

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

[12] 发明专利申请公开说明书

C12N 15/12
C07K 14/47 C07K 16/18
C12N 5/10 C12N 15/63
G01N 33/68 G01N 33/574
C12Q 1/68 A61K 38/17
A61K 39/395 A61K 31/70

[21] 申请号 99807965.0

[43] 公开日 2001 年 8 月 8 日

[11] 公开号 CN 1307636A

[22] 申请日 1999.5.28 [21] 申请号 99807965.0

[30] 优先权

[32] 1998.5.29 [33] DE [31] 19824230.1

[86] 国际申请 PCT/EP99/03716 1999.5.28

[87] 国际公布 WO99/63079 德 1999.12.9

[85] 进入国家阶段日期 2000.12.27

[71] 申请人 安娜·斯塔津斯基-波伊兹

地址 德国法兰克福

[72] 发明人 安娜·斯塔津斯基-波伊兹

西尔维亚·科特奇安

海克·汉德罗-梅特兹玛彻

[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

代理人 陈文平

权利要求书 4 页 说明书 23 页 附图页数 1 页

[54] 发明名称 新的子宫内膜异位相关基因

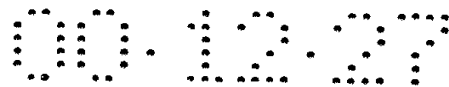
[57] 摘要

本发明涉及侵袭性疾病如子宫内膜异位相关基因、所述基因编码的多肽、抗该多肽的抗体和该核酸、多肽和抗体的药物学应用。

ISSN 1008-4274

权利要求书

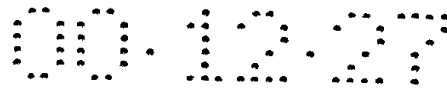
1. 一种核酸，其特征在于它含有
 - (a) SEQ ID NO. 1、3 和/或 5 所示核苷酸序列、其组合或其编码蛋白质的片段，
 - (b) 遗传密码简并性范围内与 (a) 中序列相应的核苷酸序列，
 - (c) 在严谨条件下与 (a) 和/或 (b) 中的序列杂交的核苷酸序列，条件是该核苷酸序列与 EMBL EST 数据库中登录号为 Z98886、Ac003017、Aa453993、AL023586 和 Aa452856 的序列不同。
2. 根据权利要求 1 的核酸，其特征在于它含有 SEQ ID NO. 1、3 和/或 5 所示核苷酸序列的编码蛋白质的片段。
3. 根据权利要求 1 的核酸，其特征在于它与 SEQ ID NO. 1、3 和/或 5 所示核苷酸序列或其片段具有 80% 同源性。
4. 根据权利要求 1-3 之任一项的核酸，其特征在于它编码与侵袭过程有关的多肽或其片段。
5. 包含根据权利要求 1-4 之任一项的核苷酸序列的修饰核酸或核酸类似物。
6. 多肽，其特征在于它是由根据权利要求 1-4 之任一项的核酸编码的，其中不考虑权利要求 1 的条件。
7. 根据权利要求 6 的多肽，其特征在于它
 - (a) 含有 SEQ ID NO. 2 或 4 所示氨基酸序列或
 - (b) 与根据 (a) 的氨基酸序列具有 70% 以上的同源性。



8. 包含根据权利要求 6 或 7 的氨基酸序列的修饰多肽。
9. 肽，其特征在于它是 SEQ ID NO. 2 或 4 所示氨基酸序列的至少 10 个氨基酸的片段。
10. 载体，其特征在于它包含至少一个拷贝的根据权利要求 1-4 之任一项目的核酸或其片段。
11. 根据权利要求 10 的载体，其特征在于它便于在适当宿主细胞中表达核酸。
12. 细胞，其特征在于它已被根据权利要求 1-4 之任一项目的核酸或根据权利要求 10 或 11 的载体转化。
13. 抗权利要求 6-8 之任一项目的多肽或抗根据权利要求 9 的肽的抗体。
14. 根据权利要求 13 的抗体，其特征在于它是抗所述完整多肽或抗其选自 SEQ ID NO. 4 第 1-330 位氨基酸片段的片段。
15. 药物组合物，其特征在于它包含作为活性成分的：
 - (a) 根据权利要求 1-5 的核酸，其中不考虑权利要求 1 的条件，
 - (b) 根据权利要求 10 或 11 的载体，
 - (c) 根据权利要求 12 的细胞，
 - (d) 根据权利要求 6-8 之任一项目的多肽，
 - (e) 根据权利要求 9 的肽和/或
 - (f) 根据权利要求 13 或 14 的抗体。
16. 根据权利要求 15 的组合物，其特征在于它还含有药学上的传统载体、赋形剂和/或添加物。

