



(12) 发明专利申请

(10) 申请公布号 CN 101885758 A

(43) 申请公布日 2010. 11. 17

(21) 申请号 200910136343. X

G01N 33/86 (2006. 01)

(22) 申请日 2009. 05. 11

C12Q 1/68 (2006. 01)

G01N 33/574 (2006. 01)

(71) 申请人 中国医学科学院肿瘤医院

地址 100021 北京市朝阳区潘家园南里 17 号

(72) 发明人 赵晓航 何平 刘飞 孙玉琳
刘芳 周兰萍 谢鹏 张金强
安怀杰 李潜

(74) 专利代理机构 中国国际贸易促进委员会专
利商标事务所 11038

代理人 罗菊华

(51) Int. Cl.

C07K 7/08 (2006. 01)

C12N 15/12 (2006. 01)

C12N 15/63 (2006. 01)

C12N 1/19 (2006. 01)

C12N 1/21 (2006. 01)

C12N 5/10 (2006. 01)

C12P 21/02 (2006. 01)

G01N 33/53 (2006. 01)

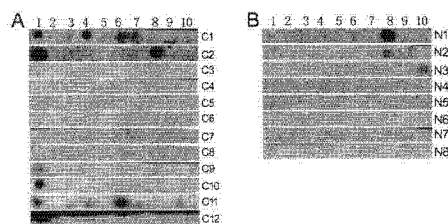
权利要求书 1 页 说明书 13 页 附图 1 页

(54) 发明名称

一种自身抗原表位、其编码基因及其用途

(57) 摘要

本发明涉及免疫化学、蛋白质组学和肿瘤免疫领域。具体地涉及一种自身抗原表位、其编码基因及其用途。更具体地, 涉及人组织蛋白酶 D (cathepsin D) 抗原表位区及其编码基因, 以及这种多肽在鉴定胰腺癌患者血清中自身抗体中的用途。



1. 抗原表位,其选自下述序列之一:
 - (a) SEQ ID NO:1 或 11 所示的氨基酸序列;
 - (b) 具有作为组织蛋白酶 D 的抗原表位的功能的、且由与 SEQ ID NO:1 或者 11 所示的氨基酸序列具有 95% 以上同源性的氨基酸序列构成的氨基酸序列;和
 - (c) 具有作为组织蛋白酶 D 的抗原表位的功能的、且由经过取代、缺失或添加一个或几个氨基酸的 SEQ ID NO:1 或 11 所示的氨基酸序列构成的氨基酸序列。
2. 编码权利要求 1 的抗原表位的基因,其选自下述序列之一:
 - (a) SEQ ID NO:12 所示的编码 SEQ ID NO:1 的核苷酸序列;
 - (b) SEQ ID NO:13 所示的编码 SEQ ID NO:11 的核苷酸序列;
 - (c) 编码下述氨基酸序列的核苷酸序列:所述氨基酸序列是具有作为组织蛋白酶 D 的抗原表位的功能的、且由与 SEQ ID NO:1 所示的氨基酸序列具有 95% 以上同源性的氨基酸序列构成的氨基酸序列;
 - (d) 编码下述氨基酸序列的核苷酸序列:所述的氨基酸序列是具有作为组织蛋白酶 D 的抗原表位的功能、且由经过取代、缺失或添加一个或几个氨基酸的 SEQ ID NO:1 所示的氨基酸序列构成的氨基酸序列;和
 - (d) 因为密码子简并性而分别与 SEQ ID NO:12 和 13 不同、但是编码 SEQ ID NO:1 或 11 所示氨基酸序列的核苷酸序列。
3. 含有权利要求 2 的基因的重组载体。
4. 用权利要求 4 的重组载体转化、转导或转染得到的遗传工程化宿主细胞。
5. 权利要求 1 的抗原表位或权利要求 2 的基因的用途,其用于制备下述制剂之一:
 - (a) 组织蛋白酶 D 的抗原表位;
 - (b) 组织蛋白酶 D 的胰腺癌特异的抗原表位;
 - (c) 检测人组织蛋白酶 D 的抗体;和
 - (d) 疫苗。
6. 权利要求 5 的应用,其特征在于权利要求 1 所述的抗原表位或权利要求 2 的基因在制备用于检测胰腺癌血清中自身抗体的制剂中的用途。
7. 权利要求 5 的应用,其特征在于权利要求 1 所述的抗原表位或权利要求 2 的基因在制备用于检测胰腺癌血清中自身抗体的试剂盒中的用途。
8. 胰腺癌检测方法,其包括使用权利要求 1 的抗原表位或权利要求 2 的基因。

一种自身抗原表位、其编码基因及其用途

技术领域

[0001] 本发明涉及免疫化学、蛋白质组学和肿瘤免疫学领域。具体地涉及一种自身抗原表位、其编码基因及其用途。更具体地,涉及人组织蛋白酶 D(cathepsin D) 抗原表位区及其编码基因,以及这种多肽在鉴定胰腺癌患者血清中自身抗体中的用途。

背景技术

[0002] 自身抗体是指抗自身细胞内、细胞表面和细胞外抗原的免疫球蛋白,自身抗体可以是生理性的,也可以是病理性的。生理性自身抗体普遍存在,主要用于清除衰老或死亡的细胞。当机体免疫功能失调,自身抗体的滴度超出一定阈值时,就成为病理性的了,例如,自身免疫疾病就是一种涉及自身抗体的慢性疾病。对于临床研究和诊断而言,结合病人的病史和体检,自身抗体阳性将有诊断自身免疫病的意义。(滕庆,自身免疫性疾病自身抗体检查方法及其评价,中国实用儿科杂志,2003,(02))。

[0003] 最近几年研究人员发现血清中的自身抗体不是自身免疫病的唯一特征,自身抗体也可以出现在免疫系统发生异常改变的其他疾病,如肿瘤 (Belousov, P. V. 等人, (2008) Biochemistry (Mosc) 73, 562-72) 中。目前已经在多种人类肿瘤中检测到自身抗体,包括肺癌、大肠癌、乳腺癌、前列腺癌、肝癌、卵巢癌、胰腺癌和副肿瘤性神经综合征等。

