



(12) 发明专利申请

(10) 申请公布号 CN 114277117 A

(43) 申请公布日 2022.04.05

(21) 申请号 202011040855.9

(22) 申请日 2020.09.28

(71) 申请人 上海市生物医药技术研究院
地址 200032 上海市徐汇区斜土路2140号
申请人 上海交通大学医学院附属第九人民
医院

(72) 发明人 牛振民 陈振琦 黄薇

(74) 专利代理机构 上海专利商标事务所有限公
司 31100

代理人 陶启长

(51) Int. Cl.

C12Q 1/6883 (2018.01)

C12N 15/11 (2006.01)

G16B 20/20 (2019.01)

权利要求书2页 说明书11页
序列表2页

(54) 发明名称

唇腭裂相关SNP标记及其应用

(57) 摘要

本发明提供与唇腭裂相关的SNP标记rs144318134,及其在评估唇腭裂患病风险或诊断唇腭裂中的应用。

1. 一种分离的来自哺乳动物的核酸分子,所述核酸分子含SNP标记rs144318134,其中,rs144318134为T或G,

优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G,

更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

2. 如权利要求1所述的核酸分子,其特征在于,

所述核酸分子长度为11bp-1000bp,和/或

所述核酸分子的核苷酸序列至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

3. 用于检测SNP标记rs144318134的引物,其中,rs144318134为T或G,

优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G,

更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

4. 如权利要求3所述的引物,其特征在于,

所述引物的扩增产物至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体,和/或

所述引物含有:(1)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物或其片段杂交的序列,所述片段含有该扩增产物自5'末端起第201位核苷酸,(2)与SEQ ID NO:1或其片段杂交的序列,所述片段含有该扩增产物自5'末端起第201位核苷酸,(3)SEQ ID NO:4或5所示的核苷酸序列或与之具有至少90%序列相同性的突变体,(4)(1)或(2)或(3)的互补序列。

5. 用于检测SNP标记rs144318134的探针,其中,rs144318134为T或G,

优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G,

更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

6. 如权利要求5所述的探针,其特征在于,

所述探针至少识别SEQ ID NO:1第196-206位核苷酸,和/或

所述探针具有:(1)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物杂交的序列,杂交的区域含有该扩增产物自5'末端起第201位核苷酸,(2)与SEQ ID NO:1杂交的序列,杂交的区域包含SEQ ID NO:1第201位核苷酸,(3)(1)或(2)的互补序列。

7. 一种记载有SNP标记rs144318134的介质,其中,rs144318134为T或G,所述介质用于与DNA测序数据比对以确定SNP标记rs144318134的基因型,

优选地,所述介质记载有权利要求1或2所述的核酸分子的序列,

优选地,所述介质是印有所述序列的卡片,例如纸质、塑料、金属、玻璃卡片,

优选地,所述介质是存储有所述序列和计算机程序的计算机可读介质,当所述计算机程序被处理器执行时,实现下述步骤:将样品的测序数据与介质记载的序列比较,从而获得所述样品中SNP标记rs144318134的基因型。

8. 一种试剂盒,其含有用于检测SNP标记rs144318134的试剂和/或记载有SNP标记

rs144318134的介质,其中,rs144318134为T或G,

优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G,更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G,

优选地,所述试剂包含权利要求3或4所述的引物和/或权利要求5或6所述的探针和任选的权利要求1或2所述的核酸分子,所述介质如权利要求7所述。

9. 检测SNP标记rs144318134的试剂和/或记载有SNP标记rs144318134的介质在制备用于评估唇腭裂风险或诊断唇腭裂的试剂盒中的用途,其中,rs144318134为T或G,

优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G,更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G,

优选地,所述试剂检测哺乳动物基因组中的SNP标记rs144318134,从而评估所述哺乳动物的后代患有唇腭裂的风险或诊断其唇腭裂,或所述试剂检测哺乳动物胎儿基因组中的SNP标记rs144318134,从而评估所述胎儿患有唇腭裂的风险或诊断其唇腭裂。

10. 如权利要求8所述的用途,其特征在于,所述试剂包含权利要求3或4所述的引物和/或权利要求5或6所述的探针和任选的权利要求1或2所述的核酸分子,所述介质如权利要求7所述。

唇腭裂相关SNP标记及其应用

技术领域

[0001] 本发明涉及唇腭裂相关SNP标记及其应用。

背景技术

[0002] 唇腭裂是一种胚胎发育过程中面部各突起融合障碍所形成的先天性发育畸形,是最常见的先天性疾病之一。在一般情况下,发病率约为1/800。根据种族、民族、地理位置和所处环境的不同,发病率也有所差异。亚洲和美洲印第安人患病率最高,通常为1/500,白种人患病率约在1/1000,非洲人群患病率最低,约1/2500,这表示不同种族之间易感性基因的相对贡献度也不同。我国为唇腭裂高发国家发病率高达1.82%,总发生率占我国出生缺陷的14.01%。唇腭裂患者后代发病危险为普通人的20-30倍。

[0003] 口腔颌面部的发育开始为胚胎发育的第三周。在胚胎第七周,嗅窝底破裂而形成鼻孔,左右侧上颌突与外侧鼻突相连形成鼻孔底及上唇,两侧内侧鼻突相连形成鼻小柱、人中及前颌,两侧下颌突也向中线生长并相连而形成下颌与下唇。在胚胎发育的第12周,前端与鼻中隔相连部分骨化形成硬腭,后端不与鼻中隔相连部分无骨质发生,即为软腭。胎儿的口和鼻逐渐具备成人的形态。

