



## (12)发明专利申请

(10)申请公布号 CN 106929568 A

(43)申请公布日 2017.07.07

(21)申请号 201611005880.7

(22)申请日 2010.10.06

(30)优先权数据

61/278,510 2009.10.07 US

(62)分案原申请数据

201080055282.5 2010.10.06

(71)申请人 弗·哈夫曼-拉罗切有限公司

地址 瑞士巴塞尔

(72)发明人 T·W·伯伦斯 R·R·格拉姆

(74)专利代理机构 北京市中咨律师事务所

11247

代理人 史文静 黄革生

(51)Int.Cl.

C12Q 1/68(2006.01)

权利要求书2页 说明书50页

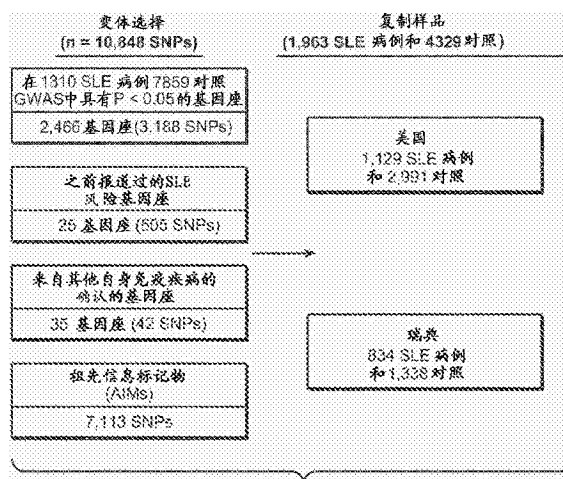
序列表4页 附图14页

### (54)发明名称

用于治疗、诊断和监控狼疮的方法

### (57)摘要

提供了鉴别、诊断和预测狼疮的方法,包括一些狼疮的亚表型(subphenotype),以及提供了治疗狼疮的方法,包括患者的一些亚群。提供的方法基于一组与系统性红斑狼疮(SLE)风险基因座相关的等位基因,所述风险基因座包括BLK、TNIP1、PRDM1、JAZF1、UHRF1BP1、IL10、IFIH1、CFB、CEC16A、IL12B和SH2B3,所述风险基因座对SLE风险有作用。还提供了鉴别有效的狼疮治疗剂的方法,和预测对狼疮治疗剂的响应性的方法。



1. 在对象中鉴定狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且其中对象怀疑患有狼疮。

2. 在对象中鉴定狼疮的方法,方法包括在源自对象的生物学样品中检测表4所述的至少一个SLE风险基因座中变异的存在,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。

3. 预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在SLE风险基因座中包括变异,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,其中BLK基因座中变异的存在表示对象对治疗剂的响应性。

4. 预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表4所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的单核苷酸多态性 (SNP) 位置对应的核苷酸位置上,其中在至少一个基因座中变异的存在表示对象对治疗剂的响应性。

5. 预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表6所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的单核苷酸多态性 (SNP) 位置对应的核苷酸位置上,其中在至少一个基因座中变异的存在表示对象对治疗剂的响应性。

6. 在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中:

(a) 生物学样品已知包括或怀疑包括了包含BLK基因座中的变异的核酸;

(b) 变异发生在与单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶;和

(c) BLK基因座中变异的存在是对象中狼疮的诊断或预测。

7. 在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在,其中:

(a) 生物学样品已知包括或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异的核酸;

(b) 在至少一个基因座中的变异包括表4所述的SNP或位于与表4所述的SNP对应的核苷酸位置上;和

(c) 在至少一个基因座中变异的存在是对象中狼疮的诊断或预测。

8. 在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表6所述的至少一个SLE风险基因座中变异的存在,其中:

(a) 生物学样品已知包括或怀疑包括了包含表6所述的至少一个SLE风险基因座中的变异的核酸;

(b) 在至少一个基因座中的变异包括表6所述的SNP或位于与表6所述的SNP对应的核苷酸位置上;和

(c) 在至少一个基因座中变异的存在是对象中狼疮的诊断或预测。

9. 治疗对象的狼疮病症的方法,所述对象中已知在与SLE风险基因座中单核苷酸多态性(SNP)对应的核苷酸位置上存在遗传变异,其中SNP是rs922483(SEQ ID NO:13),SLE风险基因座是BLK,且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,方法包括向对象施用治疗病症有效的治疗剂。

10. 治疗对象的狼疮病症的方法,所述对象中已知在表4所述的至少一个SLE风险基因座中的与表4所述的单核苷酸多态性(SNP)对应的核苷酸位置上存在遗传变异,方法包括向对象施用治疗病症有效的治疗剂。

## 用于治疗、诊断和监控狼疮的方法

[0001] 本申请为2010年10月6日提交的、发明名称为“用于治疗、诊断和监控狼疮的方法”的PCT申请PCT/US2010/051589的分案申请,所述PCT申请进入中国国家阶段的日期为2012年6月6日,申请号为201080055282.5。

[0002] 相关申请的交叉引用

[0003] 本申请要求于2009年10月7日提交的美国临时申请号61/278,510的优先权,其通过引用全文整合到本文中。

[0004] 序列表

[0005] 本申请含有通过EFS-Web按ASCII格式提交的序列表,通过引用全部整合到本文中。所述ASCII拷贝创建于2010年9月20日,名为P4325R1W.txt,大小为57,896字节。

[0006] 领域

[0007] 提供了鉴别、诊断、预测和评估出现狼疮的风险的方法,以及治疗狼疮的方法。还提供了用于鉴别有效的狼疮治疗剂和预测对狼疮治疗剂的响应性的方法。

[0008] 背景

[0009] 狼疮是自身免疫疾病,据估计其几乎影响一百万美国人,主要是年龄在20-40岁之间的女性。狼疮涉及攻击结缔组织的抗体。狼疮的主要形式是系统性狼疮(系统性红斑狼疮;SLE)。SLE是具有强遗传及环境成分的慢性自身免疫疾病(见例如Hochberg MC,Dubois' Lupus Erythematosus.第5版.,Wallace DJ,Hahn BH,编辑Baltimore:Williams和Wilkins (1997);Wakeland EK等,Immunity 2001;15(3):397-408;Nath SK等,Curr.Opin.Immunol.2004;16(6):794-800;D' Cruz等,Lancet(2007),369:587-596)。已知多种其他形式的狼疮,包括但不限于皮肤红斑狼疮(CLE)、狼疮性肾炎和新生儿期狼疮。

[0010] 随着它从攻击皮肤和关节进展至攻击内脏,包括肺、心脏和肾(主要关注肾病),未治疗的狼疮可以是致死的,从而使得早期和准确诊断狼疮和/或评估罹患狼疮的风险尤其关键。狼疮主要表现为一系列发作(flare-up),具有少量疾病现象或无疾病现象的间隔期。通过尿中蛋白尿的量测量的肾损伤是与SLE中的致病性相关的最尖锐的损伤区域之一,并解释了该疾病的至少50%的死亡率和发病率。

[0011] 在临床上,SLE是表征为高亲和力自身抗体(autoAb)的异质性障碍(heterogeneous disorder)。自身抗体在SLE的发病机理中发挥重要作用,且该疾病多样的临床现象是由包含抗体的免疫复合物在血管中沉积导致肾、脑和皮肤中的炎症引起的。自身抗体还具有促进溶血性贫血和血小板减少的直接致病作用。SLE与抗核抗体、循环免疫复合物的产生和补体系统的激活相关。SLE在20至60岁之间的妇女中具有约1/700的发病率。SLE可以影响任意器官系统,并可以引起严重的组织损伤。SLE中存在大量具有不同特异性的自身抗体。SLE患者通常产生具有抗DNA、抗Ro和抗血小板特异性,且能够起始该疾病的临床特征,如肾小球肾炎、关节炎、浆膜炎、新生儿完全性心脏传导阻滞和血液学异常的自身抗体。这些自身抗体还可能与中枢神经系统障碍相关。Arbuckle等描述了自身抗体在SLE的临床发作前的出现(Arbuckle等N.Engl.J.Med.349(16):1526-1533(2003))。狼疮(包括SLE)不容易确诊,导致医生采取多因子征兆和基于症状的分类方法(Gill等人,American

Family Physician 68(11):2179-2186(2003))。

[0012] 复杂的自身免疫病(例如狼疮)的临床管理中一个最大的难题是在患者中准确的早期鉴别疾病。近年来,已实施了许多品系和候选基因的研究,鉴别了对SLE易感性有作用的遗传因子。携带II类HLA等位基因DRB1\*0301和DRB1\*1501的单元型明显与疾病和核自身抗原的抗体存在相关。参见例如,Goldberg MA等人,Arthritis Rheum.19(2):129-32(1976);Graham RR等人,Am J Hum Genet.71(3):543-53(2002);和Graham RR等人,Eur J Hum Genet.15(8):823-30(2007)。最近,发现了作为SLE的显著风险因子的干扰素调控因子5(IRF5)和转录的信号转导子和激活子4(STAT)的变体。参见例如,Sigurdsson S等人,Am J Hum Genet.76(3):528-37(2005);Graham RR等人,Nat Genet.38(5):550-55(2006);Graham RR等人,Proc Natl Acad Sci USA 104(16):6758-63(2007);和Remmers EF等人,N Engl J Med.357(10):977-86(2007)。鉴别IRF5和STAT4作为SLE风险基因因为这样的概念提供了支持,即,在某些情况下,I型干扰素(IFN)通路在SLE疾病的致病性中扮演了重要的角色。I型IFN存在于SLE病例的血清中,IFN的生产与包含Ab和核酸的免疫复合物的存在相关(综述在Ronnlblom等人,J Exp Med 194:F59(2001)中;还参见Baechler EC等人,Curr Opin Immunol.16(6):801-07(2004);Banchereau J等人,Immunity 25(3):383-92(2006);Miyagi等人,J Exp Med 204(10):2383-96(2007))。大部分的SLE病例都在血细胞中表现出显著的I型IFN基因表达“签名”(Baechler等人,Proc Natl Acad Sci USA 100:2610(2003);Bennett等人,J Exp Med 197:711(2003)),并具有升高的IFN诱导型细胞因子和趋化因子的血清水平(Bauer等人,PLoS Med 3:e491(2006))。含有天然DNA和RNA的免疫复合物刺激树枝状细胞和B细胞表达的toll-样受体(TLR)7和9,产生I型干扰素,后者进一步刺激免疫复合物形成(综述在Marshak-Rothstein等人,Annu Rev Immunol 25,419(2007)中)。

[0013] 此外,已实施了多项研究,鉴别用于诊断和预兆目的的可靠的生物标记物。然而,尚未鉴别出任何临床确认的诊断标记物(例如,生物标记物),使医生或其他人能够准确的定义SLE的病理生理学方面、临床活性、对治疗的响应或出现疾病的风险,尽管已鉴别了多个被认为对SLE易感性有贡献的候选基因和等位基因(变体)。例如,已报道了至少13个普通的等位基因对欧洲血统个体的SLE风险有作用(Kyogoku等人,Am J Hum Genet 75(3):504-7(2004);Sigurdsson等人,Am J Hum Genet 76(3):528-37(2005);Graham等人,Nat Genet 38(5):550-55(2006);Graham等人,Proc Natl Acad Sci U S A 104(16):6758-63(2007);Remmers等人,N Engl J Med 357(10):977-86(2007);Cunninghame Graham等人,Nat Genet 40(1):83-89(2008);Harley等人,Nat Genet 40(2):204-10(2008);Hom等人,N Engl J Med 358(9):900-9(2008);Kozyrev等人,Nat Genet 40(2):211-6(2008);Nath等人,Nat Genet 40(2):152-4(2008);Sawalha等人,PLoS ONE 3(3):e1727(2008))。推断为因果关系的等位基因已知有HLA-DR3、HLA-DR2、FCGR2A、PTPN22、ITGAM和BANK1(Kyogoku等人,Am J Hum Genet 75(3):504-7(2004);Kozyrev等人,Nat Genet 40(2):211-6(2008);Nath等人,Nat Genet 40(2):152-4(2008)),而IRF5、TNFSF4和BLK的风险单元型可能通过影响mRNA和蛋白质的表达水平而对SLE起作用(Sigurdsson等人,Am J Hum Genet 76(3):528-37(2005);Graham等人,Nat Genet 38(5):550-55(2006);Graham等人,Proc Natl Acad Sci U S A 104(16):6758-63(2007);Cunninghame Graham等人,Nat Genet 40(1):

83-89 (2008) ;Hom等人,N Engl J Med 358 (9) :900-9 (2008) )。STAT4、KIAA1542、IRAK1、PXX和其他基因 (例如BLK) 的因果等位基因尚未确定 (Remmers等人,N Engl J Med 357 (10) :977-86 (2007) ;Harley等人,Nat Genet 40 (2) :204-10 (2008) ;Hom等人,N Engl J Med 358 (9) :900-9 (2008) ;Sawalha等人,PLoS ONE 3 (3) :e1727 (2008) )。这些以及其它与狼疮相关的遗传变异还描述在国际专利申请号PCT/US2008/064430 (国际公开号W0 2008/144761) 中。虽然这些遗传变异对SLE风险和迄今为止描述过的疾病的各个方面有重要贡献,但仍然需要确定关于遗传变异对例如SLE的显著临床异质性的作用的更多信息。

[0014] 因此,拥有其他基于分子的诊断方法是非常有利的,所述诊断方法可用于客观的鉴别患者中疾病的存在和/或分类疾病、定义狼疮的病理生理学方面、临床活性、对治疗的响应、预后,和/或出现狼疮的风险。此外,具有与疾病的各种临床的和/或病理生理学的和/或其他生物学指示物相关的基于分子的诊断标记物是有利的。因而,对鉴别与狼疮以及其他自身免疫病相关的新型风险基因座和多态性存在持续的需求。此类相关性极大的有利于鉴别患者中狼疮的存在或确定出现疾病的易感性。此类相关性还有利于鉴别狼疮的病理生理学方面、临床活性、对治疗的响应,或预后。此外,与此类相关性相关的统计学和生物学显著性和可重复的信息可用作鉴别特定患者亚群的工作中的组成部分,所述患者亚群预期将显著的受益于特定治疗剂的治疗,例如在此类特定的狼疮患者亚群中治疗受益的治疗剂或在临床研究中表现出治疗受益的治疗剂。

[0015] 本文描述的发明满足了上述要求,并提供了其他益处。

[0016] 本文引用的所有参考文献,包括专利申请和出版物,出于任何目的都通过引用全文整合到本文中。

[0017] 概述

[0018] 本文提供的方法至少部分的基于这一点发现,即一组与SLE相关并造成疾病风险的新型基因座 (SLE风险基因座)。此外,提供了与这些SLE风险基因座相关的一组等位基因。还包括了在与增加SLE风险的生物学效应相关的BLK基因座内的因果等位基因。此外,提供了与其他自身免疫病和增加的SLE风险相关的风险基因座。

[0019] 在一个方面,提供了在对象中鉴别狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,其中对象怀疑患有狼疮。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定中的方法。

[0020] 在另一个方面,提供了在对象中鉴别狼疮的方法,方法包括在源自对象的生物学样品中检测表4所述的至少一个SLE风险基因座中变异的存在,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP的位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。在一个实施方案中,至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。在一个实施方案中,至少一个基因座的变异包括表4所述的SNP。在某些实施方案中,检测到在表

4所述的至少一个SLE风险基因座中变异的存在(其中,在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上)与BLK SLE风险基因座中变异的存在(其中BLK基因座中的变异发生在与SNP的位置对应的核苷酸位置上,其中SNP是rs922483(SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)组合。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定中的方法。

[0021] 在仍然另一个实施方案中,提供了在对象中鉴别狼疮的方法,方法包括在源自对象的生物学样品中检测表6所述的至少一个SLE风险基因座中变异的存在,其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP的位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。在一个实施方案中,至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在一个实施方案中,至少一个基因座中的变异包括表6所述的SNP。在某些实施方案中,检测到在表6所述的至少一个SLE风险基因座中变异的存在(其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上)与BLK SLE风险基因座中变异的存在(其中BLK基因座中的变异发生在与SNP的位置对应的核苷酸位置上,其中SNP是rs922483(SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)组合。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定中的方法。

[0022] 在仍然另一个方面,提供了在对象中鉴别狼疮的方法,方法包括在源自对象的生物学样品中检测表4所述的至少一个SLE风险基因座中变异的存在和表6所述的至少一个SLE风险基因座中的变异的存在,其中在每个基因座中的变异分别发生在与表4和表6所述的各基因座的SNP位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。在某些实施方案中,在至少三个基因座、或至少四个基因座、或至少五个基因座、或至少七个基因座或至少10个基因座中检测到变异。在一个实施方案中,表4所述的至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10,而表6所述的至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在一个实施方案中,表4所述的至少一个基因座中的变异和表6所述的至少一个基因座中的变异分别包括表4和表6所述的SNP。在某些实施方案中,检测到表4所述的至少一个SLE风险基因座中变异的存在(其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上),和表6所述的至少一个SLE风险基因座中变异的存在(其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上),与BLK SLE风险基因座中变异的存在组合(其中,BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483(SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定中的方法。

[0023] 在一个方面,提供了预测患狼疮的对象对狼疮治疗剂的响应性的方法,方法包括

确定对象是否在SLE风险基因座中包括变异,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,其中BLK SLE风险基因座中变异存在表示对象对治疗剂的响应性。

[0024] 在另一个方面,提供了预测患狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表4所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上,其中在至少一个基因座中变异存在表示对象对治疗剂的响应性。在某些实施方案中,对象在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中包括变异。在一个实施方案中,至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。在一个实施方案中,至少一个基因座中的变异包括表4所述的SNP。在某些实施方案中,方法包括确定对象是否包括在表4所述的至少一个SLE风险基因座中的变异(其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上),与BLK SLE风险基因座中的变异(其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)组合,其中表4所述的至少一个基因座中变异的存在和BLK基因座中变异的存在表示对象对治疗剂的响应性。

[0025] 在仍然另一个方面,提供了预测患狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表6所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上,其中在至少一个基因座中变异的存在表示对象对治疗剂的响应性。在某些实施方案中,对象在至少两个基因座、或至少三个基因座、或至少四个基因座或五个基因座中包括变异。在一个实施方案中,至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在一个实施方案中,至少一个基因座中的变异包括表6所述的SNP。在某些实施方案中,方法包括确定对象是否包括在表6所述的至少一个SLE风险基因座中的变异(其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上),与BLK SLE风险基因座中的变异(其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)组合,其中表6所述的至少一个基因座中变异的存在和BLK基因座中变异的存在表示对象对治疗剂的响应性。

[0026] 在其他方面,提供了预测患狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表4所述的至少一个SLE风险基因座中包括变异(其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上),以及在表6所述的至少一个SLE风险基因座中包括变异(其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上),其中在表4所述的至少一个基因座中变异的存在和在表6所述的至少一个基因座中变异的存在表示对象对治疗剂的响应性。在某些实施方案中,对象在至少三个基因座、或至少四个基因座、或至少五个基因座、或至少7个基因座、或至少十个基因座中包括变异。在一个实施方案中,表4所述的至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10,表6所述的至少一个基因座选自IFIH1、CFB、

CLEC16A、IL12B和SH2B3。在一个实施方案中,表4所述的至少一个基因座中的变异和表6所述的至少一个基因座中的变异分别包括表4和表6所述的SNP。在某些实施方案中,方法包括确定对象是否包括在表4所述的至少一个SLE风险基因座中的变异(其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上),和在表6所述的至少一个SLE风险基因座中的变异(其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上),以及BLK SLE风险基因座中的变异(其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶)组合,其中表4所述的至少一个基因座中变异的存在和表6所述的至少一个基因座中变异的存在以及BLK基因座中变异的存在表示对象对治疗剂的响应性。

[0027] 在仍然另一个方面,提供了在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0028] 在仍然另一个方面,提供了在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上;并且在至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。在某些实施方案中,方法包括检测在表4所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组合,其中:生物学样品已知或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表4所述的至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上,并且BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表4所述的至少一个基因座中变异的存在和在BLK基因座中变异的存在是对象中狼疮的诊断或预测。

[0029] 在仍然另一个方面,提供了在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表6所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了包含表6所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上;并且在至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,方法包括检测在表6所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组合,其中:生物学样品已知或怀疑包括了包含表6所述的至少一个

SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表6所述的至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上,并且BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表6所述的至少一个基因座中变异的存在和在BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0030] 在仍然另一个方面,提供了在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在和在表6所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了:包含表4所述的至少一个SLE风险基因座中的变异和表6所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异分别包括表4和表6所述SNP或位于与表4和表6所述SNP对应的核苷酸位置上;以及在表4所述的至少一个基因座中变异的存在和在表6所述的至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少三个基因座、或至少四个基因座、或至少五个基因座、或至少七个基因座、或至少十个基因座中检测到变异。在一个实施方案中,表4所述的至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10,表6所述的至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,方法包括检测在表4所述的至少一个SLE风险基因座中变异的存在,和在表6所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组合,其中:生物学样品已知或怀疑包括了:包含表4所述的至少一个SLE风险基因座中的变异和表6所述的至少一个SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表4所述的至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上,并且在表6所述的至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上,并且在BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表4所述的至少一个基因座中变异的存在和在表6所述的至少一个基因座中变异的存在,和在BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0031] 在另一个方面,提供了在对象中帮助诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0032] 在仍然另一个方面,提供了在对象中帮助诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上;并且在至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。在某些实施方案中,方法包括检测在表4所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组

合,其中:生物学样品已知或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表4所述的至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上,并且BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表4所述的至少一个基因座中变异的存在和在BLK基因座中变异的存在是对象中狼疮的诊断或预测。

[0033] 在仍然另一个方面,提供了在对象中帮助诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表6所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了包含表6所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上;并且在至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,方法包括检测在表6所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组合,其中:生物学样品已知或怀疑包括了包含表6所述的至少一个SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表6所述的至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上,并且BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表6所述的至少一个基因座中变异的存在和在BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0034] 在仍然另一个方面,提供了在对象中帮助诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在和在表6所述的至少一个SLE风险基因座中变异的存在,其中:生物学样品已知或怀疑包括了:包含表4所述的至少一个SLE风险基因座中的变异和表6所述的至少一个SLE风险基因座中的变异的核酸;在至少一个基因座中的变异分别包括表4和表6所述SNP或位于与表4和表6所述SNP对应的核苷酸位置上;以及在表4所述的至少一个基因座中变异的存在和在表6所述的至少一个基因座中变异的存在是对象的狼疮的诊断或预测。在某些实施方案中,在至少三个基因座、或至少四个基因座、或至少五个基因座、或至少七个基因座、或至少十个基因座中检测到变异。在一个实施方案中,表4所述的至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10,表6所述的至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,方法包括检测在表4所述的至少一个SLE风险基因座中变异的存在,和在表6所述的至少一个SLE风险基因座中变异的存在,以及与BLK SLE风险基因座中变异的存在组合,其中:生物学样品已知或怀疑包括了:包含表4所述的至少一个SLE风险基因座中的变异和表6所述的至少一个SLE风险基因座中的变异和BLK基因座中的变异的核酸,在表4所述的至少一个基因座中的变异包括表4所述SNP或位于与表4所述SNP对应的核苷酸位置上,并且在表6所述的至少一个基因座中的变异包括表6所述SNP或位于与表6所述SNP对应的核苷酸位置上,并且在BLK基因座中的变异发生在与SNP位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且在表4所述的至少一个基因座中变异的存在和在表6所述的至少一个基因座

中变异的存在,和在BLK基因座中变异的存在是对象的狼疮的诊断或预测。

[0035] 在一个方面,提供了治疗对象的狼疮病症的方法,其中,已知遗传变异存在于与SLE风险基因座中的SNP对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),而SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,方法包括向对象施用有效治疗病症的治疗剂。

[0036] 在另一个方面,提供了治疗对象的狼疮病症的方法,已知所述对象中的遗传变异在表4所述的至少一个SLE风险基因座中的与表4所述SNP对应的核苷酸位置上存在,方法包括向对象施用治疗病症有效的治疗剂。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0037] 在另一个方面,提供了治疗对象的狼疮病症的方法,已知所述对象中的遗传变异在表6所述的至少一个SLE风险基因座中的与表6所述SNP对应的核苷酸位置上存在,方法包括向对象施用治疗病症有效的治疗剂。在一个实施方案中,至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0038] 在另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用在对对象中治疗病症有效的治疗剂,所述对象在与SLE风险基因座中的SNP对应的核苷酸位置上具有遗传变异,其中SNP是rs922483 (SEQ ID NO:13),而SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0039] 在另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用在对对象中治疗病症有效的治疗剂,所述对象在表4所述的至少一个SLE风险基因座中的与表4所述SNP对应的核苷酸位置上具有遗传变异。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0040] 在另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用在对对象中治疗病症有效的治疗剂,所述对象在表6所述的至少一个SLE风险基因座中的与表6所述SNP对应的核苷酸位置上具有遗传变异。在一个实施方案中,至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0041] 在仍然另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用在至少一个临床研究中表现出有效治疗所述病症的治疗剂,其中所述研究中所述治疗剂被施用给至少5个人类对象,所述对象各自在与SLE风险基因座中的SNP对应的核苷酸位置上具有遗传变异,其中SNP是rs922483 (SEQ ID NO:13),而SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0042] 在仍然另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用至少一个临床研究中表现出有效治疗所述病症的治疗剂,其中所述研究中所述治疗剂被施用给至少5个人类对象,所述对象各自在与表4所述的至少一个SLE风险基因座中的表4所述SNP对应的核苷酸位置上具有遗传变异。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0043] 在仍然另一个方面,提供了治疗有狼疮病症的对象的方法,方法包括向对象施用至少一个临床研究中表现出有效治疗所述病症的治疗剂,其中所述研究中所述治疗剂被施用给至少5个人类对象,所述对象各自在与表6所述的至少一个SLE风险基因座中的表6所述SNP对应的核苷酸位置上具有遗传变异。在一个实施方案中,至少一个SLE风险基因座选

自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0044] 在另一个方面,提供了包括制备狼疮治疗剂的方法,包括包装治疗剂与向对象施用治疗剂的说明书,所述对象患有或被认为患有狼疮,且在与SLE风险基因座中的SNP对应的位置具有遗传变异,其中SNP是rs922483 (SEQ ID NO:13),SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0045] 在仍然另一个方面,提供了包括制备狼疮治疗剂的方法,包括包装治疗剂与向对象施用治疗剂的说明书,所述对象患有或被认为患有狼疮,且在与表4所述的至少一个SLE风险基因座中的表4所述SNP对应的位置具有遗传变异。

[0046] 在仍然另一个方面,提供了包括制备狼疮治疗剂的方法,包括包装治疗剂与向对象施用治疗剂的说明书,所述对象患有或被认为患有狼疮,且在与表6所述的至少一个SLE风险基因座中的表6所述SNP对应的位置具有遗传变异。

[0047] 在一个方面,提供了选择用于用狼疮治疗剂治疗的患有狼疮的患者的方法,方法包括检测在与SLE风险基因座中的SNP对应的核苷酸位置遗传变异的存在,其中SNP是rs922483 (SEQ ID NO:13),且SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0048] 在其他方面,提供了选择用于用狼疮治疗剂治疗的患有狼疮的患者的方法,方法包括检测在与表4所述的至少一个SLE风险基因座中的表4所述SNP对应的核苷酸位置处遗传变异的存在。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。在一个实施方案中,在至少一个基因座的变异包括表4所述的SNP。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0049] 在其他方面,提供了选择用于用狼疮治疗剂治疗的患有狼疮的患者的方法,方法包括检测在与表6所述的至少一个SLE风险基因座中的表6所述SNP对应的核苷酸位置处遗传变异的存在。在某些实施方案中,在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。在一个实施方案中,至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。在一个实施方案中,在至少一个基因座的变异包括表6所述的SNP。在一个实施方案中,检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0050] 在另一个方面,提供了评估对象是否处于出现狼疮的风险的方法,方法包括在从对象获得的生物学样品中检测指示了出现狼疮的风险的遗传标签 (genetic signature) 的存在,其中所述遗传标签包括一组至少三个SNP,每个SNP发生于表4和/或表6所述的SLE风险基因座中。在某些实施方案中,遗传标签包括一组至少4个SNP、或至少5个SNP、或至少7个SNP、或至少10个SNP。在一些实施方案中,SLE风险基因座选自TNIP1、PRDM1、JAZF1、

UHRF1BP1、IL10、IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,遗传标签还包括SLE风险基因座中的SNP,其中SNP是rs922483 (SEQ ID NO:13),且SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0051] 在另一个方面,提供了诊断对象中的狼疮的方法,方法包括在从所述对象获得的生物学样品中检测指示狼疮的遗传标签的存在,其中所述遗传标签包括一组至少三个SNP,每个SNP发生于表4和/或表6所述的SLE风险基因座中。在某些实施方案中,遗传标签包括一组至少4个SNP、或至少5个SNP、或至少7个SNP、或至少10个SNP、或至少15个SNP、或至少20个SNP、或至少30个SNP。在一个实施方案中,SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1、IL10、IFIH1、CFB、CLEC16A、IL12B和SH2B3。在某些实施方案中,遗传标签还包括SLE风险基因座中的SNP,其中SNP是rs922483 (SEQ ID NO:13),且SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

## 附图简介

[0052] 图1显示了某些SNP的靶向复制研究的实验设计概况,以鉴别实施例1所述的其他SLE风险基因座。

[0053] 图2显示了SLE中的新的全基因组的显著关联,和鉴别实施例1所述的TNIP1 (A)、PRDM1 (B)、JAZF1 (C)、UHRF1BP1 (D) 和IL10 (E) 内的新风险基因座。(F) 实施例1中所述情况和对照重复样品的独立SNP的P值直方图;虚线表示零分布下的预期结果密度。

[0054] 图3显示了在实施例1所述的原始GWAS中,通过P值分层的meta-分析 (meta-analysis) 中达到候选 ( $P < 1 \times 10^{-5}$ ) 和验证 ( $P < 5 \times 10^{-8}$ ) 状态的变体百分比。

[0055] 图4显示了在实施例2所述的BLK启动子区中的连锁不平衡块 (linkage disequilibrium block; 显示为 $r^2$ )。图4公开了‘C>T-rs922483’为SEQ ID NO:13。

[0056] 图5显示了如实施例2所述,具有各种单元型的BLK启动子区的荧光素酶报告基因表达测定的结果。(A) 在BJAB细胞中,SNP rs922483 C>T (SEQ ID NO:13); (B) 在Daudi细胞中,SNP rs922483 C>T (SEQ ID NO:13); (C) 在BJAB细胞中,SNP rs1382568 A>C/G>C; (D) 在Daudi细胞中,SNP rs1382568 A>C/G>C; (E) 在BJAB细胞中,SNP rs4840568 G>A; (F) 在Daudi细胞中,SNP rs4840568 G>A;所显示的数据表示三次重复测定的平均值 $\pm$ 平均值的标准误差;点条显示了标在图左侧的单元型的结果;阴影线的条:风险单元型22-ACT;空条:无风险单元型22-GAC; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , ns=无显著性 (t检验)。图5A-F公开了‘22 (GT) 重复’为SEQ ID NO:15。图5C-F还公开了‘rs922483 C>T’为SEQ ID NO:13。