[0004] 自身抗体识别的抗原物质即为自身抗原,自身抗原是自身抗体产生的前提,自身抗原可以是完全正常的组织成分,也可以是已改变了的细胞成分,比如肿瘤发生的早期阶段,肿瘤细胞内或者血液循环中就会出现自身抗原,尤其是在肿瘤发生的早期阶段,肿瘤细胞内或者血液循环中会产生肿瘤相关抗原,但是其难以被常规的方法检测到,但是机体针对这些自身抗原所产生的自身抗体作为免疫系统对肿瘤存在信号的生物放大而长期存留在肿瘤细胞内或者血液循环中 (Purcell, A 等人, (2004) Mol Cell Proteomics 3, 193-208),这种自身抗体可以识别那些难以直接检测到的仅发生了微弱变化的肿瘤相关抗原 (Hanash, S. (2003) Nat Biotechnol 21, 37-8)。

[0005] 另外一些抗原,例如细胞周期蛋白,不但自身在肿瘤中高表达,并且可以诱导机体产生针对细胞周期蛋白的自身抗体,这些自身抗体对于肿瘤的诊断具有较好的灵敏度和特异度 (Covini, G 等人, (1997) Hepatology 25, 75-80)。因此,基于体液免疫筛选和鉴定的自身抗体和肿瘤抗原,将有利于挖掘肿瘤早期筛查和诊断的潜在的生物标志物。

[0006] 发明概述

[0007] 如前所述,自身抗体作为免疫系统对肿瘤或者疾病存在信号的生物放大,相对而言,可以长期存在于肿瘤细胞内或者血液循环中,利用这种现象,研究人员可以通过检验自身抗体的存在来对肿瘤的发生、发展、患病的风险等进行评估和预测,由此可以对疾病相关蛋白进行研究,更进一步地,能够实现包括肿瘤在内的疾病的早期检测和预防。

[0008] 但是目前对于如何捕获自身抗体尤其是体液中的自身抗体还没有很满意的工具,因为在体液中的一些自身抗体的含量往往非常低,而作为捕获自身抗体尤其是捕获肿瘤特异性的自身抗体的自身抗原来源有限。虽然目前本领域的技术人员通过 CTL 识别法、MHC

结合的多肽洗脱法等获得了一些自身抗原,例如这些自身抗原都可以在网络数据库中查询(<http://www2.licr.org/CancerImmunomeDB/>)得到,但是随着高通量蛋白质组学技术的不断发展,这些抗原远远不能满足日益增长的高通量操作的需要。

[0009] 而且完整抗原的制备通常需要克隆、过表达、分离以及纯化多个步骤。虽然最近已发展了高通量的生产和纯化目的蛋白的技术(ZhuH等人, Global analysis of protein activities using proteomechips. Science, 2001, 293(5537):2101-5),但是如何制备高质量并且有功能活性的完整的重组抗原依然是研究人员面临的主要挑战。另外,制备完整抗原的成本非常昂贵(Andresen H等人, Functional peptide microarrays for specific and sensitive antibody diagnostics. Proteomics, 2006, 6(5):1376-84)。

[0010] 因此,本发明的目的之一就是提供可以替代完整自身抗原的自身抗原表位,尤其是具有肿瘤特异性的代表完整自身抗原的自身抗原表位,通过提供这种抗原表位,可以减少前期投入成本和提高自身抗原在高通量筛选中的利用效率,例如在容量有限的蛋白质芯片中承载更多的有代表性的抗原表位进而实现高通量筛选、鉴别,更高效地捕获自身抗体,进而为疾病尤其是肿瘤的研究提供更好的分子水平的研究工具,对自身抗体进行更好的研究,进而为研发抗肿瘤药物提供新的作用靶点。

[0011] 基于上述动机,申请人进行了深入的研究,考虑到肿瘤相关抗原通常都是血液中较高丰度的蛋白质,容易被质谱分析捕获和鉴定。同时,过表达的肿瘤相关抗原容易刺激机体的免疫系统,产生自身抗体。因此,可以通过质谱技术鉴定肿瘤特异性抗原,合成代表肿瘤特异性抗原的抗原表位,然后与含有自身抗体的肿瘤患者血清进行免疫杂交,就可以而鉴定能够结合自身抗体的自身抗原表位。

[0012] 按照这一发明构思,本发明的申请人采用已建立的器官无血清培养体系,收集并浓缩了来自于正常胰腺和胰腺癌组织的分泌蛋白质。随后分别采用 Shotgun LC-MS/MS 和 SDS-PAGE HDMS 两种平行的蛋白质组学分析策略,分析了正常胰腺组织和胰腺癌组织的分泌蛋白质组,并确定了胰腺癌组织特异的分泌蛋白质组。

[0013] 在 Shotgun LC-MS/MS 策略中,申请人将胰腺癌及其邻近正常胰腺,以及正常胰腺组织各自的多样本等量混合的分泌蛋白进行胰酶消化,将酶切产物经过反相高效液相色谱分离,线性洗脱获得的组分直接进行质谱鉴定,通过 SEQUEST 软件解析质谱数据,获得正常胰腺组织分泌蛋白质组或者胰腺癌组织特异的分泌蛋白质组。

[0014] 申请人还采用了平行于 Shotgun LC-MS/MS 策略的 SDS-PAGE HDMS 以提高鉴定数据的覆盖度和可信度。具体地,申请人通过 SDS-PAGE HDMS 技术策略对 2 例胰腺癌、1 例正常胰腺以及 3 例正常胰腺组织等量混合的分泌蛋白进行分析,之后应用 Mascot 搜索二级质谱离子数据库,获得正常胰腺组织分泌蛋白质组或者胰腺癌组织特异的分泌蛋白质组。

[0015] 申请人将通过 Shotgun LC-MS/MS 技术策略三次重复鉴定得到的蛋白;或经过 SDS-PAGE HDMS 技术策略鉴定一次以上的蛋白确定为高可信度胰腺癌组织特异的分泌蛋白质。结果申请人共确定了包含 246 个蛋白的正常胰腺组织分泌蛋白质组和包含 729 个蛋白的胰腺癌分泌蛋白质组。经过差减分析,发现 159 个蛋白同时出现于正常胰腺组织分泌蛋白质组和胰腺癌分泌蛋白质组,因而最终确定了包含 570 个高可信度的蛋白的胰腺癌特异分泌蛋白质组。

