAGGCCCCGGGGCGCACGGTCT	88
chr2:32494387-32494389	>1	0.049615576	TTGGAAGCTCGGGGGGCGGAGATCCCGCGCCCGCCCGCGGCTCGGGCGG	89
chr2:31655107-31655109	-1 < x < 0	4.80E-39	CTGCACAGCCCTGCTCCGCCCTCGCGGCCCTGCACGGTGTCTCGGTGCT	90
chr2:51831832-51831834	0 < x < 1	2.84E-32	TCCGGCCGTGACCCCTCTGGGTGCCGCCCTGCTCTTGTGCTGCTCTGCT	91
chr2:64665545-64665547	-1 < x < 0	4.33E-08	CTTACAGTAAGCCTAAAGAGTAGGCGTTATCTCCGTTTATGGTGTAACTGG	92
chr2:72450823-72450825	0 < x < 1	3.75E-51	GCCTGTCCGAACGCGACGCTGGCCATCGCGCGCGGGCGCGGCAAGGGCCGCG	93
chr20:905463-905465	-1 < x < 0	0.259671131	CTGCAGGGCATCCGGGTGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	94
chr20:3395529-3395531	-1 < x < 0	1.93E-16	AAAAGCCAGGGGAGAGGGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	95
chr20:3833805-3833807	-1 < x < 0	7.58E-41	AGTAAGGAGCTGCTCTCATGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCT	96
chr20:32463049-32463051	-1 < x < 0	3.95E-24	TTTTACTATTGTATTCAATTAAGCTCGGAGCATGTATTGAGTGCCTACTGTG	97
chr20:37906191-37906193	0 < x < 1	1.04E-24	GCCGGGGCGCAGGGTTCCAGCTCTCGTCAAGTGGCTTGTCCCTCGCTCAG	98
chr20:40194554-40194556	-1 < x < 0	2.50E-11	CTGTAGTAGCTCTGTCTCCAGGCCTCGGAGTCAGCTGAATGTCACCCCTT	99
chr20:43366718-43366720	0 < x < 1	0.013883014	GCCGATCGCGCCCGCTGGAGGTGCGGGGGAACTCCCGCGCGCGCTCGGG	100
chr20:44942213-44942215	>1	6.74E-26	GGCAGGGCCCGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT	101
chr20:46098568-46098570	0 < x < 1	5.41E-54	TCGGTGCCGCCCGCTCGGGCTCTGCGCGCGGCGCACGCGCTGCGGGTGACAGA	102

[0357]

chr20:49086890-49086892	-1<x<0	1.65E-84	CTGGCACACACGCCCCACGACCCACGCGTTTGACAGGCACACGGGGCCACCGTCTC	103
chr20:49844561-49844563	-1<x<0	3.24E-27	TGAAATTTTCGAGTCTCGGAGACGGCGTGATGCTGGGGGTCTCGAAAGCCT	104
chr20:49940551-49940553	-1<x<0	9.74E-50	ACGTGCCCATGGGGGATAATAAAGTTAATGAAACCCCTTGACAGAGTACTC	105
chr20:49972926-49972928	-1<x<0	4.55E-08	CAGCAGCTCCTCATTATTCCTACAGCCGACCCCTACGTAGTACTATTAT	106
chr20:54301453-54301455	0<x<1	7.23E-89	ACCACGCCCCGGGGCGAGCGTCGGGGGGTCGGAGCCGAGCAGCTGCCG	107
chr20:53921003-53921005	-1<x<0	1.50E-09	CCCTCAGGAACCTCTCCAGTCCAGCTCCGCTGCCATCAGGTCCACGG	108
chr20:57347921-57347923	-1<x<0	1.02E-54	ACCAGCCGCCAGCCGTGACCTTCCGGGCTGTCTCGGGCTCTGGGCGC	109
chr20:57621119-57621121	-1<x<0	8.85E-13	AGCGCAGGACACCCCTGCAGCTCGGCGCGGGCGGTGCCAAGTCTGGG	110
chr21:31923268-31923270	0<x<1	4.88E-17	GGCGCAGGGCCGGTGGAAACCTGCGGGTCACTAGGCTGGCGAAGGCGA	111
chr21:43490796-43490798	0<x<1	1.07E-75	ACGCGCTGGCCCCGGCTGCCCGCGGCTCGGGAGAGCATCCCTGCCGCTGC	112
chr22:46374563-46374565	>1	9.37E-07	TTGGACGGGAGGCGCTCGGGCGCGGAGCAACGAGCAGCAGCAGCGTGG	113
chr23:4004092-4004094	0<x<1	2.33E-19	TTAGGTGTACCCGCGCCGCTCCGCGGAGCCGTGTGCAATGCTATCCGA	114
chr23:29966538-29966540	-1<x<0	2.55E-29	AGGGGAGGTGACATGGGAATTAACGCCCTTCCCTCAGCCAGTGCCAAAT	115
chr23:31144202-31144204	-1<x<0	7.13E-08	GAACCCCACTGCTCACCAGCCGAGGCGCCACCTTGAAATGCCCCTGTTAT	116
chr23:31144337-31144339	-1<x<0	1.68E-29	GAGCCACCGAACTACTTGAATCACCCGCGGGCAGGCGCTGTCCAACATT	117
chr23:37987526-37987528	-1<x<0	0.102124346	CGGGCGGGAATCCGCCGCCCTCGCCCCGGCGCTAGACCCGGGGCGC	118
chr23:50123623-50123625	0<x<1	2.15E-99	CCAGAACTACGGCAACTCGCGCGGGTGCAGCTGGAAGGGGAGTCCG	119
chr24:12518910-12518912	0<x<1	9.79E-97	CTCTTCTCGAGATCCATGCTCGCCGTGACGGAAGGAGCCAGATGTCT	120
chr24:17354043-17354045	-1<x<0	5.21E-41	TGGGTAGGTGAGAGCAGGCGAGGGGCGCAAGTTGCGGTGCCAGACTGCT	121
chr24:21051024-21051026	>1	1.16E-98	ACGTGCTCTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCC	122
chr24:24968751-24968753	-1<x<0	9.40E-88	GTCTCTGGCGAGGGAAGTTACAGACGCCAAGCCAGAGTCTGGGGTGCCCA	123
chr24:43790837-43790839	0<x<1	2.77E-39	CGCCGGCCCCAGGAGCGCGCCGTACGACCCGCGCTCGCGTAGCGCTCCG	124
chr24:45869679-45869681	0<x<1	1.98E-82	AACAAAGGCGACGTGCGCGCGCGCACGCGGTCTCAACACTGCTTTTGTCTT	125
chr24:46167206-46167208	-1<x<0	4.99E-21	CTAATGTTGACCCGATGGCGTGGCGCGCGCGGGGTGCATCCGTCAGGTCA	126
chr25:16764068-16764070	-1<x<0	6.84E-40	CGCTGAGCATCTACTACTGCGGCGAATTAATCCACATTTCTCTGCTCAT	127
chr25:18016409-18016411	0<x<1	1.39E-49	AGAGCCAGGCGCGCGCGCGAGCGCGGTGCACGTGCACACAGGGGGCGCC	128
chr25:47511184-47511186	0<x<1	4.90E-64	CCGGACCCGGAACGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT	129
chr26:1578502-1578504	-1<x<0	7.75E-09	TCCGATGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT	130
chr26:5490242-5490244	-1<x<0	1.62E-87	TCCCAGCGTTTCAGTCTGAGTGTGCGGTGGAAGCCCGCGCGCGCGCGCGCAT	131
chr26:6348316-6348318	-1<x<0	2.12E-06	GGAGGTGCAGACGCTTGCCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT	132
chr26:6353216-6353218	-1<x<0	2.32E-10	ATCCAAACTAGTGGCAGAGTACCGGACGCGCGCGCGCGCGCGCGCGCGCGCAT	133
chr26:9207175-9207177	-1<x<0	1.03E-07	CCCGGAGTACTGGCTCTTCGCCCGCGGAGCAGCGCGCGCGCGCGCGCGCGCGCAT	134
chr26:10691576-10691578	-1<x<0	1.57E-17	TGTATTAACCTCATAAATGATATACGGGATAGAAAGGACCTGCCCTGTCTCC	135
chr26:16852803-16852805	-1<x<0	7.42E-21	CGGTGGGGTGCAGCCACTGAGCCACGAAGGAGTCCAGCTTGGAGGCAGGGG	136
chr26:22918623-22918625	-1<x<0	3.70E-22	TAAATCCATTCTGGTTGTAGAAATCTGCTGCAGGCTATTGTGATAGGATA	137