[0004] 在怀孕的第4周到第10周,由于受到某些因素的影响,使面突的正常发育与融合受到阻碍时,可使胎儿发生相应的唇、腭裂畸形。目前医学认为可能的致病因素有:1、遗传因素:部分患儿直系或旁系亲属中有类似畸形发生。2、环境因素:包括感染和损伤:怀孕初期(2个月左右)的母亲感染过病毒,如流感、风疹;母体怀孕期间患有如贫血、糖尿病、严重营养不良障碍等慢性疾病;怀孕期间服用某些药物如镇静药、抗癫痫药及激素类药等;怀孕期间母体接受过大剂量X线照射。

[0005] 唇腭裂一般分为综合征性(30%)和非综合征性(70%)。综合征性唇腭裂是由染色体异常或单基因突变引起的,一般以多发畸形的综合征形式出现。非综合征性唇腭裂表现单独畸形,是由遗传背景的易感性和环境因素相互作用造成的,为多基因遗传病,也是唇腭裂基因研究的重点。通过基因层面的检测可以高效快速估计疾病的发病概率,但现有技术中尚无非综合征性唇腭裂与基因型之间的关系。所以,本领域需要一种有效估计唇腭裂发病概率的方法。

发明内容

[0006] 本发明提出了一种与唇腭裂风险相关、能够有效用于唇腭裂风险评估或诊断唇腭裂的SNP标记,利用该SNP标记设计引物和探针,对样本DNA进行检测,从而实现对唇腭裂风险的评估,有效地解决孕前、产前唇腭裂评估的难题。

[0007] 具体而言,本发明第一方面提供一种分离的来自哺乳动物的核酸分子,所述核酸分子含SNP标记rs144318134,其中,rs144318134为T或G。

[0008] 在一个或多个实施方案中,所述核酸分子长度为11bp-1000bp、11bp-500bp、11bp-200bp、11bp-100bp、11bp-50bp或11bp-30bp;

[0009] 在一个或多个实施方案中,所述核酸分子是FGFRL1基因的片段。

[0010] 在一个或多个实施方案中,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。

[0011] 优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0012] 在一个或多个实施方案中,所述哺乳动物是人。

[0013] 在一个或多个实施方案中,所述核酸分子是SEQ ID NO:1的包含第201位核苷酸的片段。优选地,所述核酸分子的核苷酸序列至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。在一个或多个实施方案中,所述核酸分子的核苷酸序列包括SEQ ID NO:1或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0014] 本发明还提供用于检测SNP标记rs144318134的引物,其中,rs144318134为T或G。

[0015] 在一个或多个实施方案中,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。

[0016] 优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0017] 在一个或多个实施方案中,所述哺乳动物是人。

[0018] 在一个或多个实施方案中,所述引物的扩增产物包含SEQ ID NO:1第201位核苷酸。在一个或多个实施方案中,所述引物的扩增产物至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。在一个或多个实施方案中,所述引物的扩增产物包括SEQ ID NO:1或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0019] 在一个或多个实施方案中,所述引物含有:(1)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物或其片段杂交的序列,所述片段含有该扩增产物自5'末端起第201位核苷酸,(2)与SEQ ID NO:1或其片段杂交的序列,所述片段含有该扩增产物自5'末端起第201位核苷酸,(3)SEQ ID NO:4或5所示的核苷酸序列或与之具有至少90%序列相同性的突变体,(4)(1)或(2)或(3)的互补序列。

[0020] 本发明还提供用于检测SNP标记rs144318134的探针,其中,rs144318134为T或G。

[0021] 在一个或多个实施方案中,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。

[0022] 优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0023] 在一个或多个实施方案中,所述哺乳动物是人。

[0024] 在一个或多个实施方案中,探针具有:(1)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物杂交的序列,杂交的区域含有该扩增产物自5'末端起第201位核苷酸,(2)与SEQ ID NO:1杂交的序列,杂交的区域包含SEQ ID NO:1第201位核苷酸,(3)(1)或(2)的互补序列。所述第201位核苷酸是T或G。示例性地,所述探针具有(1)与SEQ ID NO:2或SEQ ID NO:3杂交的序列,杂交的区域包含SEQ ID NO:2或SEQ ID NO:3第201位核苷酸,(2)(1)的互补序列。

[0025] 在一个或多个实施方案中,探针选自以下的一种或多种:(1) 识别SEQ ID NO:1的片段的探针,所述片段包含SEQ ID NO:1第201位核苷酸,所述碱基是T或G,(3) (1) 或(2)的互补序列。优选地,所述探针选自以下的一种或多种:(1) 识别SEQ ID NO:2的片段的探针,所述片段包含SEQ ID NO:3第201位核苷酸,所述碱基是T,(2) 识别SEQ ID NO:3的片段的探针,所述片段包含SEQ ID NO:3第201位核苷酸,所述碱基是G,(3) (1) 或(2)的互补序列。在一个或多个实施方案中,所述探针至少识别SEQ ID NO:1-3中任一的第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0026] 本发明还提供记载有SNP标记rs144318134的介质,其中,rs144318134为T或G。所述介质用于与DNA测序数据比对以确定SNP标记rs144318134的基因型。

[0027] 在一个或多个实施方案中,所述介质记载有核酸分子的序列,所述核酸分子如本文第一方面所述。

[0028] 在一个或多个实施方案中,所述核酸分子是SEQ ID NO:1的包含第201位核苷酸的片段。优选地,所述核酸分子的序列至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。在一个或多个实施方案中,所述核酸分子的序列包括SEQ ID NO:1或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0029] 在一个或多个实施方案中,所述介质是印有所述序列的卡片,例如纸质、塑料、金属、玻璃卡片。