[0057] 图6显示了如实施例2所述,具有18 (GT) 重复 (SEQ ID NO:14) 或22 (GT) 重复 (SEQ ID NO:15) 和在Daudi细胞中的SNP rs1382568 A>C/G>C的BLK启动子区的荧光素酶报告基因表达测定的结果。所显示的数据表示两次重复测定的平均值 $\pm$ 平均值的标准误差;ns=无显著性 (t检验)。图6公开了‘18 (GT) 重复’为SEQ ID NO:14、‘22 (GT) 重复’为SEQ ID NO:15,和“rs922483 C>T”为SEQ ID NO:13。

[0058] 图7显示了SNP,rs922483 (SEQ ID NO:13) 的序列,以及如实施例2所述BLK基因座的因果等位基因的SNP中的位置。因果等位基因的位置用粗体括号显示;C/T变异用粗体表示。

**[0059] 发明详述**

**[0060]** 除非另有指明,本发明的实施将利用本领域技术内的常规分子生物学(包括重组技术)、微生物学、细胞生物学、生物化学和免疫学技术。诸如“Molecular Cloning: A Laboratory Manual”,第二版(Sambrook等,1989);“Oligonucleotide Synthesis”(M.J.Grait编辑,1984);“Animal Cell Culture”(R.I.Freshney编辑,1987);“Methods in Enzymology”(Academic Press, Inc.);“Current Protocols in Molecular Biology”(F.M.Ausubel等编辑,1987和定期更新);“PCR: The Polymerase Chain Reaction”,(Mullis等编辑,1994)的文献中充分解释了这类技术。此外,可以用本领域已知的标准技术产生用于本发明的引物、寡核苷酸和多核苷酸。

**[0061]** 除非另作定义,本文所使用的技术和科学术语具有与本发明所属领域的普通技术人员的通常理解相同的意义。例如,Singleton等,Dictionary of Microbiology and Molecular Biology第二版,J.Wiley&Sons (New York, N.Y.1994) 和March, Advanced Organic Chemistry Reactions, Mechanisms and Structure第四版, John Wiley&Sons (New York, N.Y.1992) 为本领域技术人员提供了用于本申请中的许多术语的一般指导。

**[0062] 定义**

**[0063]** 出于解释本说明书的目的,应用下列定义,恰当时,以单数形式使用的术语也包括复数形式,反之亦然。如本说明书和所附权利要求中使用的,除非上下文另外明确指出,单数形式(“a”、“an”和“the”)包括复数形式。因而,例如,“蛋白质”的称谓包括复数蛋白质;“细胞”的称谓包括细胞的混合物,等等。在下文所述定义与通过引用整合到本文中的任何文件中的定义冲突的情况下,下文所述的定义适用。

**[0064]** 本文使用的“狼疮”或“狼疮病症”是一般涉及攻击结缔组织的抗体的自身免疫疾病或障碍。狼疮的主要形式是系统性狼疮,系统性红斑狼疮(SLE),其包括皮肤SLE和亚急性皮肤SLE,以及其他类型的狼疮(包括肾炎、肾外、脑炎、儿科、非肾、盘状和脱毛)。

**[0065]** 本文互换使用的术语“多核苷酸”或“核酸”指任意长度的核苷酸聚合物,且包括DNA和RNA。该核苷酸可以是脱氧核糖核苷酸、核糖核苷酸、修饰核苷酸或碱基和/或它们的类似物,或可以通过DNA或RNA聚合酶掺入聚合物中的任意底物。多核苷酸可以包含修饰核苷酸,如甲基化核苷酸和它们的类似物。若存在,对核苷酸结构的修饰可以在组装聚合物之前或之后进行。核苷酸序列可以由非核苷酸成分中断。可以在聚合作用之后如通过与标记成分缀合来进一步修饰多核苷酸。其他类型的修饰包括,例如“帽”;用类似物取代一个或多个天然存在的核苷酸;核苷酸间修饰,例如具有不带电荷的键(例如甲基磷酸酯、磷酸三酯、磷酸胺、氨基甲酸酯(cabamate)等)和具有带电荷的键(例如硫代磷酸酯、二硫代磷酸酯等)的那些、包含悬垂部分例如蛋白质(例如核酸酶、毒素、抗体、信号肽、聚-L-赖氨酸等)的那些、具有嵌入剂(例如吡啶、补骨脂素等)的那些、含有螯合剂(例如金属、放射性金属、硼、氧化金属等)的那些、含有烷化剂(alkylator)的那些、具有修饰键(例如 $\alpha$ 异头核酸等)的那些、以及未修饰形式的多核苷酸。此外,通常存在于糖类中的任意羟基可以例如通过磷酸酯基、磷酸基来取代,通过标准保护基来保护,或活化以制备与其他核苷酸的其他键,或可以与固相支持体缀合。可以磷酸化或用胺或1至20个碳原子的有机帽化基团部分取代5'和3'端的OH。还可以将其他羟基衍生为标准保护基团。多核苷酸还可以包含一般为本领域已知的类似物形式的核糖或脱氧核糖糖类,其包括例如2'-O-甲基-2'-O-烯丙基,2'-氟-或2'-

叠氨基-核糖,碳环糖类似物, $\alpha$ -异头糖类,差向异构糖类如阿拉伯糖、木糖或来苏糖,吡喃糖类,呋喃糖类,景天庚酮糖,无环类似物和无碱基核苷类似物如甲基核苷。可以通过其他连接基团来取代一个或多个磷酸二酯键。这些其他连接基团包括但不限于其中通过P(O)S(“硫代酯(thioate)”)、P(S)S(“二硫代酯(dithioate)”)、(O)NR<sub>2</sub>(“酰胺化物”)、P(O)P、P(O)OR'、CO或CH<sub>2</sub>(“formacetal”)取代磷酸的实施方案,其中R或R'独立地是H或任选地包含醚(—O—)键的取代或非取代烷基(1-20C)、芳基、链烯基、环烷基、环烯基或araldyl。无需多核苷酸中的所有键都相同。前述适用于本文提到的所有多核苷酸,包括RNA和DNA。

[0066] 本文所用的“寡核苷酸”指长度至少为约7个核苷酸且长度少于约250个核苷酸的短的单链多核苷酸。寡核苷酸可以是合成的。术语“寡核苷酸”和“多核苷酸”并不互相排斥。上文对多核苷酸的描述同等地和完全地适用于寡核苷酸。

[0067] 术语“引物”指单链多核苷酸,其能够与核酸杂交,且一般通过提供自由3'-OH基团来允许互补核酸的聚合作用。

[0068] 术语“遗传变异”或“核苷酸变异”指核苷酸序列中相对于参考序列(例如常见和/或野生型序列和/或主要等位基因的序列)的改变(例如一个或多个核苷酸的插入、缺失、倒位或取代,如单核苷酸多态性(SNP))。除非另有指明,该术语还包括该核苷酸序列的互补序列中对应的改变。在一个实施方案中,遗传变异是体细胞多态性。在一个实施方案中,遗传变异是种系多态性。

[0069] “单核苷酸多态性”或“SNP”指DNA中的单个碱基位置,在该位置上,不同的等位基因或其他核苷酸存在于群体中。SNP位置之前和之后通常具有高度保守的等位基因序列(例如在群体的少于1/100或1/1000的成员中不同的序列)。就每个SNP位置上的等位基因而言,个体可以是纯合的或杂合的。

[0070] 术语“氨基酸变异”指氨基酸序列中相对于参考序列的改变(例如一个或多个氨基酸的插入、取代或缺失,如内部缺失或N端或C端截短)。

[0071] 术语“变异”指核苷酸变异或氨基酸变异。

[0072] 术语“对应于SNP位置的核苷酸位置上的遗传变异”、“对应于SNP位置的核苷酸位置上的核苷酸变异”及其语法变通形式指多核苷酸序列中相对对应的DNA位置上的遗传变异,该位置在基因组中由该SNP占据。除非另作指明,该术语还包括该核苷酸序列的互补序列中对应的变异。

[0073] 术语“阵列”或“微阵列”指可杂交的阵列元件,优选多核苷酸探针(例如寡核苷酸)在底材上的有序排列。该底材可以是固体底材,如载玻片,或半固体底材,如硝酸纤维素膜。

[0074] 术语“扩增”指产生参考核酸序列或其互补序列的一个或多个拷贝的方法。扩增可以是线形扩增或指数式扩增(例如PCR)。“拷贝”并非意味着相对于模板序列的完美的序列互补性或同一性。例如,拷贝可以包含核苷酸类似物,如脱氧肌苷、有意的序列改变(如通过包含可与模板杂交但不完全互补的序列的引物引入的序列改变)和/或扩增过程中出现的序列错误。

[0075] 术语“等位基因特异的寡核苷酸”指与包含核苷酸变异(一般是取代)的靶核酸区域杂交的寡核苷酸。“等位基因特异的杂交”指在使等位基因特异的寡核苷酸与其靶核酸杂交时,该等位基因特异的寡核苷酸中的核苷酸特异性地与该核苷酸变异碱基配对。能够对特定核苷酸变异进行等位基因特异的杂交的等位基因特异的寡核苷酸称为对该变异“特异

的”。

[0076] 术语“等位基因特异的引物”指等位基因特异的寡核苷酸,其是引物。

[0077] 术语“引物延伸测定”指这样的测定,其中将核苷酸加入核酸中,产生直接或间接检测的更长的核酸或“延伸产物”。可以加入核苷酸来延伸该核酸的5'或3'末端。

[0078] 术语“等位基因特异的核苷酸掺入测定”指这样的引物延伸测定,其中使引物(a)在处于核苷酸变异3'或5'的区域与靶核酸杂交,并(b)通过聚合酶延伸,从而将与该核苷酸变异互补的核苷酸掺入延伸产物中。

[0079] 术语“等位基因特异的引物延伸测定”指这样的引物延伸测定,其中使等位基因特异的引物与靶核酸杂交并延伸。

[0080] 术语“等位基因特异的寡核苷酸杂交测定”指这样的测定,其中(a)使等位基因特异的寡核苷酸与靶核酸杂交,并(b)直接或间接检测杂交。

[0081] 术语“5'核酸酶测定”指这样的测定,其中等位基因特异的寡核苷酸与靶核酸的杂交允许杂交探针的核酸降解切割(nucleolytic cleavage),产生可检测的信号。

[0082] 术语“应用分子信标的测定”指这样的测定,其中等位基因特异的寡核苷酸与靶核酸的杂交产生检测信号水平,其高于由自由寡核苷酸发射的检测信号水平。

[0083] 术语“寡核苷酸连接测定”指这样的测定,其中使等位基因特异的寡核苷酸和第二寡核苷酸彼此相邻地在靶核酸上杂交和连接在一起(直接地或通过间插核苷酸间接地),并直接或间接检测该连接产物。

[0084] 术语“靶序列”、“靶核酸”或“靶核酸序列”一般指疑似或已知其中包含核苷酸变异的的多核苷酸序列,包括通过扩增产生的这种靶核酸的拷贝。

[0085] 术语“检测”包括任意检测手段,包括直接和间接检测。

[0086] 术语“SLE风险基因座”和“确认的SLE风险基因座”指任一表4和表6中标出的基因座和BLK基因座。

[0087] 术语“SLE风险等位基因”和“确认的SLE风险等位基因”指存在于SLE风险基因座中的变异。这类变异包括但不限于单核苷酸多态性、插入和缺失。表4和表6中显示了某些示例性SLE风险等位基因。

[0088] 如本文所使用,具有出现狼疮的“风险”的对象可以具有或不具有可检测的疾病或疾病症状,且在本文所述的治疗方法之前可以已显示或尚未显示可检测的疾病或疾病症状。“处于风险”表示对象具有一个或多个风险因子,如本文所述和本领域已知,风险因子是与狼疮的患有相关的可测量的参数。有一个或多个这些风险因子的对象比无一个或多个这些风险因子的对象具有更高的出现狼疮的概率。

[0089] 本文用术语“诊断”来指分子或病理状态、疾病或病症的鉴定和分类。例如,“诊断”可以指狼疮病症的特定类型,例如SLE的鉴定。“诊断”还可以指狼疮的特定亚型的分类,例如通过所涉及的组织/器官(狼疮性肾炎),通过分子特征(例如表征为特定基因或核酸区域中的遗传变异的亚群)。

[0090] 本文用术语“帮助诊断”来指帮助进行关于狼疮的特定类型的症状或病症的存在或性质的临床测定的方法。例如,帮助诊断狼疮的方法可以包括测量来自个体的生物学样品中缺乏一个或多个SLE风险基因座或SLE风险等位基因的存在。

[0091] 本文用术语“预测(prognosis)”来指可归因于自身免疫障碍的疾病症状,包括例

如自身免疫疾病如狼疮的复发、突发 (flaring) 和药物抗性的可能性的预测。本文用术语“预测 (prediction)”来指患者将有利地或不利地对药物或药物组反应的可能性。

[0092] 如本文所使用,“治疗”指尝试改变所治疗的个体或细胞的天然过程的临床干预,并可以在临床病理过程之前或期间进行。希望的治疗效应包括预防疾病或其病症或症状的发生或复发、减轻该疾病的病症或症状、减少该疾病的任何直接或间接的病理结果、降低疾病进行速率、改善或缓和疾病状态和达到缓解或改善的预测。在一些实施方案中,本发明的方法和组合物用于延迟出现疾病或障碍的尝试中。

[0093] “有效量”指对达到希望的治疗或预防结果有效的量(所需的剂量和时间)。治疗剂的“治疗有效量”可以根据诸如个体的疾病状态、年龄、性别和体重,及抗体在个体中引起希望的反应的能力的因素而不同。治疗有效量还是这样的量,其中治疗上的有益作用胜于治疗剂的任何毒性或有害作用。“预防有效量”指对达到希望的预防结果有效的量(所需的剂量和时间)。由于在疾病之前或疾病的早期阶段在对象中使用预防剂量,预防有效量通常但并非必然低于治疗有效量。

[0094] “个体”、“对象”或“患者”是脊椎动物。在某些实施方案中,该脊椎动物是哺乳动物。哺乳动物包括但不限于灵长类(包括人和非人灵长类)和啮齿类动物(例如小鼠和大鼠)。在某些实施方案中,哺乳动物是人。

[0095] 如本文所使用,“患者亚群”及其语法变通形式指表征为具有一种或多种独特的可测量和/或可鉴定的特征的患者亚群,该特征将该患者亚群从其所属的更广的疾病分类中的其他患者区分开来。这类特征包括疾病亚类(例如SLE、狼疮性肾炎)、性别、生活方式、健康史、所涉及的器官/组织、治疗史等。

[0096] “对照对象”指尚未诊断为患有狼疮或狼疮病症,且不患有与狼疮或狼疮病症相关的任何病征或症状的健康对象。

[0097] 本文所用的术语“样品”指获自或源自目的对象的组合物,其包含待例如根据物理、生化、化学和/或生理学特征表征和/或鉴定的细胞和/或其他分子实体。例如,短语“生物学样品”或“疾病样品”及其变通形式指获自预期或已知其包含待表征的细胞和/或分子实体的目的对象的任意样品。

[0098] “组织或细胞样品”意指一批获自对象或患者组织的类似的细胞。该组织或细胞样品的来源可以是实体组织,如来自新鲜的、冷冻的和/或保藏的器官或组织样品或活组织检查或吸出物(aspirate);血液或任意血液组分;体液,如脑脊液、羊膜液、腹膜液或组织液;来自该对象的妊娠或发育中的任意时间的细胞。组织样品还可以是原代或培养的细胞或细胞系。可选地,组织或细胞样品获自疾病组织/器官。组织样品可以包含并非在自然界中天然与组织混合的化合物,如防腐剂、抗凝剂、缓冲剂、固定剂、养分、抗生素等。本文所用的“参考样品”、“参考细胞”、“参考组织”、“对照样品”、“对照细胞”或“对照组织”指获自己知或认为未患有正用本发明的方法或组合物鉴定的疾病或病症的来源的细胞或组织。在一个实施方案中,参考样品、参考细胞、参考组织、对照样品、对照细胞或对照组织获自正用本发明的组合物或方法在其中鉴定疾病或病症的同一对象或患者的身体健康部分。在一个实施方案中,参考样品、参考细胞、参考组织、对照样品、对照细胞或对照组织获自个体的身体健康部分,该个体不是正用本发明的组合物或方法在其中鉴定疾病或病症的对象或患者。

[0099] 为了本文的目的,组织样品的“切片”意指单部分或单块组织样品,例如从组织样

品切下的组织或细胞的薄片。若理解了本发明包括在形态学水平和分子水平二者上分析或就蛋白质和核酸二者分析组织样品的同一切片的方法,则理解了可以取得组织样品的多个切片并按照本发明进行分析。

[0100] “相关”或“相关的”意指以任意方式将第一分析或流程的表现和/或结果与第二分析或流程的表现和/或结果相比较。例如,可以将第一分析或流程的结果用于进行第二流程,和/或可以用第一分析或流程的结果来确定是否应进行第二分析或流程。就基因表达分析或流程的实施方案而言,可以用基因表达分析或流程的结果来确定是否应进行具体的治疗方案。

[0101] 在本文中使用时,词“标记”指直接或间接与诸如核酸探针或抗体的试剂缀合或融合,并便于其与之缀合或融合的试剂的检测的化合物或组合物。标记本身可以是可检测的(例如放射性同位素标记或荧光标记),或者在酶标记的情况下,其可以催化可检测的底物化合物或组合物的化学改变。

[0102] “药剂”是治疗疾病、障碍和/或病症的活性药物。在一个实施方案中,该疾病、障碍和/或病症是狼疮或其症状或副作用。

[0103] 在按照本发明使用时,术语对特定治疗剂或治疗选择的“提高的抗性”意指对标准剂量的药物或对标准治疗流程的降低的反应。

[0104] 在按照本发明使用时,术语对特定治疗剂或治疗选择的“降低的敏感性”意指对标准剂量的治疗剂或对标准治疗流程的降低的反应,其中可以通过提高治疗剂剂量或治疗强度来补偿(至少部分补偿)降低的反应。

[0105] 可以用显示对患者的益处的任意终点来评估对象的“预测反应”及其变通形式,所述终点非限制性地包括(1)疾病进展的某种程度的抑制,包括减慢和完全停止;(2)疾病发作次数和/或症状的减少;(3)病灶大小的减小;(4)疾病细胞浸润入相邻的周围器官和/或组织的抑制(即降低、减慢或完全停止);(5)疾病扩散的抑制(即降低、减慢或完全停止);(6)自身免疫应答的减少,其可以但并非必然导致疾病病灶的消退或脱离;(7)与该障碍相关的一种或多种症状某种程度的减轻;(8)治疗后显示无疾病的长度的增加;和/或(9)治疗后在给定时间点的降低的死亡率。

[0106] 本文所用的“狼疮治疗剂”、“对治疗狼疮有效的治疗剂”及其语法变通形式指在以有效量提供时,已知、临床上显示或临床医生预期在患有狼疮的对象中提供治疗益处的活性剂(agent)。在一个实施方案中,该短语包括作为临床上接受的活性剂由厂家销售,或以其他方式由有执照的临床医生使用的任意活性剂,预期在以有效量提供时,其将在患有狼疮的对象中提供治疗效应。在一个实施方案中,狼疮治疗剂包括非类固醇抗炎药(NSAID),其包括乙酰水杨酸(例如阿司匹林)、布洛芬(Motrin)、萘普生(Naprosyn)、吲哚美辛(Indocin)、萘丁美酮(Relafen)、托美丁(Tolectin),及包括在治疗上等同的活性成分及其制剂的任意其他实施方案。在一个实施方案中,狼疮治疗剂包括对乙酰氨基酚(例如Tylenol)、皮质类固醇或抗疟剂3(例如氯喹、羟基氯喹)。在一个实施方案中,狼疮治疗剂包括免疫调节药物(例如硫唑嘌呤、环磷酰胺、氨甲喋呤、环孢菌素)。在一个实施方案中,狼疮治疗剂是抗B细胞剂(例如抗CD20(例如利妥昔单抗)、抗CD22)、抗细胞因子剂(例如抗肿瘤坏死因子 $\alpha$ 、抗白细胞介素-1-受体(例如阿那白滞素(anakinra))、抗白细胞介素10、抗白细胞介素6受体、抗干扰素 $\alpha$ 、抗B淋巴细胞刺激物)、共刺激抑制剂(例如抗CD154、CTLA4-Ig(例

如abatacept))、B细胞能量调节剂(例如LJP 394(例如阿贝莫司(abetimus)))。在一个实施方案中,狼疮治疗剂包括激素治疗(例如DHEA)和抗激素疗法(例如抗促乳素剂溴麦角环肽)。在一个实施方案中,狼疮治疗剂是提供免疫吸附的活性剂、是抗补体因子(例如抗C5a)、T细胞接种、用T细胞受体 $\zeta$ 链转染细胞或肽疗法(例如靶向抗DNA独特型的edratide)。

[0107] 本文所用的具有“销售批准”或已“批准作为治疗剂”的治疗剂或这些短语的语法变通形式指由相关政府实体(例如联邦、州或地方管理机构、部门、办公室)批准、许可、注册或授权由和/或通过和/或代表商业实体(例如盈利实体)销售用于治疗特定障碍(例如狼疮)或患者亚群(例如患有狼疮性肾炎的患者,特定种族、性别、生活方式、疾病风险谱的患者等)的活性剂(例如药物制剂、药剂的形式)。相关政府实体包括例如美国食品及药品管理局(FDA)、欧洲药品管理局(EMA)及其等同机构。

[0108] “抗体”(Ab)和“免疫球蛋白”(Ig)指具有相似结构特征的糖蛋白。抗体显示对具体抗原的结合特异性,而免疫球蛋白包括抗体和一般缺乏抗原特异性的其他抗体样分子二者。后一类型的多肽例如由淋巴系统按低水平和由骨髓瘤按提高的水平产生。

[0109] 术语“抗体”和“免疫球蛋白”在最广泛的意义中可互换使用,并包含单克隆抗体(例如全长或完整的单克隆抗体)、多克隆抗体、单价抗体、多价抗体、多特异性抗体(例如双特异性抗体,只要它们显示希望的生物活性),且还可以包括某些抗体片段(如本文更详细地描述)。抗体可以是嵌合抗体、人抗体、人源化抗体和/或亲和力成熟抗体。

[0110] 本文将术语“全长抗体”、“完整抗体”和“全抗体”可互换地使用来指处于其基本上完整的形式抗体,而不是下文所定义的抗体片段。该术语尤其指具有包含Fc区的重链的抗体。

[0111] “抗体片段”包含完整抗体的部分,尤其是包含其抗原结合区。抗体片段的实例包括Fab、Fab'、F(ab')<sub>2</sub>和Fv片段;双抗体;线性抗体;单链抗体分子;和从抗体片段形成的多特异性抗体。

[0112] 木瓜蛋白酶消化抗体产生两个相同的抗原结合片段(称为“Fab”片段,每个片段具有单个抗原结合部位)和剩余的“Fc”片段(其名称反映了它易于结晶的能力)。胃蛋白酶处理产生具有两个抗原结合部位且仍然能够交联抗原的F(ab')<sub>2</sub>片段。

[0113] “Fv”是包含完整抗原结合部位的最小抗体片段。在一个实施方案中,双链Fv种类由紧密非共价结合的一个重链可变结构域和一个轻链可变结构域的二聚体组成。Fv的6个CDR共同赋予抗体抗原结合特异性。但是,甚至单个可变结构域(或仅包含3个对抗原特异的CDR的半个Fv)也具有识别和结合抗原的能力,虽然以低于整个结合部位的亲和力。

[0114] Fab片段包含重链可变结构域和轻链可变结构域,且还包含轻链的恒定结构域和重链的第一恒定结构域(CH1)。Fab'片段与Fab片段不同在于在包含一个或多个来自抗体较链区的半胱氨酸的重链CH1结构域的C端加入了几个残基。Fab'-SH是本文对其中恒定结构域的半胱氨酸残基具有自由巯基的Fab'的命名。最初作为其间具有铰合部半胱氨酸的Fab'片段对产生F(ab')<sub>2</sub>。还已知抗体片段的其他化学偶联。

[0115] 本文所用的术语“单克隆抗体”指获自基本上同质的抗体群体的抗体,即除了可以以较少的量存在的可能的突变,例如天然存在的突变外,包含单种抗体的群体是同一的。因此,修饰词“单克隆”表示该抗体不是分离的抗体的混合物的性状。在某些实施方案中,这种单克隆抗体通常包括含有结合靶标的多肽序列的抗体,其中该结合靶标的多肽序列是通过

包括从多个多肽序列选择单个结合靶标的多肽序列的方法获得的。例如,该选择方法可以从多个克隆,如杂交瘤克隆、噬菌体克隆或重组DNA克隆的库选择唯一的克隆。应理解,可以进一步改变所选择的结合靶标的序列,例如以改善对该靶标的亲和力、以人源化该结合靶标的序列、以改善其在细胞培养物中的产生、以降低其在体内的免疫原性、以产生多特异性抗体等,且包含该改变的结合靶标的序列的抗体也是本发明的单克隆抗体。与通常包含针对不同决定簇(表位)的不同抗体的多克隆抗体制剂不同,单克隆抗体制剂的每个单克隆抗体针对抗原上的单个决定簇。除了其特异性外,单克隆抗体制剂的优势在于它们一般未受其他免疫球蛋白污染。

[0116] 修饰词“单克隆”表示该抗体获自基本上同质的抗体群体的性状,而不解释为需要通过任意特定的方法产生该抗体。例如,可以通过多种技术产生待按照本发明使用的单克隆抗体,例如杂交瘤法(例如Kohler等,Nature,256:495(1975);Harlow等,Antibodies:A Laboratory Manual,(Cold Spring Harbor Laboratory Press,第二版1998);Hammerling等,在Monoclonal Antibodies and T-Cell Hybridomas 563-681(Elsevier,N.Y.,1981)中)、重组DNA法(见例如美国专利号4,816,567)、噬菌体展示技术(见例如Clackson等,Nature,352:624-628(1991);Marks等,J.Mol.Biol.222:581-597(1992));Sidhu等,J.Mol.Biol.338(2):299-310(2004);Lee等,J.Mol.Biol.340(5):1073-1093(2004);Fellouse,Proc.Natl.Acad.Sci.USA 101(34):12467-12472(2004);和Lee等,J.Immunol.Methods 284(1-2):119-132(2004))和用于在具有编码人免疫球蛋白序列的部分或全部人免疫球蛋白基因座或基因的动物中产生人或人样抗体的技术(见例如W098/24893;W096/34096;W096/33735;W091/10741;Jakobovits等,Proc.Natl.Acad.Sci.USA 90:2551(1993);Jakobovits等,Nature 362:255-258(1993);Bruggemann等,Year in Immunol.7:33(1993);美国专利号5,545,807、5,545,806、5,569,825、5,625,126、5,633,425、5,661,016;Marks等,Bio.Technology 10:779-783(1992);Lonberg等,Nature 368:856-859(1994);Morrison,Nature 368:812-813(1994);Fishwild等,Nature Biotechnol.14:845-859(1996);Neuberger,Nature Biotechnol.14:826(1996);及Lonberg和Huszar,Intern.Rev.Immunol.13:65-93(1995))。

[0117] 本文的单克隆抗体特异性地包括“嵌合”抗体,其中部分重链和/或轻链与源自特定物种或隶属于特定抗体种类或亚类的抗体中对应的序列相同或同源,而链的其余部分与源自另一物种或隶属于另一抗体种类或亚类的抗体中对应的序列相同或同源,以及这类抗体的片段,只要它们显示希望的生物活性(美国专利号4,816,567;和Morrison等,Proc.Natl.Acad.Sci.USA 81:6855-9855(1984))。

[0118] “人源化”形式的非人(例如鼠)抗体是包含源自非人免疫球蛋白的最小序列的嵌合抗体。在一个实施方案中,人源化抗体是这样的人免疫球蛋白(受体抗体),其中通过来自非人物种(供体抗体)如小鼠、大鼠、兔或非人灵长类的具有希望的特异性、亲和力和/或容量的高变区的残基取代来自该受体高变区的残基。在一些情况下,通过对应的非人残基取代该人免疫球蛋白的构架区(FR)残基。此外,人源化抗体可以包含不见于该受体抗体或供体抗体中的残基。可以进行这些修饰来进一步精制抗体性能。一般而言,人源化抗体将包含至少1个,且通常2个可变结构域的基本上全部,其中全部或基本上全部高变环对应于非人免疫球蛋白的高变环,而全部或基本上全部FR是人免疫球蛋白序列的FR。人源化抗体还将

可选地包含至少部分免疫球蛋白恒定区 (Fc), 通常是人免疫球蛋白恒定区。进一步的详情见 Jones 等, *Nature* 321:522-525 (1986); Riechmann 等, *Nature* 332:323-329 (1998); 和 Presta, *Curr.Op.Struct.Biol.* 2:593-596 (1992)。还见以下综述文章和其中引用的参考文献: Vaswani 和 Hamilton, *Ann.Allergy,Asthma&Immunol.* 1:105-115 (1998); Harris, *Biochem.Soc.Transactions* 23:1035-1038 (1995); Hurle 和 Gross, *Curr.Op.Biotech.* 5:428-433 (1994)。

[0119] “人抗体”是包含对应于由人产生和/或已用本文所公开的用于产生人抗体的技术中的任一种产生的抗体的氨基酸序列的氨基酸序列的抗体。这类技术包括筛选人衍生的组合文库, 如噬菌体展示文库 (见例如 Marks 等, *J.Mol.Biol.*, 222:581-597 (1991) 和 Hoogenboom 等, *Nucl.Acids Res.*, 19:4133-4137 (1991)); 用人骨髓瘤和小鼠-人杂骨髓瘤 (heteromyeloma) 细胞系来产生人单克隆抗体 (见例如 Kozbor *J.Immunol.*, 133:3001 (1984); Brodeur 等, *Monoclonal Antibody Production Techniques and Applications*, 55-93 页 (Marcel Dekker, Inc., New York, 1987); 和 Boerner 等, *J.Immunol.*, 147:86 (1991)); 和在能够在无内源性免疫球蛋白产生的情况下产生人抗体的所有组成成分的转基因动物 (例如小鼠) 中产生单克隆抗体 (见例如 Jakobovits 等, *Proc.Natl.Acad.Sci USA*, 90:2551 (1993); Jakobovits 等, *Nature*, 362:255 (1993); Bruggermann 等, *Year in Immunol.*, 7:33 (1993))。人抗体的此定义特异性排除了包含来自非人动物的抗原结合残基的人源化抗体。

[0120] “亲和力成熟的”抗体是与不具有这些改变的亲本抗体相比, 在其一个或多个 CDR 中具有导致该抗体对抗原的亲和力改善的一个或多个改变的抗体。在一个实施方案中, 亲和力成熟的抗体对靶抗原具有纳摩尔或甚至皮摩尔亲和力。通过本领域已知的方法产生亲和力成熟的抗体。Marks 等, *Bio/Technology* 10:779-783 (1992) 描述了通过 VH 和 VL 结构域改组的亲和力成熟。Barbas 等 *Proc Natl.Acad.Sci.USA* 91:3809-3813 (1994); Schier 等 *Gene* 169:147-155 (1995); Yelton 等 *J.Immunol.* 155:1994-2004 (1995); Jackson 等, *J.Immunol.* 154 (7):3310-9 (1995); 和 Hawkins 等, *J.Mol.Biol.* 226:889-896 (1992) 描述了 HVR 和/或构架残基的随机诱变。