[0358]

chr26:30413129-30413131	-1<x<0	5.47E-11	AGAGGCCAGGTGGCTGACTGTTGACGGCCAGGGCAGCAGAGCCCCCTCCTGT	138
chr27:1856425-1856427	0<x<1	1.16E-19	CTCAGAGGCACGGTCTGAGGCCACCGGTTGCTAAGGCTTCACGGTTTCAGC	139
chr27:9830220-9830222	-1<x<0	0.000162475	TTGGAGTCTGGCCTTCTAGAACTCGGTTCAAATGCTTTTGTCTCTGGGAG	140
chr27:28156834-28156836	0<x<1	9.27E-45	GAAATGCTGTGCTGAAGATGCTCCAGCGGAGGGGGCCCGGGCCGACGTCG	141
chr27:29250528-29250530	0<x<1	4.00E-55	AGGGGACCCCCCGCCGCGCTGCGCGCTTACCGCCAGCTCGTCGGGCT	142
chr27:29307058-29307060	-1<x<0	1.72E-26	CTCGCCCTTCTGTCTCCGACTGGCTGCGGCTCCCGCCAGGCACAACTGCACTC	143
chr27:29307097-29307099	-1<x<0	2.72E-25	CAACATGCAACTCGGGTGTCTCGGATGTGTTAAGGCCAATGTAATTC	144
chr27:41569781-41569783	-1<x<0	3.06E-39	TCTCGTGTTCATAGGAAGAAATTCGGGAAGAGATAGAACAGCAGTCAATT	145
chr27:45037864-45037866	0<x<1	2.77E-85	CCAGCGCCGCCAGCCCTGCGCCCGCGACACAAAGATGCGGGCATTC	146
chr28:5936067-5936069	-1<x<0	6.92E-08	CCTGGATGGCAACTTTGAACGCTGCTGAGGACAGAAACAATCGGAGAAG	147
chr28:8994532-8994534	0<x<1	6.45E-45	GAGCCGCCAAGTTAGAAGACTGTGCTGGGGGCCACACGGTGTCCGGG	148
chr28:13976453-13976455	-1<x<0	3.48E-34	GCCTGTCTCCACCTCTTGGTCCGCTGTTATGGTTCTCTGGGAATGG	149
chr29:9541477-9541479	-1<x<0	3.89E-12	ACTAGATTTGAATCTATGATGACCGGGACTACTACTCTACCTTTC	150
chr29:19479611-19479613	0<x<1	3.87E-66	TGGGAGGTGGGAATCGGTACACACGGTGGCAGTTGTGTACAGGCAG	151
chr29:32189910-32189912	-1<x<0	2.42E-54	GCAGGTGTGGGACAGCCATTGGCCGTGTAGAACCGAGGTTCACCGTTTG	152
chr29:35629270-35629272	-1<x<0	1.54E-14	TGCGTCACTCCCTTGTAGGAATCCCGATGATGTTGATGATCATCAGAG	153
chr29:35629292-35629294	-1<x<0	4.54E-41	TCCGATGATGTTGATGATGATCATTACGAGGACGGTAATAAGATTGTACCCG	154
chr3:31677748-31677750	0<x<1	1.40E-89	GCTGAGCCCGCAGCTGCGGCGCGGACGCCGTGCGACGCCCGCGGGCTGG	155
chr3:46852141-46852143	0<x<1	7.92E-18	CTATGGCGGATATTGCAAGCAAGACGACCCGAGGCACAGCCAAATTCACAT	156
chr3:46852146-46852148	0<x<1	3.80E-11	GGCGATATTGCAAGCAAGACGACCCGAGGCACAGCCAAATTCACATGTTTG	157
chr3:46843750-46843752	>1	1.05E-109	CCAGGGCCACAGACCTGGAAGTCCGCGGAGCCCGGACGACATCTGAACTTATGTCTC	158
chr3:61539374-61539376	-1<x<0	0.000110942	AGATTCAAGGCCCTTACCTGGGCGGATCATCTGAACTTATGTCTC	159
chr3:70831746-70831748	>1	5.57E-17	CTACCTGTGCGGGAGGCGAGCACCGGGCGCCGACGTCGCGCCGCGGG	160
chr3:72353348-72353350	-1<x<0	3.36E-46	GGGACACCTACTCTGGCCCCCGCACGCGGAGACAGACAGGTGCTGCAAA	161
chr30:8031058-8031060	0<x<1	8.37E-45	GGCCGTCGGGCGGGGCGGCTCACCGCGCGCGCTGGAGGGCGCGGCGG	162
chr30:8031060-8031062	0<x<1	4.70E-34	CCGTCGGGCGGGGCGGCTACCGCGCGCGCTGGAGGGCGCGGCGCGG	163
chr30:8640557-8640559	-1<x<0	1.10E-14	CGGGGACGGGATGTGGCATCACACGCTGGACGAAAGTTGGGTGCGGTAG	164
chr32:10044870-10044872	-1<x<0	3.77E-08	GTGAGACGGAATGTGTGATGTACGTAAACAGGTTTTAAGAAATGTGTAC	165
chr32:17605966-17605968	-1<x<0	2.27E-18	TAAAGATTAAATATCAGAAAGCCACCGTTTCTTTAGCCTCCCGTGACGAA	166
chr32:35345144-35345146	0<x<1	5.23E-46	GTGAAACGAGGGCCCTCCACCCCCCGGATGAAAGATGGCAGTTGTCTGA	167
chr32:29642986-29642988	-1<x<0	4.41E-42	CTCTTGGTGGATGAGATGTTGACGGGATTCGTCCTTACTTTCTGGGAG	168
chr32:39585142-39585144	-1<x<0	0.108871384	GAGGCTGTGCGAGGTGAGGCTGTGGCGGGCGGGAGGGCGGGGACCCG	169
chr32:40571018-40571020	-1<x<0	1.37E-64	TTTGGGCTTAGCCAAAGGTACACATCGGCGTTGTTTGTGATTTGTGATCTA	170
chr33:18234603-18234605	0<x<1	0.04858715	GAGGAGCCCGATGAGGGCGCCCCCGCGCCAGGTAAAGCGGCGCTC	171
chr33:23247947-23247949	-1<x<0	8.71E-28	GACAAAGGATGGAATACCTACAAACGGTGATTCTGATCTTTTCAAGGGTG	172

[0359]

chr33:23473772-23473774	0 < x < 1	5.14E-07	ACAGGCAACATCGTGGGCTTCTACCGGGGCACTTCTCTCTCTGCCCCGCA	173
chr33:23473856-23473858	0 < x < 1	1.85E-12	GCCAGAGGTATCCACAGCCAGCTGCTGTGTGCTCACACACACAGCGTCTCA	174
chr33:26512692-26512694	> 1	3.95E-49	CGGCTGCCGCGCCGCGAGTCGTGCCGCCGCGCTCTGCCAGGCCGACCTGG	175
chr33:26512695-26512697	0 < x < 1	2.71E-54	CTGCCGCGCCGCGAGTCGTGCCGCCGCGCTCTGCCAGGCCGACCTGGAGC	176
chr33:26512711-26512713	> 1	7.18E-24	CGTCCGCGCGCGCTCTGCCAGGCCGACCTGGAGCGGAAGCTGAAGCCGAA	177
chr33:19810904-19810906	0 < x < 1	3.02E-58	GCGCTCTCGCCGCTCTCGGTGGCCGGAGAACAAAGCCCGGCGCTGGCC	178
chr33:28241609-28241611	-1 < x < 0	7.57E-60	GGAGGCGAGCTGGGAGTACTTCCCACGCCACATGAGCCTTTGTGTGAAAAT	179
chr33:28241833-28241835	-1 < x < 0	1.89E-67	AATTACTTTCACCTTGTGTGAGGTGCTGCTGTCTCTGTCTACCCACAGG	180
chr33:29201592-29201594	-1 < x < 0	8.73E-65	CTGCTGCCCGACGGAAATAGGAGCGCGGACGCTGCCAGAGTGGGACAGTCC	181
chr33:29201578-29201580	-1 < x < 0	7.87E-56	GTGGCCCTGTACCCTGGTGGCCGACGGAATAGGAGCGCGGAGCCTTGCCA	182
chr34:20742132-20742134	-1 < x < 0	1.32E-18	TCATCTACGCTGGGAGATGAATATCGGTCTGTACAGCCGAGGCTGGAAT	183
chr35:23852932-23852934	0 < x < 1	7.91E-33	ACAGCGCTGCGCTCTCTTGGACGACACGCGCTGGACATCAGCTCCGA	184
chr35:23853673-23853675	0 < x < 1	1.12E-17	CCACCTTCCCGGCGACGCCACGCGGCGGCGGCGCTCCAGCGCACAGCAG	185
chr35:23853863-23853865	> 1	1.01E-15	CGCTACGCGCTCTGTGTGCTCCGCGGCTCCAGCGCACAGCAGCAG	186
chr35:23853763-23853765	0 < x < 1	1.66E-57	GCTGACGCGCTCTTACGCGCGGCGGCGACGCTGTGGAAACGACGCGCA	187
chr36:641903-641905	-1 < x < 0	2.02E-46	GAGAGGGGGAAGAAATGTGTCTCCCTGAGTCTCTTTTGTGCTTCTAGTTCTA	188
chr36:8405088-8405090	0 < x < 1	1.18E-47	CGGCTCCGCGCGGACGCGGCGACCGCCAGCGCGCTCCGAGCGGAGAG	189
chr36:18662589-18662591	-1 < x < 0	1.78E-29	CCCGTGTGCTCCAAACACAGCAGCGCTCGGTTTGTGCGGATGTTGTCAG	190
chr37:2007181-2007183	-1 < x < 0	4.37E-40	TTCATCATTTGATGATAGTTTTCACGACCTTGAGAGTCTCTCTGTAAAAT	191
chr37:26349705-26349707	0 < x < 1	2.71E-23	GCAACGCAATCGCGAGAACCTTCTCCGAAGGGCGGACCCCCCGACGGC	192
chr37:26735403-26735405	-1 < x < 0	2.68E-24	CCCTTTTTACACATGAACAAACGCGCTACAGAGAGGAATAACCTCCCCAG	193
chr38:14762928-14762930	0 < x < 1	7.07E-51	GCCACGCGCGCGCGCCAGCCCCACGACAGCAGCAGCAGAGTCCGACAGCA	194
chr38:21317970-21317972	-1 < x < 0	4.36E-17	GGAGTCTAGGATGGCAGTGTGACTCCGAGAAAGTAGGACAGCGCGGAGAGG	195
chr4:8252080-8252082	0 < x < 1	7.58E-11	CCGGGCTCCAGGCGGGGCGCGCTGCGGGGCGGACGCGCGCGCTACTG	196
chr4:23880514-23880516	-1 < x < 0	0.034364027	CGCCGACCCCGGAACCCCGGTGGCGTCACTGTGCGCTCCGCGCGGAGCA	197
chr4:25509815-25509817	-1 < x < 0	9.11E-34	ATGGCTCTCTGTCTATCTGGGACCCGCTCCACCCCATCAGAGCCAAATGA	198
chr4:36661557-36661559	-1 < x < 0	2.89E-22	TATTACAGTGTCTGAGTCAACGCGCTTTTCACTCAATGGAATTCCTCT	199
chr4:36180013-36180015	0 < x < 1	2.94E-96	GCGGCTGGGGGCGCGGAGGGCGCGCTCGCACCGCGGACGCTGCAGTTC	200
chr4:43217432-43217434	0 < x < 1	2.41E-62	GTGTGTGACAGTTAGCTGTGGCCGACGGCTGTGTGGGCGCTGGGCGCAAGC	201
chr5:21480151-21480153	0 < x < 1	2.64E-26	TTGAGCTGCTGGAGCAGCAGCCGTCGGCCACCTAGCAGGAGTGTCTTCGCA	202
chr5:32347978-32347980	-1 < x < 0	0.375939729	ACACGATAGCCCTGGCCCGGTCCCGGCGCACCCCGCCAGCAGCCCGATGC	203
chr5:32711703-32711705	0 < x < 1	0.161513412	GGTATCCGCTCCCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCGCCG	204
chr5:32884319-32884321	-1 < x < 0	1.10E-34	CCAGCTTGTGTGGGTCTCCACACCGCTGTGTATGAATGGGACTAAATTT	205
chr5:55848493-55848495	-1 < x < 0	8.10E-13	GACCTGTGCCCTCAGCTGGCCTCAGTGCATGGAATGTTGCTGTGCCCTGT	206
chr5:56981423-56981425	-1 < x < 0	3.11E-54	CTGCCGGGCTGGCCGGGCGCGCTTCGTGGTCTGCACTTACCCACGCTGCCA	207