[0030] 在一个或多个实施方案中,所述介质是存储有所述序列和计算机程序的计算机可读介质,当所述计算机程序被处理器执行时,实现下述步骤:将样品的测序数据与介质记载的序列比较,从而获得所述样品中SNP标记rs144318134的基因型。

[0031] 本发明还提供一种评估唇腭裂风险或诊断唇腭裂的试剂盒,其含有用于检测SNP标记rs144318134的试剂和/或介质,其中,rs144318134为T或G。

[0032] 在一个或多个实施方案中,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。

[0033] 优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0034] 在一个或多个实施方案中,所述哺乳动物是人。

[0035] 在一个或多个实施方案中,所述试剂包含选自以下的一种或多种:本文任一实施方案所述引物和/或本文任一实施方案所述探针和任选的本文任一实施方案所述的核酸分子。

[0036] 在一个或多个实施方案中,所述试剂盒还包含选自以下一种或多种的物质:PCR缓冲液、聚合酶、dNTP、限制性内切酶、酶切缓冲液、荧光染料、荧光淬灭剂、荧光报告剂、外切核酸酶、碱性磷酸酶、内标、对照物。

[0037] 本发明还提供检测SNP标记rs144318134的试剂或介质在制备用于评估唇腭裂风险或诊断唇腭裂的试剂盒中的用途,所述试剂包含引物、探针和任选的核酸分子,所述介质是记载有SNP标记rs144318134的序列的介质。

[0038] 在一个或多个实施方案中,所述试剂检测哺乳动物基因组中的SNP标记rs144318134,其中,rs144318134为T或G,从而评估所述哺乳动物的后代患有唇腭裂的风险或诊断唇腭裂。其中,在rs144318134位点分别为TT纯合子和TG杂合子的亲本的后代有50%

的可能性在该位点含有G,从而发生唇腭裂;在rs144318134位点均为TG杂合子的两个亲本的后代有75%的可能性在该位点含有G,从而发生唇腭裂。

[0039] 所述检测通过检测哺乳动物的样品实现。所述样品包括哺乳动物的外周血。

[0040] 在一个或多个实施方案中,所述试剂检测哺乳动物胎儿基因组中的SNP标记rs144318134,其中,rs144318134为T或G,从而评估所述胎儿是否患有唇腭裂。其中,胎儿基因组的rs144318134位点含有G则胎儿患有唇腭裂。在一个或多个实施方案中,所述检测通过检测妊娠哺乳动物的样品实现。所述样品包括妊娠哺乳动物的外周血、脐带血、羊水等。

[0041] 在一个或多个实施方案中,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。

[0042] 优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0043] 在一个或多个实施方案中,所述哺乳动物是人。

[0044] 在一个或多个实施方案中,所述试剂包含选自以下的一种或多种:本文所述的核酸分子、本文所述的引物、本文所述的探针。

[0045] 在一个或多个实施方案中,所述介质为本文任一实施方案所述的介质。

[0046] 本发明还提供一种用于评估唇腭裂风险或诊断唇腭裂的方法,包括:(1)获取哺乳动物对象或哺乳动物胎儿对象的样品中rs144318134的基因型;(2)任选地与参照样品比较;(3)若哺乳动物对象的rs144318134为GG或TG则该哺乳动物对象的后代患有唇腭裂的风险增加,或,若哺乳动物胎儿对象的rs144318134为GG或TG则该胎儿对象患有唇腭裂。

[0047] 在一个或多个实施方案中,所述方法在步骤(1)之前还包括:样品中基因组DNA的抽提、质检和/或储存。

[0048] 在一个或多个实施方案中,所述样品包括哺乳动物的外周血。

[0049] 在一个或多个实施方案中,所述样品包括妊娠哺乳动物的外周血、脐带血、羊水等。

[0050] 本发明还提供一种用于评估唇腭裂风险或诊断唇腭裂的的装置,所述装置包括存储器、处理器以及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现以下步骤:(1)获取哺乳动物对象或哺乳动物胎儿对象的样品中rs144318134的基因型;(2)任选地与参照样品比较;(3)若哺乳动物对象的rs144318134为GG或TG则该哺乳动物对象的后代患有唇腭裂的风险增加,或,若哺乳动物胎儿对象的rs144318134为GG或TG则该胎儿对象患有唇腭裂。

[0051] 在一个或多个实施方案中,所述方法在步骤(1)之前还包括:样品中基因组DNA的抽提、质检和/或储存。

[0052] 在一个或多个实施方案中,所述样品包括哺乳动物的外周血。

[0053] 在一个或多个实施方案中,所述样品包括妊娠哺乳动物的外周血、脐带血、羊水等。

具体实施方式

[0054] 发明人通过筛选,发现SNP标记rs144318134的特定基因型与唇腭裂相关,因此,通过检测rs144318134可以在孕前和产前评估后代发生唇腭裂的风险。经比对发现,该SNP突

变改变FGFRL1基因第502位氨基酸,读码框从TAT(编码氨基酸为Y)变为终止密码子TAG,属于获得终止子的突变(stop_gained),严重影响基因功能。在公共数据库中,rs144318134位点突变频率小于万分之一,属于及其罕见的突变。