- 17.根据权利要求 6-8 之任一项的多肽或该多肽的片段作为免疫原在产生抗体中的应用。
- 18.根据权利要求 15 或 16 的组合物在侵袭过程相关疾病的诊断中的应用。
- 19.根据权利要求 15 或 16 的组合物在侵袭过程相关疾病的易感性的诊断中的应用。
- 20.根据权利要求 15 或 16 的组合物在侵袭过程相关疾病的治疗或预防中的应用。
- 21.根据权利要求 15 或 16 的组合物在子宫内膜异位症的诊断、治疗或预防中的应用。
- 22.根据权利要求 15 或 16 的组合物在肿瘤性疾病的诊断、治疗或预防中的应用。
- 23.根据权利要求 15 或 16 的组合物作为基因治疗剂的应用。
- 24.根据权利要求 15 或 16 的组合物作为反义抑制剂的应用。
- 25.根据权利要求 15 或 16 的组合物在胚胎植入中的应用。
- 26.根据权利要求 15 或 16 的组合物在鉴定权利要求 6-8 之任一项的多肽的抑制剂和/或能结合该多肽的分子的抑制剂中的应用。
- 27.根据权利要求 6-8 之任一项的多肽或该多肽的片段在检测样品中抗子宫内膜异位相关蛋白或其片段的抗体中的应用。

28.根据权利要求 13 或 14 的抗体或该抗体的片段在检测子宫内膜异位相关蛋白或其片段中的应用。



说 明 书

新的子宫内膜异位相关基因

本发明涉及侵袭过程例如子宫内膜异位相关的基因、其编码的多肽、抗该肽的抗体和该核酸、多肽和抗体的药学应用。

子宫内膜异位是女性第二大常见疾病，定义为在子宫外出现子宫内膜细胞。子宫内膜异位影响约五分之一的生殖期女性和二分之一具有生育问题的女性。

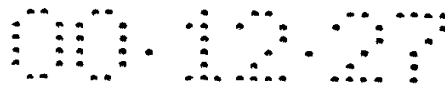
在正常情况下，子宫内膜仅在子宫内出现。在子宫内膜异位时，在子宫外例如在子宫外表面上、肠上甚至在胰腺或肺中出现组织学表型类似子宫内膜的组织。尽管这些子宫内膜异位病灶位于子宫外，但在月经过程中也会出血，因此它们受女性周期激素的影响。因为在该周期中类似子宫内膜的子宫内膜异位病灶经历体积变化，所以根据部位不同这些变化可引起疼痛。而且肌体对子宫内膜异位细胞发生炎症反应，这也引起疼痛。而且，炎症在卵巢和输卵管区导致粘附，结果引起染病女性所谓的机械不孕。然而，很明显，即使没有出现粘附，在子宫内膜异位中也会释放可降低染病女性生育力的信息。

考虑到它们的病理学特性，子宫内膜异位细胞可归为正常细胞和肿瘤细胞之间的类型：一方面它们不显示赘生行为，然而另一方面，象转移性肿瘤细胞一样，它们能通过生物体内器官的边界并长入其它器官，也就是说它们显示侵袭行为。因此，文献中子宫内膜异位细胞定义为“良性肿瘤细胞”，尽管至今这种类型细胞中仍未发现原癌基因的肿瘤特异突变。

因为子宫内膜异位的发病机理仍不完全清楚，所以对子宫内膜异位相关疾病的治疗和预防仍没有有效的选择。

本发明的目的是鉴定在侵袭过程中起作用并可能与子宫内膜异位的病理生理学表型相关的新基因。

根据本发明，该目标通过鉴定、克隆和表征编码多肽的称为子宫内膜



异位相关基因的基因来实现。该基因序列是通过差异显示 RT-PCR (Liang 和 Pardee, 科学 (Science) 257 (1992), 967-971) 发现的。为此, 在子宫内膜细胞系的侵袭性和非侵袭性变体之间进行相互比较。在该过程中, 发现一个子宫内膜细胞侵袭性变体特异的 cDNA 序列。发现长度为 4kb 的相关 RNA。从 cDNA 噬菌体库分离的相应 cDNA 具有一个 302 氨基酸的开放阅读框 (ORF)。