[0360]

chr5:61645225-61645227	>1	4.38E-60	TAAAGGGCGATGCCGACACCGAGGCGCTGGCGTGGGTCCACCGGCGTCGC	208
chr5:67427228-67427230	0 < x < 1	3.63E-14	CGCGCCCCGCGAGCGCGGGCCCCGGGGACCCCGCGCGCGCGCCCC	209
chr5:81484511-81484513	-1<x<0	3.60E-10	GGTTGCTGAGAGAGAAATGCTGACCGGATCTTCTATATTCATTCTAACTTG	210
chr6:89151114-89151116	-1<x<0	5.73E-49	TCCAGACCTAGCAGTTGAGAGGATCCGAGTTCGAATCTGTCTTTGCCATTC	211
chr6:9019848-9019850	-1<x<0	3.88E-05	CGTATCCTAACTGATCCCAAGCTGCGCTGACTGGAAGTCTTGTTCCTCA	212
chr6:10398666-10398668	0 < x < 1	1.45E-24	GGAGCGAGCGAGTCGCCAGAGCCGCGGAGACCCAGCTGAGCCCAACCCACT	213
chr6:12827818-12827820	-1<x<0	6.58E-07	GGCTCTCCCTCACACCCACTGACTCGGCTCTGAACTCAGGATGGTCCCTT	214
chr6:26928260-26928262	-1<x<0	0.100532489	ACGCCGTCTCAGTCTCACGAGACGACCGCGGAGATTTCAAAGGAGCTTTT	215
chr6:29756505-29756507	0 < x < 1	7.43E-20	TGCACGACCAACGCGAGCTCCCAACCGTGGTCTTGCCCCCTGGCTCCCTCCT	216
chr6:38452376-38452378	-1<x<0	3.46E-13	CAGGGTCTGCGTGGTTTCCCCATCACGTCCACCCAGTTGCCCTAGTGACTGT	217
chr6:40186467-40186469	0 < x < 1	1.66E-20	TGGACCTTCCCTGTGACCCCCCTCCGTGCCGCGGCCCTCTCCTGAAGAGGG	218
chr6:41080691-41080693	-1<x<0	3.21E-15	CAGCGGTACAAAGTGCTTCAAGGAGCGGCGTGTACCAAAGGCCATATCGGCT	219
chr6:41485415-41485417	-1<x<0	1.57E-05	GGCGTCCATGGCAGTGCCGGCCCCACGAGGACGGGGCAGATAAGGGGGCGG	220
chr6:41661247-41661249	-1<x<0	2.79E-41	TCCTGGCGCAATAAAAAGAGACCCCCGTGAGGAAAGCCCCCAGCACTCAT	221
chr6:45773846-45773848	-1<x<0	3.10E-187	GAGAAAGACGTCCCTGTCGCGAAGCCGTTCGACGCGCGGCCCTGGGGCGG	222
chr6:60675998-60676000	>1	1.33E-31	GATAGCTTCTCGGCTCGGTGCGGGCGGCTGCTGAGCGCGGAGCGGCGCGG	223
chr7:8160746-8160748	-1<x<0	6.40E-06	TGTCCAGATCCTCAGAGTAAACAAACGAAGTTGTGAGTTGGGTAAATGGTTG	224
chr7:9749742-9749744	-1<x<0	7.07E-22	GATGAAATAGCAACCAGGTTGTTGGCGACCATGAAAGGCAAAACCTCACAGGG	225
chr7:29867181-29867183	-1<x<0	7.92E-16	TTGAGAGGTCCAGCACCAAAATGTTACGTGATTCAGAAGTGGACATCTAAACAA	226
chr7:38420583-38420585	0 < x < 1	6.92E-09	CAGGGACCTCGGCCCTCCCTCCAGCGCGCGGGACACAGACTCCGGGTGCCT	227
chr7:38136719-38136721	-1<x<0	1.80E-10	CGCTCCCGGAGGTGTTACCGACCCCGCGCACACACACCCCGTACGTGCCA	228
chr7:41785316-41785318	-1<x<0	5.87E-19	AGGGCTGGGAAGATTGGCAGCCCTTCGGAGTTATGCGTGAAGGGAGTCCAT	229
chr7:41980137-41980139	-1<x<0	1.81E-39	CCACAGCATCAGCCCAGGAGAGGCACGGTTATTGACAGGTTACACCCACAGG	230
chr8:14259939-14259941	-1<x<0	9.05E-30	ATTTAGATGCCAATCTTTGAAGGCGTTGGAATGAACACCCCTTCATTCC	231
chr8:4122009-4122011	0 < x < 1	2.05E-33	CTCAGGAAACGGGGTCGGACCGGGCGAGCAACAAGGGGTTCACAGTCGGGAG	232
chr8:15508656-15508658	-1<x<0	7.24E-09	ACCTTGCCCTCCTTGGCAGGCCAGCGCTCCGGAGGCCCTAAAGAAACGGAC	233
chr8:17226001-17226003	-1<x<0	6.70E-31	AAGCATCTGCTGCTCTTTTCGTGCTCTGCTAAGGGCTTGTAAGGAGCC	234
chr8:33509086-33509088	0 < x < 1	2.46E-82	GGGCCAGTGAGTGGGCTGCGGGCCGGGAGGGCGGGCGGGCGCGCGGGCG	235
chr8:67527837-67527839	0 < x < 1	2.39E-96	GCGTGGCGCCGAAAGCGTGCCTGCGCGCCAGCTTGCGCCCTTCTGCCC	236
chr8:67627712-67627714	-1<x<0	9.60E-76	ATTTCTGAGTGCCCACTCTGTGCCCGACATGGGTACACAGCTCCCCCAGG	237
chr9:933885-933887	-1<x<0	9.12E-21	CACCTGTGCTAGTCCAGCTGATGGCGCAGGTGCAGCCCAACACTCAGGGGG	238
chr9:1046331-1046333	-1<x<0	1.35E-20	AGGAAAGGGGGCGAGGGCATGCTTCGAGTTTGGATCCGAGGTCTTCAGGGG	239
chr9:2360284-2360286	0 < x < 1	1.95E-08	CGGAGCCACGACCCCGGGGAAACCGAATTTAAGATGAAGGTATTTTCATT	240
chr9:3516451-3516453	-1<x<0	6.18E-49	CGCAGCCCGTGGCCCGCTTCGACTGCTGCTCAAGCTGGTGCATCACTACAT	241
chr9:4719064-4719066	0 < x < 1	5.92E-33	TTCTCGCTGAACCCGGCGCGCTGGCGGTGAGTGGCGGGGTTCGGGGCAC	242