[0121] “封闭性抗体”或“拮抗抗体”是抑制或降低其所结合的抗原生物活性的抗体。某些封闭性抗体或拮抗抗体部分或完全抑制该抗原的生物活性。

[0122] 本文将“小分子”或“小的有机分子”定义为具有低于约 500 道尔顿的分子量的有机分子。

[0123] 在本文中使用时, 词“标记”指可检测的化合物或组合物。标记本身可以是可检测的 (例如放射性同位素标记或荧光标记), 或者在酶标记的情况下, 其可以催化产生可检测产物的底物化合物或组合物的化学改变。可以作为检测标记的放射性核素包括例如 I-131、I-123、I-125、Y-90、Re-188、Re-186、At-211、Cu-67、Bi-212 和 Pd-109。

[0124] “分离的”生物分子, 如核酸、多肽或抗体是已鉴定并从其天然环境的至少一个组分分开和/或回收的生物分子。

[0125] 本文提到“约”某个值或参数时包括 (描述) 了涉及该值或参数本身的实施方案。例如, 提到“约 X”的描述包括了“X”的描述。

[0126] 一般技术

[0127] 本文提供与狼疮相关的核苷酸变异。这些变异提供了用于狼疮和/或倾向于或有助于狼疮的出现、持续和/或进展的生物标记。因此,本文所公开的发明用于多种背景(setting)中,例如用于涉及狼疮诊断和治疗的方法和组合中。

[0128] 在某些实施方案中,方法涉及预测,即可归因于自身免疫障碍的疾病症状,包括例如自身免疫疾病如狼疮的复发、突发(flaring)和药物抗性的可能性的预测。在一个实施方案中,该预测涉及那些反应的程度。在一个实施方案中,该预测涉及患者是否将在治疗,例如用特定治疗剂治疗后存活或改善且在一定时期内无疾病复发和/或其概率。通过选择最适合于任意特定患者的治疗模式,可以在临床上使用本发明的预测方法作出治疗决定。本发明的预测方法是预测患者是否可能有利地对治疗方案(如给定的治疗方案,包括例如给定的治疗剂或组合的施用、手术干预、类固醇治疗等)反应,或该患者是否可能在治疗方案之后长期存活的有价值的工具。SLE的诊断可以根据目前的American College of Rheumatology (ACR) 标准。可以通过一条British Isles Lupus Activity Group's (BILAG) “A” 标准或两条BILAG “B” 标准来定义活性疾病。修改自Tan等“The Revised Criteria for the Classification of SLE” Arth Rheum 25 (1982) 的用于诊断SLE的一些病征、症状或其他指示物可以是面颊疹,如颊疹、盘状疹或红色隆起斑块;光敏感性,如对阳光反应,导致皮疹的发展或增加;口腔溃疡,如鼻或口中的溃疡,通常无痛;关节炎,如涉及两个或多个外周关节的非侵蚀性关节炎(关节周围的骨骼未受破坏的关节炎);浆膜炎、胸膜炎或心包炎;肾功能障碍,如尿中过量的蛋白质(大于0.5gm/天或在测试棒上为3+) 和/或细胞管型(源自尿和/或白细胞和/或肾小管细胞的异常成分);神经学病征、症状或其他指示物,癫痫发作(惊厥),和/或无药物情况下的精神病,或已知引起这些效应的代谢紊乱;和血液学病征、症状或其他指示物,如溶血性贫血或白细胞减少(白细胞计数低于4,000细胞/mm<sup>3</sup>) 或淋巴细胞减少(少于1,500淋巴细胞/mm<sup>3</sup>) 或血小板减少(少于100,000血小板/mm<sup>3</sup>)。白细胞减少和淋巴细胞减少一般必须在两个或多个时刻检测。血小板减少一般必须在缺乏已知诱导血小板减少的药物的情况下检测。本发明不限制于狼疮的这些病征、症状或其他指示物。

[0129] 遗传变异的检测

[0130] 以上方法中任一种的核酸可以是基因组DNA、从基因组DNA转录的RNA或从RNA产生的cDNA。核酸可以源自脊椎动物,例如哺乳动物。如果它是直接从该来源获得的或者如果它是见于该来源中的核酸的拷贝,则将核酸称为“源自”特定来源。

[0131] 核酸包括该核酸的拷贝,例如产生自扩增的拷贝。在某些情况下扩增可以是想要的,例如为了获得希望的量的材料用于检测变异。然后可以对扩增子进行变异检测方法,如下文所述的方法,以测定变异是否存在于该扩增子中。

[0132] 可以通过本领域技术人员已知的某些方法来检测变异。这些方法包括但不限于DNA测序;引物延伸测定,包括等位基因特异的核苷酸掺入测定和等位基因特异的引物延伸测定(例如等位基因特异的PCR、等位基因特异的连接链反应(LCR) 和缺口LCR);等位基因特异的寡核苷酸杂交测定(例如寡核苷酸连接测定);切割保护测定,其中用保护免遭切割物质来检测核酸双链体中的错配碱基;MutS蛋白质结合分析;比较变体和野生型核酸分子迁移率的电泳分析;变性梯度凝胶电泳(DGGE,如例如Myers等(1985) Nature 313:495中);RNase在错配碱基对处切割的分析;异源双链体DNA的化学或酶切割的分析;质谱分析法(例如MALDI-TOF);遗传位元分析法(genetic bit analysis,GBA);5' 核酸酶测定(例如

**TaqMan®**) ; 和利用分子信标的测定。下文进一步详细讨论了这些方法中的某些。

[0133] 可以通过用本领域公知的技术对靶核酸进行分子克隆和测序来实现该靶核酸中变异的检测。备选地, 可以用扩增技术如聚合酶链反应 (PCR) 来直接来自肿瘤组织的基因组DNA制备物扩增靶核酸序列。然后可以测定所扩增的序列的核酸序列, 并从中鉴定变异。扩增技术为本领域公知, 例如聚合酶链反应描述于Saiki等, *Science* 239:487, 1988; 美国专利号4,683,203和4,683,195中。

[0134] 还可以用本领域已知的连接酶链反应来扩增靶核酸序列。见例如Wu等, *Genomics* 4:560-569 (1989)。此外, 还可以用称为等位基因特异PCR的技术来检测变异 (例如取代)。见例如Ruano和Kidd (1989) *Nucleic Acids Research* 17:8392; McClay等 (2002) *Analytical Biochem.* 301:200-206。在此技术的某些实施方案中, 使用等位基因特异的引物, 其中该引物的3' 端核苷酸与该靶核酸中的特定变异互补 (即能够与之特异性碱基配对)。若该特定变异不存在, 则观察不到扩增产物。还可以用Amplification Refractory Mutation System (ARMS) 来检测变异 (例如取代)。ARMS描述于例如欧洲专利申请公开号0332435中和Newton等, *Nucleic Acids Research*, 17:7, 1989中。

[0135] 用于检测变异 (例如取代) 的其他方法包括但不限于 (1) 等位基因特异的核苷酸掺入测定, 如单碱基延伸测定 (见例如Chen等 (2000) *Genome Res.* 10:549-557; Fan等 (2000) *Genome Res.* 10:853-860; Pastinen等 (1997) *Genome Res.* 7:606-614; 和Ye等 (2001) *Hum. Mut.* 17:305-316); (2) 等位基因特异的引物延伸测定 (见例如Ye等 (2001) *Hum. Mut.* 17:305-316; 和Shen等 *Genetic Engineering News*, 卷23, 2003年3月15日), 包括等位基因特异的PCR; (3) 5' 核酸酶测定 (见例如De La Vega等 (2002) *Bio Techniques* 32:S48-S54 (描述 **TaqMan®** 测定); Ranade等 (2001) *Genome Res.* 11:1262-1268; 和Shi (2001) *Clin. Chem.* 47:164-172); (4) 利用分子信标的测定 (见例如Tyagi等 (1998) *Nature Biotech.* 16:49-53; 和Mhlanga等 (2001) *Methods* 25:463-71); 和 (5) 寡核苷酸连接测定 (见例如Grossman等 (1994) *Nuc. Acids Res.* 22:4527-4534; 专利申请公开号US 2003/0119004 A1; PCT国际公开号W0 01/92579 A2; 和美国专利号6,027,889)。

[0136] 还可以通过错配检测法来检测变异。错配是并非100%互补的杂交核酸双链体。完全互补的缺乏可以由缺失、插入、倒位或取代引起。错配检测法的一个实例是描述于例如Faham等, *Proc. Natl Acad. Sci. USA* 102:14717-14722 (2005) 和Faham等, *Hum. Mol. Genet.* 10:1657-1664 (2001) 中的错配修复检测 (MRD) 测定。错配切割技术的另一实例是RNase保护法, 其描述于Winter等, *Proc. Natl. Acad. Sci. USA*, 82:7575, 1985和Myers等, *Science* 230:1242, 1985中。例如, 本发明的方法可以涉及与人野生型靶核酸互补的标记的核糖核酸探针的使用。将核糖核酸探针与源自组织样品的靶核酸复性 (杂交) 在一起, 随后用能够检测双链体RNA结构中的一些错配的酶RNase A消化。若RNase A检测到错配, 则它在该错配位点处切割。因此, 在电泳凝胶基质上分离复性的RNA制备物时, 若RNase A已检测到并切割了错配, 则将看到比核糖核酸探针与mRNA或DNA的全长双链体RNA小的RNA产物。核糖核酸探针无需是靶核酸的全长, 而可以是靶核酸的部分, 只要它包含疑似具有变异的位置。

[0137] 可以以类似的方式用DNA探针来检测错配, 例如通过酶或化学切割。见例如Cotton

等,Proc.Natl.Acad.Sci.USA,85:4397,1988;和Shenk等,Proc.Natl.Acad.Sci.USA,72:989,1975。备选地,可以通过错配双链体相对于配对双链体的电泳迁移率变动来检测错配。见例如Cariello,Human Genetics,42:726,1988。可以在杂交前用核糖核酸探针或DNA探针扩增疑似包含变异的靶核酸。还可以用Southern杂交检测靶核酸中的改变,尤其是如果该改变是总体重排(gross rearrangement),如缺失和插入。

[0138] 可以用用于靶核酸或周围标记基因的限制性片段长度多态性(RFLP)探针来检测变异,例如插入或缺失。还可以通过靶核酸的克隆、测序和扩增来检测插入和缺失。还可以用单链构象多态性(SSCP)分析来检测等位基因的碱基改变变体。见例如Orita等,Proc.Natl.Acad.Sci.USA 86:2766-2770,1989和Genomics,5:874-879,1989。

[0139] 微阵列是多重技术,通常使用排列的一系列数千个核酸探针与例如cDNA或cRNA样品在高严谨条件下杂交。通常通过检测荧光团-、银-或化学发光-标记的靶,来检测和定量探针-靶杂交,确定靶中的核酸序列的相对丰度。在典型的微阵列中,探针通过与化学基质的共价键(通过环氧硅烷、氨基硅烷、赖氨酸、聚丙烯酰胺或其它)附着于固体表面。固体表面是例如玻璃、硅片或显微珠(microscopic bead)。各种微阵列是可商购的,包括由例如Affymetrix,Inc.和Illumina,Inc.生产的那些。

[0140] 可以用本领域技术人员已知的某些方法获得生物学样品。可以从脊椎动物,且尤其是哺乳动物获得生物学样品。通常用组织活检检查获得代表性肿瘤组织块。备选地,可以以已知或认为其包含目的肿瘤细胞的组织或流体形式直接获得肿瘤细胞。例如,可以通过切除、支气管镜检查术、细针抽吸、支气管刷检,或从唾液、胸膜液或血液获得肺癌病灶的样品。可以从肿瘤样品或其他身体样品如尿、唾液或血清(肿瘤细胞从肿瘤脱落并出现在这类身体样品中)检测靶核酸(或所编码的多肽)中的变异。通过筛选这类身体样品,可以达到疾病如癌症的简单早期诊断。此外,通过针对靶核酸(或所编码的多肽)中的变异测试这类身体样品,可以更容易地监测治疗的进展。此外,本领域已知用于针对肿瘤细胞富集组织制备物的方法。例如,可以从石蜡或恒冷箱切片分离组织。还可以通过流式细胞术或激光捕获显微解剖从正常细胞分离癌细胞。

[0141] 确定受试者或组织或细胞样品包含本文公开的遗传变异之后,考虑可以对该受试者施用有效量的适当的狼疮治疗剂,以在该受试者中治疗该狼疮病症。可以由熟练的从业者在哺乳动物中进行本文所述的多种病理病症的诊断。允许例如在哺乳动物中诊断或检测狼疮的诊断技术是本领域可得的。

[0142] 可以按照已知的方法施用狼疮治疗剂,如作为快速浓注或通过在规定时间内连续输注静脉内施用,通过肌内、腹膜内、intracerebrospinal、皮下、关节内、滑膜内、鞘内、口腔、局部或吸入途径。任选地,可以通过使用多种市售器械的微泵输注来进行施用。

[0143] 可以根据经验来测定施用狼疮治疗剂的有效剂量和时间表,且进行这类测定在本领域技术范围内。可以利用单个或多个剂量。例如,单独使用的干扰素抑制剂的有效剂量或量可以在从每天约1mg/kg体重至约100mg/kg体重的范围内。可以以本领域已知的方式,例如按Mordenti等,Pharmaceut.Res.,8:1351(1991)中所公开进行剂量的物种间增减。

[0144] 在利用狼疮治疗剂的体内施用时,取决于施用途径,正常剂量可以从每天约10ng/kg至至多100mg/kg哺乳动物体重或更多变动,优选约1μg/kg/天至10mg/kg/天。文献中提供了关于特定剂量和递送方法的指导;见例如美国专利号4,657,760、5,206,344或5,225,

212.预期不同的制剂将对不同治疗化合物和不同障碍有效,靶向一个器官或组织的施用例如可以使得必需以不同于另一器官或组织的方式递送。

[0145] 考虑还可以在该方法中利用其他疗法。该一种或多种其他疗法可以包括但不限于对所讨论的障碍施用类固醇和护理方案的其他标准。考虑可以将这类其他疗法用作从例如定向狼疮治疗剂分开的活性剂。

#### [0146] 试剂盒

[0147] 对于上文描述或提示过的申请中的应用,还提供了试剂盒或制备品。此类试剂盒可包括载体工具,所述载体工具被划分从而以紧密密封来接受一个或多个容器工具,例如小瓶、管等,每个容器工具包括用于方法中的一种分离的元件。例如,一个容器工具可包括探针,所述探针被或可以被可检测的标记。此类探针可以是特异性针对包含SLE风险基因座的多核苷酸的多核苷酸。其中,试剂盒利用核酸杂交来检测靶核酸,试剂盒还可以具有包含用于扩增靶核酸序列的一个或多个核苷酸的容器,和/或包含报道工具的容器,该报道工具如生物素结合蛋白质,如与诸如酶标记、荧光标记或放射性同位素标记的报道分子结合的抗生物素蛋白或链霉抗生物素蛋白。

[0148] 试剂盒通常包括上述容器和一个或多个其他容器,所述其他容器包括从市场和用户的观点出发需要的材料,包括缓冲液、稀释剂、过滤器、针头、注射器,和指导使用的说明书。标签可以位于容器上,提示组合物是用于特定的疗法或非治疗性应用,还可以为体内或体外用途提示说明,例如上述的那些用途。

[0149] 试剂盒中的其他任选组分包括一种或多种缓冲液(例如,封闭缓冲液、洗涤缓冲液、底物缓冲液等),其他试剂例如被酶标记化学改变的底物(例如,色原)、表位挽回溶液(epitope retrieval solution)、对照样品(阳性和/或阴性对照)、对照玻片等。其他组分是酶,例如包括但不限于核酸酶、连接酶或聚合酶。

#### [0150] 销售方法

[0151] 还提供了用于销售狼疮治疗剂或其可药用组合物的方法,其包括向目标听众宣传、告知和/或说明该治疗剂或其药物组合物在治疗已从其获得样品显示存在本文所公开的遗传变异的狼疮患者或患者亚群中的用途。

[0152] 销售一般是通过非人媒介的付费传播,其中确认广告客户并控制信息。为了本文的目的,销售包括宣传、公关活动、产品放置、赞助、包销和促销。此术语还包括赞助的出现在任意印刷传播媒介中的信息公告,其设计用于呼吁大量读者,以说服、告知、促进、刺激,或以其他方式改变对偏好的购买、支持或批准本发明的模式的行为。

[0153] 可以通过任意手段来实现诊断方法的销售。用于传递这些信息的销售媒介的实例包括电视、广播、电影、杂志、报纸、网络和广告牌,包括广告,其是出现在广播媒介中的信息。

#### [0154] 具体地,本发明涉及以下实施方案:

[0155] 1、在对象中鉴定狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性(SNP)的位置对应的核苷酸位置上,其中SNP是rs922483(SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,并且其中对象怀疑患有狼疮。

[0156] 2、在对象中鉴定狼疮的方法,方法包括在源自对象的生物学样品中检测表4所述

的至少一个SLE风险基因座中变异的存在,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的单核苷酸多态性(SNP)的位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。

[0157] 3、实施方案2的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。

[0158] 4、实施方案2的方法,其中至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0159] 5、实施方案2的方法,其中至少一个基因座中的变异包括表4所述的SNP。

[0160] 6、在对象中鉴定狼疮的方法,方法包括在源自对象的生物学样品中检测表6所述的至少一个SLE风险基因座中变异的存在,其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的单核苷酸多态性(SNP)的位置对应的核苷酸位置上,并且其中对象怀疑患有狼疮。

[0161] 7、实施方案6的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。

[0162] 8、实施方案6的方法,其中至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0163] 9、实施方案6的方法,其中至少一个基因座中的变异包括表6所述的SNP。

[0164] 10、实施方案1的方法,其包括在表4所述的至少一个SLE风险基因座中检测其它变异,其中在至少一个基因座中的其它变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0165] 11、实施方案1的方法,其包括在表6所述的至少一个SLE风险基因座中检测其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0166] 12、实施方案10的方法,其包括在表6所述的至少一个SLE风险基因座中检测其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0167] 13、实施方案2的方法,其包括在表6所述的至少一个SLE风险基因座中检测其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0168] 14、实施方案1-13的任一项的方法,其中检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0169] 15、预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在SLE风险基因座中包括变异,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性(SNP)的位置对应的核苷酸位置上,其中SNP是rs922483(SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,其中BLK基因座中变异的存在表示对象对治疗剂的响应性。

[0170] 16、预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在

表4所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表4所述的至少一个基因座的单核苷酸多态性 (SNP) 位置对应的核苷酸位置上,其中在至少一个基因座中变异的存在表示对象对治疗剂的响应性。

[0171] 17、实施方案16的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。

[0172] 18、实施方案16的方法,其中至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0173] 19、实施方案16的方法,其中至少一个基因座中的变异包括表4所述的SNP。

[0174] 20、预测有狼疮的对象对狼疮治疗剂的响应性的方法,方法包括确定对象是否在表6所述的至少一个SLE风险基因座中包括变异,其中在至少一个基因座中的变异发生在与表6所述的至少一个基因座的单核苷酸多态性 (SNP) 位置对应的核苷酸位置上,其中在至少一个基因座中变异的存在表示对象对治疗剂的响应性。

[0175] 21、实施方案20的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。

[0176] 22、实施方案20的方法,其中至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0177] 23、实施方案20的方法,其中至少一个基因座中的变异包括表6所述的SNP。

[0178] 24、实施方案15的方法,其包括确定对象是否在表4所述的至少一个SLE风险基因座中包括其它变异,其中在至少一个基因座中的其它变异发生在与表4所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0179] 25、实施方案15的方法,其包括确定对象是否在表6所述的至少一个SLE风险基因座中包括其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0180] 26、实施方案24的方法,其包括确定对象是否在表6所述的至少一个SLE风险基因座中包括其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0181] 27、实施方案16的方法,其包括确定对象是否在表6所述的至少一个SLE风险基因座中包括其它变异,其中在至少一个基因座中的其它变异发生在与表6所述的至少一个基因座的SNP位置对应的核苷酸位置上。

[0182] 28、在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测SLE风险基因座中变异的存在,其中SLE风险基因座是BLK,其中:

[0183] (a) 生物学样品已知包括或怀疑包括了包含BLK基因座中的变异的核酸;

[0184] (b) 变异发生在与单核苷酸多态性 (SNP) 的位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶;和

[0185] (c) BLK基因座中变异的存在是对象中狼疮的诊断或预测。

[0186] 29、在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表4所述的至少一个SLE风险基因座中变异的存在,其中:

[0187] (a) 生物学样品已知包括或怀疑包括了包含表4所述的至少一个SLE风险基因座中的变异的核酸;

[0188] (b) 在至少一个基因座中的变异包括表4所述的SNP或位于与表4所述的SNP对应的核苷酸位置上;和

[0189] (c) 在至少一个基因座中变异的存在是对象中狼疮的诊断或预测。

[0190] 30、实施方案29的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或至少五个基因座、或至少十个基因座、或至少13个基因座、或26个基因座中检测到变异。

[0191] 31、实施方案29的方法,其中至少一个基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0192] 32、实施方案29的方法,其中至少一个基因座中的变异包括表4所述的SNP。

[0193] 33、在对象中诊断或预测狼疮的方法,方法包括在源自对象的生物学样品中检测在表6所述的至少一个SLE风险基因座中变异的存在,其中:

[0194] (a) 生物学样品已知包括或怀疑包括了包含表6所述的至少一个SLE风险基因座中的变异的核酸;

[0195] (b) 在至少一个基因座中的变异包括表6所述的SNP或位于与表6所述的SNP对应的核苷酸位置上;和

[0196] (c) 在至少一个基因座中变异的存在是对象中狼疮的诊断或预测。

[0197] 34、实施方案33的方法,其中在至少两个基因座、或至少三个基因座、或至少四个基因座、或五个基因座中检测到变异。

[0198] 35、实施方案33的方法,其中至少一个基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0199] 36、实施方案33的方法,其中至少一个基因座中的变异包括表6所述的SNP。

[0200] 37、实施方案28的方法,其中:

[0201] (a) 生物学样品已知包括或怀疑包括了包含表4所述的至少一个SLE风险基因座中的其它变异的核酸;

[0202] (b) 在至少一个基因座中的其它变异包括表4所述的SNP或位于与表4所述的SNP对应的核苷酸位置上;和

[0203] (c) 在至少一个基因座中其它变异的存在是对象中狼疮的诊断或预测。

[0204] 38、实施方案28的方法,其中:

[0205] (a) 生物学样品已知包括或怀疑包括了包含表6所述的至少一个SLE风险基因座中的其它变异的核酸;

[0206] (b) 在至少一个基因座中的其它变异包括表6所述的SNP或位于与表6所述的SNP对应的核苷酸位置上;和

[0207] (c) 在至少一个基因座中其它变异的存在是对象中狼疮的诊断或预测。

[0208] 39、实施方案37的方法,其中:

[0209] (a) 生物学样品已知包括或怀疑包括了包含表6所述的至少一个SLE风险基因座中的其它变异的核酸;

[0210] (b) 在至少一个基因座中的其它变异包括表6所述的SNP或位于与表6所述的SNP对

应的核苷酸位置上;和

[0211] (c) 在至少一个基因座中其它变异的存在是对象中狼疮的诊断或预测。

[0212] 40、实施方案29的方法,其中:

[0213] (a) 生物学样品已知包括或怀疑包括了包含表6所述的至少一个SLE风险基因座中的其它变异的核酸;

[0214] (b) 在至少一个基因座中的其它变异包括表6所述的SNP或位于与表6所述的SNP对应的核苷酸位置上;和

[0215] (c) 在至少一个基因座中其它变异的存在是对象中狼疮的诊断或预测。

[0216] 41、实施方案28-40的任一项的方法,其中检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0217] 42、治疗对象的狼疮病症的方法,所述对象中已知在与SLE风险基因座中单核苷酸多态性(SNP)对应的核苷酸位置上存在遗传变异,其中SNP是rs922483 (SEQ ID NO:13), SLE风险基因座是BLK,且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶,方法包括向对象施用治疗病症有效的治疗剂。

[0218] 43、治疗对象的狼疮病症的方法,所述对象中已知在表4所述的至少一个SLE风险基因座中的与表4所述的单核苷酸多态性(SNP)对应的核苷酸位置上存在遗传变异,方法包括向对象施用治疗病症有效的治疗剂。

[0219] 44、治疗对象的狼疮病症的方法,所述对象中已知在表6所述的至少一个SLE风险基因座中的与表6所述的单核苷酸多态性(SNP)对应的核苷酸位置上存在遗传变异,方法包括向对象施用治疗病症有效的治疗剂。

[0220] 45、治疗具有狼疮病症的对象的方法,方法包括向对象施用治疗对象中病症有效的治疗剂,所述对象在SLE风险基因座中具有遗传变异,其中SLE风险基因座是BLK,其中BLK基因座中的变异发生在与单核苷酸多态性(SNP)位置对应的核苷酸位置上,其中SNP是rs922483 (SEQ ID NO:13),且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0221] 46、治疗具有狼疮病症的对象的方法,方法包括向对象施用治疗对象中病症有效的治疗剂,所述对象在表4所述至少一个SLE风险基因座中的与表4所述的单核苷酸多态性(SNP)对应的核苷酸位置上具有遗传变异。

[0222] 47、实施方案46的方法,其中至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0223] 48、治疗具有狼疮病症的对象的方法,方法包括向对象施用治疗对象中病症有效的治疗剂,所述对象在表6所述至少一个SLE风险基因座中的与表6所述的单核苷酸多态性(SNP)对应的核苷酸位置上具有遗传变异。

[0224] 49、实施方案48的方法,其中至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0225] 50、方法,其包括制备狼疮治疗剂和包装所述治疗剂与向对象施用所述治疗剂的说明书,所述对象具有或被认为具有狼疮,且在与SLE风险基因座中的单核苷酸多态性(SNP)对应的位置具有遗传变异,其中SNP是rs922483 (SEQ ID NO:13), SLE风险基因座是

BLK,且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0226] 51、方法,其包括制备狼疮治疗剂和包装所述治疗剂与向对象施用所述治疗剂的说明书,所述对象具有或被认为具有狼疮,且在表4所述至少一个SLE风险基因座中的与表4所述的单核苷酸多态性(SNP)对应的位置上具有遗传变异。

[0227] 52、方法,其包括制备狼疮治疗剂和包装所述治疗剂与向对象施用所述治疗剂的说明书,所述对象具有或被认为具有狼疮,且在表6所述至少一个SLE风险基因座中的与表6所述的单核苷酸多态性(SNP)对应的位置上具有遗传变异。

[0228] 53、选择用于用狼疮治疗剂治疗的患狼疮患者的方法,其包括检测在SLE风险基因座中的单核苷酸多态性(SNP)对应的位置上遗传变异的存在,其中SNP是rs922483 (SEQ ID NO:13),SLE风险基因座是BLK,且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0229] 54、选择用于用狼疮治疗剂治疗的患狼疮患者的方法,其包括检测在表4所述的至少一个SLE风险基因座中与表4所述的单核苷酸多态性(SNP)对应的核苷酸位置上遗传变异的存在。

[0230] 55、实施方案54的方法,其中至少一个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。

[0231] 56、选择用于用狼疮治疗剂治疗的患狼疮患者的方法,其包括检测在至少一个表6所述的SLE风险基因座中与表6所述的单核苷酸多态性(SNP)对应的核苷酸位置上遗传变异的存在。

[0232] 57、实施方案56的方法,其中至少一个SLE风险基因座选自IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0233] 58、实施方案53-57的任一项的方法,其中检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0234] 59、评估对象是否处于出现狼疮的风险的方法,方法包括在从对象获得的生物学样品中检测指示出现狼疮的风险的遗传标签的存在,其中所述遗传标签包括一组至少三个单核苷酸多态性(SNP),每个SNP发生在表4和表6所述的SLE风险基因座中。

[0235] 60、实施方案59的方法,其中遗传标签包括一组至少4个SNP、或至少5个SNP、或至少7个SNP、或至少10个SNP、或至少15个SNP、或至少20个SNP、或至少30个SNP。

[0236] 61、实施方案59的方法,其中每个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1、IL10、IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0237] 62、实施方案58的方法,其中遗传标签还包括SLE风险基因座中的SNP,其中SNP是rs922483 (SEQ ID NO:13),且SLE风险基因座是BLK,其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0238] 63、诊断对象的狼疮的方法,方法包括在从对象获得的生物学样品中检测指示狼疮的遗传标签的存在,其中所述遗传标签包括一组至少三个单核苷酸多态性(SNP),每个SNP发生在表4和表6所述的SLE风险基因座中。

[0239] 64、实施方案63的方法,其中遗传标签包括一组至少4个SNP、或至少5个SNP、或至少7个SNP、或至少10个SNP、或至少15个SNP、或至少20个SNP、或至少30个SNP。

[0240] 65、实施方案63的方法,其中每个SLE风险基因座选自TNIP1、PRDM1、JAZF1、UHRF1BP1、IL10、IFIH1、CFB、CLEC16A、IL12B和SH2B3。

[0241] 66、实施方案63的方法,其中遗传标签还包括SLE风险基因座中的SNP,其中SNP是rs922483 (SEQ ID NO:13),且SLE风险基因座是BLK,且其中变异是在人的8号染色体的染色体11389322位的胸腺嘧啶。

[0242] 67、实施方案59-66的任一项的方法,其中检测包括进行选自引物延伸测定;等位基因特异性引物延伸测定;等位基因特异性核苷酸掺入测定;等位基因特异性寡核苷酸杂交测定;5' 核酸酶测定;应用分子信标的测定;和寡核苷酸连接测定的方法。

[0243] 以下是本发明的方法和组合物的实施例。应理解,由于上文提供的一般描述,可以实施多种其他实施方案。

## 实施例

[0244] 在所有实施例中,通过编号标注对某些出版物的引用,所述编号在实施例部分的最后有完整的目录信息。

[0245] 实施例1

[0246] 鉴别SLE的新型风险基因座

[0247] 方法和对象

[0248] 对象

[0249] 之前已描述了SLE病例、用于全基因组关联性扫描 (GWAS) 的样品以及来自New York Health Project (NYHP) 收藏中心 (Mitchell等, J Urban Health 81 (2):301-10 (2004)) 的对照的选择和基因型分型 (Hom等, N Engl J Med 358 (9):900-9 (2008))。如下文详述,该SLE病例由3个病例系列组成:a) 来自NIH/NIAMS资助的贮藏处Autoimmune Biomarkers Collaborative Network (ABCn) (Bauer等, PLoS medicine 3 (12):e491 (2006)) 的338个病例和来自Multiple Autoimmune Disease Genetics Consortium (MADGC) (Criswell等, Am J Hum Genet 76 (4):561-71 (2005)) 的141个病例;b) 来自University of California San Francisco (UCSF) Lupus Genetics Project (Seligman等, Arthritis Rheum 44 (3):618-25 (2001); Remmers等, N Engl J Med 357 (10):977-86 (2007)) 的613个病例;和c) 来自University of Pittsburgh Medical Center (UPMC) (Demirci等, Ann Hum Genet 71 (Pt 3):308-11 (2007)) 的335个病例和来自The Feinstein Institute for Medical Research的8个病例。对照是来自NYHP标本收藏的1861个样本、来自公开可得的iControlDB数据库 (可在Illumina Inc. 获得) 的1722个样本和来自公开可得的National Cancer Institute Cancer Genetic Markers of Susceptibility (CGEMS) 项目 (可在URL:cgems(dot)cancer(dot)gov下获得) 的4564个样本。