[0016] 申请人将胰腺癌特异性肽段定为胰腺癌自身抗原的“抗原表位库”,其来自于以上

蛋白质组质谱技术鉴定的 570 个胰腺癌组织特异分泌蛋白的所有肽段。

[0017] 之后申请人用 Autospot 阵列合成仪在活化纤维素膜表面合成上述肽库所包含的部分肽段,得到含有部分肽的多肽阵列。采用该阵列对胰腺癌患者血清进行筛选,终于得到了可以用于检测自身抗体的一个抗原表位,并意外地发现该表位是组织蛋白酶 D 的第 33-50 位的氨基酸,其氨基酸序列为:RRTMSEVGGSVEDLIAKG(SEQ ID NO:1)。该氨基酸的编码基因序列为:5'-CGCCGGACCATGTCGGAGGTTGGGGGCTCTGTGGAGGACCTGATTGCCAAAGGC-3'(SEQ ID NO:12),在通过 Pepscan 技术进行深入鉴定时,发现该抗原表位中 RTMSEVGGSVEDLIA(SEQ ID NO:11)具有相同的抗原表位功能,其编码序列为:5'CGGACCATGTCGGAGGTTGGGGGCTCTGTGGAGGACCTGATTGCC3'(SEQ ID NO:13),由此提供了胰腺癌特异性的自身抗原的抗原表位,完成了本发明。

附图说明:

[0018] 图 1 为 SEQ ID NO:1 的自身抗体检测图,其中 (A) 为针对 12 例胰腺癌血清检测结果。(B) 为针对 8 例健康人血清检测结果。

[0019] 图 2 为采用 Pepscan 技术验证人组织蛋白酶 D 的片段是否具有实施例 3 获得的自身抗原表位的免疫结合活性的试验结果。其中 (A) 和 (B) 为与含自身抗原的胰腺癌患者的血清杂交的结果,(C) 为与健康对照的血清杂交的结果。

[0020] 发明详述

[0021] 本发明的一个方面,提供了下述内容:

[0022] 1. 抗原表位,其选自下述序列之一:

[0023] (a) SEQ ID NO:1 或 11 所示的氨基酸序列;

[0024] (b) 具有作为组织蛋白酶 D 的抗原表位的功能的、且由与 SEQ ID NO:1 或者 11 所示的氨基酸序列具有 95% 以上同源性的氨基酸序列构成的氨基酸序列;和

[0025] (c) 具有作为组织蛋白酶 D 的抗原表位的功能的、且由经过取代、缺失或添加一个或几个氨基酸的 SEQ ID NO:1 或 11 所示的氨基酸序列构成的氨基酸序列。

[0026] 2. 编码项目 1 的抗原表位的基因,其选自下述序列之一:

[0027] (a) SEQ ID NO:12 所示的编码 SEQ ID NO:1 的核苷酸序列;

[0028] (b) SEQ ID NO:13 所示的编码 SEQ ID NO:11 的核苷酸序列;

[0029] (c) 编码下述氨基酸序列的核苷酸序列:所述氨基酸序列是具有作为组织蛋白酶 D 的抗原表位的功能的、且由与 SEQ ID NO:1 所示的氨基酸序列具有 95% 以上同源性的氨基酸序列构成的氨基酸序列;

[0030] (d) 编码下述氨基酸序列的核苷酸序列:所述的氨基酸序列是具有作为组织蛋白酶 D 的抗原表位的功能、且由经过取代、缺失或添加一个或几个氨基酸的 SEQ ID NO:1 所示的氨基酸序列构成的氨基酸序列;和

[0031] (d) 因为密码子简并性而分别与 SEQ ID NO:12 和 13 不同、但是编码 SEQ ID NO:1 或 11 所示氨基酸序列的核苷酸序列。

[0032] 3. 含有项目要求 2 的基因的重组载体。

[0033] 4. 用项目 4 的重组载体转化、转导或转染得到的遗传工程化宿主细胞。

[0034] 5. 项目 1 的抗原表位或项目 2 的基因的用途,其用于制备下述制剂之一:

[0035] (a) 组织蛋白酶 D 的抗原表位 ;

[0036] (b) 组织蛋白酶 D 的胰腺癌特异的抗原表位 ;

[0037] (c) 检测人组织蛋白酶 D 的抗体 ;和

[0038] (d) 疫苗。

[0039] 6. 项目 5 的应用,其特征在于项目 1 所述的抗原表位或项目 2 的基因在制备用于检测胰腺癌血清中自身抗体的制剂中的用途。

[0040] 7. 项目 5 的应用,其特征在于项目 1 所述的抗原表位或项目 2 的基因在制备用于检测胰腺癌血清中自身抗体的试剂盒中的用途。

[0041] 8. 胰腺癌检测方法,其包括使用项目 1 的抗原表位或项目 2 的基因。

[0042] 由于氨基酸序列的特殊性,任何含有 SEQ ID NO :1 所示氨基酸序列的多肽片段或突变体,如其保守性变体、生物活性片段或衍生物,只要该多肽的片段或多肽变体与前述氨基酸序列同源性在 95%以上,均属于本发明保护范围之列。

[0043] 具体地,所述改变可包括氨基酸序列中氨基酸的缺失、插入或替换 ;其中,对于变体的保守性改变,所替换的氨基酸具有原氨基酸相似的结构或化学性质,如用亮氨酸替换异亮氨酸,变体也可具有非保守性改变。

[0044] 所述自身抗原表位可以是重组多肽、天然多肽或合成多肽,可以是天然纯化的产物,或是化学合成的产物,或使用重组技术从原核或真核宿主(例如:细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。

[0045] 本发明的另外一个方面,还涉及编码 SEQ ID NO :1 所示的自身抗原表位的基因。所述基因可含有 SEQ ID NO :12 所示多核苷酸同源性 70-100%的核苷酸序列。