[0361]

chr9:5338779-5338781	-1<x<0	0.159924272	TCCCGCCCTTCCCTACCTTCTCGCCGACCTGTCAAGGCGTCTCGGGCGCCT	243
chr9:5481047-5481049	-1<x<0	0.359052165	GCCTTCCCACTCAGCGCGCGGCCCGGCTTCCCGCCCAACGGCCCGGCCA	244
chr9:10371698-10371700	-1<x<0	1.46E-18	GGAAGGTGCTGAGAGCATCTGGCGCGGGGCTGGCTGTCCCCAAGCTGGT	245
chr9:12522577-12522579	0<x<1	0.261625965	TACGGGAGGAGATTAGGGTGATTCCCGTTGGCTCTGGCCAGGCCGTACGGCC	246
chr9:12642800-12642802	0<x<1	1.38E-29	GCGAGGCTGCGAGGCTGCGAGGCTGCGAGGCGGGCGGCCCGTGGTGTAC	247
chr9:18277673-18277675	0<x<1	2.12E-34	GCCTCGCAGCCCTGGCGCGCGCTCGCAGATGGAGCGGACGGCGGATGG	248
chr9:19654152-19654154	0<x<1	0.005118485	AAGTGTGAGTCCACCACGTGTCCGCGGAGGTCTCCTCGGGGATCTCCACCA	249
chr9:19703027-19703029	-1<x<0	4.28E-20	AAGCCCTCCCCCAAGCCAAGGACCCGTCACCCATCCCCCTCCACCTCCCTA	250
chr9:25415871-25415873	0<x<1	2.46E-50	GTGAGGATGCTGCTGAGCTTGCGATCGAAGGAGCCACTCATCCTGGCGAGCG	251
chr9:24344496-24344498	0<x<1	1.12E-83	AATTAAAGGGTTATGGGCAGACAGCGTGCGGGGTCTCAGAAAGGAAAATTG	252
chr9:34421146-34421148	0<x<1	3.96E-39	GACTCGGGCGGCCCTGCAGACAGACCGACGACGCCCGACCTCTGCCCGACGG	253
chr9:35986072-35986074	0<x<1	7.61E-27	CAGCCAGGAGCAGTAGCCGAGCCATCGGAGCTGGGAGAGACGCGCCAGGGC	254
chr9:35999030-35999032	0<x<1	0.109805996	GGCAGCGCGACAGCGGCCGCCACCGGCCCTCGCCGCTGCCCACTCGGC	255
chr9:47196375-47196377	0<x<1	4.17E-30	CCATAGCGGCTGCCAGCTCGTCCCGAAGCGGACCATGGTTCCCGGGCGCC	256
chr9:47612682-47612684	-1<x<0	5.29E-27	GCTCTGGGCTGGGGACTCAGCAGACGCTCCCAACCCGCTAGCAGGACCT	257
chr9:50976682-50976684	0<x<1	1.00E-43	CCGAGTGATGACGCTGATTGGTCCGGCCGATTGTGGGGTGATGTACGG	258
chr9:49248845-49248847	-1<x<0	2.14E-28	GGTCACTGAGCTCTGGTTTCATCAGCAGGAGGAACTTACTCAGCTCGAGGAA	259
chr9:54685917-54685919	-1<x<0	1.65E-21	CTCTGGCCGCTGTGGCCCTCCTCTGGCGTCTCCGAGACAAGCAGGCTCGCAT	260
chr9:54993941-54993943	-1<x<0	1.24E-27	CTGCCCTGTCCCAACAAGCTACCTCCCGAGTGAAATACCCATGCGCTGAGCCC	261

表 4——前 3 个时钟

位点	时钟系数	p 值 (Pr(> t))	序列	SEQ ID NO:
(截距)	-0.737834998	0.000221316		
chr2.32494387.32494389	37.1425602	1.55E-05	TTGGAAGCTCGGGGGCGGAGATCCCCGCGCCCGCCCGGCCCTCGGGGCGG	89
chr22.46374563.46374565	25.17478169	1.08E-09	TTGGACGCGGAGGCCGTCCGGGGCCGAGCAACGAGCAGCACAGCACGGTGG	113
chr6.45773846.45773848	20.26108043	1.13E-193	GAGAAAGGACGTCCCTGTGCGGAAGCCGTTCCGAGCGCGCGGCCCTGGGGCGG	222

[0362]

表 5——前 5 个时钟

位点	时钟系数	p 值 (Pr(> t))	序列	SEQ ID NO:
(截距)	-0.727254445	0.000153021		
chr2.32494387.32494389	37.1612328	6.43E-06	TTGGAAGCTCGGGGGCGGAGATCCCGCGCCGCCGGCGGCGCTCGGGGCGG	89
chr22.46374563.46374565	22.39755562	1.62E-08	TTGGACGCGGAGGCCGTCCGGGCGCCGAGCAACGAGCAGCACAGCACGGTGG	113
chr6.45773846.45773848	17.93314936	1.51E-140	GAGAAGGACGTCCCTGTTCGCGAAGCCGTTTCGACGCGCGGCCCTGGGGGCG	222
chr5.61645225.61645227	13.1422123	3.10E-09	TAAAGGGCGATGCCGCGACCCAGGCGCTGGCGTGGGGTCCACCGGCGTCGC	208
chr3.70831746.70831748	16.98136108	3.58E-06	CTACCTGTTCGGCGGAGGCGAGCACCGGGCGCCCGCACGTCTCGGCCCGGCGG	160

表 6——前 10 个时钟

位点	时钟系数	p 值 (Pr(> t))	序列	SEQ ID NO:
(截距)	3.098969668	1.79E-15		
chr2.32494387.32494389	43.24386632	3.39E-12	TTGGAAGCTCGGGGGCGGAGATCCCGCGCCGCCGGCGGCGCTCGGGGCGG	89
chr22.46374563.46374565	18.0837229	1.05E-09	TTGGACGCGGAGGCCGTCCGGGCGCCGAGCAACGAGCAGCACAGCACGGTGG	113
chr6.45773846.45773848	9.927654715	6.75E-58	GAGAAGGACGTCCCTGTTCGCGAAGCCGTTTCGACGCGCGGCGCTCGGGGCG	222
chr5.61645225.61645227	8.930766415	7.00E-08	TAAAGGGCGATGCCGCGACCCAGGCGCTGGCGTGGGTCCACCGGCGTCGC	208
chr3.70831746.70831748	13.04994007	1.98E-06	CTACCTGTTCGGCGGAGGCGAGCACCGGGCGCCCGCACGTCTCGGCCGCGGG	160
chr15.10856498.10856500	16.51970306	4.70E-10	CCCCAGCCGGGCCGGTCTGGGTTTCCGACGGAGGGCGGGCTGCGCGGCGTC	55
chr16.8886545.8886547	9.736863844	1.16E-14	GGCGGGTCTCCCGTGC GCGCACGGCGCGGGCGGCGTCTCCGCGGGGCTCC	61
chr3.46843750.46843752	6.421789458	7.38E-22	CCAGGGCCACAGACCTGGAAGTCCCGGGGAGCCCGGAGCAGCAGCTCTCAAG	158
chr20.44942213.44942215	6.783887287	8.61E-09	GGCAGGGCCGCCAGGCGCCGCCCGGTTGAAGAAGTCTCGGCGCTGCGCT	101
chr20.57347921.57347923	-10.03483995	3.52E-40	ACCAGCCGCCACGCCGTGACCTTTCGGGGCTGTTCCTGCGGCTCTGGGCGC	109

[0363]