[0250] 1310个SLE病例和7859个对照的全基因组数据集

[0251] 我们之前描述了SLE病例样品的选择和基因型分型 (Hom等, N Engl J Med 358 (9):900-9 (2008))。如通过自我报告确定并通过基因型分型确认,所有SLE病例都是欧洲血统的北美人。通过医疗记录审查 (94%) 或通过主治风湿病学家书写的标准文件 (6%) 在所有病例中确认了SLE的诊断 (达到4条或多条American College of Rheumatology [ACR] 定义的标准 [Hochberg等, Arthritis Rheum 40 (9):1725 [1997]])。这些病例系列的临床资料

在其他地方给出 (Seligman等, Arthritis Rheum 44 (3):618-25 (2001); Criswell等, Am J Hum Genet 76 (4):561-71 (2005); Bauer等, PLoS medicine 3 (12):e419 (2006); Demirci等, Ann Hum Genet 71 (Pt 3):308-11 (2007); Remmers等, N Engl J Med 357 (10):977-86 (2007))。之前描述了NYHP样品的基因型分型和选择 (Hom等, N Engl J Med 358 (9):900-9 (2008))。表1描述了按位点组织的有效样品数。

[0252] 用下文所述的软件程序PLINK和EIGENSTRAT内的分析模块进行样品和SNP过滤 (还见Purcell等, Am J Hum Genet 81 (3):559-75 (2007); Price等, Nat Genet 38 (8):904-09 (2006))。将全基因组SNP数据用于此研究中, 以便于病例和对照的紧密匹配, 提供已确认的和疑似的SLE基因座上的基因型。

[0253] 表1. 在全基因组和按位点组织的复制研究中分析的样品数

|        |                     | 病例<br>收集  | N    | 对照<br>收集               | N    |
|--------|---------------------|-----------|------|------------------------|------|
| [0254] | 发现样品                |           |      |                        |      |
|        | GWAS <sup>a</sup>   |           | 1310 |                        | 7859 |
|        | 复制样品                |           |      |                        |      |
|        | US <sup>b</sup>     | PROFILE   | 415  | NYHP                   | 776  |
|        |                     | UMN       | 366  | ALZ                    | 2215 |
|        |                     | UCSF      | 284  |                        |      |
|        |                     | UPMC      | 52   |                        |      |
|        |                     | JHU       | 12   |                        |      |
|        |                     | Total     | 1129 |                        | 2991 |
|        | Sweden <sup>c</sup> | Umeå      | 244  | Umeå                   | -    |
|        |                     | Uppsala   | 145  | Uppsala                | 132  |
|        |                     | Stockholm | 270  | Stockholm <sup>d</sup> | 1112 |
|        |                     | Lund      | 155  | Lund                   | 94   |
|        |                     | 共计        | 834  | 共计                     | 1338 |
| 共计     |                     | 3273      |      | 12188                  |      |

[0255] <sup>a</sup>描述了全基因组关联性扫描的样品 (Hom, G. 等人, N Engl J Med 358:900-9 (2008))。<sup>b</sup>美国同期组群的独立SLE病例来自PROFILE SLE联合会、加州大学三番分校 (UCSF) (Thorburn, C.M. 等人, Genes Immun 8:279-87 (2007))、匹兹堡大学医学中心 (UPMC)、明尼苏达大学 (UMN), 和约翰霍普金斯大学 (JHU)。美国对照来自纽约健康项目 (Gregersen等人) 和阿兹海默病例和对照来自匹茨堡大学和NCRAD。<sup>c</sup> SLE病例和对照来自Stockholm、Karolinska、Solna、Uppsala、Lund和Umeå, Sweden。<sup>d</sup>使用Illumina 317K SNP阵列将来自Stockholm的823例对照基因分型。按“方法”中所述, 对这些样品中的SNP进行归因和分析。

[0256] 定制SNP阵列

[0257] 用通过了下文所述的质量控制测量的10,848个SNP设计定制的阵列。完整的阵列具有12,864个SNP, 但2016个SNP未满足质量控制测量, 剩下10,848个SNP进入分析。定制的阵列由3,188个SNP组成, 所述SNP是基于在SLE全基因组的关联性扫描中的标称 $P < 0.05$ 而选择的, 505个SNP来自25个之前报道过的SLE风险基因座, 42个SNP是在针对确认的风险等位基因进行文献检索后从其他自身免疫病中选择的, 而7113个SNP用于确定和控制群体亚

结构。后组包括已用于定义大陆群体差异的SNP (Kosoy, R. 等人, Hum. Mutat. 30:69-78 (2009)) 和欧洲群体亚结构富集的SNP (Tian, C. 等人, PLoS Genet 4, e4 (2008))。定制的阵列由Illumina, Inc. 使用它的iSelect Custom BeadChip和我们为通过了下文所述的质量控制过滤的SNP提供的rs标识号来生产。

[0258] 质量控制和插补(imputation)

[0259] 关于美国数据,在所述定制的Illumina芯片上将共计1,464个U.S. 病例和3,078个美国对照基因分型,所述芯片在本文中也称为定制的12K芯片。使用严格的质量控制(QC) 标准确保最终的分析中包括高质量的数据。具体而言,a) 排除了116个具有>5% 丢失数据的个体,和b) 基于隐含的亲缘关系(cryptic relatedness) 和基于状态同源 (Identical by State; IBS) 状态 ( $PI_{\text{Hat}} > 0.15$ ) 的重复样品,排除了279个个体。仅包括这样的SNP,所述SNP具有a) <5% 丢失数据,b) 哈迪-温伯格平衡 (HWE)  $p$  值  $> 1 \times 10^{-6}$ , c) 小等位基因频率 (MAF)  $> 0.01\%$  和d) 在关于病例和对照的差异丢失缺失的测试中具有 $p$  值  $> 1 \times 10^{-5}$  的SNP。还检验了SNP的批次效果。在应用上述过滤后,1144个病例和3003个对照和11024个SNP的最终组可用于分析。使用PLINK实施所有QC测试 (Purcell 等人, Am J Hum Genet 81 (3):559-75 (2007))。

[0260] 关于瑞典人的数据,在定制的12K芯片上基因分型过的888个病例和527个对照的组可用于分析。分析中还掺入了在Illumina, Inc. 317K Human HapMap SNP珠子阵列 (在本文中也称为317K阵列) 上基因分型的分开的1115瑞典人对照的组合。按下列步骤组合两套数据集。首先,生成12K和317K数据之间的6,789SNP的重叠数据集。使用该数据集检验瑞典人复制组的隐含的亲缘关系和重复样品。结果排除了313个样品 ( $PI_{\text{Hat}} > 0.15$ )。在质量控制检查后,就分析运送在定制的12K芯片上基因分型的863个病例和523个对照和在317K Illumina芯片上基因分型的831个对照。第二,用317K阵列插补 (见下文) 基因分型过的831个瑞典人对照,产生更大的重叠SNP集。在剩余的SNP中,通过插补捕获了4605个SNP。最终11394个重叠SNP的最终集进行分析。使用上述相同的阈值过滤该数据集中的SNP。剩下的未被插补(imputation) 捕获的1250个SNP仅在12K芯片上基因分型的瑞典人样品初始集合中进行分析。

[0261] 使用II期HapMap CEU样品作为参照,使用MACH (基于Markov Chain的单元型分型软件程序,可获得自URL [sph.umich.edu/csg/abecasis/MACH](http://sph.umich.edu/csg/abecasis/MACH)) 插补831个用317K阵列基因分型的瑞典人对照。II期HapMap CEU指来自人类单元型项目的样品,已知是从“II期”数据释放的具有北欧和西欧祖先 (CEU) 的犹他州居民的人类单元型项目 (还参见Li 等人, Am J Hum Genet 579 at 2290 (2006))。在插补前,对317K SNP应用严格的质量控制检查。插补中包括了293,242个通过了下列标准 (1)  $MAF > 1\%$ , (2) 缺失率  $< 5\%$  和 (3)  $HWE \ p \text{ 值} > 1 \times 10^{-6}$  的标记物子集。在插补后,抛弃了具有低插补质量的SNP,即,MACH报道的 $R^2_{\text{Hat}}$  ( $RSQR_{\text{HAT}} < 0.40$ )。重叠的11394个标记物集可用于分析。考虑到插补中的不确定性,分析中使用概率分值而非读取基因型 (genotype call)。

[0262] 为了插补全基因组关联性研究样品,用于meta-分析 (meta-analysis) 的基因型数据来自1310个用Illumina 550K全基因组SNP平台基因分型的SLE病例 (参见Hom, G. 等人, N Engl J Med 358:900-9 (2008))。之前描述了SLE病例样品的选择和基因分型 (Hom, G. 等人, N Engl J Med 358:900-9 (2008))。除上述3583个对照外 (Hom, G. 等人, N Engl J Med 358:

900-9 (2008)), 获得批准后还包括4564个来自公众可获得的癌症遗传学易感性标记物 (Cancer Genetics Markers of Susceptibility (CGEMS)) 项目的对照样品 (获得自URL: [cgems.cancer.gov](http://cgems.cancer.gov))。使用上述数据质量控制过滤 (Hom, G. 等人, N Engl J Med 358:900-9 (2008)) 检验整个7859个对照的样品。之后使用IMPUTE版本1 (获得自URL [www.stats.ox.ac.uk/~marchini/software/gwas/impute.html](http://www.stats.ox.ac.uk/~marchini/software/gwas/impute.html)) 推断使用HapMap Phase II CEU样品作为参照的基因分型 (Marchini, J. 等人, Nat. Genet. 39:906-913 (2007))。使用SNPTEST (获得自URL [www.stats.ox.ac.uk/~marchini/software/gwas/snpTEST\\_v1.1.4.html](http://www.stats.ox.ac.uk/~marchini/software/gwas/snpTEST_v1.1.4.html)) 生成关联性统计 (Marchini, J. 等人, Nat. Genet. 39:906-913 (2007))。具体而言, 使用相加模型 (SNPTEST中的Frequentist 1选项) 生成关联性统计, 对插补基因型的不确定性进行调整 (SNPTEST中的-proper选项)。关联性统计的排序列表用于选择所述的复制区域。

[0263] 复制样品的群体分层 (Population Stratification)

[0264] 对于每个复制组 (replication cohort), 使用祖先信息标记物校正可能的群体分层。使用通过了严格质量控制标准的5486个不相关的祖先信息标记物的子集, 用EIGENSTRAT软件来推断前十位的主要遗传变异组分 (Price等人, Nat Genet 38 (8):904-09 (2006))。从每个样品集去除离群值 (定义为 $\sigma > 6$ )。具体而言, 分别从美国组去除了27个遗传离群值, 而从瑞典组去除了45个离群值。在美国和瑞典复制集合中都观察到了沿着前两个特征向量的一定程度的群体分层。为了校正病例-对照分层, 使用以下策略之一: (1) 对可获得基因型数据的美国复制数据集和瑞典复制数据集应用掺入在EIGENSTRAT中的Cochran-Armitage检验统计量的校正; (2) 在分析插补的瑞典人数据中, 使用主要组分作为逻辑回归模型中的协变。

[0265] 关联性分析

[0266] 对于美国人数据, 在实施未校正的自由度1的等位基因关联性测试后, 观察到测试统计值的一些膨胀 (inflation) (PLINK (Purcell, S. 等人, Am J Hum Genet 81:559-75 (2007)))。为了校正美国人样品中的群体分层, 使用5486个不相关的祖先信息标记物进行主要组分分析 (EIGENSTRAT)。首先, 去除遗传离群值 (定义为 $\sigma > 6$ )。其次, 计算1129个病例和2991个对照中的每个基因分型的SNP的Cochran-Armitage趋势卡方测试统计值, 再使用前四个特征向量调节EIGENSTRAT中的每个SNP的测试统计值。计算基于每个SNP的测试统计值的双尾p值。在校正了群体分层后, 美国人样品的 $\lambda_{gc}$ 是1.05。

[0267] 对于瑞典人数据, 使用5486个在12K样品以及额外的Illumina 317K对照中基因分型过的祖先信息标记物, 检查瑞典组的隐藏的群体分层。在去除遗传离群值后, 有834个病例和515个对照在定制的12K芯片上基因分型, 和823个对照在Illumina 317K芯片上基因分型。将EIGENSTRAT中执行的测试统计值校正用于两个Illumina阵列之间的6789个SNP的重叠组。为了校正在12K样品中基因分型并在Illumina 317K样品中插补的4605个SNP组的分层, 使用了上述确定的前四个特征向量作为SNPTEST中执行的逻辑回归模型的协变异, 因为EIGENSTRAT并非意图用于插补的基因型数据。一小组未被Illumina 317K SNP插补所捕获的1250个标记物仅在定制的12K芯片上基因分型的834个病例和515个对照中进行分析。在校正了群体分层后, 瑞典人样品的 $\lambda_{gc}$ 是1.10。

[0268] Meta-分析

[0269] 使用加权的z分数方法进行meta-分析。为了组合不同组的结果,将等位基因定位在美国国立生物技术信息中心 (NCBI) 的人基因组36号参照序列的正向链上,以避免与C/G和A/T SNP相关的歧义。NCBI的人基因组参照序列可获得自URL:[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)。还参见Pruitt等人,Nucl.Acids Res.35 (database issue):D61-D65 (2007)。考虑相对于任意的参照等位基因的效应方向,将每个组的P值转化为z分数。通过取每个组的样本大小的平方根,然后将其总和除以总样本大小的平方根,加权每个z分数,计算z分数的加权总和。将瑞典人和美国人复制组的组合z分数转化为单尾p值。将meta-分析的z分数转化为两尾p值,并计算关联性的证据。认为越过阈值 $5 \times 10^{-8}$ 的SNP与SLE压倒性的相关。没有通过全基因组显著性的、具有小于 $1 \times 10^{-5}$ 的组合p值的基因座被认为是强候选物。使用可免费获得的METAL软件包(可获得自URL:[www.sph.umich.edu/csg/abecasis/Metal](http://www.sph.umich.edu/csg/abecasis/Metal))进行meta-分析方法。为了计算汇总的比值比,使用METAL软件执行的Cochran-Mantel-Haenszel (CMH) 方法。比值比的计算是相对于每个SNP的风险等位基因的。此外,相对于每个SNP的风险等位基因计算对照中的加权平均等位基因频率。

[0270] 百分比方差解释 (Percent Variance Explained)

[0271] 对于之前与SLE相关的SNP和在我们的复制研究中具有小于 $1 \times 10^{-5}$ 的meta p值的SNP,计算方差解释的百分比。使用易患性阈值模型 (liability threshold model),其中假设SLE具有通常以均值0和方差1分布的潜在易患性 (liability) 分数。假设SLE在普通群体中的流行率为0.1%。为了计算每个基因型的阈值,使用对照中的等位基因频率和与我们分析中的比值比 (OR) 对应的效应大小。

[0272] 相互作用分析

[0273] 为了寻找顶部信号之间的上位效应,在表2、4和6中汇编了所有SNP的列表,并使用PLINK中执行的上位性选项,对每个复制组进行相互作用分析。为了获得更大的统计功效,实施纯病例分析 (case-only analysis)。在校正了测试数后,在 $p < 0.05$ 的显著性水平未发现任何SNP-SNP相互作用。

[0274] 条件分析

[0275] 在每个表现出与SLE强关联性的基因组区域,选择表现出最强信号的SNP。使用PLINK作为该SNP的条件,并寻找其他表现出与SLE强关联性的SNP。

[0276] 大规模的复制研究鉴别TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10为系统性红斑狼疮的新型风险基因座

[0277] 现有的全基因组关联性 (GWA) 和候选基因研究已经鉴别了至少15个达到全基因组显著性 ( $P < 5 \times 10^{-8}$ ) 的共同风险等位基因。这些包括对获得性免疫和自身抗体生产重要的基因 (HLA II类等位基因、BLK、PTPN22和BANK1),和在先天性免疫和干扰素信号传导中发挥作用的基因 (ITGAM、TNFAIP3、STAT4和IRF5) (Cunningham Graham, D.S. 等人, Nat. Genet. 40:83-89 (2008); Graham, R.R. 等人, Nat. Genet. 40 (9):1059-61 (2008); Graham, R.R. 等人, J Intern Med 265:680-88 (2009); Harley, J.B. 等人, Nat. Genet. 40:204-10 (2008); Hom, G. 等人, N Engl J Med 358:900-9 (2008); Kozyrev, S.V. 等人, Nat. Genet. 40:211-6 (2008); Sawalha, A.H. 等人, PLoS ONE 3:e1727 (2008); Sigurdsson, S. 等人, Am J Hum Genet 76:528-37 (2005))。为了鉴别其他的风险基因座,对2446个基因座的SNP实施靶向复制研究,所述基因座在现有的对1310个病例和7859个对照的GWAS7扫描

中表现出标称 $p$ 值 $<0.05$ 。还将来自25个之前报道过的SLE风险基因座中的SNP、来自其他自身免疫病提示的35个基因座中的42个SNP和超过7000个祖先信息标记物基因分型。实验设计的综述显示在图1中。将上述SNP掺入到Illumina定制的SNP阵列。阵列对美国 and 瑞典的独立病例和对照基因分型。使用Illumina 310K SNP阵列将823个瑞典人对照基因分型,如上文“方法”所述,分析变体。

[0278] 具体而言,如上所述,设计了由 $>12,000$ 个变体组成的定制SNP阵列,基因分型为2类独立的SLE病例和对照群体,分别来自美国(1129个SLE病例和2991个对照)和瑞典(834个SLE病例和1338个对照)。美国人对照中包括了2215个阿兹海默病的病例/对照样品,这被认为可接受作为对照,因为SLE和阿兹海默病的遗传基础预计是相互独立的。之后,应用数据质量过滤去除性能低下的样品和SNP、群体离群值和重复的/相关的个体(参见上文“方法”)。在这些质量控制测量后,检验了10848个SNP的最终集合,如图1所示。计算3735个变体的关联性统计值,并使用7113个祖先信息标记物校正群体分层(参见上文“方法”)。

[0279] 首先检验了之前报道与SLE相关的25个变体(来自23个基因座)(参见表2)。我们还发现关于21个变体( $P<0.05$ )的关联性的其他证据,包括在现有的组合数据集中达到全基因组显著性( $P<5\times 10^{-8}$ )的9个基因座。在全基因组显著的结果中的是HLA II类DR3(DRB1 \* 0301)、IRF5、TNFAIP3、BLK、STAT4、ITGAM、PTPN22、PHRF1(KIAA1542)和TNFSF4(OX40L)。该分析为9个基因座的变体提供了额外的证据,在一项之前的研究中报道了显著性的全基因组水平:HLA \* DR2、TNFAIP3(rs6920220)、BANK1、ATG5、PTTG1、PXX、FCGR2A、UBE2L3和IRAK1/MECP2。

[0280] 更早的候选基因研究鉴别了MECP2作为SLE的潜在的风险等位基因(Sawalha, A.H. 等人, PLoS ONE 3:e1727 (2008))。然而,在现有的数据集中,靠近IRAK1的SNP表现出最强的关联性证据,其中IRAK1是Toll样受体7和9信号传导的关键性基因,位于MECP2周围已鉴别的连锁不平衡区域内。最近也报道了相似的发现(Jacob, C.O. 等人, Proc.Natl.Acad.Sci.USA (2009)),还需要其他工作来确定IRAK1/MECP2基因座中的因果等位基因。我们发现了关于3个基因座——TYK2、ICAI和NMNAT2——的关联性的其他证据,这3个基因座之前表现出显著性但没有全基因组水平的关联性证据(Harley, J.B. 等人, Nat.Genet. 40:204-10 (2008); Sigurdsson, S. 等人, Am J Hum Genet 76:528-37 (2005))。对于4个之前提示的变体——LYN、SCUBE1、TLR5和LY9——在组合的数据集值没有观察到任何关联性的证据。

[0281] 为了鉴别新的SLE风险基因座,我们检验了共计3188个SNP,这些SNP来自在我们的全基因组数据集中表现出与SLE的关联性证据的2446个不同的基因座(Hom, G. 等人, N Engl J Med 358,900-9 (2008)),所述数据集包括了在1310个SLE病例和扩展的7859个对照集合中基因分型过的502033个SNP。使用该数据集,用Phase II HapMap CEU样本作为参照,插补了 $>2.1M$ 个变体(参见上文“方法”),并生成了有序排列的关联性统计值列表。针对可能包含在定制的复制阵列上,来选择 $P<0.05$ 的变体。为了有效地基因分型,鉴别了相关变体( $r^2>0.2$ )的组,再从每个最低的 $p$ 值 $<0.001$ 的组中选择至少2个SNP。对于其余的组,将组中具有最低 $p$ 值的SNP包括在内。在复制样品中,计算关联性统计值(参见“方法”),并观察到相对于预期的零分布的复制结果的显著富集。排除之前报道过的SLE风险等位基因,有134个基因座具有 $P<0.05$ (预期64,  $P=2\times 10^{-15}$ )和12个基因座具有 $P<0.001$ (预期1,  $P=1\times 10^{-9}$ ),

提示存在真阳性。

[0282] 图2A-2各自显示了在由TNIP1(图2A)、PRDM1(图2B)、JAZF1(图2C)、UHRF1BP1(图2D)和IL-10(图2E)定义的基因座周围500kb区域内,相对于x轴上的基因组位置,y轴上标绘的全基因组关联性扫描的关联性结果。在各个图2A-2E中,实心方块表示最相关的标记物的meta-分析P值。对于各个图2A-2E,来自基因组扫描的P值被标记以表示与全基因组相关的变体的LD:有点圆圈表示 $r^2 > 0.8$ ,虚线圆圈, $r^2 > 0.5$ ;有斑纹圆圈, $r^2 > 0.2$ ;空心圆圈 $r^2 < 0.2$ 。沿着图2A-2E各图的底部,在每个图下面显示了CEU HapMap中(实黑线)和已知的人类基因的重组率。在图2B(PRDM1)中,通过实心黑圆圈标出了在ATG5基因附近的之前报道过的和独立的SLE风险基因座(rs2245214)。图2F显示了在1963个病例和4329个对照复制样品中,1256个独立SNP(与阵列中的任何其他SNP的 $r^2 < 0.1$ )的P值直方图。在零分布下,图2F中用虚线标出了预期的结果密度。如图2F所示,观察到了小于 $P < 0.05$ 的结果的显著富集。

[0283] 因此,复制研究鉴别出5个新的SLE风险基因座,这5个基因座具有超过全基因组显著性阈值的组合P值( $P < 5 \times 10^{-8}$ ):TNIP1、PRDM1、JAZF1、UHRF1BP1和IL10。关于这些基因座和其他基因座的详细统计学关联性显示在下表4中。

[0284] 5q33.1上的变体rs7708392在所有3个组中与SLE显著相关,具有组合 $P = 3.8 \times 10^{-13}$ (图2A),该变体位于TNF- $\alpha$ 诱导蛋白3(TNFAIP3)的相互作用蛋白1(TNIP1)的内含子中。最近发现靠近TNIP1的变体对牛皮癣的风险有作用(Nair,R.P.等人,Nat Genet 41:199-204(2009)),然而SLE和牛皮癣变体被21kb分隔,并似乎是不同的遗传信号( $r^2 = 0.001$ )。TNIP1和TNFAIP3是相互作用的蛋白质(Heyninck,K.等人,FEBS Lett 536:135-40(2003)),然而,TNIP1在调控TNFAIP3中的精确作用尚属未知。靠近TNFAIP3的多个不同变体与SLE(Graham,R.R.等人,Nat.Genet.40(9):1059-61(2008),Musone,S.L.等人,Nat.Genet.40(9):1062-64(2008))、类风湿性关节炎(Plenge,R.M.等人,Nat Genet 39:1477-82(2007))、牛皮癣(Nair,R.P.等人,Nat Genet 41:199-204(2009))和I型糖尿病(Fung,E.Y.等人,Genes Immun 10:188-91(2009))的关联性提示,该通路在自身免疫病的调控中具有重要的作用。

[0285] 第二个确认的风险变体(rs6568431, $P = 7.12 \times 10^{-10}$ )是在具有ZNF结构域的含PR结构域的1(PRDM1,也称为BLIMP1)和APG5自体吞噬5-样(ATG5)之间的基因间区域内被鉴定的。rs6568431的信号似乎不同于之前报道过的ATG5中的SLE风险等位基因rs2245214(Harley,J.B.等人,Nat Genet 40:204-10(2008))(参见表4),因为rs6568431与rs2245214具有 $r^2 < 0.1$ ,而rs2245214在掺入了rs6568431的条件性逻辑回归后仍然与SLE显著相关( $P < 1 \times 10^{-5}$ )(图2B)。

[0286] 与另一个锌指基因并列的1(JAZF1)的启动子区是第3个新确认的SLE基因座(rs849142, $P = 1.54 \times 10^{-9}$ )(图2C)。令人感兴趣的是,同一变体之前与2型糖尿病的风险(Zeggini,E.等人,Nat Genet 40:638-45(2008))和身高差异(Johansson,A.等人,Hum Mol Genet 18:373-80(2009))相关联。靠近JAZF1的分离的前列腺癌等位基因rs10486567(Thomas,G.等人,Nat Genet 40:310-5(2008))在本研究中没有表现出任何关联性的证据。

[0287] ICBP90结合蛋白1(UHRFBP1,rs11755393, $P = 2.22 \times 10^{-8}$ )的非同义等位基因(R454Q)定义了SLE中的第4个新的风险基因座。该等位基因是在UHRF1的假定的结合伙伴中的非保守性氨基酸改变,UHRF1是与多个通路相关的转录和甲基化因子(Arita,K.等人,Nature 455:818-21(2008))。UHRFBP1风险等位基因位于延伸的连锁不平衡区域内,所述区

域涵盖了多个基因,包括小核核糖核蛋白多肽C (SNPRC),SNPRC是SLE自身抗体经常针对的RNA加工复合物的一部分。

[0288] 所鉴别的第5个新型SLE基因座是白介素-10 (IL10;rs3024505, $P=3.95 \times 10^{-8}$ ) (图2E)。IL-10是通过下调免疫应答发挥功能的重要的免疫调控细胞因子(Diveu,C.等人,Curr Opin Immunol 20:663-8 (2008)),IL-10中的变异与SLE的关联被不一致的报道过(Nath,S.K.等人,Hum Genet 118:225-34 (2005))。与SLE相关的变体与目前鉴别为对溃疡性结肠炎(Franke,A.等人,Nat Genet 40:1319-23 (2008))和I型糖尿病(Barrett,J.C.等人,Nature Genetics 41:703-707 (2009))贡献风险的SNP相同,提示在这些障碍之间可能有共享的IL10通路的病理生理学。

[0289] 在组合的复制样品中使用 $P < 1 \times 10^{-5}$ 的显著性阈值,鉴别了21个额外的SLE候选风险基因座(表4)。在meta-分析的零分布条件下,预期少于1个(0.01)基因座具有 $P < 1 \times 10^{-5}$  ( $P=8 \times 10^{-77}$ ),提示这些基因座中的若干个可能是真阳性的基因座。该列表中令人感兴趣的候选基因包括:a)干扰素调控因子8 (IRF8),其在之前的GWAS中被提示过(Graham,R.R.等人,Nat.Genet.40 (9):1059-61 (2008)),并且它的家族成员IRF5和IRF7都落入确认的SLE风险基因座中;b)TAO激酶3 (TAOK3),它是淋巴细胞中表达的激酶的错义等位基因(rs428073,N47S);c)溶酶体运输调节因子 (LYST),它的突变导致人类的Chediak-Higashi综合征,这是特征为淋巴增生性障碍的复杂疾病;和d)白介素12受体 $\beta 2$  (IL12RB2),该基因座包括了IL23R和SERPBP1,但似乎不同于自身免疫病炎症性肠病、牛皮癣和强直性脊柱炎中报道过的IL23R变体(Duerr,R.H.等人,Science 314:1461-3 (2006))。

[0290] 目前GWA研究的显著特征是在不同的复杂疾病之间共享的大量的重叠基因座(Zhernakova,A.等人,Nat Rev Genet 10:43-55 (2009))。我们测试了来自35个基因座的42个变体(表6和7),所述基因座之前报道为与SLE相关的自身免疫病风险等位基因。然而,没有单一基因座具有未调节的 $P$ 值 $< 5 \times 10^{-8}$ ,但是发现了相关等位基因的富集。从35个测试的基因座中(共计42个变体),有5个等位基因具有未调节的 $P < 0.0004$  (少于通过机率预期的一个结果, $P=4.4 \times 10^{-12}$ ),且在针对35个之前说明的基因座进行Bonferroni校正后,具有 $P < 0.05$ 。对于这5个变体中的每一个,SLE相关的等位基因匹配之前报道过的等位基因并具有相同的效应方向(表6)。观察到IFIH1的错义等位基因非常显著的关联性(rs1990760, $P=3.3 \times 10^{-7}$ ),该等位基因之前与I型糖尿病和Grave's病相关(Smyth,D.J.等人,Nat Genet 38:617-9 (2006);Sutherland,A.等人,J Clin Endocrinol Metab 92:3338-41 (2007))。还观察到补体因子B (CFB,rs641153)的错义等位基因(R32Q)的关联性,该基因位于III类HLA的区域中,是经过验证的年龄相关性黄斑变性的风险等位基因(Gold,B.等人,Nat Genet 38:458-62 (2006))。SLE风险等位基因与其他和SLE相关的HLA区域变体(DR2/DR3)没有显著的连锁不平衡(LD),在整合了DR2和DR3的条件性逻辑回归分析之后仍然是显著的。HLA是复杂的遗传区,但令人惊讶的是SNP rs641153的等位基因具有与报道过的AMD风险等位基因(Gold,B.等人,Nat Genet 38:458-62 (2006))几乎相同的保护效应。指出了5个候选疾病等位基因的其他研究。

[0291] 此外,表7提供了在其他自身免疫病中鉴别的42个变体的详细的摘要统计。令人感兴趣的是,其他自身免疫病中的显著风险因子CTLA4、IL23R、NOD2和CD40的变体似乎没有表现出任何与SLE的关联性证据。