[0046] 由于核苷酸序列的特殊性,任何 SEQ ID NO :12 所示多核苷酸变体,只要与该多核苷酸具有 95%以上的同源性,均属于本发明保护范围之列。只要所述核苷酸序列没有从实质上改变其编码的多肽的序列和功能,均包含在本发明的保护范围之内。

[0047] 本发明的一个方面,还涉及一种含有编码 SEQ ID NO :1 所示的自身抗原表位的基因的重组载体,以及用所述的重组载体转化、转导或转染得到的遗传工程化宿主细胞。

[0048] “载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒(如:腺病毒、逆转录病毒)或其他载体。在本发明中适用的载体包括但不限于:在细菌中表达的基于 T7 启动子的表达载体;在哺乳动物细胞中表达的 pcDNA3.1 载体和在昆虫中表达的来源于杆状病毒的 pFastBac 载体。

[0049] 本领域技术人员可以采用熟知的方法构建含编码本发明的抗原表位的 DNA 序列和合适的转录/翻译调控元件的表达载体。

[0050] 这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等。所述的 DNA 序列可有效连接到表达载体中的适当启动子,以指导 mRNA 合成。这些启动子的代表性例子有:大肠杆菌的 lac 或 trp 启动子;噬菌体的 PL 启动子;真核启动子包括 CMV 早期启动子、HSV 胸苷激酶启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其他一些已知的可控制基因在原核细胞或真核细胞或其病毒中表达的启动子。

[0051] 本发明的另一个方面,还涉及包含含有编码 SEQ ID NO :1 的重组载体的宿主细胞。在本发明中,编码 SEQ ID NO :1 的自身抗原表位的多核苷酸或含有该多核苷酸的重组载体可转化或转导入宿主细胞,以构成含有该核苷酸或重组载体的基因工程化宿主细胞。

[0052] “宿主细胞”指原核细胞,如细菌细胞;或是低等真核细胞,如酵母细胞;或是高等真核细胞,如哺乳动物细胞。代表性例子有:大肠杆菌,链霉菌属;细菌细胞如鼠伤寒沙门氏菌;真菌细胞如酵母;植物细胞;昆虫细胞如果蝇 S2 或 Sf9;动物细胞如 CHO、COS 或 Bowes 黑素瘤细胞等。

[0053] 在本发明中,本领域的技术人员可以用熟知的常规技术将 SEQID NO:1 的 DNA 序列或含有所述 DNA 序列的重组载体转化宿主细胞。

[0054] 当宿主为原核生物如大肠杆菌时,能吸收 DNA 的感受态细胞可在指数生长期后收获,用 CaCl_2 法处理,或者可供选择的使用 MgCl_2 。如果需要,也可用电穿孔的方法进行转化。当宿主是真核生物时,可选用包括磷酸钙共沉淀法,或者常规机械方法如显微注射、电穿孔、脂质体包装等在内的 DNA 转染法。

[0055] 下面通过实施例对本发明的方法进行详细的说明,不过本申请所要求保护的范围并不局限于所述的实施例。

具体实施方式

[0056] 实施例 1:获得肽库

[0057] 胰腺癌及其远端正常胰腺新鲜组织样本于手术离体 10 分钟内,由外科医生各切取约两个黄豆粒大小组织块,分别放入含 1%青链霉素的 L-15 培养基(购自 Invitrogen 公司)的无菌瓶中备用。在洁净工作台中,取无菌玻璃平皿,加入 5mL 洗液,将胰腺组织从无菌瓶中取出放置于平皿中洗净。

[0058] 用手术刀将胰腺癌和正常胰腺组织切成 2mm^3 左右的小块,放入已涂好胶原并划好沟槽的平皿中,4℃冰箱放置 2-3 小时以促进组织块贴壁。从冰箱取出平皿以后,小心滴加经过改良的 LHC-9 无血清培养基(购自 Invitrogen 公司,含表皮生长因子、转铁蛋白、胰岛素等),防止将组织块冲起。

[0059] 将培养皿放入培养盒中,充入含有 50% N_2 ,45% O_2 和 5% CO_2 的混合气体,密封培养盒。将培养盒放在摇床上,摇摆频率为 3-4 次/min,36.5℃,4% CO_2 孵箱中培养 48-72h。收集并浓缩了来自于正常胰腺和胰腺癌组织的分泌蛋白质。

[0060] 随后分别采用 Shotgun LC-MS/MS 和 SDS-PAGE HDMS 两种平行的蛋白质组学分析策略,分析了正常胰腺组织和胰腺癌组织的分泌蛋白质组,并确定了胰腺癌组织特异的分泌蛋白质组。

[0061] Shotgun LC-MS/MS 鉴定方法:样品为 1 例胰腺癌及其邻近正常胰腺,以及 3 例正常胰腺等量混合的分泌蛋白。

[0062] 具体地,使用胰酶在 37 度条件下过夜消化浓缩分泌蛋白。取 16 μg 酶切产物经过反相高效液相色谱 [GE Ettan MDLC(GE Healthcare, Piscataway, NJ, USA)] 分离,线性洗脱获得的组分直接进 ESI 离子源系统 [LTQ-MS/MS(Thermo Finnigan(San Jose, CA, USA))] 进行质谱鉴定,质谱数据由 SEQUEST 软件解析。

[0063] 每个培养基样本[包含组织分泌蛋白和 LHC-9 无血清培养基(购自 Invitrogen 公司,含表皮生长因子、转铁蛋白、胰岛素等)]重复分析了三次以确定鉴定方法的可重复性,并获得高可信度的分泌蛋白。

[0064] 结果三次平行分析后,分别从 3 例正常胰腺组织分泌上清中鉴定出 40 余种蛋白

(三次分析的结果分别为 45、49 和 49 种蛋白),定为组 1,从 1 例正常胰腺组织分泌上清中鉴定出 20 余种蛋白(三次分析的结果分别为 28、22 和 21 种蛋白)定为组 2,在 1 例胰腺癌组织分泌上清中鉴定出 270 余种蛋白(三次分析的结果分别为 277、287 和 271 种蛋白),定为组 3。