本发明涉及一种核酸, 其包含

(a) SEQ ID NO. 1、3 或/和 5 所示的核苷酸序列、其组合或编码蛋白的片段,

(b) 在遗传密码简并性的范围内相应于 (a) 的核苷酸序列或

(c) 在严谨条件下与 (a) 和/或 (b) 中的序列杂交的核苷酸序列。

该核酸优选编码与侵袭过程相关的多肽或其片段。

下列核苷酸序列已存在 EMBL EST 数据库中, 其登录号为: Z98886、Ac003017、AL023586、Aa52993、Aa452856。这些序列并不代表根据本发明的核酸。这些序列的前两个是分离自人脑的 DNA, 显示分别与 SEQ ID NO.1 从第 970 到约 2000 位和从第 760 到约 1450 位核苷酸的片段, 或与 SEQ ID NO.3 (在 5'端额外有 84 个碱基) 从第 1054 到 2084 位和从第 844 到约 1534 位核苷酸的片段, 有大于 90% 的一致碱基。AL023586 也是人的序列, 其与 Z98885 非常相似并与 SEQ ID NO. 1 第 970 到约 2000 位的区域有同源性。

序列 Aa452993 和 Aa452856 来源于小鼠胚胎, 显示分别与 SEQ ID NO.1 第约 1060 到约 1450 位以及从第约 24 到 440 位的核苷酸 (nt), 或 SEQ ID NO.3 核苷酸位置的第 1144 到约 1534 位以及从第约 108 到约 524 位核苷酸的碱基一致性。直到目前任何这 4 个序列均未有指出阅读框或功能。

SEQ ID NO.1 中描述的核苷酸序列含有长 302 个氨基酸的多肽相应的开放阅读框。该多肽显示在 SEQ ID NO. 2 所描述的氨基酸序列中。SEQ ID NO.3 显示一个相应于 SEQ ID NO.1 中的核苷酸序列, 但其在 5'端具有另外 84 个核苷酸。结果, 在两个核酸序列中, 彼此相应的核苷酸的位

置移动了 84 个核苷酸。因此，SEQ ID NO.3 编码的多肽在 N 端具有另外 28 个氨基酸，并且描述于具有 330 个氨基酸的 SEQ ID NO.4 中。SEQ ID NO.2 和 4 描述了天然多肽的 C-末端片段。

为了说明，请参考图 1，该图显示了根据本发明的子宫内膜异位相关基因的 cDNA 图示说明。显示了五个外显子 E1 至 E5，和用作 DDRT-PCR 探针的片段 1 (394 nt) 的位置。还显示了用作 RT-PCR 的 PCR 引物 (见实施例 4，表 1) 的位置。

图 1 未显示在 SEQ ID NO.5 中描述了其核苷酸序列的另一个外显子 4a。该 4a 外显子可能存在。如果存在，应处于外显子 4 和外显子 5 之间。这相应于 SEQ ID NO.3 中 nt1054 和 nt1055 之间的位置。因此，SEQ ID NO.1/3 与 SEQ ID NO.5 的组合产生例如一个在所述位置含有外显子 4a 序列的序列。

除了显示于 SEQ ID NO.1、3 和 5 的核苷酸序列和其组合例如在 nt1054 和 1055 间具有 SEQ ID NO.5 序列的 SEQ ID NO.3，以及在遗传密码的简并性范围内相应于这些序列之一的核苷酸序列，本发明还包括与前述序列之一杂交的核苷酸序列。根据本发明，术语“杂交”的使用见 Sambrook 等 (分子克隆实验室手册，Cold Spring Harbor Laboratory Press (1989)，1.101-1.104)。优选地，如果在 50℃，优选 55℃，尤其优选 62℃，最优选 68℃，用 1×SSC 和 0.1% SDS 洗涤 1 小时，特别优选在 55℃，优选 55℃，尤其优选 62℃，最优选 68℃，用 0.2×SSC 和 0.1% SDS 洗涤 1 小时后，仍可观察到阳性杂交信号，则将杂交称为严谨的。在这些洗涤条件下与 SEQ ID NO.1、3 和 5 描述的一或多个核苷酸序列，或在遗传密码简并性范围内相应于这些序列的核酸序列杂交的核苷酸序列是根据本发明的核苷酸序列。

根据本发明的核苷酸序列优选为 DNA。然而，也可包括例如 RNA 或核酸类似物如肽核酸。尤其优选的是，根据本发明的核酸包括 SEQ ID NO.1、3 和/或 5 所示核苷酸序列编码蛋白的片段，或与 SEQ ID NO.1、3 或 5 所示核苷酸序列有大于 80%，优选大于 90%，尤其优选大于 95% 同源性的序列，或它们的优选至少 20 个核苷酸 (nt)，尤其优选至少 50nt

的片段。如上所述，同样的原则也适用于除了 SEQ ID NO.1 或 3 的序列外还含有 SEQ ID NO.5 的序列的核酸。当两个核酸（或肽链）相比时，同源性以一致性位置的百分比给出，其中 100%同源性意味着相比较的两分子完全一致（Herder: Lexikon der Biochemie und Molekularbiologie [生物化学和分子生物学字典], Spektrum Akademischer Verlag 1995）。