[0292] 使用26个SLE风险等位基因(21个之前在表2中报道过的基因座,和5个上述新型SLE基因座),实施一些额外的分析。用确认的基因座进行成对相互作用分析,与之前的SLE(Harley, J.B.等人, Nat Genet 40:204-10(2008))和其他复杂疾病(Barrett, J.C.等人, Nat Genet 40:955-62(2008))的文献一致,没有观察到任何非累加相互作用的证据。使用条件性逻辑回归分析,没有发现多个独立的等位基因对任何单个风险基因座的风险起作用的任何证据。之后,使用Barrett等人描述的方法(Barrett, J.C.等人, Nat Genet 40:955-62(2008)),由每个确认的SLE风险等位基因估算了方差解释的百分比。HLA-DR3、IRF5和STAT4各自被评估占>1%的遗传方差,而其余的基因座分别占小于1%的方差。总而言之,这26个SLE风险基因座解释了约8%的SLE总遗传易感性。

[0293] 靶向复制GWAS结果是验证其他风险基因座的有效研究设计(Hirschhorn, J.N.等人, Nat Rev Genet 6:95-108(2005))。然而,对于缺少可接受的对于全基因组显著性的可接受的P值标准的复制结果的概率,只有极少数可利用的数据。在本研究中,复制囊括了来自原始GWAS研究的所有 $P < 0.05$ 的变体。如图3所示, GWAS研究中的P值越低,在复制meta-分析中获得候选物或验证状态的可能性越高。令人感兴趣的是,在GWAS P在0.05和0.01之间的变体组的现有研究中,没有获得任何候选物或验证的结果,尽管所述组占复制中测试的所有变体的~50%。这些结果可用于指导其他的靶向研究设计,虽然的确在计划复制工作中还需要仔细考虑原始GWAS群体的大小、复制样品大小、疾病结构和候选变体的效应量。

[0294] 这些数据进一步提供了证据,证明在对于免疫系统的继发性和先天性防御功能重要的基因中的共同变异,对于确立出现SLE的风险是重要的。尽管每个鉴别的等位基因仅对整体遗传风险的一部分负责,这些和其他正在进行的研究为狼疮的发病机制提供了新的视角,为药物发现和研发提示了新的靶和通路。

[0295]

表2. 之前报道过的SLE风险基因座的复制结果

| P 值                                 |         |                 |                         |                         |                         |                         |                      |        |      |                  |  | 风险等位基因 |  |
|-------------------------------------|---------|-----------------|-------------------------|-------------------------|-------------------------|-------------------------|----------------------|--------|------|------------------|--|--------|--|
| SNP                                 | Chr     | 关键区域            | GWAS                    | US                      | 瑞典                      | 组合                      | 基因座                  | 风险等位基因 | 频率   | OR (95% C.I.)    |  |        |  |
| 现有数据集中具有 $P < 5 \times 10^{-8}$ 的变体 |         |                 |                         |                         |                         |                         |                      |        |      |                  |  |        |  |
| rs3135394 <sup>a</sup>              | 6p21.32 | 32,027-32,874   | 7.8 × 10 <sup>-22</sup> | 1.8 × 10 <sup>-26</sup> | 8.3 × 10 <sup>-21</sup> | 2.0 × 10 <sup>-60</sup> | HLA-DRA <sup>b</sup> | G      | 0.10 | 1.98 (1.84-2.14) |  |        |  |
| rs7574885 <sup>a</sup>              | 2q32.2  | 191,609-191,681 | 3.0 × 10 <sup>-19</sup> | 6.4 × 10 <sup>-16</sup> | 2.7 × 10 <sup>-12</sup> | 1.4 × 10 <sup>-41</sup> | STAT4                | T      | 0.23 | 1.57 (1.49-1.69) |  |        |  |
| rs2070187 <sup>a</sup>              | 7q32.1  | 128,276-128,476 | n.a.                    | 1.4 × 10 <sup>-16</sup> | 4.1 × 10 <sup>-9</sup>  | 5.6 × 10 <sup>-24</sup> | IRF5                 | C      | 0.11 | 1.68 (1.78-1.95) |  |        |  |
| rs11860850 <sup>a</sup>             | 16p11.2 | 31,195-31,277   | 5.3 × 10 <sup>-11</sup> | 1.8 × 10 <sup>-5</sup>  | 9.2 × 10 <sup>-9</sup>  | 1.9 × 10 <sup>-20</sup> | ITGAM                | T      | 0.13 | 1.43 (1.32-1.54) |  |        |  |
| rs2736340                           | 8p23.1  | 11,331-11,488   | 5.5 × 10 <sup>-6</sup>  | 4.6 × 10 <sup>-6</sup>  | 0.0028                  | 7.9 × 10 <sup>-17</sup> | BLK                  | T      | 0.25 | 1.35 (1.27-1.43) |  |        |  |
| rs5029937 <sup>a</sup>              | 6q23.3  | 138,174-138,284 | 1.0 × 10 <sup>-4</sup>  | 2.4 × 10 <sup>-7</sup>  | 3.1 × 10 <sup>-5</sup>  | 5.3 × 10 <sup>-10</sup> | TNFAIP3              | T      | 0.03 | 1.71 (1.51-1.95) |  |        |  |
| rs2476601                           | 1p13.2  | 113,963-114,251 | 3.3 × 10 <sup>-6</sup>  | 4.5 × 10 <sup>-6</sup>  | 1.5 × 10 <sup>-5</sup>  | 3.4 × 10 <sup>-12</sup> | PTPN22               | A      | 0.10 | 1.35 (1.24-1.47) |  |        |  |
| rs4983128                           | 11p15.5 | 0,485-0,664     | 0.0021                  | 1.5 × 10 <sup>-6</sup>  | 8.7 × 10 <sup>-4</sup>  | 4.9 × 10 <sup>-3</sup>  | PIRF1                | C      | 0.67 | 1.20 (1.13-1.27) |  |        |  |
| rs2205960                           | 1q25.1  | 171,454-171,523 | 9.5 × 10 <sup>-6</sup>  | 0.030                   | 6.7 × 10 <sup>-4</sup>  | 6.3 × 10 <sup>-5</sup>  | TNFSF4               | T      | 0.23 | 1.22 (1.15-1.30) |  |        |  |
| 之前报道 $P < 5 \times 10^{-8}$ 的变体     |         |                 |                         |                         |                         |                         |                      |        |      |                  |  |        |  |
| rs9271366 <sup>a</sup>              | 6p21.32 | 32,446-32,695   | 0.0079                  | 7.4 × 10 <sup>-4</sup>  | 8.3 × 10 <sup>-5</sup>  | 1.4 × 10 <sup>-7</sup>  | HLA-DRA <sup>b</sup> | G      | 0.16 | 1.26 (1.18-1.36) |  |        |  |
| rs6920220 <sup>a</sup>              | 6q23.3  | 138,000-138,048 | 9.9 × 10 <sup>-4</sup>  | 5.2 × 10 <sup>-4</sup>  | 0.049                   | 4.0 × 10 <sup>-7</sup>  | TNFAIP3              | A      | 0.21 | 1.17 (1.10-1.25) |  |        |  |
| rs2269368                           | Xq28    | 152,743-152,943 | 2.5 × 10 <sup>-5</sup>  | n.a.                    | 0.0049                  | 7.5 × 10 <sup>-7</sup>  | IRAK1/MECP2          | T      | 0.14 | 1.11 (1.01-1.22) |  |        |  |
| rs2431099                           | 5q33.3  | 159,813-159,821 | 1.5 × 10 <sup>-6</sup>  | 0.16                    | 0.047                   | 1.6 × 10 <sup>-4</sup>  | PTTG1                | G      | 0.52 | 1.15 (1.09-1.22) |  |        |  |
| rs5754217                           | 22q11.2 | 20,240-20,315   | 0.0080                  | 8.4 × 10 <sup>-4</sup>  | 0.018                   | 2.3 × 10 <sup>-5</sup>  | UBE3L3               | T      | 0.19 | 1.20 (1.13-1.27) |  |        |  |
| rs2245214 <sup>a</sup>              | 6q21    | 106,749-106,876 | 0.032                   | 4.3 × 10 <sup>-4</sup>  | 0.35                    | 1.2 × 10 <sup>-5</sup>  | ATG5                 | G      | 0.37 | 1.15 (1.09-1.21) |  |        |  |
| rs10516487                          | 4q24    | 102,930-103,134 | 0.097                   | 0.091                   | 0.0015                  | 8.3 × 10 <sup>-4</sup>  | BANK1                | G      | 0.70 | 1.11 (1.04-1.18) |  |        |  |
| rs2176082 <sup>a</sup>              | 3p14.3  | 58,214-58,443   | 0.010                   | 0.012                   | 0.0031                  | 1.2 × 10 <sup>-4</sup>  | PXK                  | A      | 0.28 | 1.17 (1.10-1.25) |  |        |  |
| rs1801274                           | 1q23.3  | 159,724-159,746 | 4.1 × 10 <sup>-4</sup>  | n.a.                    | n.a.                    | 4.1 × 10 <sup>-4</sup>  | FCGR2A               | G      | 0.50 | 1.16 (1.09-1.20) |  |        |  |
| 之前报道 $P > 5 \times 10^{-8}$ 的变体     |         |                 |                         |                         |                         |                         |                      |        |      |                  |  |        |  |
| rs280519 <sup>a</sup>               | 19p13.2 | 10,387-10,430   | 7.1 × 10 <sup>-4</sup>  | n.a.                    | 0.036                   | 7.4 × 10 <sup>-5</sup>  | TYK2                 | A      | 0.48 | 1.13 (1.06-1.21) |  |        |  |
| rs10156091                          | 7p21.3  | 8,134-8,154     | 0.095                   | 0.0031                  | 8.7 × 10 <sup>-4</sup>  | 6.5 × 10 <sup>-4</sup>  | ICA1                 | T      | 0.10 | 1.16 (1.06-1.27) |  |        |  |
| rs3022013                           | 1q25.3  | 181,538-181,670 | 0.26                    | 2.05 × 10 <sup>-6</sup> | 2.8 × 10 <sup>-4</sup>  | 0.0015                  | NKAIN2               | T      | 0.60 | 1.09 (1.03-1.16) |  |        |  |
| rs7629816                           | 8q12.1  | 56,985-57,025   | 0.49                    | 0.76                    | 0.19                    | 0.17                    | LYN                  | A      | 0.79 | 1.05 (0.96-1.17) |  |        |  |
| rs2071725                           | 22q13.2 | 41,908-41,970   | 0.63                    | 0.34                    | 0.29                    | 0.30                    | SCUBE1               | G      | 0.86 | 1.09 (0.98-1.20) |  |        |  |
| rs5744168 <sup>a</sup>              | 1q41    | n.a.            | n.a.                    | 1.00                    | 0.40                    | 0.67                    | TLR5                 | G      | 0.94 | 1.02 (0.94-1.12) |  |        |  |
| rs509749                            | 1q23.3  | 158,993-159,067 | 0.64                    | 0.94                    | 0.93                    | 0.76                    | LYN                  | G      | 0.96 | 1.01 (0.91-1.12) |  |        |  |

关键区域定义为这样的最小区域，所述区域含有在 HapMap CEU 群体中  $r^2 > 0.4$  的变体并接 HG18 坐标 (Mb) 报告所述关键区域。从指示的病例/对照群体计算 P 值 (GWAS: 1310 例和 7859

[0296]

个对照; U.S.: 1129 例和 2991 个对照, 瑞典: 834 例和 1338 个对照, 组合: 3273 例和 12,188 个对照样品), 按方法中描述的计算组合 P 值。风险等位基因是相对于+参照链报道的。风险等位基因频率是对照染色体中的频率。比值比是上文“方法”中描述的组合比值比。<sup>a</sup> 指示按“方法”中所述在 GWAS 样品中插补并在复制样品中直接基因分型的标记物。<sup>b</sup> rs3135394 与 HLA\*DR3 (DRB1\*0301) 等位基因具有  $r^2 = 0.87$ 。<sup>c</sup> rs9271366 与 HLA\*DR2 (DRB1\*1501) 等位基因具有  $r^2 = 0.97$ 。参见表 3 的扩大汇总统计。 N.A. = 不可获得; 由于未能通过 QC 测量 (TYK2、FCGR2A 和 IRAK1/MECP2), 或者所述特定的变体在全基因组的阵列中不存在 (TLR5 和 IRF5)。然而, rs2070197 (IRF5 区) 与 rs10488631 强烈的连锁不平衡 (LD), 在基因组扫描中具有  $P = 2 \times 10^{-11}$ 。

[0297]

表 3. 之前报道过的SLE风险基因座的复制研究中的等位基因频率

| GWAS       |         |                 |             |        |      |       |       |        |       | US    |      |       |       | 瑞典 |  |  |  |
|------------|---------|-----------------|-------------|--------|------|-------|-------|--------|-------|-------|------|-------|-------|----|--|--|--|
| 等位基因频率     |         |                 |             | 等位基因频率 |      |       |       | 等位基因频率 |       |       |      |       |       |    |  |  |  |
| SNP        | Chr     | 关键区域            | 基因座         | A1/A2  | OR   | 对照    | 病例    | OR     | 对照    | 病例    | OR   | 对照    | 病例    |    |  |  |  |
| rs3135394  | 6p21.32 | 32,027-32,874   | HLA-DP3     | G/A    | 1.89 | 0.102 | 0.166 | 2.33   | 0.090 | 0.188 | 2.27 | 0.135 | 0.239 |    |  |  |  |
| rs7574865  | 2q32.2  | 191,608-191,681 | STAT4       | T/G    | 1.57 | 0.233 | 0.314 | 1.54   | 0.217 | 0.300 | 2.03 | 0.208 | 0.347 |    |  |  |  |
| rs2070197  | 7q32.1  | 128,276-128,476 | IRF5        | C/T    | n.a. | n.a.  | n.a.  | 1.82   | 0.105 | 0.176 | 2.08 | 0.125 | 0.226 |    |  |  |  |
| rs11860650 | 16p11.2 | 31,195-31,277   | ITGAM       | T/C    | 1.50 | 0.130 | 0.177 | 1.27   | 0.142 | 0.175 | 1.84 | 0.112 | 0.178 |    |  |  |  |
| rs2736340  | 8p23.1  | 11,331-11,488   | BLK         | T/C    | 1.31 | 0.246 | 0.299 | 1.47   | 0.236 | 0.313 | 1.25 | 0.262 | 0.307 |    |  |  |  |
| rs5029937  | 6q23.3  | 138,174-138,284 | TNFAIP3     | T/G    | 1.57 | 0.034 | 0.051 | 1.84   | 0.033 | 0.059 | 1.88 | 0.034 | 0.065 |    |  |  |  |
| rs2476601  | 1p13.2  | 113,963-114,251 | PTPN22      | A/G    | 1.35 | 0.098 | 0.116 | 1.48   | 0.083 | 0.119 | 1.47 | 0.120 | 0.167 |    |  |  |  |
| rs2245214  | 6q21    | 106,749-106,876 | ATG5        | G/C    | 1.10 | 0.370 | 0.393 | 1.31   | 0.353 | 0.416 | 1.05 | 0.407 | 0.420 |    |  |  |  |
| rs4963128  | 11p15.5 | 0,485-0,664     | PHRF1       | C/T    | 1.16 | 0.673 | 0.698 | 1.28   | 0.660 | 0.712 | 1.27 | 0.885 | 0.734 |    |  |  |  |
| rs2205860  | 1q25.1  | 171,454-171,523 | TNFSF4      | T/G    | 1.25 | 0.230 | 0.269 | 1.18   | 0.225 | 0.255 | 1.38 | 0.233 | 0.280 |    |  |  |  |
| rs9271366  | 6p21.32 | 32,446-32,695   | HLA-DP2     | G/A    | 1.16 | 0.172 | 0.192 | 1.41   | 0.142 | 0.188 | 1.39 | 0.157 | 0.204 |    |  |  |  |
| rs6920220  | 6q23.3  | 138,000-138,048 | TNFAIP3     | A/G    | 1.19 | 0.206 | 0.234 | 1.28   | 0.190 | 0.231 | 1.16 | 0.232 | 0.257 |    |  |  |  |
| rs2269368  | Xq28    | 152,743-152,943 | IRAK1/MECP2 | T/C    | 1.29 | 0.141 | 0.175 | n.a.   | n.a.  | n.a.  | n.a. | n.a.  | n.a.  |    |  |  |  |
| rs2431099  | 5q33.3  | 159,813-159,821 | PTTG1       | G/A    | 1.20 | 0.522 | 0.568 | 1.09   | 0.515 | 0.536 | 1.14 | 0.541 | 0.578 |    |  |  |  |
| rs5754217  | 22q11.2 | 20,240-20,315   | USE2L3      | T/G    | 1.16 | 0.188 | 0.213 | 1.23   | 0.191 | 0.225 | 1.22 | 0.231 | 0.268 |    |  |  |  |
| rs2176382  | 3p14.3  | 58,214-58,443   | PXK         | A/G    | 1.13 | 0.284 | 0.309 | 1.21   | 0.274 | 0.314 | 1.22 | 0.308 | 0.351 |    |  |  |  |
| rs280519   | 19p13.2 | 10,387-10,430   | TYK2        | A/G    | 1.16 | 0.477 | 0.507 | n.a.   | n.a.  | n.a.  | 1.15 | 0.476 | 0.511 |    |  |  |  |
| rs1801274  | 1q23.3  | 159,724-159,746 | FCGR2A      | G/A    | 1.16 | 0.500 | 0.537 | n.a.   | n.a.  | n.a.  | n.a. | n.a.  | n.a.  |    |  |  |  |
| rs10156091 | 7p21.3  | 8,134-8,154     | JCA1        | T/C    | 1.12 | 0.098 | 0.104 | 1.26   | 0.105 | 0.129 | 1.19 | 0.095 | 0.110 |    |  |  |  |
| rs10516487 | 4q24    | 102,930-103,134 | BANK1       | G/A    | 1.08 | 0.694 | 0.712 | 1.10   | 0.698 | 0.716 | 1.20 | 0.722 | 0.758 |    |  |  |  |
| rs2022013  | 1q25.3  | 181,538-181,670 | MMNAT2      | T/C    | 1.05 | 0.599 | 0.609 | 1.22   | 0.580 | 0.627 | 1.04 | 0.618 | 0.627 |    |  |  |  |
| rs7829816  | 8q12.1  | 56,985-57,025   | LYN         | A/G    | 1.04 | 0.786 | 0.795 | 1.03   | 0.783 | 0.789 | 1.07 | 0.817 | 0.827 |    |  |  |  |
| rs2071725  | 22q13.2 | 41,908-41,970   | SCUBE1      | G/A    | 1.03 | 0.859 | 0.870 | 1.12   | 0.859 | 0.873 | 1.01 | 0.887 | 0.899 |    |  |  |  |
| rs5744168  | 1       | n.a.            | TLR5        | G/A    | n.a. | n.a.  | n.a.  | 0.93   | 0.947 | 0.943 | 1.19 | 0.920 | 0.932 |    |  |  |  |
| rs509749   | 1q23.3  | 158,993-159,067 | LY9         | G/A    | 1.02 | 0.425 | 0.428 | 1.00   | 0.429 | 0.430 | 0.99 | 0.453 | 0.451 |    |  |  |  |

关键区域在此定义为这样的最小区域，所述区域含有在 HapMap CEU 群体中  $r^2 > 0.4$  的变体并按 HG18 坐标报告所述关键区域。从指示的病例/对照群体计算等位基因频率 (GWAS: 1310 例和 7859 个对照; U.S.: 1129 例和 2991 个对照; 瑞典: 834 例和 1338 个对照; 组合: 3273 例和 12,188 个对照样品)。等位基因是相对于+参照链报道的，所有的数据都参考等位基因 1(A1)。列举了每个群体的比值比(OR)。

[0298]

表4. 组合数据集的新型SLE风险基因座

| 表 4. 组合数据集中的新型SLE风险基因座               |      |                 |                        |                        |                         |          |        |      |                  |
|--------------------------------------|------|-----------------|------------------------|------------------------|-------------------------|----------|--------|------|------------------|
| SNP                                  | Chr. | 关键区域            | P 值                    |                        |                         | 基因座      | 风险等位基因 | 频率   | OR (95% C.I.)    |
|                                      |      |                 | US                     | 瑞典                     | 组合                      |          |        |      |                  |
| 全基因组显著的基因座                           |      |                 |                        |                        |                         |          |        |      |                  |
| rs7708392 <sup>a</sup>               | 5    | 150,419-150,441 | 7.7 × 10 <sup>-4</sup> | 1.2 × 10 <sup>-5</sup> | 3.8 × 10 <sup>-10</sup> | TNIP1    | C      | 0.24 | 1.27 (1.10-1.35) |
| rs6568431                            | 6    | 106,675-106,705 | 0.0016                 | 0.0050                 | 7.1 × 10 <sup>-10</sup> | PRDM1    | A      | 0.38 | 1.20 (1.14-1.27) |
| rs849142 <sup>a</sup>                | 7    | 28,108-28,223   | 0.10                   | 5.4 × 10 <sup>-4</sup> | 1.5 × 10 <sup>-9</sup>  | JAZF1    | T      | 0.49 | 1.19 (1.13-1.26) |
| rs11753393 <sup>a</sup>              | 6    | 34,658-35,090   | 3.7 × 10 <sup>-4</sup> | 5.1 × 10 <sup>-4</sup> | 2.2 × 10 <sup>-6</sup>  | UHRF1BP1 | G      | 0.35 | 1.17 (1.10-1.24) |
| rs3024505                            | 1    | 205,007-205,016 | 0.062                  | 1.9 × 10 <sup>-4</sup> | 4.0 × 10 <sup>-8</sup>  | IL10     | A      | 0.16 | 1.19 (1.11-1.28) |
| 具有组合 P 值 < 1 × 10 <sup>-5</sup> 的基因座 |      |                 |                        |                        |                         |          |        |      |                  |
| rs10911363 <sup>a</sup>              | 1    | 181,672-181,816 | 1.5 × 10 <sup>-5</sup> | 0.52                   | 9.5 × 10 <sup>-3</sup>  | NCF3     | T      | 0.27 | 1.19 (1.12-1.26) |
| rs12444486 <sup>a</sup>              | 16   | 84,548-84,576   | 0.021                  | 0.026                  | 1.9 × 10 <sup>-7</sup>  | IRF8     | T      | 0.50 | 1.16 (1.10-1.23) |
| rs11013310 <sup>a</sup>              | 10   | 23,181-23,337   | 0.013                  | 0.12                   | 2.0 × 10 <sup>-7</sup>  | ARMC3    | T      | 0.21 | 1.18 (1.11-1.26) |
| rs1874791 <sup>a</sup>               | 1    | 67,563-67,887   | 0.012                  | 0.11                   | 3.4 × 10 <sup>-7</sup>  | IL12RB2  | A      | 0.18 | 1.18 (1.10-1.26) |
| rs9782355                            | 1    | 233,893-234,107 | 0.057                  | 0.12                   | 4.6 × 10 <sup>-7</sup>  | LYST     | C      | 0.74 | 1.18 (1.11-1.26) |
| rs7683337 <sup>a</sup>               | 4    | 185,805-185,914 | 0.11                   | 0.0013                 | 7.6 × 10 <sup>-7</sup>  | MLF1P    | T      | 0.82 | 1.23 (1.14-1.33) |
| rs428073                             | 12   | 117,706-117,315 | 0.22                   | 0.0079                 | 7.7 × 10 <sup>-7</sup>  | TAKK3    | T      | 0.69 | 1.18 (1.11-1.26) |
| rs497273 <sup>a</sup>                | 12   | 119,610-119,891 | 0.068                  | 0.021                  | 8.2 × 10 <sup>-7</sup>  | SPPL3    | G      | 0.65 | 1.14 (1.08-1.21) |
| 具有组合 P 值 < 1 × 10 <sup>-5</sup> 的基因座 |      |                 |                        |                        |                         |          |        |      |                  |
| rs1861525                            | 7    | 25,097-25,183   | 0.16                   | 0.0027                 | 1.9 × 10 <sup>-6</sup>  | CYCS     | G      | 0.05 | 1.27 (1.12-1.45) |
| rs921916                             | 7    | 50,193-50,205   | 0.027                  | 0.014                  | 2.0 × 10 <sup>-6</sup>  | IKZF1    | C      | 0.18 | 1.15 (1.07-1.23) |
| rs7333671                            | 13   | 73,177-73,198   | 0.14                   | 0.0027                 | 2.2 × 10 <sup>-6</sup>  | KLF12    | G      | 0.08 | 1.22 (1.11-1.34) |
| rs12992463                           | 2    | 22,312-22,464   | 0.23                   | 0.023                  | 2.6 × 10 <sup>-6</sup>  | ...      | A      | 0.50 | 1.12 (1.06-1.19) |
| rs12620999                           | 2    | 237,616-237,770 | 0.040                  | 0.45                   | 3.1 × 10 <sup>-6</sup>  | COPB3    | C      | 0.19 | 1.13 (1.06-1.21) |
| rs503425 <sup>a</sup>                | 11   | 118,079-118,198 | 3.3 × 10 <sup>-4</sup> | 0.43                   | 3.3 × 10 <sup>-6</sup>  | CDX6     | C      | 0.20 | 1.16 (1.08-1.24) |
| rs10743326 <sup>a</sup>              | 11   | 94,733-94,809   | 0.017                  | 0.21                   | 3.6 × 10 <sup>-6</sup>  | APIP     | G      | 0.59 | 1.14 (1.08-1.21) |
| rs4766321 <sup>a</sup>               | 12   | 117,835-117,863 | n.a.                   | 0.036                  | 4.6 × 10 <sup>-6</sup>  | KIAA1853 | G      | 0.67 | 1.18 (1.09-1.27) |
| rs11951576 <sup>a</sup>              | 5    | 6,741-6,866     | 0.42                   | 0.014                  | 4.6 × 10 <sup>-6</sup>  | POLR1A   | C      | 0.69 | 1.14 (1.08-1.22) |
| rs6438700                            | 3    | 123,355-123,454 | 0.23                   | 0.020                  | 5.5 × 10 <sup>-6</sup>  | CD86     | C      | 0.82 | 1.18 (1.09-1.27) |
| rs6486730 <sup>a</sup>               | 12   | 127,830-127,840 | 0.16                   | 0.049                  | 6.9 × 10 <sup>-6</sup>  | SLC15A4  | G      | 0.41 | 1.13 (1.07-1.19) |
| rs4748857 <sup>a</sup>               | 10   | 23,529-23,854   | 0.68                   | 1.3 × 10 <sup>-4</sup> | 6.9 × 10 <sup>-6</sup>  | C10orf67 | C      | 0.73 | 1.16 (1.09-1.24) |
| rs3914167 <sup>a</sup>               | 5    | 39,426-39,454   | 0.24                   | 0.0081                 | 7.6 × 10 <sup>-6</sup>  | DAB2/IC9 | G      | 0.27 | 1.15 (1.09-1.23) |

样品、关键区域、P 值、风险等位基因和比值比按表 2 的说明所定义。<sup>a</sup> 指按上文“方法”中所述，从 GWAS 样品插补的并在复制样品中直接基因分型的标记物。扩大汇总统计参见表 5。

[0299]

表5. 组合数据集中的显著SLE风险基因座的其他汇总统计。

| SNP        | Chr | 关键区域            | 基因座        | GWAS   |      |       |      | US     |      |       |      | 瑞典     |       |
|------------|-----|-----------------|------------|--------|------|-------|------|--------|------|-------|------|--------|-------|
|            |     |                 |            | 等位基因频率 |      | 病例    |      | 等位基因频率 |      | 病例    |      | 等位基因频率 |       |
|            |     |                 |            | 对照     | OR   | A1/A2 | OR   | 对照     | OR   | 对照    | OR   | 对照     | OR    |
| rs7706392  | 5   | 150,419-150,441 | TNIP1      | 0.232  | 1.28 | C/G   | 1.19 | 0.267  | 1.37 | 0.502 | 1.37 | 0.256  | 1.324 |
| rs6588431  | 6   | 106,875-106,705 | PRDM1      | 0.380  | 1.22 | A/C   | 1.22 | 0.370  | 1.18 | 0.418 | 1.18 | 0.412  | 0.451 |
| rs849142   | 7   | 26,108-28,223   | JAZF1      | 0.490  | 1.23 | T/C   | 1.23 | 0.491  | 1.25 | 0.516 | 1.25 | 0.499  | 0.552 |
| rs11755393 | 6   | 34,658-35,090   | UHRF1BP1   | 0.354  | 1.15 | G/A   | 1.15 | 0.347  | 1.27 | 0.380 | 1.27 | 0.326  | 0.376 |
| rs3024505  | 1   | 205,007-205,016 | IL10       | 0.166  | 1.28 | A/G   | 1.14 | 0.152  | 1.21 | 0.169 | 1.21 | 0.150  | 0.176 |
| rs10911363 | 1   | 181,672-181,816 | NCF2       | 0.274  | 1.21 | T/G   | 1.27 | 0.273  | 1.05 | 0.323 | 1.05 | 0.275  | 0.293 |
| rs12444486 | 18  | 84,548-84,576   | IRF8       | 0.507  | 1.19 | T/C   | 1.11 | 0.501  | 1.15 | 0.528 | 1.15 | 0.482  | 0.526 |
| rs11013210 | 10  | 23,181-23,337   | ARMC3      | 0.199  | 1.27 | A/G   | 1.17 | 0.229  | 1.13 | 0.225 | 1.13 | 0.225  | 0.243 |
| rs1874791  | 1   | 67,563-67,687   | IL13RB2    | 0.186  | 1.25 | A/G   | 1.09 | 0.188  | 1.15 | 0.202 | 1.15 | 0.146  | 0.163 |
| rs9782955  | 1   | 233,893-234,107 | LYST       | 0.737  | 1.26 | C/T   | 1.12 | 0.744  | 1.15 | 0.766 | 1.15 | 0.765  | 0.788 |
| rs7683537  | 4   | 185,805-185,914 | MLF1P      | 0.811  | 1.23 | T/C   | 1.13 | 0.832  | 1.34 | 0.843 | 1.34 | 0.834  | 0.872 |
| rs428073   | 12  | 117,796-117,315 | TAOK3      | 0.691  | 1.22 | T/C   | 1.07 | 0.694  | 1.31 | 0.708 | 1.31 | 0.682  | 0.736 |
| rs497273   | 12  | 119,610-119,891 | SPPL3      | 0.649  | 1.19 | G/C   | 1.03 | 0.663  | 1.16 | 0.670 | 1.16 | 0.618  | 0.660 |
| rs1861525  | 7   | 25,097-25,183   | CYC5       | 0.054  | 1.52 | G/A   | 1.14 | 0.041  | 1.98 | 0.047 | 1.98 | 0.017  | 0.032 |
| rs821916   | 7   | 50,193-50,205   | KZF1       | 0.187  | 1.20 | C/T   | 1.08 | 0.189  | 1.27 | 0.200 | 1.27 | 0.152  | 0.185 |
| rs7333671  | 13  | 73,177-73,198   | KLF12      | 0.085  | 1.32 | G/A   | 1.07 | 0.082  | 1.35 | 0.088 | 1.35 | 0.066  | 0.087 |
| rs12992463 | 2   | 22,312-22,464   | LOC845949  | 0.499  | 1.20 | A/C   | 1.04 | 0.510  | 1.15 | 0.521 | 1.15 | 0.495  | 0.531 |
| rs12620999 | 2   | 237,616-237,770 | CDPS9      | 0.190  | 1.27 | C/T   | 1.12 | 0.180  | 1.03 | 0.198 | 1.03 | 0.185  | 0.190 |
| rs503425   | 11  | 118,079-118,198 | DDX6       | 0.206  | 1.18 | C/T   | 1.20 | 0.194  | 1.06 | 0.224 | 1.06 | 0.198  | 0.206 |
| rs10742326 | 11  | 34,733-34,809   | APIP       | 0.585  | 1.18 | G/A   | 1.11 | 0.591  | 1.09 | 0.617 | 1.09 | 0.577  | 0.601 |
| rs4766321  | 12  | 117,835-117,883 | KIAA1853   | 0.668  | 1.22 | G/A   | n.a. | n.a.   | 1.15 | n.a.  | 1.15 | 0.662  | 0.689 |
| rs11951576 | 5   | 6,741-6,886     | POLS/SRD5A | 0.686  | 1.32 | C/T   | 1.04 | 0.691  | 1.19 | 0.700 | 1.19 | 0.689  | 0.701 |
| rs6438700  | 3   | 123,355-123,454 | CD86       | 0.823  | 1.25 | C/T   | 1.05 | 0.826  | 1.23 | 0.834 | 1.23 | 0.821  | 0.851 |
| rs6486730  | 12  | 127,830-127,840 | SLC15A4    | 0.405  | 1.19 | G/A   | 1.06 | 0.420  | 1.14 | 0.436 | 1.14 | 0.422  | 0.450 |
| rs4748857  | 10  | 23,529-23,654   | C10orf67   | 0.715  | 1.22 | C/T   | 1.10 | 0.763  | 1.35 | 0.780 | 1.35 | 0.742  | 0.791 |
| rs3914167  | 5   | 39,426-39,454   | DAB2IP3    | 0.274  | 1.19 | G/C   | 1.08 | 0.276  | 1.20 | 0.291 | 1.20 | 0.262  | 0.294 |