[0065] 3 组样本的三次平均重合率为 27%。约有 49%的蛋白鉴定到 2 次以上。

[0066] 为了提高鉴定数据的覆盖度和可信度,申请人还采用了平行于 Shotgun LC-MS/MS 策略的 SDS-PAGE HDMS 技术策略分析了 2 例胰腺癌、1 例正常胰腺以及 3 例正常胰腺组织等量混合的分泌蛋白。具体地,分泌蛋白经 SDS-PAGE 分离后,依次切下分离胶内不同的蛋白条带,胰酶酶切后分别经过 NanoLC-MS/MS on Acquity™Nano UPLCsystem 和 Synapt 高精度质谱分析。应用 Mascot 搜索二级质谱离子数据库,确定蛋白鉴定结果。

[0067] 以上组织分泌上清中分别鉴定到 179、89、515 和 252 个蛋白。由于 HDMS 为高精度的质谱仪器,用该体系鉴定到的蛋白定义为高质量蛋白。

[0068] 为了获得高可信度的正常胰腺以及胰腺癌的分泌蛋白,申请人如下限定了所鉴定的蛋白,只有所鉴定的蛋白满足以下条件时才被确定为高可信度蛋白,并可以被纳入申请人建立的分泌蛋白质组数据库:(1)Shotgun LC-MS/MS 技术策略三次重复鉴定的蛋白;或(2)SDS-PAGE HDMS 技术策略一次以上鉴定的蛋白。

[0069] 应用以上的标准,申请人确定了包含 246 个蛋白的正常胰腺组织分泌蛋白质组和包含 729 个蛋白的胰腺癌分泌蛋白质组。经过差减分析,发现 159 个蛋白同时出现于正常胰腺组织分泌蛋白质组和胰腺癌分泌蛋白质组,因而最终确定了包含 570 个高质量蛋白的胰腺癌特异分泌蛋白质组。

[0070] 申请人将质谱技术鉴定的胰腺癌特异性肽段确定为胰腺癌自身抗原“自身抗原表位库”,其来自于以上蛋白质组质谱技术鉴定的胰腺癌组织特异分泌蛋白(570 个)的所有肽段,总数约 2000-3000 个。每个肽段长度为 2-20 个氨基酸,平均 3-6 个多肽确定 1 个蛋白质。在表 1 中申请人具体列举了两种蛋白质组学技术策略中质谱鉴定的肽段。

[0071] 申请人用 Autospot 阵列合成仪(德国 INTAVIS 公司, Cologne, Germany)在活化纤维素膜表面合成上述肽,得到含有全部肽的多肽阵列。

[0072]

表 1

胰腺癌组织 1			胰腺癌组织 2	
Shotgun LC-MS/MS			SDS-PAGE/HDMS	SDS-PAGE/HDMS
第一次	第二次	第三次		
K. AIGAVPLIQGE	K. AIGAVPLI	TMSEVGGSVEDLIAK (35 - 49)	R. TMSEVGGSVEDLIAK. G (35 - 49)	
YMIPCEK. V	K. AIGAVPLIQGEY	QGEYMIPCEK		
(314-331) *	MIPCEK. V (314- . V (314-331)			
K. FDGILGMAYPR		YSQAVPAVTEGPIPEVLK (55 - 72)	K. YSQAVPAVTEGPIPEVLK. N (55 - 72)	
. I (195-205)				
		QVFGGEATK (177 - 184)	R. QVFGGEATKQPGITFIAAK. F (177 - 194)	
		FDGILGMAYPR (195 - 205)	K. QPGITFIAAK. F (185 - 194)	
		ISVNNVLPVFDNLMQK (206 - 222)	K. FDGILGMAYPR. I (195 - 205)	
		DPDAQPGGELMLGGTDSK (236 - 253)	K. VSTLPAITLK. L (332 - 341)	
		VSTLPAITLK (332 - 341)	R. VGFAEAAARL (404 - 412)	
		LSPEDYTLK (349 - 357)		
		VGFAEAAAR (404 - 411)		

*括号中的数字表示质谱鉴定氨基酸的位置。

[0073] 实施例 2 :含有自身抗体的肿瘤样品的获得

[0074] 收集患者血清来自中国医学科学院北京协和医学院肿瘤医院腹部外科 2006 年 3 月 1 日至 2008 年 12 月 30 日间,未合并其它疾病及未进行任何治疗的病理确诊的胰腺癌患者血清 12 例,男 7 例,女 5 例,年龄 41 ~ 70 岁,中位年龄 61 岁 ;来自中国医学科学院北京

协和医学院肿瘤医院防癌体检科 2007 年 8 月 1 日至 2008 年 8 月 1 日间,健康自愿者血清 8 例,其中男 5 例,女 3 例,年龄 30 ~ 64 岁,中位年龄 61 岁。所有收集病例均获患者知情同意签署及伦理委员会批准。

[0075] 实施例 3 :抗原性表位筛选及其序列确定

[0076] 检测 12 例病例确诊为胰腺导管腺癌患者血清中的自身抗体,以 8 例健康者血清为阴性对照,采用免疫印迹法检测该蛋白的抗原性。将实施例 1 中的合成多肽的纤维素膜先后置于 100%、75%和 50%的乙醇中各 5 分钟后,然后实施下述步骤 a)-h) :

[0077] a) 浸泡于 1×PBS(137mM NaCl,2.7mM KCl,10mM Na₂HPO₄,2mM KH₂PO₄, pH 7.4) 中 30 分钟 ;

[0078] b) 以 5%脱脂牛奶封闭纤维素膜,室温下封闭 2 小时 ;

[0079] c) 1×PBST(137mM NaCl,2.7mM KCl,10mM Na₂HPO₄,2mMKH₂PO₄,0.2% Tween-20, pH 7.4) 浸洗纤维素膜 1 次 ;

[0080] d) 加入用 1×PBST 按照 1 : 1000 稀释的含有自身抗体的待测血清,室温孵育 1 小时 ;

[0081] e) 1×PBST 洗 10 分钟 ×3 次 ;