[0055] 具体地,本发明涉及与唇腭裂相关的SNP标记,用于检测所述SNP标记的引物、探针和试剂盒、记载有所述SNP标记的介质,所述SNP标记、引物、探针、介质、试剂盒在评估唇腭裂风险或诊断唇腭裂中的用途,以及评估唇腭裂风险或诊断唇腭裂的方法。

[0056] 发明人发现如下与唇腭裂相关的SNP标记:rs144318134位于FGFRL1基因中,具有G和T等位基因型。同时,发明人发现rs144318134位点的G等位基因型是致病等位基因型。在一个或多个实施方案中,rs144318134为以哺乳动物(例如人)基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。

[0057] 本文中,SNP(single nucleotide polymorphism,SNP,即单核苷酸多态性)是一类分子遗传标记,主要是指基因组水平上由单个核苷酸的变异所引起的DNA序列多态性。SNP表现出的多态性通常仅涉及到单个碱基的变异,例如转换、颠换、插入和缺失等。

[0058] RS命名是目前最常用的SNP命名法,命名方法是rs+阿拉伯数字,包括前后序列,位置信息,分布频率等。根据已知SNP的rs号,可在GenBank的SNP数据库中搜索到相关的信息和其在基因组中的位置。

[0059] 本文中,唇腭裂主要是非综合征性唇腭裂。发明人发现,与rs144318134基因型为TT的个体相比,带有rs144318134等位基因型G的个体全部发生了唇腭裂。因此,通过检测亲本样品中的上述SNP,能够有效地评估后代或胎儿患有唇腭裂的风险。

[0060] 本发明提供一种分离的核酸分子,所述核酸分子含SNP标记rs144318134,其中,rs144318134为T或G。在一个或多个实施方案中,rs144318134为以哺乳动物(例如人)基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。通常,本发明所述核酸分子长度为20bp-1000bp。优选地,所述核酸分子的核苷酸序列至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。在一个或多个实施方案中,所述核酸分子包含SEQ ID NO:1或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。在一个或多个实施方案中,所述核酸分子用于检测样品SNP位点基因型的标准品和/或对照。

[0061] 术语“分离的”是指材料,例如核酸分子和/或蛋白质,其基本上不含天然存在的环境中通常伴随或与材料相互作用的组分或从其中移出。分离多核苷酸可以从它们天然存在的宿主细胞中纯化。本领域技术人员已知的常规核酸纯化方法可用于获得分离的多核苷酸。

[0062] 术语“扩增”产物是指在引物定向扩增反应过程中产生的核酸片段。引物定向扩增的一般方法包括聚合酶链反应(PCR),连接酶链反应(LCR)或链置换扩增(SDA)。如果选择PCR方法,则复制组合物可包含用于核酸复制的组分,例如:核苷酸三磷酸,具有适当序列的两种(或更多种)引物,热稳定聚合酶,缓冲液,溶质和蛋白质。在一个或多个实施方式中,扩增产物至少30bp、至少50bp、至少100bp、200bp,至少500bp,至少1kb,至少2kb,至少3kb,至少4kb,至少5kb,至少6kb,至少7kb,至少8kb,至少9kb,至少10kb或至少12kb。

[0063] 术语“核酸”、“核酸分子”、“多核苷酸”、“多核苷酸序列”、“核酸序列”和“核酸片段”在本文中可互换使用,指单链或双链形式的脱氧核糖核苷酸(DNA)或核糖核苷酸聚合物(RNA),及其互补物。核酸含有合成的、非天然的或改变的核苷酸碱基。核苷酸可以是核糖核苷酸、脱氧核糖核苷酸、或其修饰形式。本文考虑的多核苷酸的示例包括单链和双链DNA,单链和双链RNA,以及具有单链和双链DNA和RNA的混合物的杂合分子。DNA可以是编码链或非编码链。在一个或多个实施方案中,所述样品包含经片段化的基因组DNA。获取基因组DNA并片段化的方法本领域周知。在一个或多个实施方案中,所述样品包含cDNA。本发明的核酸可包含碱基(包括尿嘧啶、腺嘌呤、胸腺嘧啶、胞嘧啶、鸟嘌呤、肌苷、黄嘌呤次黄嘌呤、异胞嘧啶和异鸟嘌呤)的组合。本发明的核酸可经合成以包含非天然氨基酸修饰。本发明的核酸可通过化学合成方法或通过重组方法来获得。

[0064] DNA形式包括cDNA、基因组DNA、片段化DNA或人工合成的DNA。DNA可以是单链的或是双链的。DNA可以为任意长度,例如50-500bp,100-400bp,150-300bp或200-250bp。

[0065] 本文所述试剂主要是“引物”或“探针”。本文所述“引物”是指在核苷酸聚合作用起始时,引导合成的一种具有特定核苷酸序列的核酸分子。引物通常是人工合成的两段寡核苷酸序列,一个引物与靶区域一端的一条DNA模板链互补,另一个引物与靶区域另一端的另一条DNA模板链互补,其功能是作为核苷酸聚合作用的起始点。体外人工设计的引物被广泛用于聚合酶链反应(PCR)、qPCR、测序和探针合成等。引物可以为任意长度,例如5-200bp,10-100bp,20-800bp或25-50bp。

[0066] 本文所述“互补”、“识别”或“杂交”指引物与模板序列在严谨或高度严谨条件下杂交。本文所述核酸杂交的严谨条件本领域技术人员已知。优选地,所述条件是使得序列彼此至少约65%、70%、75%、85%、90%、95%、98%或99%同源,通常保持彼此杂交。严谨杂交条件的非限制性实例是在含有6xSSC、50mM Tris-HCl(pH7.5)、1mMEDTA、0.02%PVP、0.02%Ficoll11、0.02%BSA和500mg/ml变性鲑鱼精子DNA的高盐缓冲剂中在65℃杂交,和任选的在0.2xSSC、0.01%BSA中50℃洗涤一次或两次。