在此, 关键区域定义为这样的最小区域, 所述区域含有在 HapMap CEU 群体中  $r^2 > 0.4$  的变体并按 HG18 坐标报告所述关键区域。从指示的病例/对照群体计算等位基因频率 (GWAS: 1310 例和 7859 个对照; U.S.: 1129 例和 2991 个对照; 瑞典: 834 例和 1338 个对照; 组合: 3273 例和 12,188 个对照样品)。等位基因是相对于+参照链报道的, 所有的数据都参考等位基因 1(A1)。列举了每个群体的比值比(OR)。

[0300]

表 6. 显示与SLE相关的候选自身免疫基因座

| P 值         |         |     |                        |                        |        |                         |                         |      |      |      |                    |                                                                                                                       |
|-------------|---------|-----|------------------------|------------------------|--------|-------------------------|-------------------------|------|------|------|--------------------|-----------------------------------------------------------------------------------------------------------------------|
| SNP         | 基因座     | Chr | GWAS                   | US                     | 瑞典     | 组合                      | 组合校正                    | 风险等位 |      |      | 表型                 | 参考文献                                                                                                                  |
|             |         |     |                        |                        |        |                         |                         | 等位   | 基因   | 频率   |                    |                                                                                                                       |
| rs1990760   | IFIH1   | 2   | 3.2 × 10 <sup>-6</sup> | 0.015                  | 0.0039 | 3.34 × 10 <sup>-7</sup> | 1.12 × 10 <sup>-5</sup> | T    | 0.60 | 1.17 | T1D, Grave's       | Nat Genet 38:617-9 (2006);<br>J Clin Endocrinol Metab 92:3338-41 (2007)                                               |
| rs641153*   | CFB     | 6   | 0.0079                 | n.a.                   | 0.0011 | 1.4 × 10 <sup>-4</sup>  | 0.0049                  | G    | 0.91 | 1.30 | AMD                | Nat Genet 38:458-62 (2006)                                                                                            |
| rs12708716* | CLEC16A | 16  | 0.15                   | 1.3 × 10 <sup>-4</sup> | 0.062  | 1.6 × 10 <sup>-4</sup>  | 0.0056                  | A    | 0.64 | 1.16 | T1D, Addison's, MS | J Clin Endocrinol Metab 94:231-235 (2009);<br>J Clin Endocrinol Metab 93:3310-7 (2008);<br>Genes Immun 10:15-7 (2009) |
|             |         |     |                        |                        |        |                         |                         |      |      |      |                    |                                                                                                                       |
| rs6887695*  | IL12B   | 5   | 0.014                  | 0.04                   | 0.03   | 1.7 × 10 <sup>-4</sup>  | 0.0060                  | G    | 0.68 | 1.13 | PS, IBD            | Nat Genet 41:199-204 (2009);<br>Nat Genet 40:710-2 (2008)                                                             |
| rs17696736  | SH2B3   | 12  | 0.0036                 | 0.12                   | 0.19   | 4.0 × 10 <sup>-4</sup>  | 0.014                   | T    | 0.50 | 1.08 | T1D, Celiac, SLE   | J Clin Endocrinol Metab 93:3310-7 (2008);<br>Nat Genet 40:395-402 (2008);<br>N Engl J Med 359:2767-77 (2008)          |

表中的所有等位基因与报道过的变体相同或者与报道过的变体具有  $r^2 > 0.8$ , 都是具有相同的效应方向的相同的风险等位基因。按 HGI8 坐标报告位置 (碱基对)。样品、单个和组合 P 值、风险等位基因频率和 OR 按表 2 的说明所述。组合校正的 P 值是对 35 个之前报道过的风险基因座的 Bonferonni 校正的 P 值。其他自身免疫关联性: T1D = 1 型糖尿病; AMD = 年龄相关性黄斑变性; MS = 多发性硬化; IBD = 炎性肠病, 和 PS = 牛皮癣。扩大汇总统计和完整的测试变体列表参见表 7。" 指接上文 "方法" 中所述, 从 GWAS 样品插补的并在复制样品中直接基因分型的标记物。

[0301]

表7 (部分1) 确认的自身免疫疾病基因座

| SNP          | Chr. | 位置        | 基因座       | A1/A2 | GWAS                 |       |       | US                   |       |       |
|--------------|------|-----------|-----------|-------|----------------------|-------|-------|----------------------|-------|-------|
|              |      |           |           |       | P 值                  | OR    | 对照    | P 值                  | OR    | 对照    |
| rs1990760    | 2    | 162832297 | IFIH1     | T/C   | $3.2 \times 10^{-4}$ | 1.17  | 0.690 | 0.015                | 1.17  | 0.581 |
| rs641153     | 6    | 32022159  | CFB       | G/A   | 0.0078               | 1.22  | 0.910 | n.a.                 | n.a.  | n.a.  |
| rs12708716   | 18   | 11087374  | CLEC16A   | A/G   | 0.15                 | 1.06  | 0.635 | $1.3 \times 10^{-4}$ | 1.29  | 0.616 |
| rs6887695    | 5    | 158755223 | IL12B     | G/C   | 0.014                | 1.12  | 0.883 | 0.043                | 1.11  | 0.676 |
| rs17696736   | 12   | 110971201 | C12orf30  | G/A   | 0.0061               | 1.12  | 0.449 | 0.16                 | 1.01  | 0.459 |
| rs184504     | 12   | 110366991 | SH2B3     | T/C   | 0.0036               | 1.13  | 0.503 | 0.12                 | 1.04  | 0.493 |
| rs2612378    | 9    | 34700260  | COL21     | G/A   | 0.003                | 1.14  | 0.323 | 0.79                 | 1.02  | 0.325 |
| rs3761847    | 9    | 122730850 | TRAF1     | G/A   | 0.034                | 1.10  | 0.411 | 0.20                 | 1.12  | 0.416 |
| rs6598540    | 6    | 43666302  | VEGFA     | C/A   | $8.1 \times 10^{-4}$ | 1.22  | 0.172 | 0.90                 | 1.03  | 0.166 |
| rs547154     | 6    | 32018917  | C2        | T/G   | n.a.                 | n.a.  | n.a.  | 0.28                 | 0.94  | 0.080 |
| rs12044852   | 1    | 116883302 | CD58      | C/A   | 0.099                | 1.12  | 0.886 | 0.25                 | 1.04  | 0.903 |
| rs3542151    | 18   | 12769947  | PTPN2     | G/T   | 0.15                 | 1.09  | 0.152 | n.a.                 | n.a.  | n.a.  |
| rs6897932    | 5    | 35910332  | IL7R      | C/T   | 0.18                 | 1.06  | 0.740 | 0.25                 | 1.05  | 0.743 |
| rs2230199    | 19   | 6689387   | C3        | C/G   | n.a.                 | n.a.  | n.a.  | n.a.                 | n.a.  | 0.752 |
| rs3732378    | 3    | 39282168  | CXCR1     | G/A   | 0.13                 | 1.09  | 0.833 | 0.083                | 1.12  | 0.829 |
| rs1678542    | 12   | 56754882  | KIF5A     | C/G   | 0.051                | 1.09  | 0.624 | 0.94                 | 1.03  | 0.626 |
| rs1136287    | 17   | 1620026   | SERPINF1  | T/C   | 0.64                 | 1.02  | 0.641 | 0.19                 | 1.02  | 0.647 |
| rs12247631   | 10   | 52485603  | PRKG1     | A/C   | 0.32                 | 1.56  | 0.004 | 0.051                | 1.48  | 0.003 |
| rs2227306    | 4    | 74825919  | IL8       | T/C   | 0.036                | 1.11  | 0.407 | 0.42                 | 0.98  | 0.415 |
| rs12521868   | 5    | 131812292 | LOC441108 | T/G   | 0.093                | 1.08  | 0.424 | 0.87                 | 0.93  | 0.406 |
| rs10490924   | 10   | 124204438 | ARMS2     | G/T   | 0.15                 | 1.09  | 0.787 | 0.56                 | 1.03  | 0.786 |
| rs1793004    | 11   | 20655605  | NELL1     | G/C   | 0.81                 | 1.01  | 0.756 | 0.25                 | 1.01  | 0.751 |
| rs10225965   | 7    | 92111514  | CDK6      | C/T   | 0.033                | 1.12  | 0.602 | 0.61                 | 0.91  | 0.820 |
| rs1410996    | 1    | 194963556 | CFH       | A/G   | 0.08                 | 1.08  | 0.414 | 0.48                 | 0.94  | 0.427 |
| rs3087243    | 2    | 204447164 | CTLA4     | G/A   | 0.30                 | 1.04  | 0.567 | 0.86                 | 1.04  | 0.549 |
| rs2392239    | 12   | 54788447  | ERBB3     | T/G   | 0.72                 | 1.02  | 0.329 | 0.11                 | 1.10  | 0.330 |
| rs12722489   | 10   | 6142018   | IL2RA     | C/T   | 0.89                 | 1.01  | 0.849 | 0.851                | n.a.  | n.a.  |
| rs3332739    | 8    | 32011783  | C2        | C/G   | 0.92                 | 1.01  | 0.048 | n.a.                 | n.a.  | n.a.  |
| rs2076756    | 16   | 49314382  | NOD2      | G/A   | 0.44                 | 1.04  | 0.273 | n.a.                 | n.a.  | 0.256 |
| rs16853571   | 4    | 41447837  | PHOX2B    | C/A   | 0.51                 | 1.06  | 0.065 | 0.79                 | 0.99  | 0.065 |
| rs4810485    | 20   | 44181354  | CD40      | T/G   | 0.37                 | 1.04  | 0.264 | n.a.                 | n.a.  | n.a.  |
| rs5340799    | 8    | 152205074 | ESR1      | A/G   | 0.17                 | 1.06  | 0.646 | 0.26                 | 0.99  | 0.638 |
| rs3793784    | 10   | 50417545  | ERCC8     | C/G   | 0.54                 | 1.03  | 0.402 | 0.42                 | 0.97  | 0.406 |
| rs2240340    | 1    | 17535236  | PAD1A     | C/T   | 0.64                 | 1.02  | 0.579 | 0.13                 | 0.93  | 0.574 |
| rs7517847    | 1    | 67454257  | IL23R     | T/G   | 0.79                 | 1.01  | 0.569 | 0.86                 | 1.06  | 0.419 |
| LD中含上述基因座的变体 |      |           |           |       |                      |       |       |                      |       |       |
| rs1061170    | 1    | 194925860 | CFH       | T/C   | n.a.                 | n.a.  | n.a.  | n.a.                 | n.a.  | n.a.  |
| rs2234693    | 6    | 152206028 | ESR1      | T/C   | n.a.                 | n.a.  | n.a.  | 0.403                | 0.972 | 0.539 |
| rs1136267    | 17   | 1620026   | SERPINF1  | T/C   | 0.642                | 1.021 | 0.641 | 0.190                | 1.023 | 0.647 |
| rs3024987    | 6    | 43653085  | VEGFA     | A/G   | n.a.                 | n.a.  | n.a.  | 0.977                | 0.984 | 0.327 |
| rs3212227    | 5    | 158675528 | IL12B     | T/G   | 0.000                | 1.230 | 0.792 | 0.407                | 1.059 | 0.784 |
| rs2338698    | 10   | 53379358  | PRKG1     | C/T   | 0.626                | 1.022 | 0.338 | 0.190                | 1.051 | 0.338 |
| rs42041      | 7    | 92084690  | CDK6      | C/C   | 0.570                | 1.038 | 0.267 | 0.790                | 0.986 | 0.258 |

[0302]

表7(部分2)。确认的自身免疫疾病基因座

| SNP          | Chr | 位置        | 基因座       | A1/A2 | P 值    | OR    | 对照    | 病例    | 复制                   | 组合                   |
|--------------|-----|-----------|-----------|-------|--------|-------|-------|-------|----------------------|----------------------|
| rs1990760    | 2   | 162832287 | IFIH1     | T/C   | 0.0039 | 1.17  | 0.809 | 0.845 | $2.6 \times 10^{-4}$ | $3.3 \times 10^{-4}$ |
| rs641153     | 6   | 32022159  | CFB       | G/A   | 0.0011 | 1.47  | 0.915 | 0.939 | 0.0011               | $1.4 \times 10^{-4}$ |
| rs12708716   | 16  | 11081374  | CLEC16A   | A/G   | 0.061  | 1.14  | 0.694 | 0.725 | $2.8 \times 10^{-6}$ | $1.8 \times 10^{-6}$ |
| rs6887695    | 5   | 158755223 | IL12B     | G/C   | 0.032  | 1.16  | 0.664 | 0.700 | 0.0034               | $1.7 \times 10^{-6}$ |
| rs17596736   | 12  | 110971201 | C12orf30  | G/A   | 0.019  | 1.16  | 0.437 | 0.474 | 0.0186               | $2.7 \times 10^{-4}$ |
| rs3184504    | 12  | 110368981 | SH2B3     | T/C   | 0.19   | 1.08  | 0.487 | 0.508 | 0.042                | $4.0 \times 10^{-6}$ |
| rs2812378    | 9   | 34700260  | CCL21     | G/A   | 0.061  | 1.14  | 0.333 | 0.354 | 0.19                 | $1.8 \times 10^{-4}$ |
| rs3761847    | 9   | 12273060  | TRAF1     | G/A   | 0.12   | 1.08  | 0.456 | 0.475 | 0.053                | 0.0041               |
| rs6899540    | 6   | 4386302   | VEGFA     | C/A   | 0.68   | 1.04  | 0.176 | 0.180 | 0.73                 | 0.0051               |
| rs547154     | 6   | 32018317  | C2        | T/G   | 0.0012 | 0.85  | 0.085 | 0.057 | 0.011                | 0.011                |
| rs12044852   | 1   | 116889302 | CD58      | C/A   | 0.33   | 1.10  | 0.872 | 0.883 | 0.13                 | 0.026                |
| rs2342131    | 18  | 12769847  | PTPN2     | G/T   | 0.077  | 1.16  | 0.154 | 0.178 | 0.077                | 0.037                |
| rs6887932    | 5   | 35910332  | IL7R      | C/T   | 0.23   | 1.06  | 0.712 | 0.725 | 0.10                 | 0.038                |
| rs2230199    | 19  | 6669387   | C3        | C/G   | 0.081  | 0.85  | 0.205 | 0.179 | 0.081                | 0.061                |
| rs3732378    | 3   | 39282166  | CX3CR1    | G/A   | 0.24   | 0.90  | 0.644 | 0.629 | 0.42                 | 0.093                |
| rs1678542    | 12  | 56254962  | KIF5A     | C/G   | 0.58   | 1.04  | 0.581 | 0.589 | 0.80                 | 0.095                |
| rs1136287    | 17  | 1620026   | SERPINF1  | T/C   | 0.16   | 1.05  | 0.637 | 0.648 | 0.039                | 0.12                 |
| rs12247831   | 10  | 52485603  | PRKG1     | A/C   | 0.27   | n.a.  | 0.000 | 0.001 | 0.25                 | 0.14                 |
| rs2227306    | 4   | 74825919  | IL8       | T/C   | 0.57   | 1.04  | 0.461 | 0.463 | 0.75                 | 0.16                 |
| rs12521668   | 5   | 131812292 | LOC441108 | T/G   | 0.65   | 1.03  | 0.418 | 0.415 | 0.90                 | 0.21                 |
| rs10490924   | 10  | 124204438 | ARMS2     | G/T   | 0.58   | 0.96  | 0.768 | 0.780 | 0.89                 | 0.22                 |
| rs1793204    | 11  | 39655505  | NELL1     | G/C   | 0.24   | 1.09  | 0.725 | 0.746 | 0.10                 | 0.22                 |
| rs10225965   | 7   | 92111514  | CDK6      | C/T   | 0.68   | 0.97  | 0.797 | 0.787 | 0.52                 | 0.23                 |
| rs1410895    | 1   | 194963586 | CFH       | A/G   | 0.57   | 1.04  | 0.409 | 0.411 | 0.81                 | 0.23                 |
| rs3067243    | 2   | 204447164 | CTLA4     | G/A   | 0.43   | 1.08  | 0.814 | 0.830 | 0.59                 | 0.25                 |
| rs2282239    | 12  | 54768447  | ERBB3     | T/G   | 0.96   | 1.00  | 0.330 | 0.324 | 0.19                 | 0.27                 |
| rs12722489   | 10  | 6142018   | IL2RA     | C/T   | 0.042  | 1.19  | 0.806 | 0.835 | 0.042                | 0.31                 |
| rs932739     | 6   | 32011783  | C2        | C/G   | 0.037  | 1.36  | 0.044 | 0.059 | 0.037                | 0.31                 |
| rs2076756    | 16  | 49314362  | NOD2      | G/A   | 0.78   | 1.01  | 0.200 | 0.203 | 0.64                 | 0.37                 |
| rs16853571   | 4   | 41447887  | PHOX2B    | C/A   | 0.33   | 1.20  | 0.056 | 0.067 | 0.73                 | 0.47                 |
| rs4810485    | 20  | 44181354  | CD40      | T/G   | 0.85   | 0.99  | 0.246 | 0.249 | 0.85                 | 0.47                 |
| rs9340799    | 6   | 152205074 | ESR1      | A/G   | 0.60   | 1.04  | 0.680 | 0.694 | 0.57                 | 0.49                 |
| rs3793784    | 10  | 50417545  | ERCC8     | C/G   | 0.23   | 1.08  | 0.385 | 0.385 | 0.96                 | 0.61                 |
| rs2240340    | 1   | 17535226  | PADI4     | C/T   | 0.91   | 0.98  | 0.594 | 0.585 | 0.19                 | 0.64                 |
| rs7517847    | 1   | 67454257  | IL23R     | T/G   | 0.33   | 0.98  | 0.531 | 0.525 | 0.48                 | 0.80                 |
| L3中全上述基因座的变体 |     |           |           |       |        |       |       |       |                      |                      |
| rs1061170    | 1   | 194925960 | CFH       | T/C   | 0.339  | 0.914 | 0.404 | 0.382 | 0.339                | 0.339                |
| rs2234693    | 6   | 152205028 | ESR1      | T/C   | 0.719  | 0.993 | 0.556 | 0.555 | 0.375                | 0.375                |
| rs1136287    | 17  | 1620026   | SERPINF1  | T/C   | 0.159  | 1.051 | 0.636 | 0.648 | 0.059                | 0.118                |
| rs3024987    | 6   | 43853065  | VEGFA     | A/G   | 0.662  | 1.006 | 0.293 | 0.294 | 0.847                | 0.847                |
| rs3212227    | 5   | 158675528 | IL12B     | T/G   | 0.636  | 1.042 | 0.810 | 0.817 | 0.342                | $4 \times 10^{-4}$   |
| rs233898     | 10  | 53379358  | PRKG1     | C/T   | 0.799  | 1.010 | 0.316 | 0.309 | 0.368                | 0.341                |
| rs42041      | 7   | 92084680  | CDK6      | G/C   | 0.049  | 0.913 | 0.268 | 0.239 | 0.173                | 0.661                |

[0303]

按 HG18 坐标报告位置 (碱基对)。从指示的病例/对照群体计算 P 值 (GWAS: 1310 例和 7859 个对照; U.S.: 1129 例和 2991 个对照; 瑞典: 834 例和 1338 个对照; 组合: 3273 例和 12,188 个对照样品), 按上述 “方法” 所述计算组合 P 值。复制 P 值指组合的 U.S. 和瑞典人样品的 meta P 值。等位基因是相对于 + 参照链报道的, 所有的数据都参考等位基因 1 (A1)。列举了每个群体的比值比 (OR)。

[0304] 实施例2

[0305] 对BLK的因果等位基因的重新测序和鉴别

[0306] 如上所述, BLK被鉴别为达到全基因组显著性 ( $P < 5 \times 10^{-8}$ ) 的与SLE相关的风险基因座。为了进一步表征该关联性的遗传基础和鉴别因果等位基因, 我们对BLK基因座进行重新测序研究和下文所述的报告基因表达测定。

[0307] 对于重新测序研究, 对DNA中的BLK基因座的所有13个外显子和2.5kb上游启动子序列重新测序, 所述DNA分离自由NIH/NIAMS资助的资源库——自身免疫病生物标记物合作网络 (ABCoN) 的192名患者 (Bauer等人, *PLoS medicine* 3 (12): e491 (2006)), 和纽约癌症项目 (NYCP) 的96名对照个体 (Mitchell等人, *J. Urban Health* 81: 301-10 (2004))。基因组DNA是在测序前根据生产商的规程扩增的全基因组 (Qiagen, Valencia, CA., 货号150045)。

[0308] 重新测序的结果显示, 在BLK基因的编码区发现了17个突变 (10个非同义的, 7个同义的) (表8)。这些突变无一在病例中表现出比对照显著更高的频率。非同义突变的整体频率在病例 (14/191) 中没有显著高于对照 (7/96)。

[0309] 此外, 在BLK的非编码区鉴别出多个常见突变 (显示在表9中)。三个SNP (rs4840568、rs1382568 [三等位的SNP (A/C/G); C等位基因之前被鉴别为风险等位基因] 和rs922483 (SEQ ID NO: 13)) 表现出与之前GWAS中鉴别的基因座 (Hom等人, *N Engl J Med* 358: 900-09 (2008)) 的 $r^2 > 0.5$ 的关联性 (rs13277113, 比值比, 1.39,  $P = 1 \times 10^{-10}$ )。图4显示了使用Haploview (软件可自URL [www.broadinstitute.org/haploview/haploview](http://www.broadinstitute.org/haploview/haploview) 免费获得; 参见Barrett J.C.等人, *Bioinformatics* 21: 263-65 (2005)) 生成的、在BLK的启动子区域中的连锁不平衡 (LD) 块 (block) (以 $r^2$ 显示)。图的上部显示了BLK启动子区的示意图, 标出了所鉴别的SNP的相对位置。所列举的SNP之间的 $r^2$ 值显示在框中。两个SNP之间的LD强度用 $r^2$ 值表示, 提供在每个框中。在图的上部用黑粗体标出了从GWAS鉴别的基因座 (rs13277113) 和重新测序鉴别的三个SNP (rs4840568、rs1382568和rs922483 (SEQ ID NO: 13))。

[0310] 该重新测序研究没有揭示BLK编码区中的任何常见变异。然而, 鉴别出启动子区的三个常见变体 (rs4840568、rs1382568和rs922483 (SEQ ID NO: 13)) 作为BLK与增加的SLE风险相关联的生物学效应的一个或多个潜在因果等位基因。将每个这类变异用于下文详述的荧光素酶报告基因测定, 进一步表征所述关联性。

[0311] 表8. BLK编码区中的突变

| 外显子     | 氨基酸<br>改变<br>(蛋白质:<br>NP_001706.2) | 核苷酸<br>改变<br>(mRNA:<br>NM_001715.2) | dbSNP      | dbSNP | 接合性 | 非同义的 | 病例<br>(n=191) | 对照<br>(n=96) |
|---------|------------------------------------|-------------------------------------|------------|-------|-----|------|---------------|--------------|
| 2       | 39P>L                              | 697_C>T                             | N/A        | N/A   | 杂合的 | 是    | 1             | 0            |
| 4       | 71A>T                              | 792_G>A                             | rs55758736 | N/A   | 杂合的 | 是    | 6             | 4            |
| 4       | 75R>R                              | 806_G>T                             | N/A        | N/A   | 杂合的 | 否    | 1             | 0            |
| 4       | 86Q>Q                              | 839_G>A                             | rs56185487 | N/A   | 杂合的 | 否    | 2             | 0            |
| 6       | 131R>W                             | 972_C>T                             | N/A        | N/A   | 杂合的 | 是    | 2             | 0            |
| 6       | 137Q>Q                             | 992_G>A                             | N/A        | N/A   | 杂合的 | 否    | 0             | 1            |
| 7       | 180R>H                             | 1120_G>A                            | N/A        | N/A   | 杂合的 | 是    | 1             | 0            |
| 7       | 190S>S                             | 1151_C>T                            | N/A        | N/A   | 纯合的 | 否    | 0             | 1            |
| 8       | 237P>P                             | 1292_C>T                            | N/A        | N/A   | 杂合的 | 否    | 1             | 1            |
| 8       | 238R>Q                             | 1294_G>A                            | N/A        | N/A   | 杂合的 | 是    | 1             | 0            |
| 10      | 325K>T                             | 1555_A>C                            | N/A        | N/A   | 杂合的 | 是    | 2             | 0            |
| 10      | 327D>V                             | 1561_A>T                            | N/A        | N/A   | 杂合的 | 是    | 0             | 1            |
| 10      | 331R>I                             | 1573_G>T                            | N/A        | N/A   | 杂合的 | 是    | 1             | 0            |
| 11      | 359R>C                             | 1656_C>T                            | N/A        | N/A   | 杂合的 | 是    | 0             | 1            |
| 12      | 425L>P                             | 1855_T>C                            | N/A        | N/A   | 杂合的 | 是    | 0             | 1            |
| 13      | 464L>L                             | 1973_G>A                            | N/A        | N/A   | 杂合的 | 否    | 1             | 0            |
| 13      | 474R>R                             | 2003_C>T                            | N/A        | N/A   | 杂合的 | 否    | 2             | 0            |
| 非同义突变频率 |                                    |                                     |            |       |     |      | 14/191        | 7/96         |

[0313] 表9.BLK非编码区中的常见变异('rs922483'公开为SEQ ID NO:13)

| 核苷酸<br>改变 | 染色体位置    | 染色体 | dbSNP      | 相关<br>等位<br>基因 | 病例、对照<br>比例计数<br>(191,96) | 病例、对照<br>频率  | 卡方     | P 值    | r <sup>2</sup> (与rs<br>13277113比) |
|-----------|----------|-----|------------|----------------|---------------------------|--------------|--------|--------|-----------------------------------|
| C>A       | 11386986 | 8   | rs2251056  | A              | 321:61, 150:42            | 0.840, 0.781 | 3.027  | 0.0819 | 0.066                             |
| C>T       | 11387925 | 8   | N/A        | T              | 8:374, 1:191              | 0.021, 0.005 | 2.05   | 0.1522 | 0.006                             |
| T>C       | 11388088 | 8   | rs2736344  | C              | 329:53, 149:43            | 0.861, 0.776 | 6.662  | 0.0098 | 0.072                             |
| C>G       | 11388131 | 8   | rs9694294  | G              | 323:59, 149:43            | 0.846, 0.776 | 4.225  | 0.0398 | 0.078                             |
| T>C       | 11388309 | 8   | rs1382567  | T              | 199:183, 93:99            | 0.521, 0.484 | 0.684  | 0.4083 | 0.335                             |
| G>A       | 11388429 | 8   | rs4840568  | A              | 125:257, 41:151           | 0.327, 0.214 | 8.033  | 0.0046 | 0.852                             |
| A>C       | 11388630 | 8   | rs1382568  | C              | 121:261, 38:154           | 0.317, 0.198 | 9.01   | 0.0027 | 0.889                             |
| A>G       | 11388631 | 8   | rs1382568  | A              | 200:182, 94:98            | 0.524, 0.490 | 0.59   | 0.4423 | 0.331                             |
| C>T       | 11389322 | 8   | rs922483   | T              | 137:245, 43:149           | 0.359, 0.224 | 10.768 | 0.001  | 0.741                             |
| A>G       | 11389466 | 8   | N/A        | G              | 320:62, 148:44            | 0.838, 0.771 | 3.794  | 0.0514 | 0.059                             |
| C>T       | 11404079 | 8   | N/A        | C              | 274:108, 128:64           | 0.717, 0.667 | 1.56   | 0.2117 | 0.037                             |
| G>C       | 11404447 | 8   | N/A        | G              | 358:24, 176:16            | 0.937, 0.917 | 0.829  | 0.3626 | 0.018                             |
| C>T       | 11404452 | 8   | N/A        | T              | 55:327, 23:169            | 0.144, 0.120 | 0.637  | 0.4249 | 0.038                             |
| T>C       | 11404502 | 8   | N/A        | C              | 91:291, 45:147            | 0.238, 0.234 | 0.01   | 0.9186 | 0.001                             |
| T>C       | 11443842 | 8   | N/A        | T              | 196:184, 80:112           | 0.516, 0.417 | 5.019  | 0.0251 | 0.083                             |
| T>C       | 11451476 | 8   | N/A        | C              | 315:65, 150:38            | 0.829, 0.798 | 0.818  | 0.3657 | 0.016                             |
| A>G       | 11452981 | 8   | rs4841557  | A              | 158:222, 63:129           | 0.416, 0.328 | 4.135  | 0.042  | 0.187                             |
| C>T       | 11453006 | 8   | rs4841558  | C              | 157:223, 63:129           | 0.413, 0.328 | 3.897  | 0.0484 | 0.181                             |
| G>A       | 11455795 | 8   | rs1042695  | A              | 145:237, 59:133           | 0.380, 0.307 | 2.915  | 0.0878 | 0.179                             |
| G>A       | 11456175 | 8   | N/A        | A              | 335:45, 157:35            | 0.882, 0.818 | 4.325  | 0.0375 | 0.019                             |
| G>A       | 11456176 | 8   | N/A        | G              | 340:42, 157:35            | 0.890, 0.818 | 5.758  | 0.0164 | 0.015                             |
| C>T       | 11456182 | 8   | rs4841561  | T              | 146:234, 59:133           | 0.384, 0.307 | 3.282  | 0.07   | 0.177                             |
| C>T       | 11458793 | 8   | rs10097015 | T              | 160:222, 68:122           | 0.419, 0.358 | 1.967  | 0.1608 | 0.161                             |
| C>T       | 11459203 | 8   | rs1042689  | T              | 144:238, 62:130           | 0.377, 0.323 | 1.622  | 0.2028 | 0.161                             |
| G>A       | 11459455 | 8   | N/A        | G              | 213:169, 98:94            | 0.558, 0.510 | 1.145  | 0.2845 | 0.2                               |
| T>C       | 11459540 | 8   | N/A        | T              | 212:170, 100:92           | 0.555, 0.521 | 0.6    | 0.4385 | 0.219                             |