表 7——前 20 个时钟

位点	时钟系数	p 值 (Pr(> t))	序列	SEQ ID NO.
(截距)	1.705577846	9.79E-05		
chr2..32494387.32494389	41.01314123	1.17E-12	TTGGAAGCTCGGGGGCGGAGATCCCGCGCCGCCCGCGGCTCGGGGCGG	89
chr22.46374563.46374565	14.42030898	1.59E-07	TTGGACGCGGAGGCCGTCCGGGCGCCGAGCAACGAGCAGCACAGCACGGTGG	113
chr6.45773846.45773848	8.210986049	9.42E-44	GAGAAAGACGTCCCTGTGCGGAAGCCGTTCGACGCGCGGCCCTGGGGGCG	222
chr5.61645225.61645227	6.475736938	3.13E-05	TAAAGGGCGATGCCGCGACCCAGGCGCTGGCGTGGGTCCACCGGCGTCGC	208
chr3.70831746.70831748	9.913046295	0.000100547	CTACCTGTGCGGGGAGGCGAGCGAGCACCGGGCGCCCGCACGTGCGCGCGCGG	160
chr15.10856498.10856500	12.95145963	1.72E-07	CCCCAGCCGGCCGGCTCGGGTTTCCGACGGAGGGCCGGGCTGCGCGGCGTC	55
chr16.8886545.8886547	7.761075839	4.37E-11	GGCGGGCTCCCCGTGCGCGCACGGCCGCGGCCGACCTCTCCGCGCGGGCTCC	61
chr3.46843750.46843752	5.373772366	5.08E-17	CCAGGGCCACAGACCTGGAAGTCGCCGGGAGCCCGGCAGCACTCCTCTCAAG	158
chr20.44942213.44942215	4.217839864	0.000131313	GGCAGGGCCGCCAGGCGCCGCCGCCCGGTGAAGAAAGTCTCGGCGCTGCGCT	101
chr20.57347921.57347923	-9.131635543	3.73E-38	ACCAGCCGCCAGCCGTGACCTTTCCGGGCCCTGTTCCTGCGGCTCTGGGCGC	109
chr24.21051024.21051026	2.892412963	3.58E-05	ACGTTGCTCTTCGTGCTCCTGCCCCTCGCTGGTCGTGCAGCTGCTCAGCTTCC	122
chr33.26512711.26512713	3.736067928	0.052011151	CGTCGCCGCCCGCGCTCCTGCCAGGCCGACCTGGAGCGGAAGCTGAAGCCGAA	177
chr23.37987526.37987528	-9.998502923	0.005079137	CGGGGCGGGAAATCCGCCGCCCGCGCTCGCCCCCGGGCCCTAGACCCGGGCGC	118
chr5.32347978.32347980	-7.154140717	0.034125216	ACACGATAGCCCCCTGGCCCCGGTCCCCCGGCCACCCCGCCAGCAGCCCGATGC	203
chr6.60675998.60676000	4.070607232	0.00085343	GATAGCTTCCTCGGCTCGGTGCGGGCGGCTGCTGAGCGGCGGAGCGGCGCGG	223
chr19.44679975.44679977	5.332567614	0.01743534	GCCCATCGCTGCGGCCCTCGGTCTCCGGGGAGAGGGAAGTAACCTCCGAGC	87
chr10.7411293.7411295	2.93908127	2.87E-07	CAGCACCGGGCATCTGGACAGCAACCGGAGAGCGGAGGGTGTCTCTCCTGGT	14
chr33.26512692.26512694	2.029402014	0.057498714	CGGCTGCCGCGCCCGCAGTCGTCGCCGCGCGCTCCTGCCAGGCCGACCTGG	175
chr1.121182146.121182148	5.132841728	6.00E-05	CTGGCCCCGGAAGGCGCAGGGCATCTCGATGTGCTCTCCCTCCCGGACTGCA	6
chr17.62156690.62156692	13.76171543	0.02258792	CGGCAGAGGAGCTGGGCTTGGGTGCCCGTGACCTTGGGGAAGTAGCGTTTCC	7