[0067] 本发明引物用于检测SNP。具体地,所述引物用于检测SNP标记rs144318134,其中,rs144318134为T或G。优选地,rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。更优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。所述引物的扩增产物至少包括SEQ ID NO:1第196-206位核苷酸或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0068] 在一些实施方式中,引物选自(1)SEQ ID NO:4或5所示的核苷酸序列或与之具有至少90%序列相同性的突变体,(2)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物或其片段杂交的序列,所述片段含有该扩增产物自5'末端起第201位核苷酸,(3)(1)或(2)的互补序列。

[0069] 本发明也可采用探针来检测本发明所述的SNP。本文所述“探针”是识别目的序列(与目的序列互补)的核酸序列(DNA或RNA)。探针通过分子杂交与目的基因结合,产生杂交信号,从而显示目的基因。探针可以包括整个目的序列,也可以是目的序列的片段。探针可以是DNA,也可以是由之转录而来的RNA。通常,探针带有检测标记,例如荧光标记。所述荧光标记包括但不限于FAM、CY5和VIC。本领域知晓适用于本文探针的荧光标记以及将其与探针

连接的方法。

[0070] 本文中,探针具有(1)与SEQ ID NO:1杂交的序列,杂交的区域包含SEQ ID NO:1第201位核苷酸,(2)与以哺乳动物基因组DNA为模板采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物杂交的序列,杂交的区域含有该扩增产物自5'末端起第201位核苷酸,(3) (1)或(2)的互补序列。所述第201位核苷酸是T或G。示例性地,所述探针具有(1)与SEQ ID NO:2或SEQ ID NO:3杂交的序列,杂交的区域包含SEQ ID NO:2或SEQ ID NO:3第201位核苷酸,(2) (1)的互补序列。

[0071] 在一个或多个实施方案中,探针选自以下的一种或多种:(1)识别SEQ ID NO:1的片段的探针,所述片段包含SEQ ID NO:1第201位核苷酸,所述碱基是T或G,(3) (1)或(2)的互补序列。优选地,所述探针选自以下的一种或多种:(1)识别SEQ ID NO:2的片段的探针,所述片段包含SEQ ID NO:3第201位核苷酸,所述碱基是T,(2)识别SEQ ID NO:3的片段的探针,所述片段包含SEQ ID NO:3第201位核苷酸,所述碱基是G,(3) (1)或(2)的互补序列。优选地,探针至少识别SEQ ID NO:1-3中任一的包括第196-206位核苷酸的片段或与其具有90%序列相同性并且第201位核苷酸为T或G的变体。

[0072] 本文术语“变体”或“突变体”是指与参照序列相比,通过一个或多个核苷酸的插入、缺失或取代使核酸序列发生变化同时保留其与其他核酸杂交能力的多核苷酸。本文任一实施方案所述的突变体包括与参照序列具有至少70%,优选至少80%,优选至少85%,优选至少90%,优选至少95%,优选至少97%的序列相同性并保留参照序列的生物学活性的核苷酸序列。可采用例如NCBI的BLASTn计算两条比对的序列之间的序列相同性。突变体还包括在参照序列的和核苷酸序列中具有一个或多个突变(插入、缺失或取代)、同时仍保留参照序列生物学活性的核苷酸序列。所述多个突变通常指1-8个以内、例如1-5个或1-3个。取代可以是嘌呤核苷酸与嘧啶核苷酸之间的取代,也可以是嘌呤核苷酸之间或嘧啶核苷酸之间的取代。取代优选是保守性取代。例如,在本领域中,用性能相近或相似的核苷酸进行保守性取代时,通常不会改变多核苷酸的稳定性和功能。保守性取代例如嘌呤核苷酸之间的(A与G)的互换,嘧啶核苷酸之间的(T或U与C)的互换。因此,在本发明多核苷酸中用来自同一残基替换一个或几个位点,将不会在实质上影响其活性。当提及与本发明所述的引物(如SEQ ID NO:7-10)或探针具有至少75%、至少80%、至少85%、至少90%、至少95%、至少97%序列相同性的突变体时,优选地,这些突变体在严谨条件下可与包含SEQ ID NO:1-3的相对应的DNA序列杂交。

[0073] 在某些实施方案中,本发明还提供记载有SNP标记rs144318134的介质,其中,rs144318134为T或G。所述介质用于与DNA测序数据比对以确定SNP标记rs144318134的基因型。示例性的介质记载本文所述核酸分子的序列。

[0074] 本文所述介质是任何能提供所述miRNA的序列信息的介质,例如印有所述序列的卡片,例如纸质、塑料、金属、玻璃卡片;存储有所述序列的计算机可读介质。该计算机可读介质还存储有计算机程序,计算机可读介质,当所述计算机程序被处理器执行时,实现下述步骤:将样品的测序数据与介质记载的序列比较,从而获得所述样品中SNP标记rs144318134的基因型。

[0075] 本文提供在孕前和产前对后代或胎儿患有唇腭裂风险的评估。孕前诊断或孕前检查是指怀孕之进行身体检查,以评估生育出健康的婴儿可能性,从而实现优生。孕前诊断主