[0315] 实施荧光素酶报告基因测定,研究三个SNP——rs4840568、rs1382568和rs922483 (SEQ ID NO:13) 对BLK介导的基因表达的作用。使用携带风险或无风险的单元型的个体的基因组DNA扩增BLK的上游序列(-2256至+55bp)。将每个PCR产物克隆到pCR2.1-TOPO载体 (Invitrogen, Carlsbad, CA; 货号K4500-01) 中,然后亚克隆到pGL4荧光素酶报告基因载体 (Promega, Madison, WI; 货号E6651) 中。使用携带无风险单元型的构建体作为模板用于诱变 (Stratagene, La Jolla, CA; 货号10519-5), 产生各种单元型。

[0316] 用于PCR扩增的引物如下:正向:CCACCTCTCTCCGCCTTCTCAT (SEQ ID NO.:1);反向:TTTCATGGCTTGTGGCTTTCTGCC (SEQ ID NO.:2)。用于诱变的引物列举在下表10中。

[0317] 表10. 诱变引物列表

| 引物                                 | 序列                                                      |
|------------------------------------|---------------------------------------------------------|
| rs4840568_G>A 正向                   | GATCCAAGACTATGAAGAGAGAAGAGAGAGCCCCAC<br>(SEQ ID NO.: 3) |
| rs4840568_G>A 反向                   | GTGGGCTCTCTCTTCTCTCTTCATAGTCTTGGATC<br>(SEQ ID NO.: 4)  |
| rs1382568_A>C 正向                   | CCAGACACCACTCACCCCTCTAGATGTTGGGAT<br>(SEQ ID NO.: 5)    |
| rs1382568_A>C 反向                   | ATCCCAACATCTAGAGGGGTGAGTGGTGTCTGG<br>(SEQ ID NO.: 6)    |
| rs1382568_A>G 正向                   | CCAGACACCACTCACCGCTCTAGATGTTGGGAT<br>(SEQ ID NO.: 7)    |
| rs1382568_A>G 反向                   | ATCCCAACATCTAGAGCGGTGAGTGGTGTCTGG<br>(SEQ ID NO.: 8)    |
| rs1382568_G>A 正向                   | CCAGACACCACTCACCACTCTAGATGTTGGGAT<br>(SEQ ID NO.: 9)    |
| rs1382568_G>A 反向                   | ATCCCAACATCTAGAGTGGTGAGTGGTGTCTGG<br>(SEQ ID NO.: 10)   |
| rs922483_C>T (SEQ ID NO: 13)<br>正向 | CGGGGGTGCTGCTACCTCTGTCTGC<br>(SEQ ID NO.: 11)           |
| rs922483_C>T (SEQ ID NO: 13)<br>反向 | GCAGACAGAGGTAGCAGCACCCCCG<br>(SEQ ID NO.: 12)           |

[0319] 使用Renilla荧光素酶对照报告子载体pRL-TK (Promega, Madison, WI; 货号E2241) 进行归一化。将细胞系BJAB (具有B细胞特征的持续淋巴样细胞系 (骨髓来源的), 缺少 Epstein-Barr病毒基因组, 来自三个人的淋巴瘤; Klein等人, Proc. Natl. Acad. Sci. USA 71:3283-86 (1974)) 或Daudi细胞系 (美国典型培养物收藏中心 (ATCC) 货号CCL-213) 用于转染。对于每次转染, 使用**Amaxa® Nucleofector®**装置 (Lonza, Walkersville Inc., Walkersville, MD (Lonza Group Ltd., Switzerland); 货号AAD-1001), 用5μg的每种载体DNA转染 $5 \times 10^6$ 细胞。Daudi细胞使用细胞系**Nucleofector®**试剂盒L (Lonza, 货号VCA-1005) 和**Nucleofector®**装置程序A-030。BJAB细胞使用细胞系**Nucleofector®**试剂盒V (Lonza, 货号VCA-1005) 和**Nucleofector®**装置程序T-020。所有的转染都重复两次或重复三次的进行。转染后, 细胞在37℃孵育16小时。在孵育后, 收获细胞, 根据生产商的说明使用**Dual-Luciferase®**报告子测定体系 (Promega, Madison, WI; 货号E1960) 测量荧光素酶活性。

[0320] 在上述荧光素酶报告子测定体系中测量的每种SNP: rs4840568、rs1382568和rs922483 (SEQ ID NO: 13) 对BLK介导的基因表达的作用显示在图5中。将诱变产生的不同单元型与无风险的 (野生型) 单元型22-GAC (图5A-F各个中的空心条) 和风险的单元型22-ACT (图5A-F各个中的阴影条) 比较。

[0321] 图5A和5B显示了SNP rs922483 (C>T) (SEQ ID NO: 13) 在BJAB (图5A) 和Daudi细胞 (图5B) 中都导致对BLK介导的基因表达的显著作用。与无风险的单元型22-GAC (空心条) 相比, 单元型22-ACT在两种细胞系中都表现出转录活性降低几乎50%。含T等位基因的单元型持续表现出比含C等位基因的单元型更低的活性。在BJAB细胞中进行了5次独立的实验, 在Daudi细胞中进行了6次独立的实验。显示的数据表示重复三次的测定中的平均值 $\pm$ 平均值的标准误差 (s.e.m.); \* $p < 0.05$ 、\*\* $p < 0.01$ 、\*\*\* $p < 0.001$  (t检验)。

[0322] 图5C和5D显示了SNP rs1382568 (A>C/G>C) 在每种细胞系中都没有导致对BLK介

导的表达的任何显著作用。与无风险的单元型22-GAC(空心条)相比,两种单元型22-GCC和22-GGC(有点的条)都表现出相似水平的荧光素酶活性。在BJAB细胞中进行了5次独立的实验,在Daudi细胞中进行了6次独立的实验。显示的数据表示重复三次的测定中的平均值 $\pm$ s.e.m.; \* $p < 0.05$ 、\*\* $p < 0.01$ 、\*\*\* $p < 0.001$ 、ns=不显著(t检验)。

[0323] 图5E和5F显示了SNP rs4840568 (G>A) 在BJAB细胞或Daudi细胞中都没有导致对BLK介导的表达的显著作用。在BJAB细胞中,单元型22-AAC(有点的条)与无风险的单元型22-GAC(空心条)之间的差异没有统计学显著性(图5E),但在Daudi细胞中是统计学显著的(图5F)。考虑到单元型22-ACC(有点的条)相比无风险的单元型22-GAC(空心条)没有表现出任何荧光素酶活性的缺陷,A等位基因成为因果等位基因的可能性极大的降低(图5F)。显示的数据表示重复三次的测定中的平均值 $\pm$ s.e.m.; \* $p < 0.05$ 、\*\* $p < 0.01$ 、\*\*\* $p < 0.001$ 、ns=不显著(t检验)。

[0324] 之前显示了在BLK启动子上游区域中的(GT)重复可以作为BLK基因表达的增强子发挥作用(Lin等人,J Biol Chem 270:25968(1995))。因此,我们还测试了(GT)重复的长度是否可以影响BLK启动子的转录活性。为了实施这些实验,使用上述对策,选择来自同时携带了18(GT)重复(SEQ ID NO:14)或22(GT)重复(SEQ ID NO:15)的个体的基因组DNA样品进行克隆。测序五个载体,验证它们含有正确长度的(GT)重复。如图6所示,在荧光素酶报告子测定中,含18(GT)重复(SEQ ID NO:14)的单元型表现出与含22(GT)重复(SEQ ID NO:15)的单元型相似的转录活性水平。显示的数据表示重复两次的测定中的平均值 $\pm$ s.e.m.; ns=不显著(t检验)。

[0325] 总而言之,BLK重新测序工作的这些结果和荧光素酶报告基因测定的结果表示SNP rs922483 (C>T) (SEQ ID NO:13) 是导致减少的BLK转录的因果等位基因,所述BLK转录减少的生物学效应是与增加的SLE风险相关的。此外,结果显示,rs922483的T等位基因(SEQ ID NO:13)使BLK介导的基因表达水平降低了50%。

[0326] 令人感兴趣的是,注意到rs922483(SEQ ID NO:13)位于BLK的第一外显子的进化保守区域中,在可能的人转录起始位点中。人Inr基序的共同序列鉴别为YYANWYY(IUPAC核苷酸代码)。Juven-Gershon等人,Dev.Biol.339:225-229(2010)。在SNP rs922483(SEQ ID NO:13)中,Inr区中的第二个碱基相对于共同基序被改变。因此,SLE风险单元型Inr序列是CTACCTC,而“野生型”单元型Inr序列是CCACCTC。提示修饰保守的Inr基序中的第二个碱基可能改变了TFIID转录复合物的亲和力,导致所观察到的上述转录中的差异。

## 序列表

<110> GENENTECH, INC. et al.

<120> 用于治疗、诊断和监控狼疮的方法

<130> P4325R1-W0

<140>

<141>

<150> 61/278,510

<151> 2009-10-07

<160> 15

<170> PatentIn版本3.5

<210> 1

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成引物

<400> 1

ccacctctct tccgcctttc tcat

24

<210> 2

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成引物

<400> 2

tttcatggct tgtggcctttc tgcc

24

<210> 3

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成引物

<400> 3

gatccaagac tatgaagaga gaagagagag cccac

35

<210> 4

<211> 35

<212> DNA

<213> 人工序列

|                                        |    |
|----------------------------------------|----|
| <220>                                  |    |
| <223> 人工序列的描述:合成引物                     |    |
| <400> 4                                |    |
| gtgggctctc tcttctctct tcatagtctt ggate | 35 |
| <210> 5                                |    |
| <211> 33                               |    |
| <212> DNA                              |    |
| <213> 人工序列                             |    |
| <220>                                  |    |
| <223> 人工序列的描述:合成引物                     |    |
| <400> 5                                |    |
| ccagacacca ctcacccctc tagatgttgg gat   | 33 |
| <210> 6                                |    |
| <211> 33                               |    |
| <212> DNA                              |    |
| <213> 人工序列                             |    |
| <220>                                  |    |
| <223> 人工序列的描述:合成引物                     |    |
| <400> 6                                |    |
| atcccaacat ctagaggggt gagtgggtgc tgg   | 33 |
| <210> 7                                |    |
| <211> 33                               |    |
| <212> DNA                              |    |
| <213> 人工序列                             |    |
| <220>                                  |    |
| <223> 人工序列的描述:合成引物                     |    |
| <400> 7                                |    |
| ccagacacca ctcaccgctc tagatgttgg gat   | 33 |
| <210> 8                                |    |
| <211> 33                               |    |
| <212> DNA                              |    |
| <213> 人工序列                             |    |
| <220>                                  |    |
| <223> 人工序列的描述:合成引物                     |    |
| <400> 8                                |    |
| atcccaacat ctagagcggg gagtgggtgc tgg   | 33 |
| <210> 9                                |    |
| <211> 33                               |    |
| <212> DNA                              |    |

|                                                           |    |
|-----------------------------------------------------------|----|
| <213> 人工序列                                                |    |
| <220>                                                     |    |
| <223> 人工序列的描述:合成引物                                        |    |
| <400> 9                                                   |    |
| ccagacacca ctcaccactc tagatgttgg gat                      | 33 |
| <210> 10                                                  |    |
| <211> 33                                                  |    |
| <212> DNA                                                 |    |
| <213> 人工序列                                                |    |
| <220>                                                     |    |
| <223> 人工序列的描述:合成引物                                        |    |
| <400> 10                                                  |    |
| atcccaacat ctagagtggg gagtggtgtc tgg                      | 33 |
| <210> 11                                                  |    |
| <211> 25                                                  |    |
| <212> DNA                                                 |    |
| <213> 人工序列                                                |    |
| <220>                                                     |    |
| <223> 人工序列的描述:合成引物                                        |    |
| <400> 11                                                  |    |
| cgggggtgct gctacctctg tctgc                               | 25 |
| <210> 12                                                  |    |
| <211> 25                                                  |    |
| <212> DNA                                                 |    |
| <213> 人工序列                                                |    |
| <220>                                                     |    |
| <223> 人工序列的描述:合成引物                                        |    |
| <400> 12                                                  |    |
| gcagacagag gtagcagcac ccccg                               | 25 |
| <210> 13                                                  |    |
| <211> 52                                                  |    |
| <212> DNA                                                 |    |
| <213> 智人                                                  |    |
| <400> 13                                                  |    |
| ctctgatcgc agaccggggg tgctgcyacc tctgtctgct gccggcagaa ag | 52 |
| <210> 14                                                  |    |
| <211> 36                                                  |    |
| <212> DNA                                                 |    |
| <213> 人工序列                                                |    |