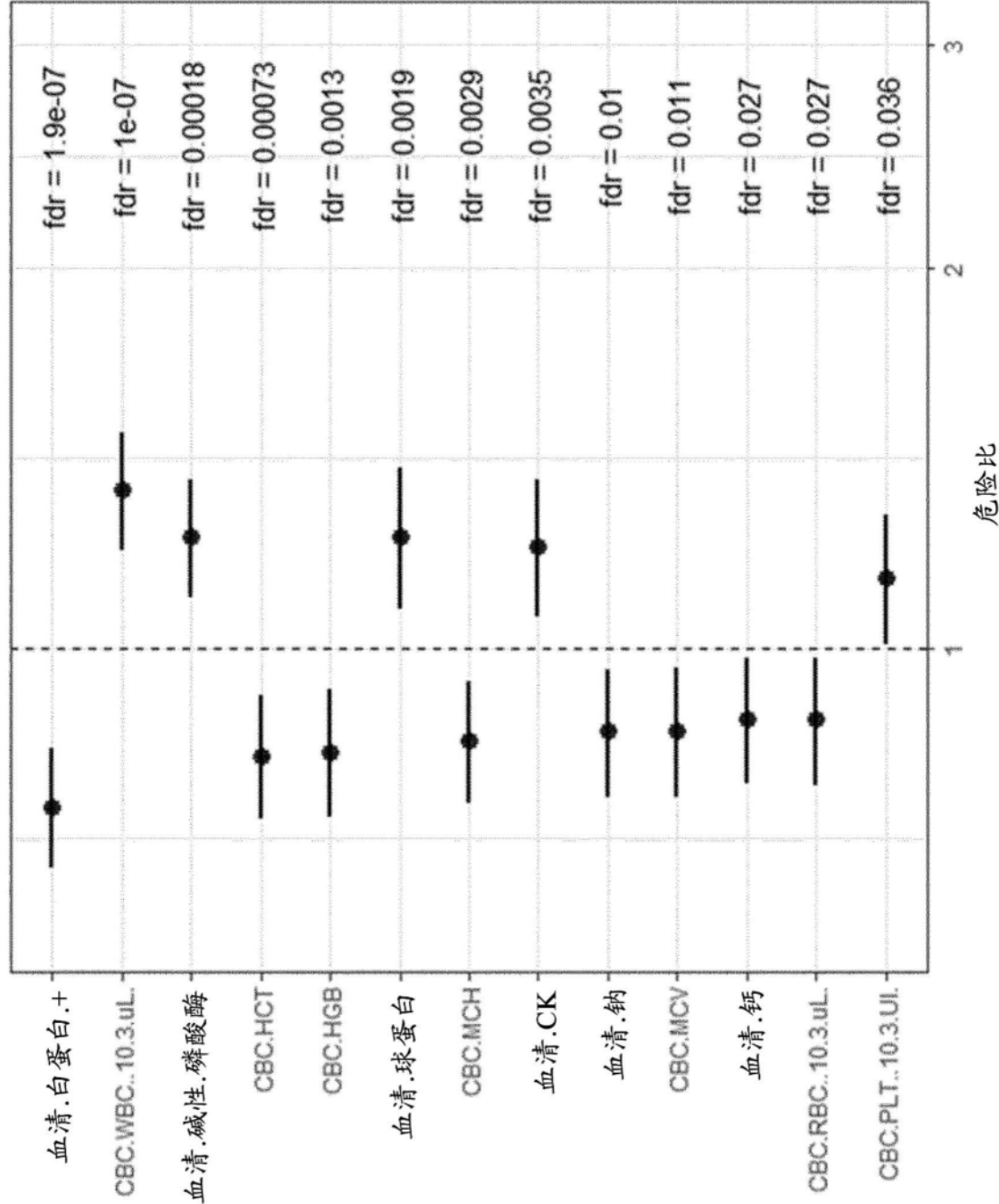


图1

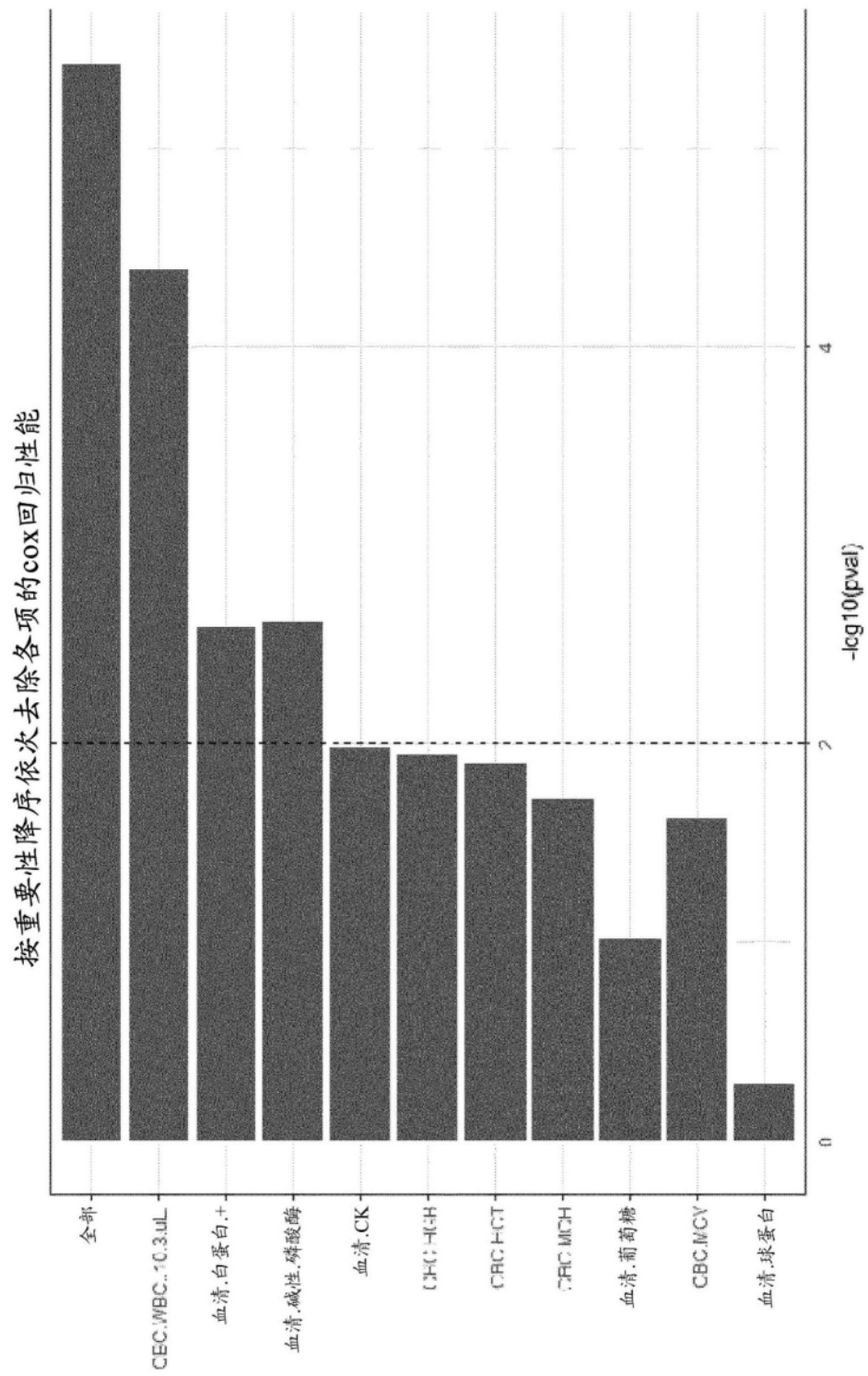


图2

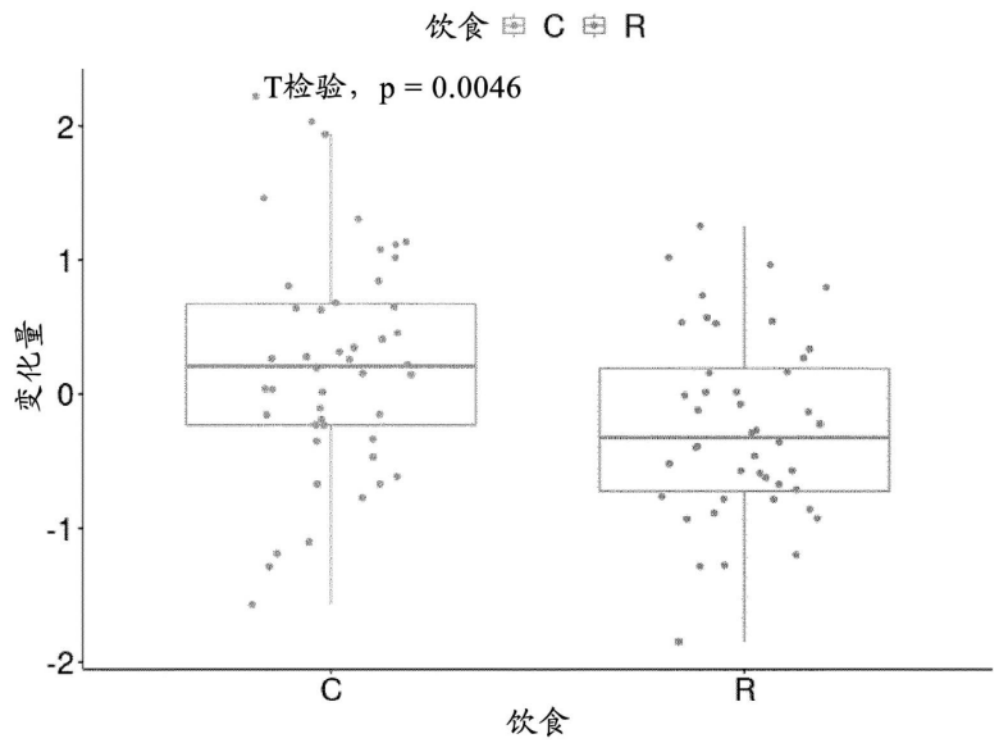


图3

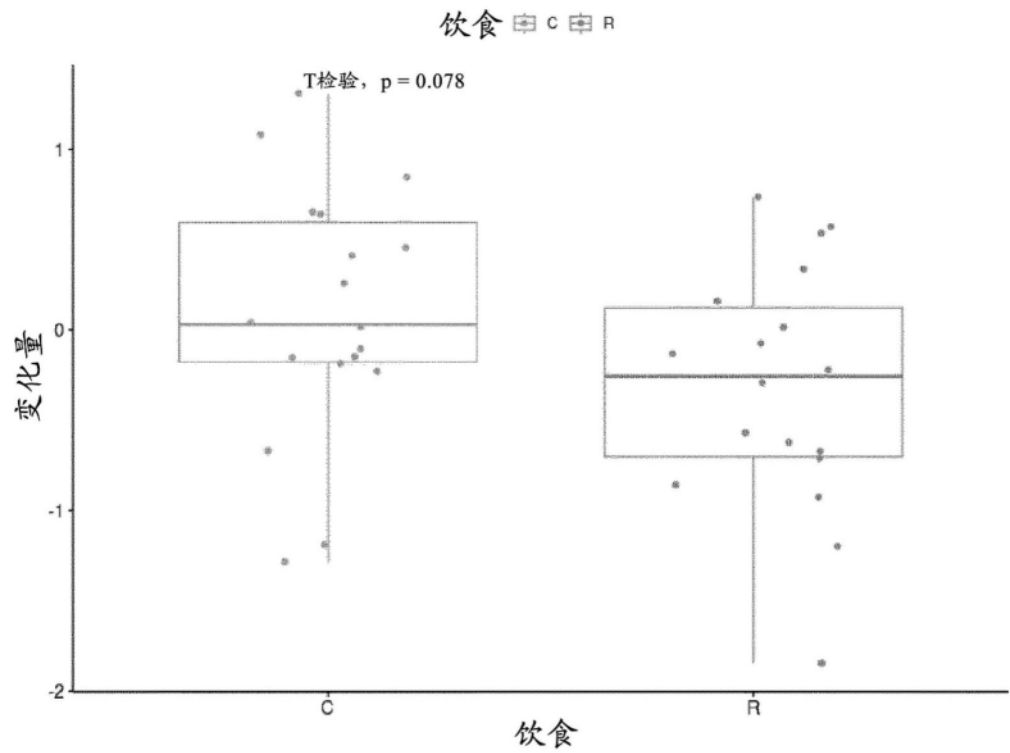


图4