[0082] f) 加入 1 : 5000 稀释 HRP 标记羊抗人 IgG 抗体 (美国 JacksonImmunoResearch Laboratories 公司产品,批号 :109-005-003, West Grove, PA, USA), 室温孵育 1 小时 ;

[0083] g) 1×PBST 洗 10 分钟 ×3 次 ;

[0084] h) 用吸水纸去除膜表面液体后, ECL 显色, X 光片曝光, 显影、定影、扫描。

[0085] 结果参见表 2, 可知 12 例胰腺癌患者中, 5 例血清均能与 SEQ ID NO :1 所示序列有结合反应, 阳性率为 41.7% (5/12) ; 1 例血清能与 SEQ ID NO :3 有结合反应, 阳性率为 8.3% (1/12) ; 1 例血清能与 SEQ ID NO :4 有结合反应, 阳性率为 8.3% (1/12) ; 1 例血清能与 SEQ ID NO :8 有结合反应, 阳性率为 8.3% (1/12)。8 例健康对照血清中, 1 例血清能与 SEQ ID NO :8 有结合反应, 阳性率为 12.5% (1/8) (图 1)。如表 1 所示, SEQ ID NO :1 更容易与胰腺癌患者血清自身抗体结合, 而且健康对照血清中未发现此种结合。

[0086] 表 2

[0087]

SEQ ID NO	合成氨基酸序列	位点	胰腺癌血清 阳性率	健康人血 清阳性率
1	RRTMSEVGGSVEDLIAKG	33-50	41.7 %	0
2	YSQAVPAVTEGPIPEVLK	55-72	0	0
3	QVFGEATKQPGITFIAAK	177-194	8.3 %	0
4	EATKQPGITFIAAKFDGI	181-198	8.3 %	0
5	IAAKFDGILGMAYPRISV	191-208	0	0
6	ISVNNVLPVFDNLMQQKL	206-223	0	0
7	DPDAQPGGELMLGGTDSK	236-253	0	0
8	AIGAVPLIQGEYMIPCEK	314-331	8.3 %	12.5 %
9	PCEKVSTLPAITLKLGGK	328-345	0	0
10	GKGYKLSPEDYTLKVSQA	344-361	0	0

[0088] 通过 Genbank 检索,发现其中 SEQ ID NO:1 所示序列是人组织蛋白酶 D (SEQ ID NO:3) 的第 33-50 位氨基酸所示序列,由此确定人组织蛋白酶 D 与胰腺癌的自身抗体产生相关,同时因为所合成的肽是 18 肽,无论是生物工程手段还是化学合成手段,其获得成本都相较完整的人组织蛋白酶 D 低,因此本发明为通量筛选或者其它公知的用于捕获、鉴定、筛选自身抗体的方法提供了很好的自身抗体表位工具肽。

[0089] 实施例 4:肿瘤特异性抗原表位的再次确定已知编码组织蛋白酶 D 的基因位于染色体 11p15,基因全长 11,238bp,编码蛋白全长 412 个氨基酸。其氨基酸序列为:

[0090] 10 20 30 40 50 60

[0091] MQPSSLLPLA LCLLAAPASA LVRIPLHKFT SIRRTMSEVG GSVEDLIAKG PVSKYSQAVP

[0092] 70 80 90 100 110 120

[0093] AVTEGPIPEV LKNYMDAQYY GEIGIGTPPQ CFTVVFDTGS SNLWVPSIHC KLLDIACWIH

[0094] 130 140 150 160 170 180

[0095] HKYNSDKSST YVKNGTSFDI HYGSGSLSGY LSQDTVSVPC QSASSASALG GVKVERQVFG

[0096] 190 200 210 220 230 240

[0097] EATKQPGITF IAAKFDGILG MAYPRISVNN VLPVFDNLMQ QKLVDQNIFS FYLSRDPDAQ

[0098] 250 260 270 280 290 300

[0099] PGGELMLGGT DSKYYKGSLS YLNVTRKAYW QVHLDQVEVA SGLTLCKEGC EAIVDTGTSL

[0100] 310 320 330 340 350 360

[0101] MVGPVDEVRE LQKAIGAVPL IQGEYMIPCE KVSTLPAITL KLGGKGYKLS PEDYTLKVSQ

[0102] 370 380 390 400 410

[0103] AGKTLCLSGF MGMDIPPPSG PLWILGDVFI GRYTTFVDRD NNRVGFAEAA RL

[0104] 应用 AutoSpot 点样体系,叠瓦式(又成为 Pepscan 技术,参见 Kopecky EM 等

人,. Mapping of FVIII inhibitor epitopes using cellulose-bound synthetic peptide arrays[J]. J Immunol Methods, 2006, 308(1-2): 90-100) 合成人组织蛋白酶 D 完整氨基酸序列的肽段阵列。

[0105] 具体地, 将人组织蛋白酶 D 蛋白完整氨基酸序列(412 个氨基酸) 按照 18 个连续氨基酸构成一个肽段, 前后两个单元相互交叠 15 个氨基酸, 依次合成 133 个肽段。

[0106] 将合成的 133 个肽段单元按照在氨基酸中的连接次序依次点样到纤维素膜上, 其中横向按照 1-20 的序号编排, 纵向按照 a-g 的顺序编排, 任何一个肽段对应的编号为纵排和横排编号的组合, 例如横排第 1 个, 纵排第 b 行, 其编号为 b1, 在该纤维素膜上, 从 a1 开始依次点样相互交叠最大的肽段, 例如 a11 和 a12 这一邻接的点样点中, a12 的前 15 个氨基酸与 a11 的后 15 个氨基酸相互重叠。

[0107] 将前述合成的涵盖人组织蛋白酶 D 蛋白完整氨基酸序列的 133 个肽段的点样点的纤维素膜分别与 5 例胰腺癌患者的混合血清、来自上述 5 例中的单个 1 例胰腺癌患者的血清, 以及 4 例健康对照的混合血清共孵育。然后, 该纤维素膜先后依次置于 100%、75% 和 50% 的乙醇中各 5 分钟, 除了所用的待测血清不同外, 其余与实施例 3 中的步骤 a) -h) 的操作相同。