<220>

<223> 人工序列的描述:合成寡核苷酸

<400> 14

gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgt

36

<210> 15

<211> 44

<212> DNA

<213> 人工序列

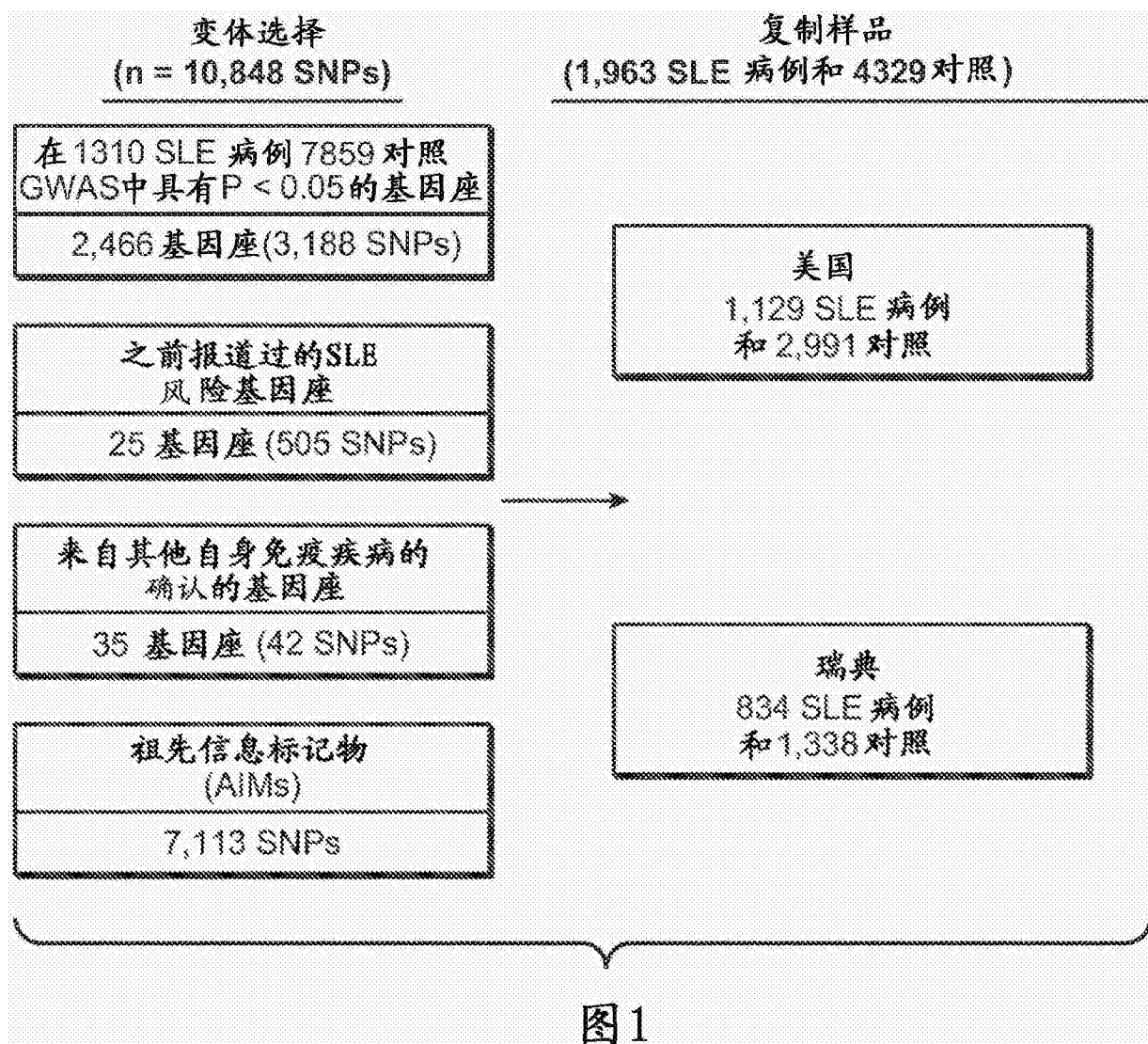
<220>

<223> 人工序列的描述:合成寡核苷酸

<400> 15

gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgt

44



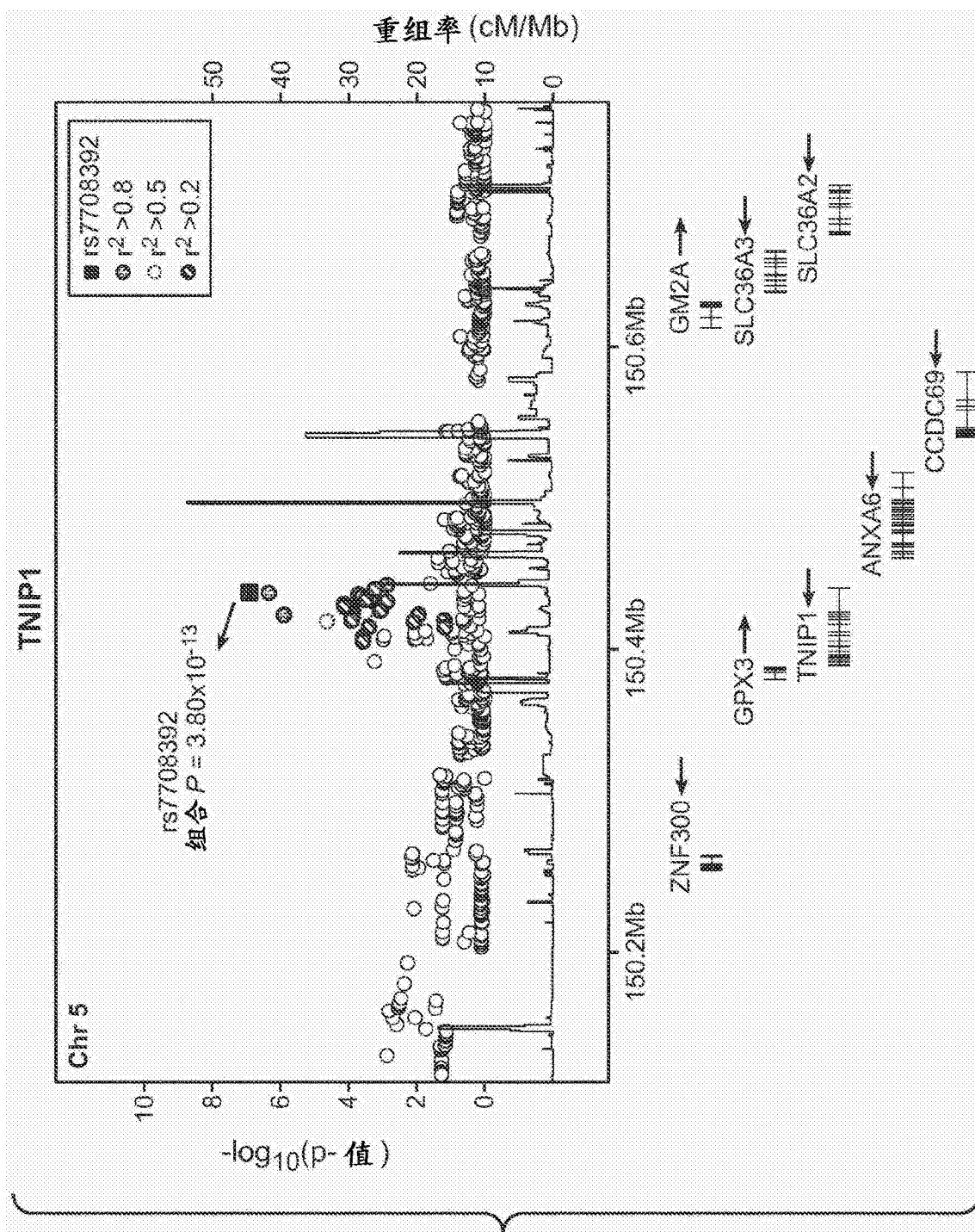


图2A

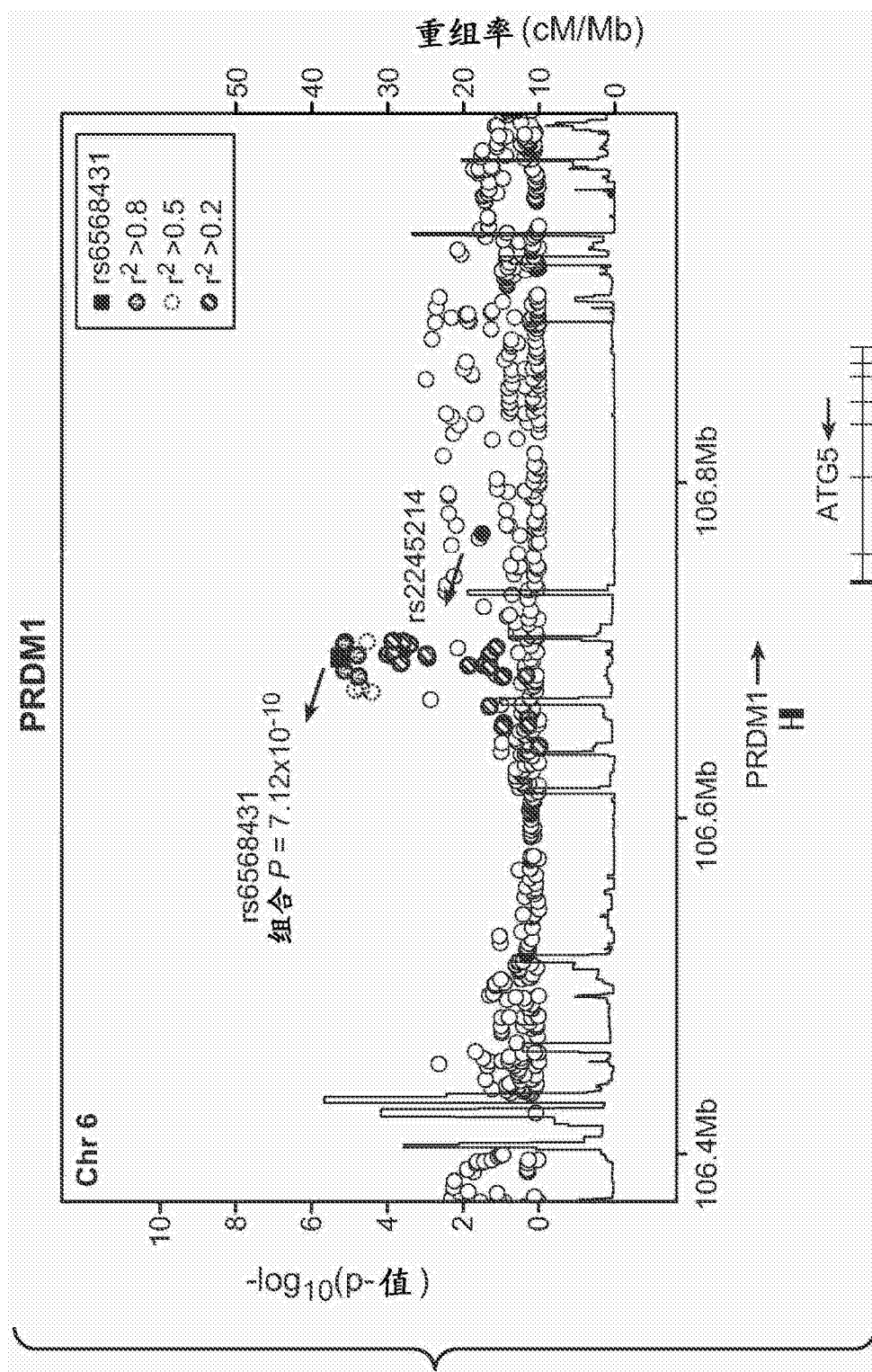


图2B

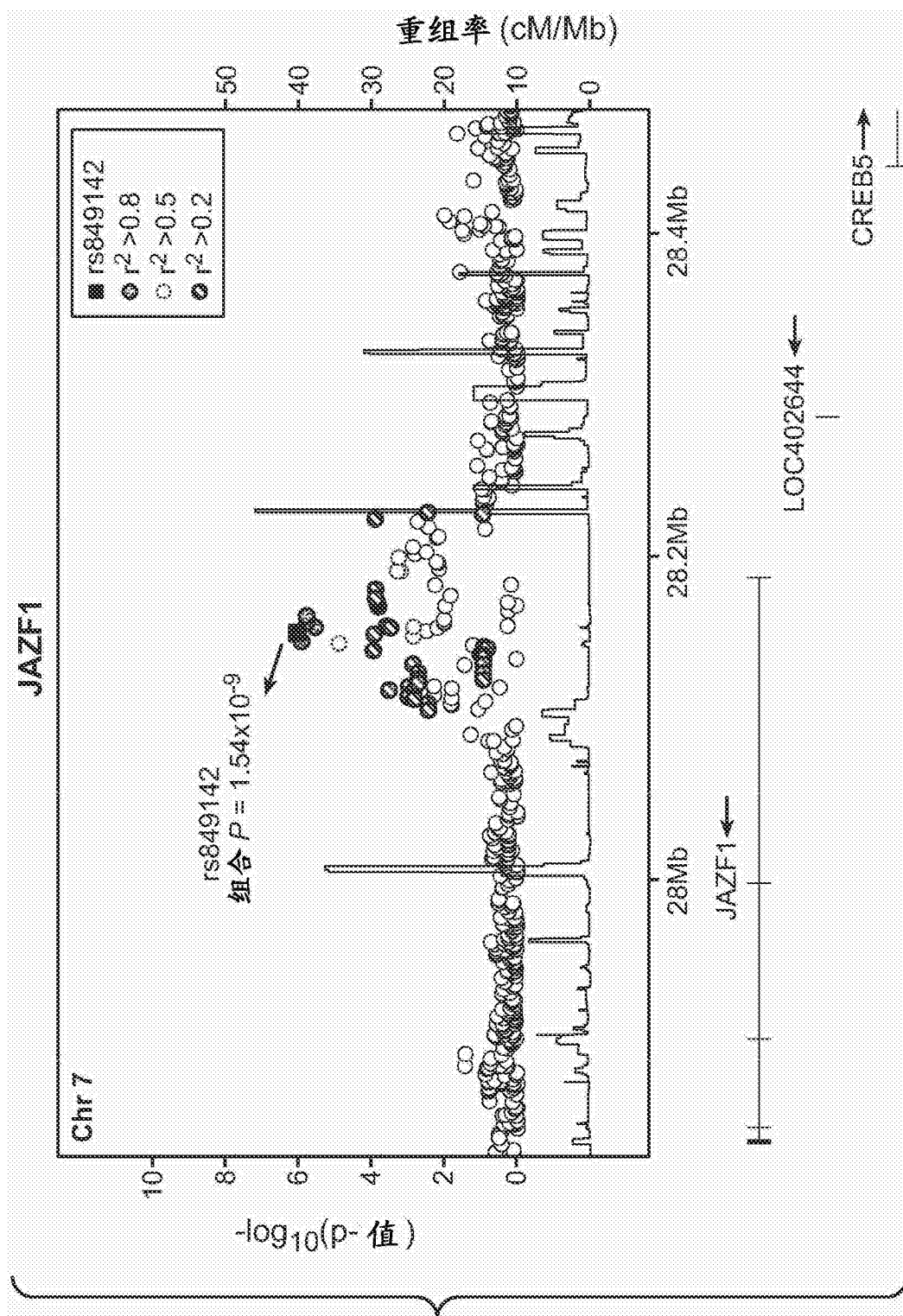


图2C

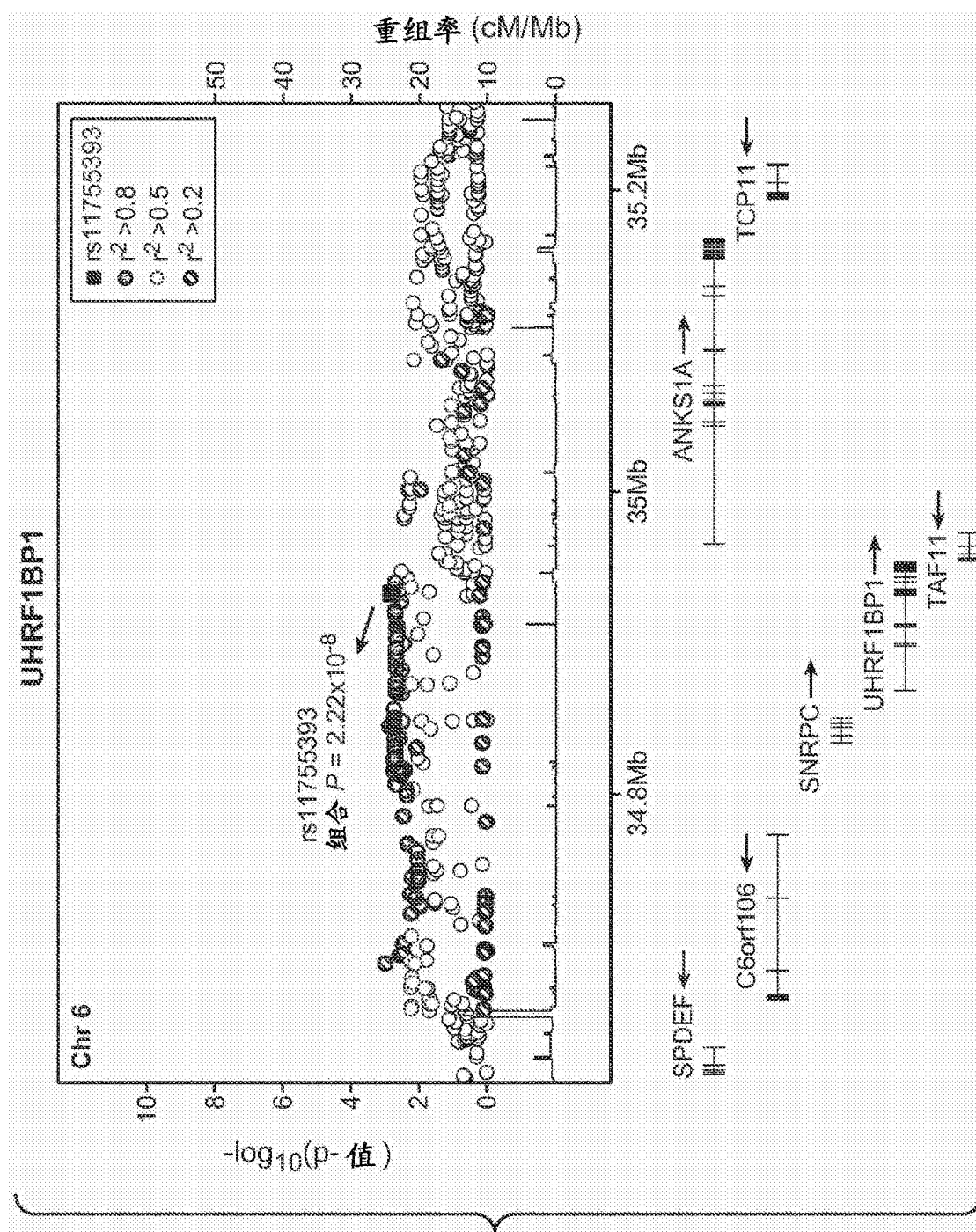


图2D

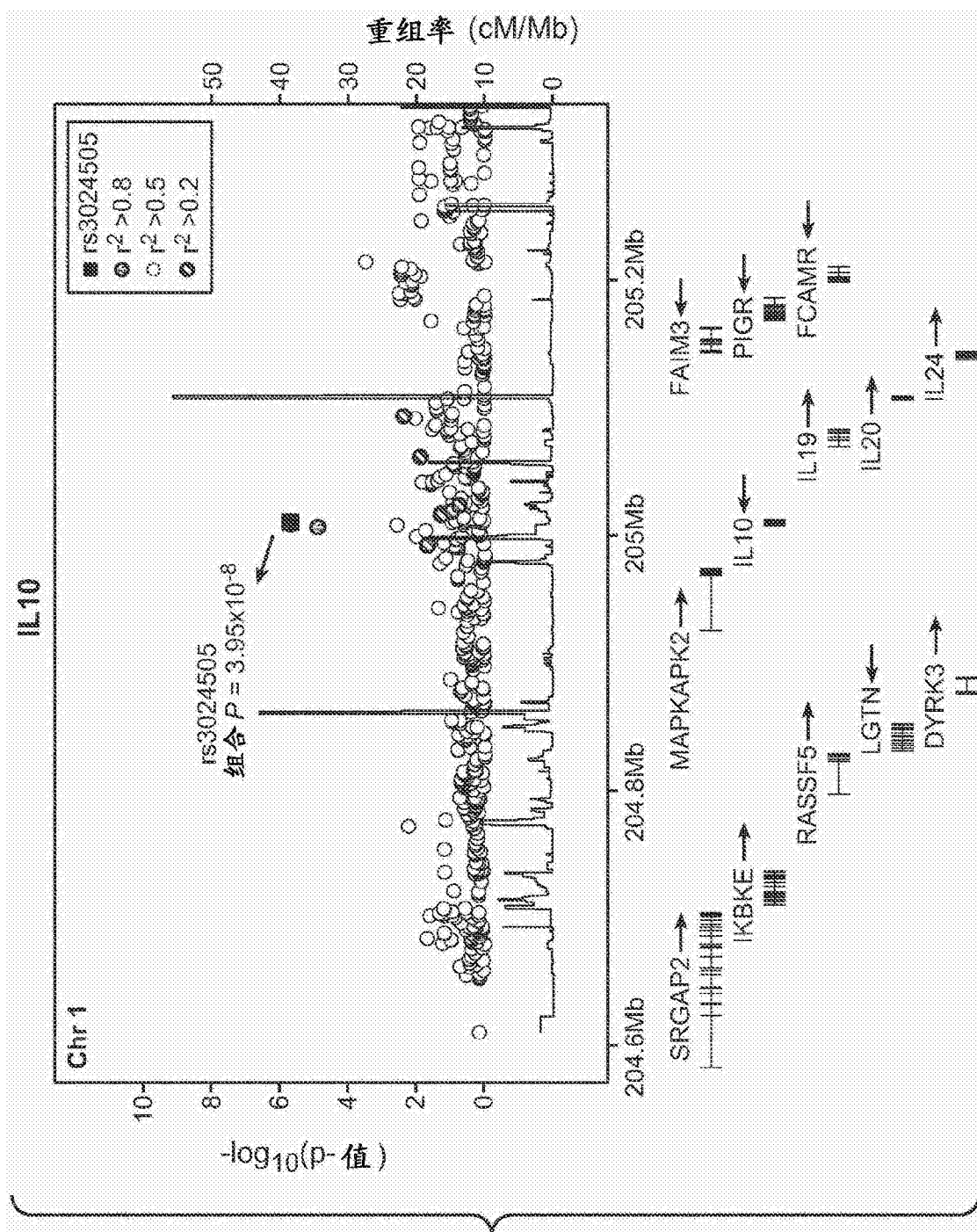


图2E

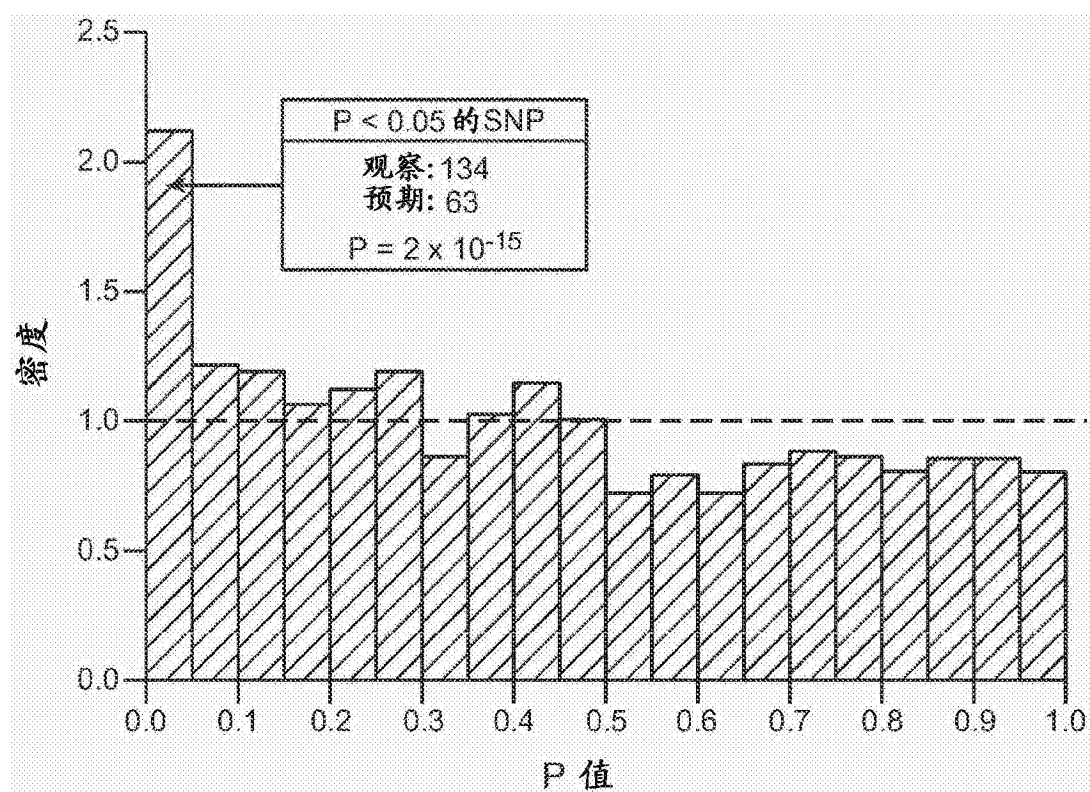


图2F

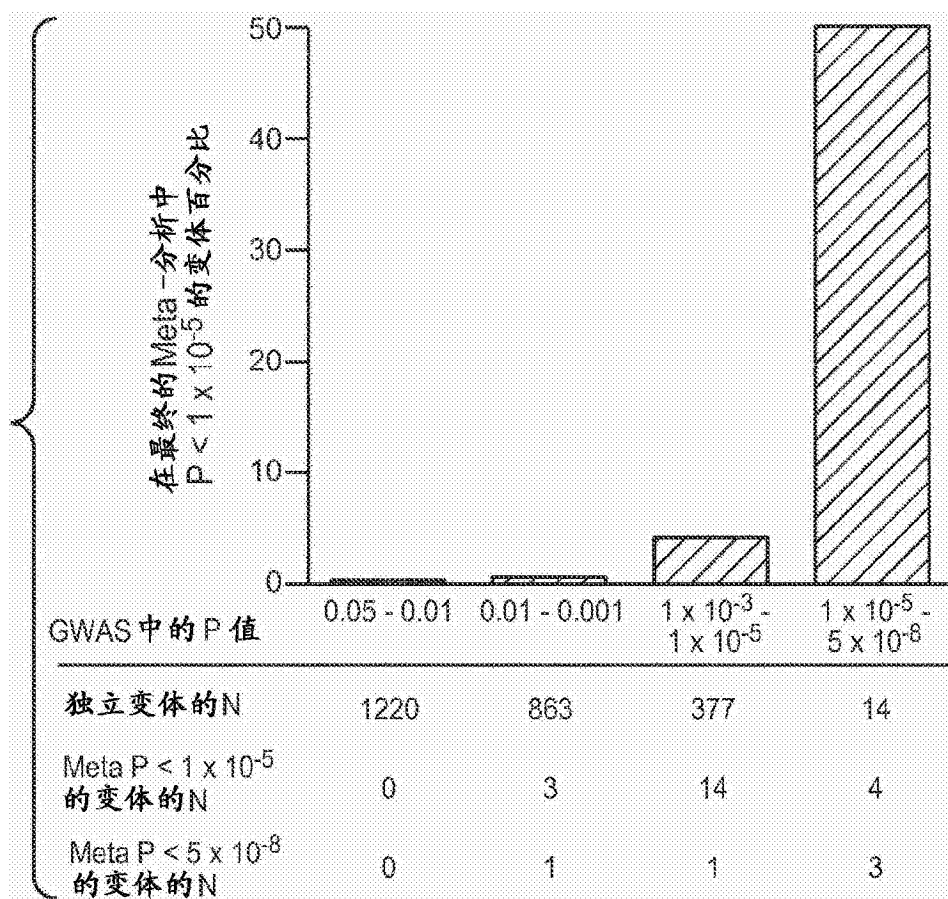


图3

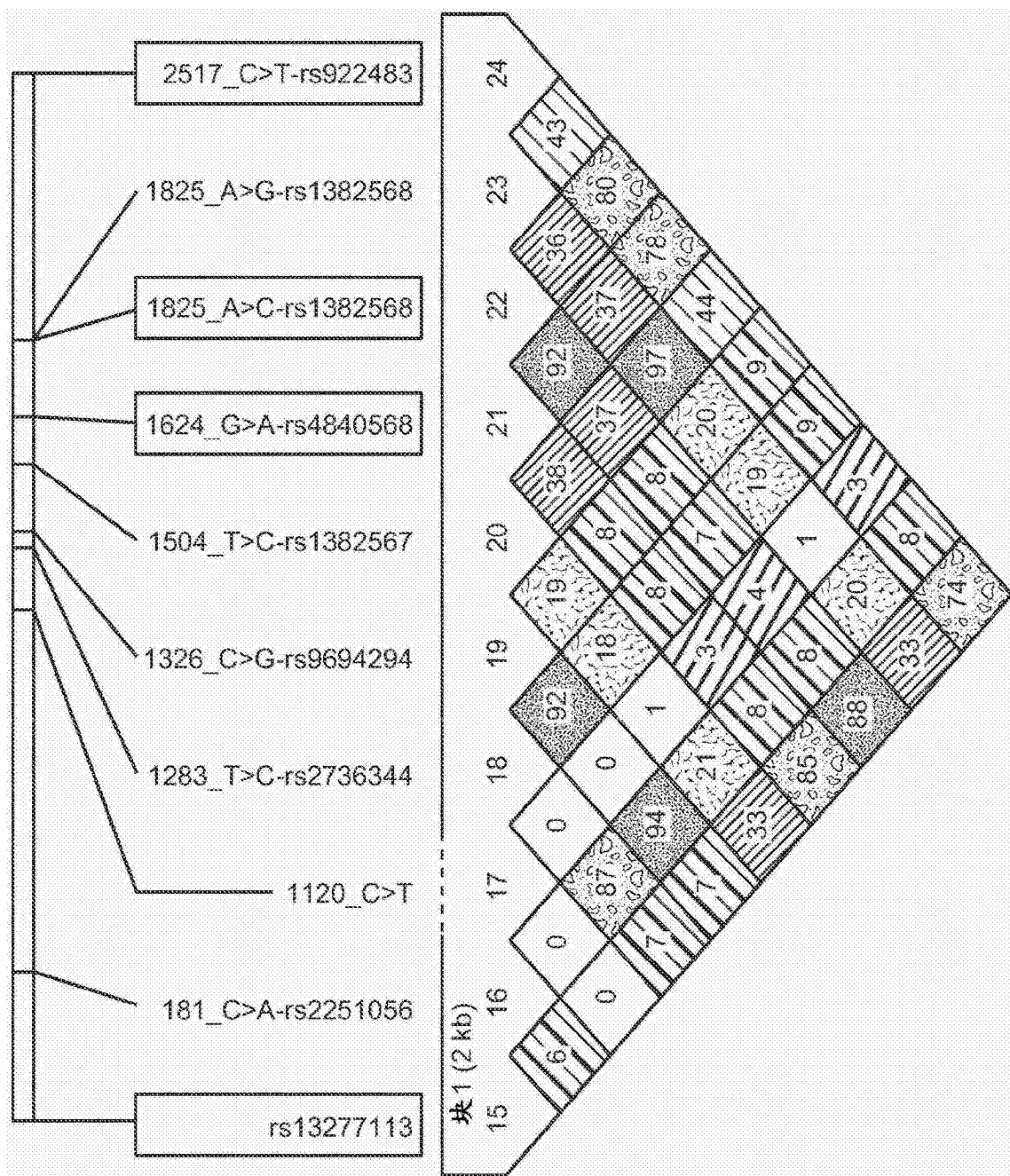


图4

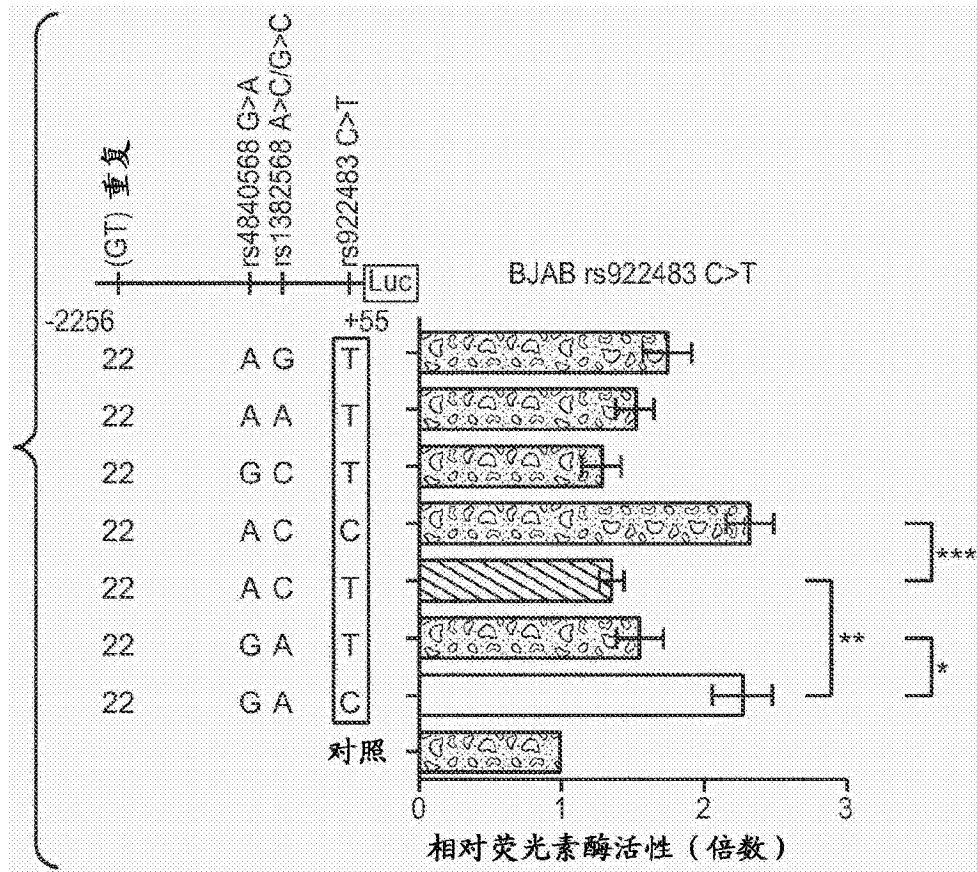


图5A

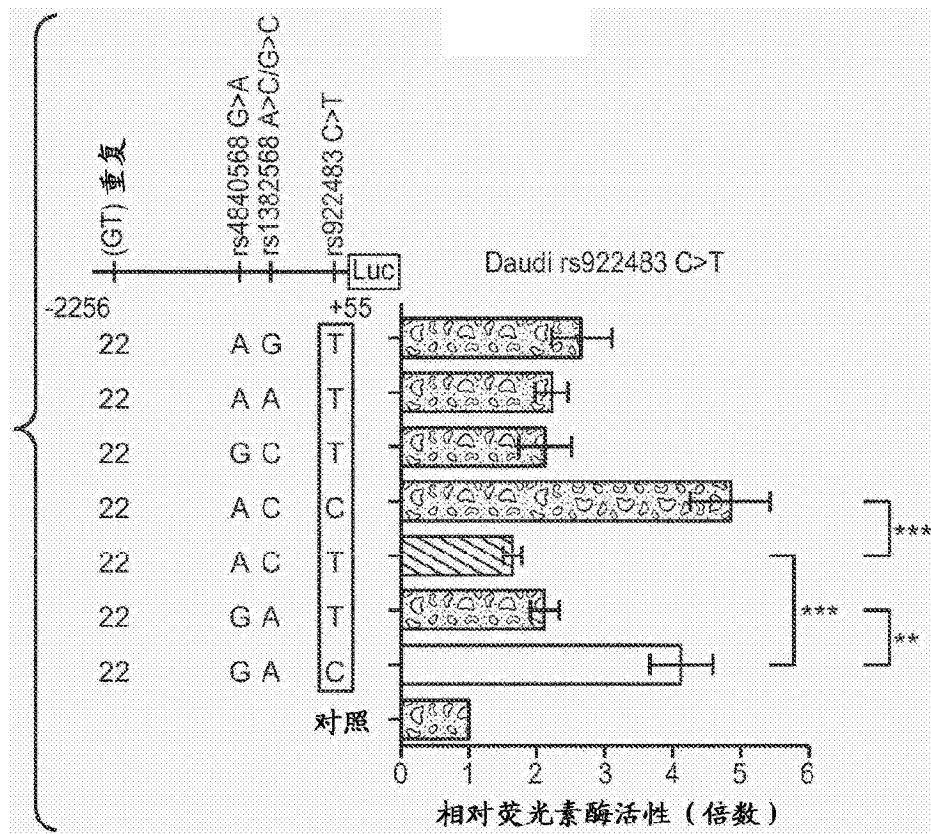


图5B

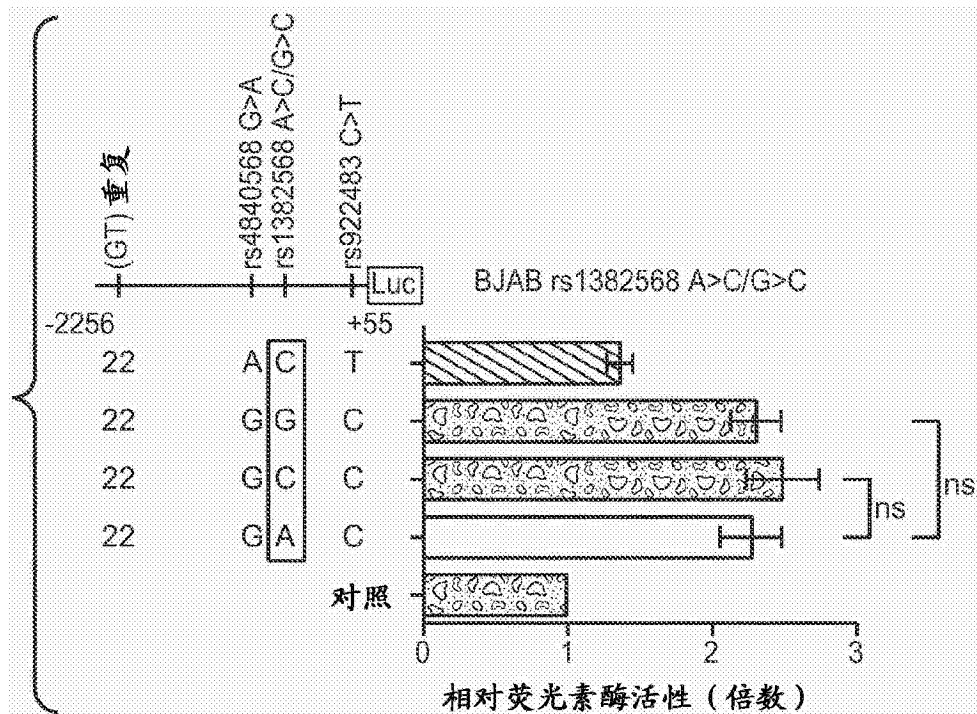


图5C

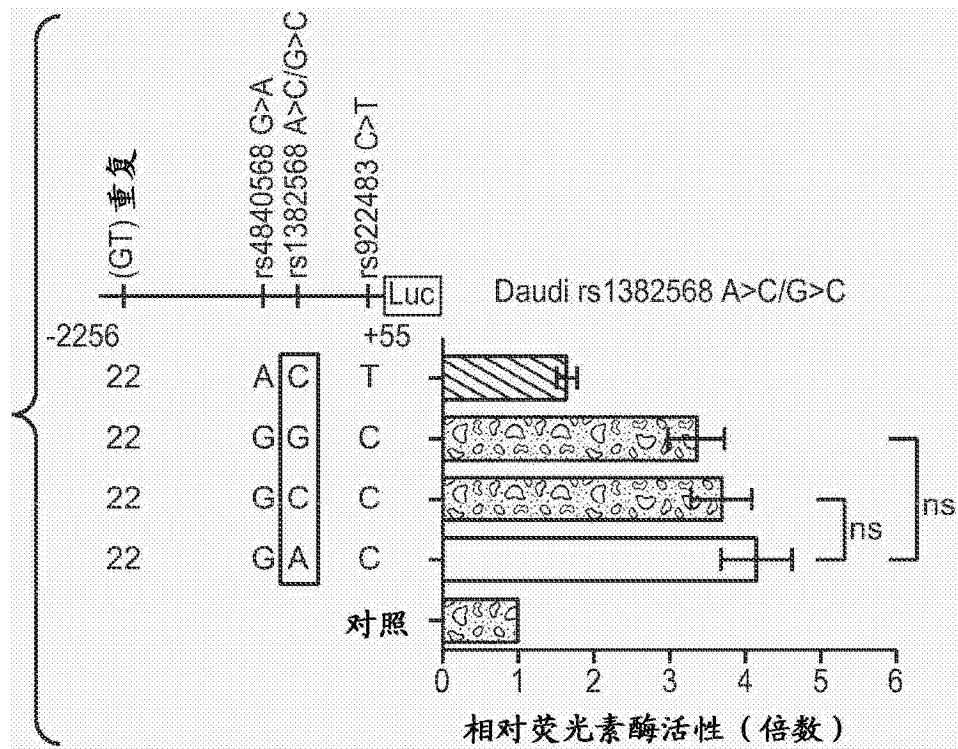


图5D

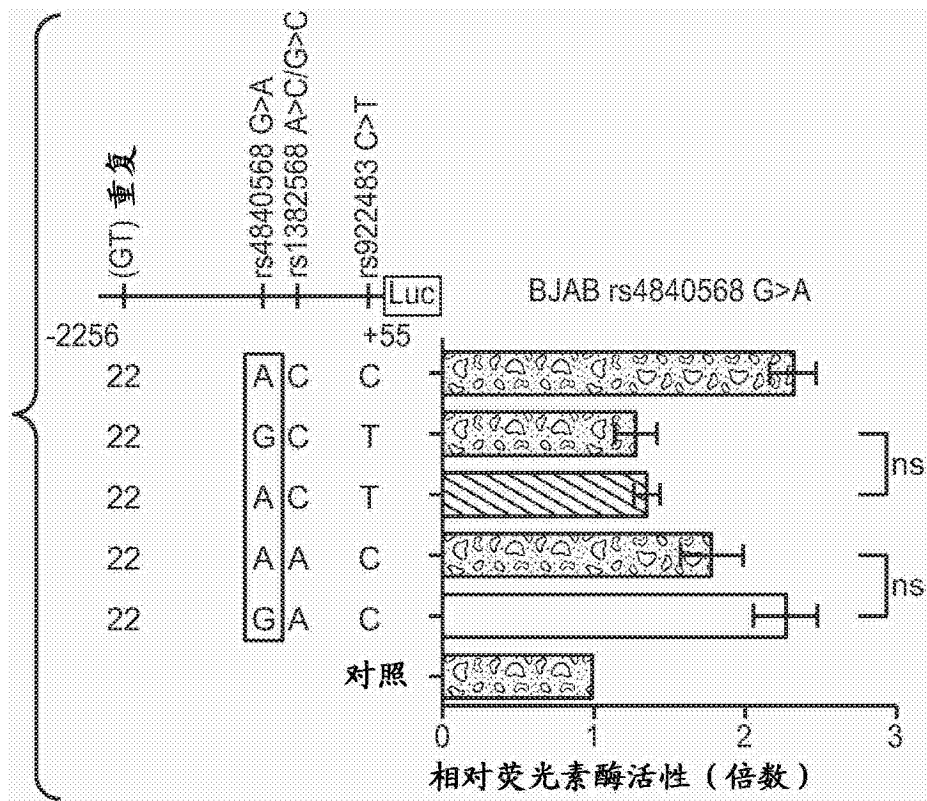


图5E

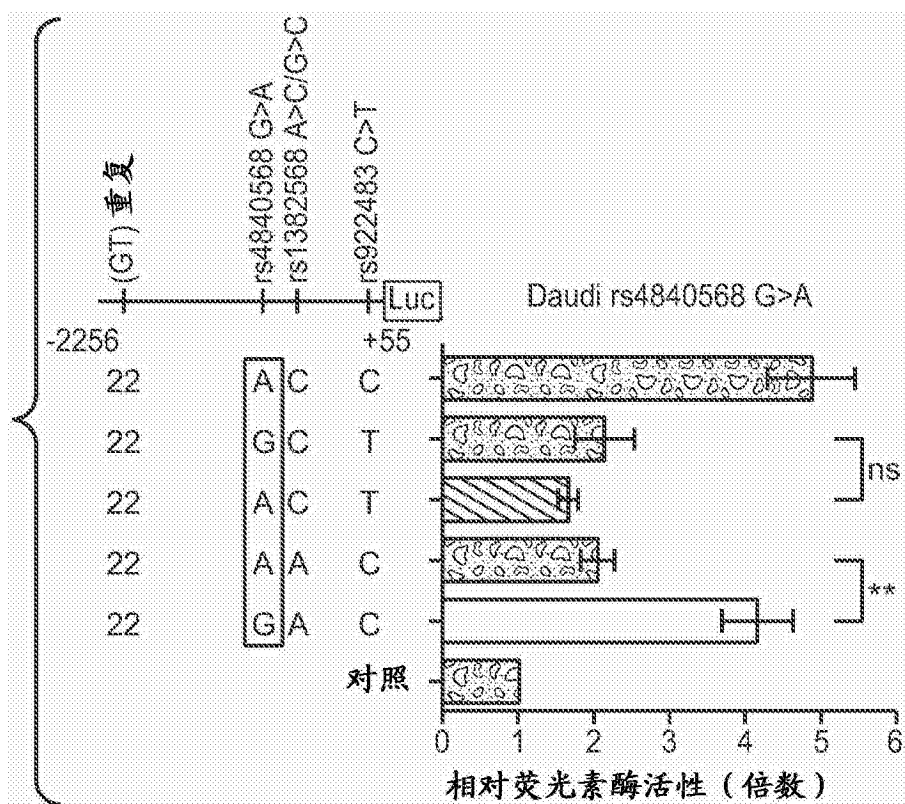


图5F

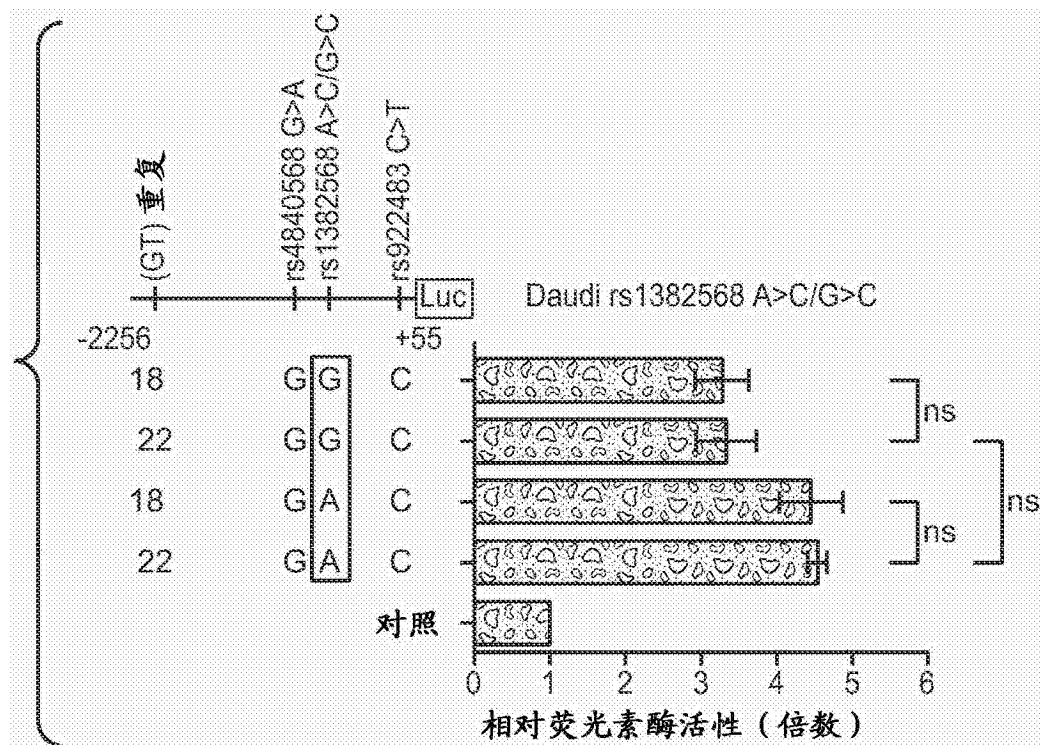


图6

**rs922483:**

CTCTGATCGCAGACCGGGGGTGCTGC[C/T]ACCTCTGTCTGCT  
GCCGGCAGAAAG

(SEQ ID No.13)

图7