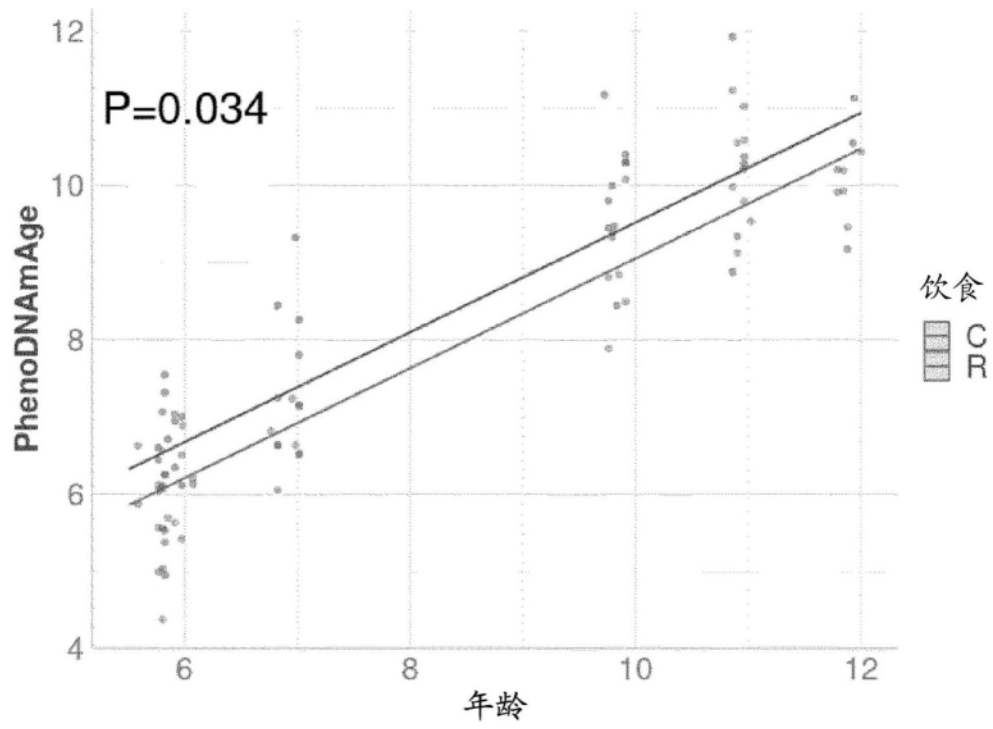


图5

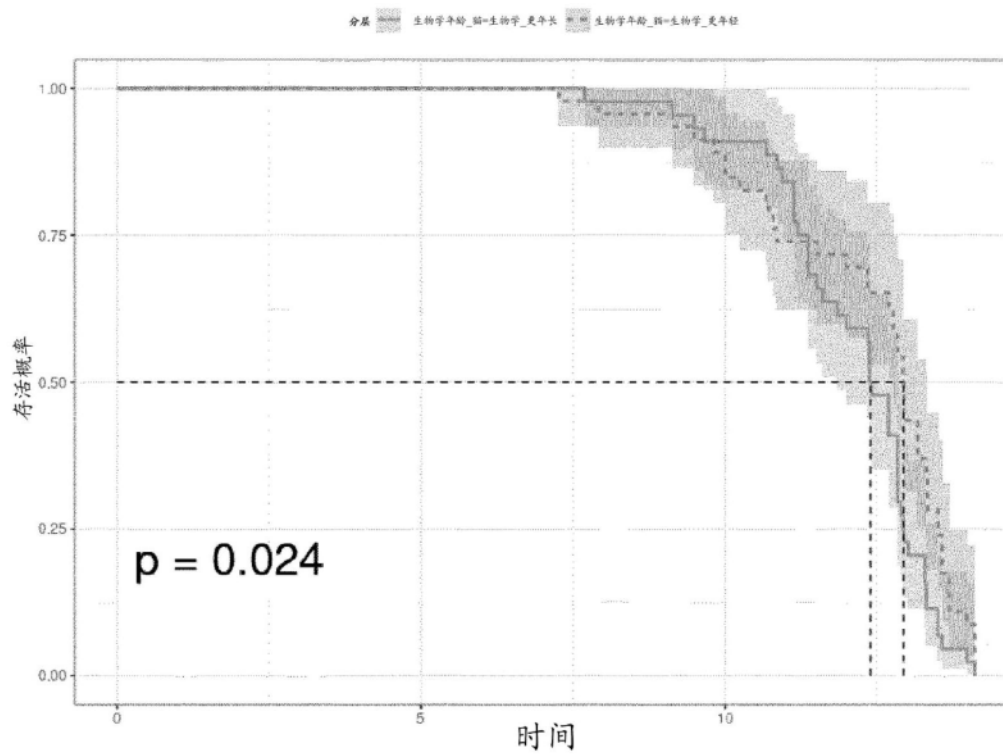


图6

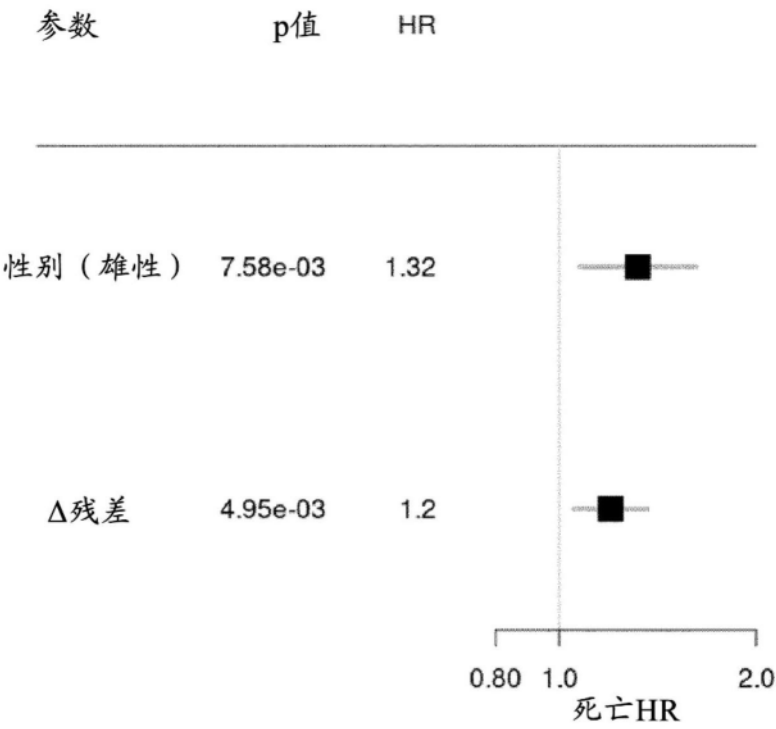


图7

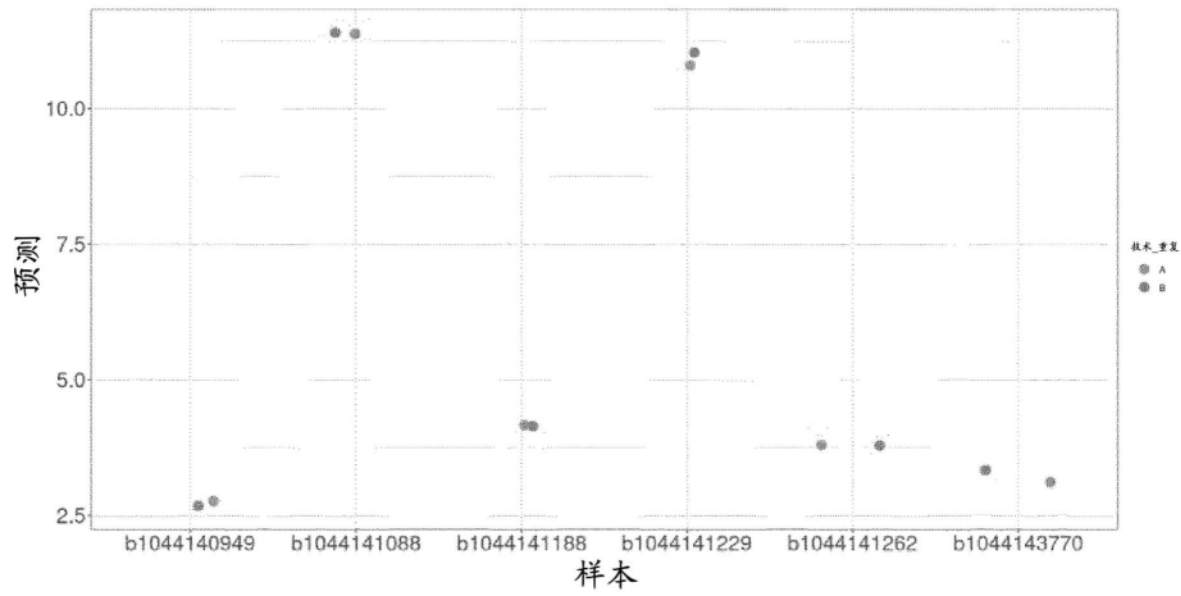


图8

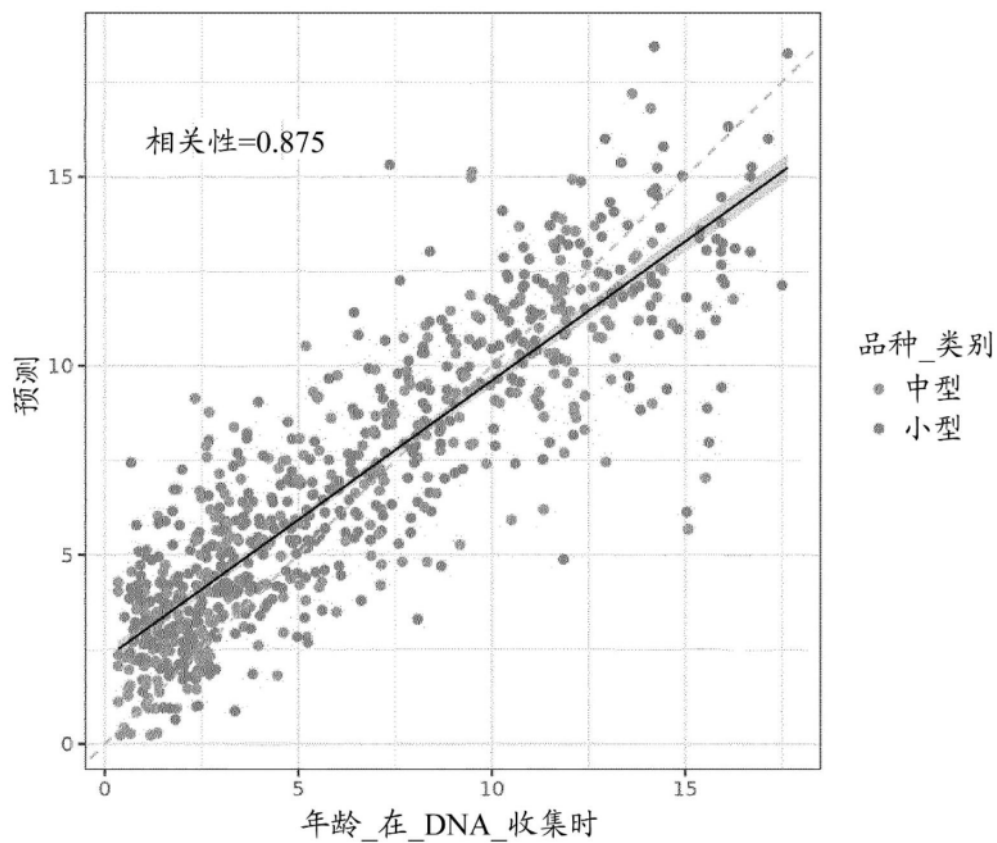