要是针对生殖系统和遗传因素所致疾病的评估。孕前检查最佳时间是怀孕前3~6个月。产前诊断是指在出生前对胚胎或胎儿的发育状态、是否患有疾病等方面进行检测诊断。

[0076] 本发明评估唇腭裂风险的方法包括检测哺乳动物基因组中的SNP标记rs144318134,其中,rs144318134为T或G,从而评估所述哺乳动物的后代患有唇腭裂的风险。所述方法进一步包括:(1)提取待测样品的DNA;(2)利用本文所述的试剂或介质,确定或定量所述DNA中的本文所述SNP标记的基因型;和(3)基于(2)的结果,评估所述哺乳动物的后代患有唇腭裂的风险。例如,通过检测亲本的上述SNP标记的基因型,从而获取后代所述标记的基因型,继而推测后代患有唇腭裂的风险。例如,亲本分别为rs144318134 TT纯合子和rs144318134 TG杂合子,其后代将有50%几率带有风险等位基因型G,该后代有至少50%的几率发生唇腭裂。例如,亲本分别为rs144318134 TG杂合子和rs144318134 TG杂合子,其后代将有75%几率带有风险等位基因型G,该后代有至少75%的几率发生唇腭裂。

[0077] 本发明另一方面提供评估唇腭裂风险的方法,包括检测哺乳动物基因组中的SNP标记rs144318134,其中,rs144318134为T或G,从而评估所述哺乳动物患有唇腭裂的风险。在一个或多个实施方案中,所述哺乳动物是动物胎儿。所述方法进一步包括:(1)提取待测样品(例如妊娠哺乳动物的外周血、脐带血或羊水)的DNA;(2)利用本文所述的试剂或介质,确定或定量所述DNA中的本文所述SNP标记的基因型;和(3)基于(2)的结果,评估动物胎儿患有唇腭裂的风险。例如,带有rs144318134等位基因型G的胎儿将罹患唇腭裂。

[0078] 本文中,提取样品中DNA的方法不受特别限制,本领域周知适用于本文的DNA提取方法。如本文所用的术语“样品”是指来自对象的任何组织或流体,其适于富集核酸。对象可以是任何活体或非活体生物,包括但不限于人、非人动物。能选择任何人或非人动物,包括但不限于哺乳动物、爬行动物、鸟类、两栖类、鱼类、有蹄类动物、反刍动物、牛科动物(如牛)、马科动物(如马)、山羊和绵羊类动物(如绵羊、山羊)、猪科动物(如猪)、羊驼类动物(如骆驼、美洲驼、羊驼)、猴子、猿(如大猩猩、黑猩猩)、熊科动物(如熊)、家禽、犬、猫、小鼠、大鼠、鱼、海豚、鲸鱼和鲨鱼。对象可为男性或女性(例如妇女、妊娠妇女)。对象可为任何年龄(如胚胎、胎儿、婴儿、儿童、成人)。

[0079] 核酸可以从任何类型的合适生物试样或样品中分离。样品或测试样品可为分离或获自对象或其部分的任何试样。试样的非限制性示例包括对象的液体或组织,包括但不限于血液或血液制品、脐带血、绒毛、羊水、脑脊液、脊髓液、洗液、活检样品、膜间液样品、细胞或其部分、女性生殖道清洗物、尿、粪便、痰、唾液、鼻黏膜、前列腺液、灌洗液、精液、淋巴液、胆汁、眼泪、汗液、母乳、乳腺体液等或其组合。在一些实施方式中,生物样品可以是血液,而有时是血浆或血清。其他合适的生物样品对于相关领域的普通技术人员来说是熟悉的。可以使用完全在临床实践者的普通知识范围内的技术获得生物样品。

[0080] 在一些实施方式中,本发明方法之前、期间或之后对核酸进行片段化或切割。本文所用的“片段化”或“切割”指使核酸分子(如核酸模板基因分子或其扩增产物)可以分成两个或更多较小核酸分子的方法或条件。核酸片段可含有重叠的核苷酸序列,这样的重叠序列可促进构建未片段化的对应核酸或其区段的核苷酸序列。在某些实施方式中,核酸可以是部分片段化的(例如,来自未完全的或中止的特异性剪切反应)或完全片段化的。这种片段化或剪切可以是序列特异性、碱基特异性或非特异性的,并且能通过任意不同方法、试剂或条件(包括例如化学、酶、物理片段化)来完成。

[0081] 本领域周知适用于本文的SNP标记检测方法,包括但不限于:测序、单链构象多态性聚合酶链式反应(PCR single strand conformation polymorphism,PCR-SSCP)、实时荧光定量PCR与高分辨率熔解曲线分析(HRM)、荧光探针法定量PCR、限制性片段长度多态性聚合酶链式反应(PCR-restriction fragment length polymorphism,PCR-RFLP)及飞行时间质谱等。本领域知晓SNP标记检测方法中所需的除引物以外的其他试剂。

[0082] 根据本发明的一些具体示例,通过对待测样品进行本文所述的SNP标记的检测来评估患有唇腭裂的风险的方法,进一步包括:提取样品中的DNA;利用本文所述引物(例如SEQ ID NO:4-5),进行DNA的荧光定量PCR,获得扩增产物;对所述扩增产物进行测序,获得所述DNA中的本文所述SNP标记的基因型;以及基于所述SNP标记的基因型,评估患有唇腭裂的风险。