[0108] 免疫印迹结果(参见图 2) 显示, 胰腺癌患者的血清与膜杂交后, 血清中的自抗体所识别的序列位于 a11、a12、c13 ~ 16、g12 和 g13(图 A 和 B) 等肽段; 健康对照血清与 NC 膜杂交后, 血清自抗体所识别的位点位于 c13 ~ 16、g12 和 g13(图 C) 等完全不同的肽段区域。产生阳性信号相邻位点的重叠氨基酸序列即为抗原线性表位。

[0109] 基于以上胰腺癌和健康对照血清杂交结果, 共确定了人组织蛋白酶 D 完整氨基酸序列中 3 个抗原线性表位: RTMSEVGGSVEDLIA (SEQ ID NO: 11; 人组织蛋白酶 D 的第 34-48 位), 其对应的编码核苷酸序列为 5' CGGACCATGTCGGAGGTTGGGGCTCTGTGGAGGACCTGATTGCC 3' (SEQ ID NO: 13); ASALGGVKV (166-174aa) 和 TVFDRDNNRVGFAEAAR (395-411aa)。其中, RTMSEVGGSVEDLIA 肽段 (34-48aa) 特异性地被胰腺癌外周血自身抗体识别 (a11 和 a12 阳性信号位点), 而且该肽段正是位于 SEQ ID NO: 1 (人组织蛋白酶 D 的第 33-50 位氨基酸) 所示序列所包含的序列, 提示 SEQ ID NO: 11 是胰腺癌特异的抗原表位中非常关键的部分, 本身也可以作为抗原表位。

[0110] 工业应用

[0111] 本发明的自身抗原表位可应用于自身抗体捕获和自身免疫疾病相关指标检测, 肿瘤相关标志物的检测等方面。在许多情况下, 与现有技术相比, 本发明的表位肽的成本更为低廉、检测精确度更高。

[0112] 序列表

[0113] <110> 中国医学科学院肿瘤研究所

[0114] <120> 一种自身抗原表位、其编码基因及其用途

[0115] <130> IDC090031

[0116] <160> 13

[0117] <170> PatentIn version 3.2

[0118] <210> 1

[0119] <211> 18

- [0120] <212>PRT
[0121] <213> 人工序列
[0122] <400>1
[0123] Arg Arg Thr Met Ser Glu Val Gly Gly Ser Val Glu Asp Leu Ile Ala
[0124] 1 5 10 15
[0125] Lys Gly
[0126] <210>2
[0127] <211>18
[0128] <212>PRT
[0129] <213> 人工序列
[0130] <400>2
[0131] Tyr Ser Gln Ala Val Pro Ala Val Thr Glu Gly Pro Ile Pro Glu Val
[0132] 1 5 10 15
[0133] Leu Lys
[0134] <210>3
[0135] <211>18
[0136] <212>PRT
[0137] <213> 人工序列
[0138] <400>3
[0139] Gln Val Phe Gly Glu Ala Thr Lys Gln Pro Gly Ile Thr Phe Ile Ala
[0140] 1 5 10 15
[0141] Ala Lys
[0142] <210>4
[0143] <211>18
[0144] <212>PRT
[0145] <213> 人工序列
[0146] <400>4
[0147] Glu Ala Thr Lys Gln Pro Gly Ile Thr Phe Ile Ala Ala Lys Phe Asp
[0148] 1 5 10 15
[0149] Gly Ile
[0150] <210>5
[0151] <211>18
[0152] <212>PRT
[0153] <213> 人工序列
[0154] <400>5
[0155] Ile Ala Ala Lys Phe Asp Gly Ile Leu Gly Met Ala Tyr Pro Arg Ile
[0156] 1 5 10 15
[0157] Ser Val
[0158] <210>6

[0159] <211>18
[0160] <212>PRT
[0161] <213> 人工序列
[0162] <400>6
[0163] Ile Ser Val Asn Asn Val Leu Pro Val Phe Asp Asn Leu Met Gln Gln
[0164] 1 5 10 15
[0165] Lys Leu
[0166] <210>7
[0167] <211>18
[0168] <212>PRT
[0169] <213> 人工序列
[0170] <400>7
[0171] Asp Pro Asp Ala Gln Pro Gly Gly Glu Leu Met Leu Gly Gly Thr Asp
[0172] 1 5 10 15
[0173] Ser Lys
[0174] <210>8
[0175] <211>18
[0176] <212>PRT
[0177] <213> 人工序列
[0178] <400>8
[0179] Ala Ile Gly Ala Val Pro Leu Ile Gln Gly Glu Tyr Met Ile Pro Cys
[0180] 1 5 10 15
[0181] Glu Lys
[0182] <210>9
[0183] <211>18
[0184] <212>PRT
[0185] <213> 人工序列
[0186] <400>9
[0187] Pro Cys Glu Lys Val Ser Thr Leu Pro Ala Ile Thr Leu Lys Leu Gly
[0188] 1 5 10 15
[0189] Gly Lys
[0190] <210>10
[0191] <211>18
[0192] <212>PRT
[0193] <213> 人工序列
[0194] <400>10
[0195] Gly Lys Gly Tyr Lys Leu Ser Pro Glu Asp Tyr Thr Leu Lys Val Ser
[0196] 1 5 10 15
[0197] Gln Ala

[0198]	<210>11	
[0199]	<211>15	
[0200]	<212>PRT	
[0201]	<213> 人工序列	
[0202]	<400>11	
[0203]	Arg Thr Met Ser Glu Val Gly Gly Ser Val Glu Asp Leu Ile Ala	
[0204]	1 5 10 15	
[0205]	<210>12	
[0206]	<211>54	
[0207]	<212>DNA	
[0208]	<213> 人工序列	
[0209]	<400>12	
[0210]	cgccggacca tgtcggaggt tgggggctct gtggaggacc tgattgccaa aggc	54
[0211]	<210>13	
[0212]	<211>45	
[0213]	<212>DNA	
[0214]	<213> 人工序列	
[0215]	<400>13	
[0216]	cggacca tgtcggaggt tgggggctct gtggaggacc tgattgcc	45