图9

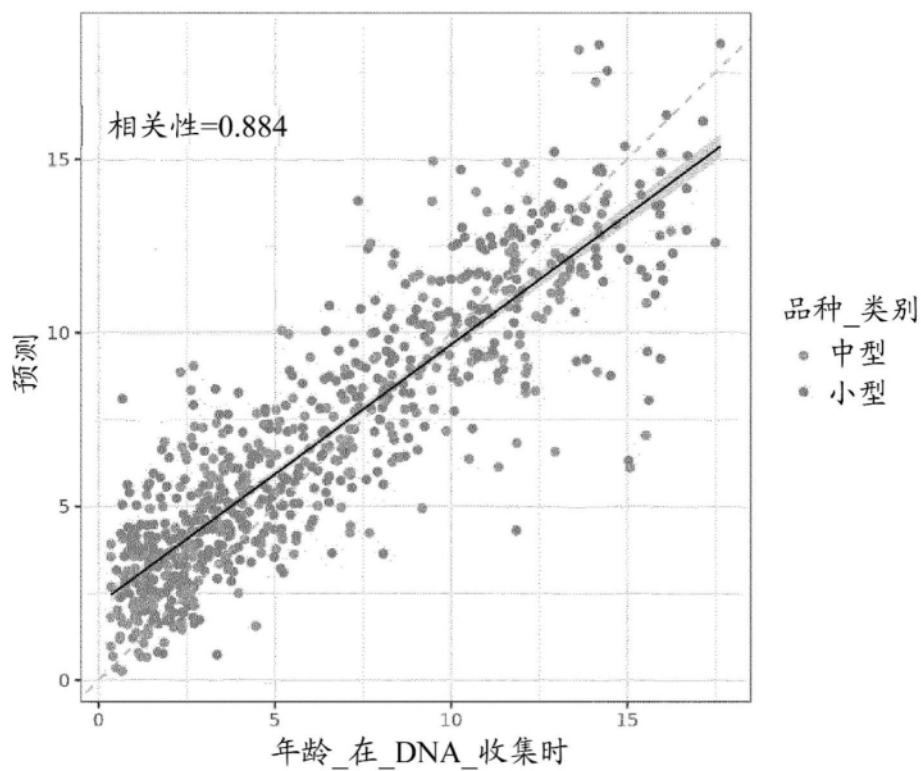


图10

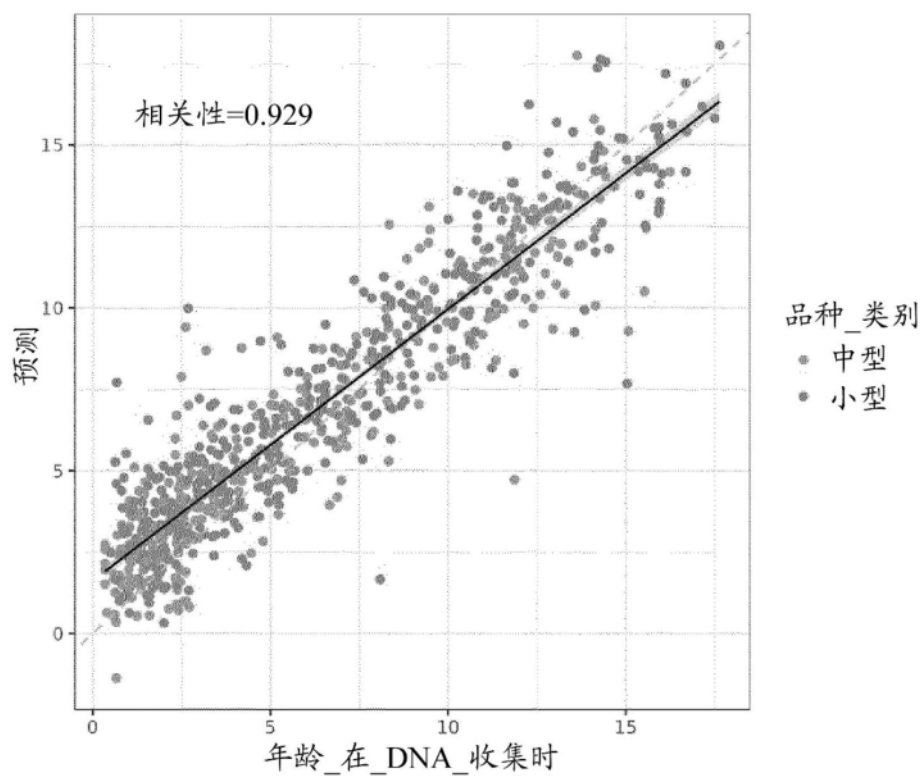


图11

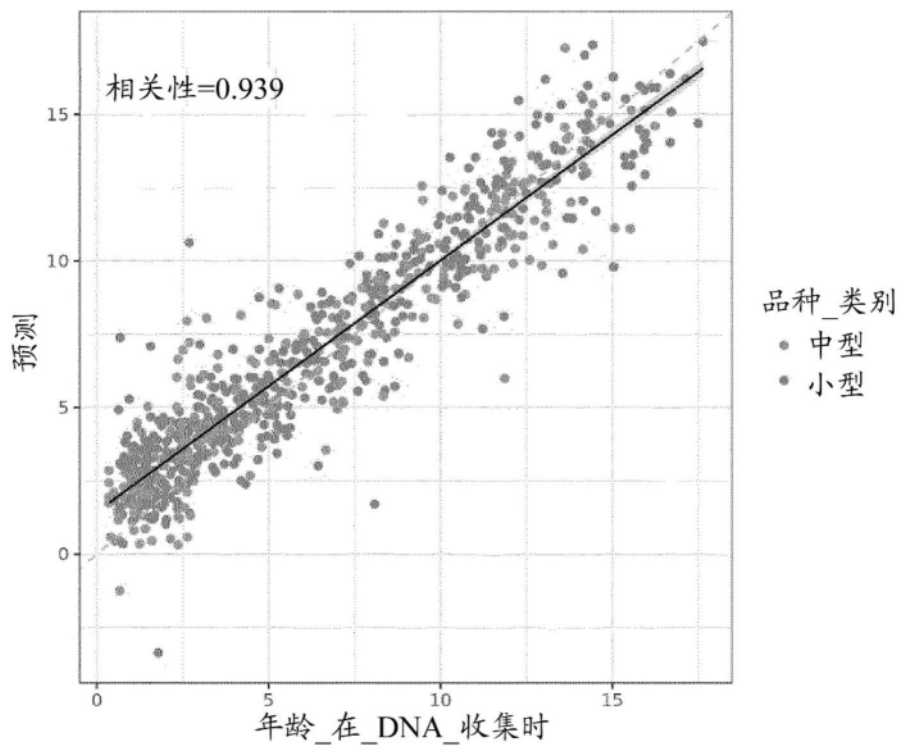


图12

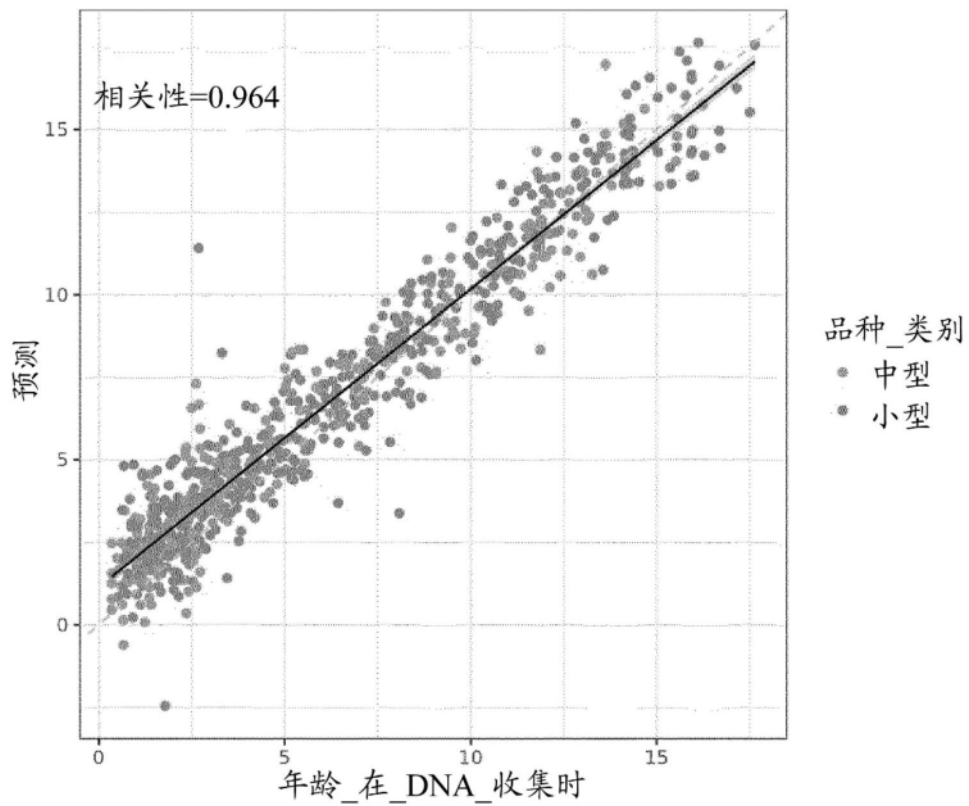


图13