[0083] 根据本发明的一些具体示例,通过对待测样品进行本文所述的SNP标记的检测来评估患有唇腭裂的风险的方法,进一步包括:提取样品中的DNA;利用本文所述引物(例如SEQ ID NO:4-5),识别扩增产物的探针,以及参照基因的引物和探针,进行DNA的荧光定量PCR,获得扩增产物;分析PCR结果,获得所述DNA中的本文所述SNP标记的基因型;以及基于所述SNP标记的基因型,评估患有唇腭裂的风险。

[0084] 本发明还提供一种试剂盒,其含有用于检测本发明所述的哺乳动物基因组中SNP标记rs144318134的试剂和/或介质,其中,rs144318134为T或G。rs144318134为以哺乳动物基因组DNA为模板,采用SEQ ID NO:4和5作为引物进行PCR扩增得到的扩增产物的自5'末端起第201位核苷酸,其为T或G。优选地,rs144318134为SEQ ID NO:1第201位核苷酸,其为T或G。所述试剂可以是本文任一实施方案所述的引物和/或探针。任选地,该试剂盒还包括本发明所述的核酸分子(即扩增产物),该核酸分子可用作内标或阳性对照。优选地,所述引物选自SEQ ID NO:4和5所示的序列或与其有至少90%相同性的序列。试剂盒中还可含有实施PCR所需的各种试剂,如缓冲液、酶、dNTP等。在优选的实施方案中,本发明的试剂盒含有SEQ ID NO:4和5所示的引物序列。

[0085] 术语“约”或“大约”表示在本领域普遍技术人员确定的特定数值的可接受误差范围内,这将取决于该数值的测量或确定方式,例如,测量体系的极限。例如,“约”可表示在1个或多个标准偏差范围内。或者,“约”可表示给定值的多至20%,或多至10%,或多至5%,或多至1%的范围。或者,尤其对于生物系统或过程而言,该术语可表示某值的一定数量级内,优选5倍内,更优选2倍内。本申请和权利要求书中描述具体值时,除非另有说明,假定术语“约”表示该具体值的可接受的误差范围内。

[0086] 除非有明确的另外说明,所有的百分数和比例/比率都是以重量计。

[0087] 除非另有说明,所有百分比和比例是基于组合物总量计算的。

[0088] 贯穿本公开内容给出的每个最大数值限制包括每个较低的数值限制,如同这些较低的数值限制在本文中明确写出一样。本公开全文中给出的每个最小数值限定包括每个较高的数值限定,如同这些较高的数值限定清楚写在这里。本公开全文中给出的每个数值范围包括落入所述较宽数值范围的每个较窄数值范围,如同本文中明确写明这些较窄数值范围。

[0089] 本文所述的值不应理解为严格限定为所述的精确的数值。相反,除非另有具体说明,所述值各自用于指所述的值和该值周围的功能等同的范围。例如,公开为“20 μ l”的值意

在表示“约20 μ l”。

[0090] 除非明确排除或者其他方式限制,否则本文所引述的每篇文献,包括任何交叉引用和相关的专利或申请,通过引用结合入本文。任何文献的引用不是对其作为本文所公开的或受权利要求书保护的发明的现有技术,或者其单独地或者与任何其它参考文献的任何组合,或者参考、提出、建议或公开任何此类发明的认可。此外,当本文中术语的任何含义或定义与以引用方式并入的文件中术语的任何含义或定义矛盾时,应当服从在本文中赋予该术语的含义或定义。

[0091] 下文将以具体实施例的方式阐述本发明。应理解,这些实施例仅仅是阐述性的,并不意图限制本发明的范围。在不脱离本公开的精神和范围的情况下,可以进行各种其他改变和修改。所附权利要求的范围包括在本公开范围内的所有这些改变和修改。实施例中未具体描述的材料、试剂和方法,均为本领域常规的材料、试剂和方法。

[0092] 实施例

[0093] 实施例1,研究对象

[0094] 病例样本全部采集自上海交通大学医学院附属第九人民医院门诊病人,患者的诊断为非综合征性唇腭裂。在知情同意的基础上,收集了190例非综合征性唇腭裂病人和480例健康的正常对照个体的外周血样本。

[0095] 实施例2,鉴定SNP

[0096] 1、DNA提取

[0097] 用常规酚氯仿法从人的外周血样本中提取DNA(也可采用商业化的试剂盒提取DNA),浓度校正至20ng/ μ l后,用于常规PCR扩增。

[0098] 2、用于PCR和测序中的引物设计

[0099] 发明人选取大量SNP,根据GenBank中的基因组序列,设计和合成引物。其中,根据GenBank中FGFRL1基因rs144318134的基因组序列,设计和合成以下引物:

[0100] 正向引物P1:5' -TGCTGGCCCTAAGTTGTACC-3' (SEQ ID NO:4)

[0101] 反向引物P2:5' -GTCCTCCATCCTCCCAGTCT-3' (SEQ ID NO:5)。

[0102] 3、基因的PCR扩增

[0103] 以提取的DNA为模板,用Taq酶,在GeneAmp 9700PCR仪上以Touchdown程序进行PCR扩增。反应条件为:94 $^{\circ}$ C预变性2分钟,94 $^{\circ}$ C变性30秒,63 $^{\circ}$ C退火40秒,72 $^{\circ}$ C延伸40秒,共10个循环,每个循环退火温度递减0.5 $^{\circ}$ C;以后94 $^{\circ}$ C变性30秒,58 $^{\circ}$ C退火40秒,72 $^{\circ}$ C延伸40秒,共30个循环;最后72 $^{\circ}$ C延伸7分钟。