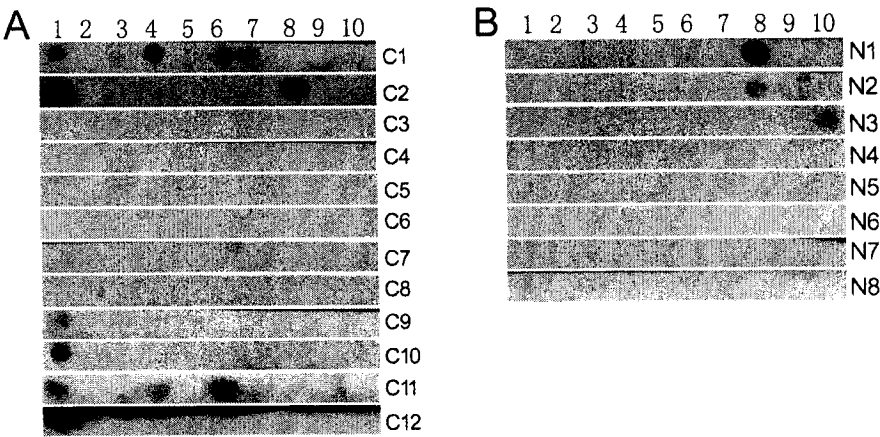


图 1

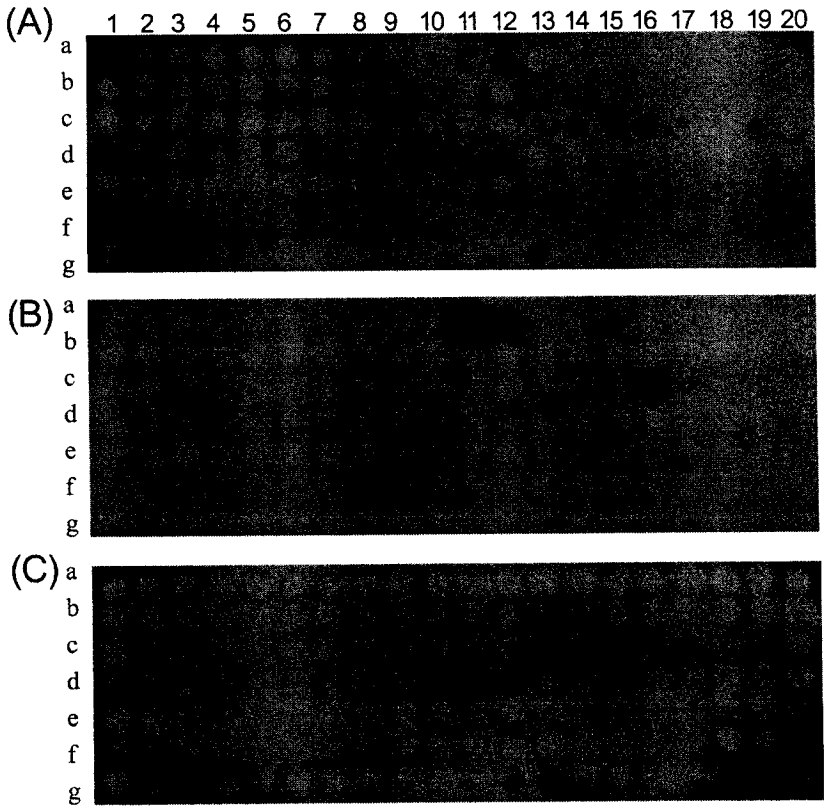


图 2

专利名称(译)	一种自身抗原表位、其编码基因及其用途		
公开(公告)号	CN101885758A	公开(公告)日	2010-11-17
申请号	CN200910136343.X	申请日	2009-05-11
[标]申请(专利权)人(译)	中国医学科学院肿瘤研究所		
申请(专利权)人(译)	中国医学科学院肿瘤研究所		
当前申请(专利权)人(译)	中国医学科学院肿瘤医院		
[标]发明人	赵晓航 何平 刘飞 孙玉琳 刘芳 周兰萍 谢鹏 张金强 安怀杰 李潜		
发明人	赵晓航 何平 刘飞 孙玉琳 刘芳 周兰萍 谢鹏 张金强 安怀杰 李潜		
IPC分类号	C07K7/08 C12N15/12 C12N15/63 C12N1/19 C12N1/21 C12N5/10 C12P21/02 G01N33/53 G01N33/86 C12Q1/68 G01N33/574		
其他公开文献	CN101885758B		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及免疫化学、蛋白质组学和肿瘤免疫领域。具体地涉及一种自身抗原表位、其编码基因及其用途。更具体地，涉及人组织蛋白酶D (cathepsin D)抗原表位区及其编码基因，以及这种多肽在鉴定胰腺癌患者血清中自身抗体中的用途。

表 1		氨基酸组成 1	氨基酸组成 2
Sequence ID: 45/MS		SDS-PAGE/IMS	SDS-PAGE/IMS
第一次	第二次	第一次	第二次
LALGAPLQEE	LALGAPPL	THREOSINYLALG (35 - 49)	LTHREOSINYLALG (35 - 49)
WMPKEL Y	LALGAPLQEEY	QCEWMPKEL	
(314-331) •	WMPKEL Y (314+ . Y (314-		
	331)		
LPRGLDMMPR		YSQMPVTPVTPPTPLA (55 - 72)	LTSQMPVTPVTPPLA W (55 - 72)
.1 (135-165)			
		QVPEGLT (177 - 184)	RQVPEGLTQVPTTALF (177 - 194)
		PRGLDMMPR (195 - 205)	LQVPTTALF (195 - 194)
		TSWNPVTPVTPVTPK (206 - 222)	LPRGLDMMPR.1 (195 - 205)
		DPNAPQEGWLGCTDS (236 - 253)	LVTSLPALTIL.1 (332 - 341)
		VSTSLPALTIL (332 - 341)	RWPSEALRI (446 - 412)
		LSPPVTL (349 - 357)	
		WPEELAR (449 - 411)	

*括号中的数字表示所提及氨基酸的位置。