[0104] PCR扩增产物经琼脂糖凝胶电泳验证,获得各基因的扩增产物。

[0105] 4、SNP的发现和检测

[0106] PCR产物经Resin树脂纯化后,用ABI-3730DNA测序仪(美国应用生物系统公司Applied Biosystems(ABI))进行测序,用Polyphred软件进行序列的判读、基因分型和SNP确认。

[0107] 发现以下与唇腭裂相关的SNP:

[0108] SEQ ID NO:1中第201位T \rightarrow G(rs144318134):

[0109] GCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGTGTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCAC TTACTGGGCCAGGCCAGTTGCTGGCCCTAAGTTGTACCCAACTCTACACAGACATCCACACACACACACAC

ACACTCTCACACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTAKCAGTGCTAGACGGCACCGTATC
TGCAGTGGGCACGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCTGCAGACGAAGGCAGGGGA
CCCATGGCGAGGAGGAATGGCCAGCACCCCAGGCAGTCTGTGTGTGAGGCATAGCCCCTGGACACACACACAGAC
ACACACACTGCCTGGATGCATGTA

[0110] 其中,序列中的“K”即为SNP位点。

[0111] 实施例3,SNP基因分型和关联分析

[0112] 用直接单向测序法进行SNP基因分型。通过测序在唇腭裂病人和正常对照组中进行分型和关联分析,结果如下表所示。

[0113] 对基因分型结果进行描述性统计分析,列联表进行卡方检验。观察基因型在病人和对照之间是否具有差异,其分析结果如下:

[0114] 结果发现,在190个病人中,有4个病人带有位点rs144318134等位基因G;而在480例对照中,无人带有此位点。在公共数据库(例如千人基因组计划数据库)中,rs144318134位点突变频率小于万分之一,属于及其罕见的突变。该突变改变FGFRL1基因第502位氨基酸,读码框从TAT(编码氨基酸为Y)变为终止密码子TAG,属于获得终止子的突变(stop_gained),严重影响基因功能。FGFRL1基因是一个已知唇腭裂基因FGFR1的关键配体,与FGFR1基因共同行使生物学功能。终止子突变可以严重破坏这个生物通路的作用,从而引发严重的唇腭裂表型。

[0115] 以上结果表明:带有rs144318134等位基因型G的个体,其唇腭裂的易感性增加,风险升高。

[0116] 本文涉及的序列如下

[0117]

序列	说明
SEQ ID NO:1	rs144318134
SEQ ID NO:2	rs144318134风险型
SEQ ID NO:3	rs144318134非风险型
SEQ ID NO:4	rs144318134的示例引物1
SEQ ID NO:5	rs144318134的示例引物2

序列表

<110> 上海人类基因组研究中心

上海交通大学医学院附属第九人民医院

<120> 唇腭裂相关SNP标记及其应用

<130> 205796

<160> 5

<170> SIPOSequenceListing 1.0

<210> 1

<211> 401

<212> DNA

<213> Artificial Sequence

<400> 1

```
gccgccctca gcgctggccc tgggtgtgggg ctgtgtgagg agcatgggtc tccggcagcc 60
ccccagcact tactgggccc aggcccagtt gctggcccta agttgtacc caaactctac 120
acagacatcc acacacacac acacacacac tctcacacac actcacacgt ggagggcaag 180
gtccaccagc acatccacta kcagtgctag acggcacctg atctgcagtg ggcacggggg 240
ggccggccag acaggcagac tgggaggatg gaggacggag ctgcagacga aggcagggga 300
cccatggcga ggaggaatgg ccagcacccc aggcagtctg tgtgtgaggc atagcccctg 360
gacacacaca cacagacaca cacactgcct ggatgcatgt a 401
```

<210> 2

<211> 401

<212> DNA

<213> Artificial Sequence

<400> 2

```
gccgccctca gcgctggccc tgggtgtgggg ctgtgtgagg agcatgggtc tccggcagcc 60
ccccagcact tactgggccc aggcccagtt gctggcccta agttgtacc caaactctac 120
acagacatcc acacacacac acacacacac tctcacacac actcacacgt ggagggcaag 180
gtccaccagc acatccacta gcagtgctag acggcacctg atctgcagtg ggcacggggg 240
ggccggccag acaggcagac tgggaggatg gaggacggag ctgcagacga aggcagggga 300
cccatggcga ggaggaatgg ccagcacccc aggcagtctg tgtgtgaggc atagcccctg 360
gacacacaca cacagacaca cacactgcct ggatgcatgt a 401
```

<210> 3

<211> 401

<212> DNA

<213> Artificial Sequence

<400> 3

```
gccgccctca gcgctggccc tgggtgtgggg ctgtgtgagg agcatgggtc tccggcagcc 60
ccccagcact tactgggccc aggcccagtt gctggcccta agttgtacc caaactctac 120
```

acagacatcc acacacacac acacacacac tctcacacac actcacacgt ggagggaag 180
gtccaccagc acatccacta tcagtgctag acggcacctg atctgcagtg ggcacggggg 240
ggccggccag acaggcagac tgggaggatg gaggacggag ctgcagacga aggcagggga 300
cccatggcga ggaggaatgg ccagcacccc aggcagtctg tgtgtgaggc atagcccctg 360
gacacacaca cacagacaca cacactgcct ggatgcatgt a 401

<210> 4

<211> 20

<212> DNA

<213> Artificial Sequence

<400> 4

tgctggccct aagttgtacc 20

<210> 5

<211> 20

<212> DNA

<213> Artificial Sequence

<400> 5

gtcctccatc ctcccagtct 20