根据本发明的核酸优选可从哺乳动物，尤其是人获得。它们可根据已知技术通过使用 SEQ ID NO.1、3 或/和 5 所示核苷酸序列的短片段作为杂交探针和/或作为扩增引物来分离。而且，根据本发明的核酸还可由化学合成制备，可使用修饰的核苷酸构件，例如适当时用 2'-O-烷基化核苷酸构件，而非传统核苷酸构件。

因此，根据本发明的核酸或其片段可用于制备优选含有标志或标记基团的引物和探针。还优选特别适合鉴定不同 mRNA 种类的跨内含子的寡核苷酸引物。

本发明还涉及如上定义的核酸编码的多肽。这些多肽优选含有

(a)SEQ ID NO.2 或 4 中所描述的氨基酸序列，或

(b)与 (a)的氨基酸序列有大于 70%，优选大于 80%，尤其优选大于 90%同源性的氨基酸序列。

除了 SEQ ID NO.2 或 4 中所描述的多肽，本发明还涉及突变蛋白、变体和其片段。由于单个氨基酸或短氨基酸片段的替换、缺失和/或插入，这些序列与 SEQ ID NO.2 或 4 中所描述的氨基酸序列不同。

术语“变体”包括子宫内膜异位蛋白天然存在的等位基因变体或拼接变体，以及重组 DNA 技术产生的蛋白（尤其是由化学合成寡核苷酸介导的体外诱变），这些蛋白就其生物和/或免疫活性而言，实质上相应于 SEQ ID NO.2 或 4 中所述的蛋白质。该术语还包括化学修饰的多肽。在肽末端和/或反应性氨基酸侧链基团中酰化例如乙酰化或酰胺化的多肽属于此类。代表 SEQ ID NO.2 或 4 所示氨基酸序列的至少 10 个氨基酸的片段的多肽片段（肽）也属于根据本发明的氨基酸序列。

本发明还涉及含有至少一个拷贝本发明的核酸的载体。该载体可以是任何携带根据本发明的 DNA 序列的原核或真核载体，其中 DNA 序列优

选连接表达信号如启动子、操作子、增强子等。原核载体的例子是染色体载体例如噬菌体和染色体外载体如质粒，尤其优选环状质粒载体。适合的原核载体的描述见例如 Sambrook 等（见上），第 1-4 章。尤其优选根据本发明的载体为真核载体，例如酵母载体，或适于高等细胞的载体例如质粒载体、病毒载体或植物载体。该类型的载体是分子生物学领域技术人员所熟知的，所以无需在此作进一步说明。具体地，该方面请参考 Sambrook 等，见上，第 16 章。

本发明还涉及含有 SEQ ID NO.1、3 或/和 5 所示序列或其组合至少 21 个核苷酸长的片段的载体。优选该片段的核苷酸序列来自于所述序列的蛋白编码区或该蛋白或多肽表达必需的区域。这些核酸尤其适合于制备治疗上可用的优选长达 50 个核苷酸的反义核酸。

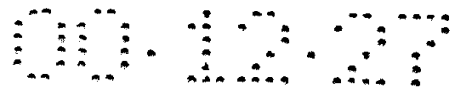
本发明进一步涉及转化了本发明的核酸或本发明的载体的细胞。该细胞可是真核也可以是原核细胞。核酸转化细胞的方法是一般的现有技术，因此无需进一步说明。细胞优选例如真核细胞，尤其是动物细胞，尤其优选哺乳动物细胞。

本发明进一步涉及抗该子宫内膜异位基因编码的多肽或其变体的抗体或该抗体的片段。该类抗体尤其优选针对所述基因编码的完整多肽或 SEQ ID NO.4 所述氨基酸序列的第 1-330 位氨基酸相应的肽序列。

根据本发明，特异地与侵袭过程相关特别是与子宫内膜异位相关的基因的鉴定、分离和表达提供了基于上述病症的疾病的诊断、治疗和预防的必要条件。

利用本发明的多肽或该多肽的片段作为免疫原可制备抗这些多肽的抗体。抗体的制备可按常规方法进行，即通过用该完整多肽或其片段免疫实验动物，之后获得产生的多克隆抗血清。根据 Köhler 和 Milstein 的方法和其发展的方法，可从实验动物的抗体生产细胞按已知方式利用细胞融合获得单克隆抗体。同样，可按已知方法生产人单克隆抗体。然后该类抗体可用于尤其是子宫内膜异位细胞组织的诊断检测或治疗。

例如，样品如体液，尤其是人体液（例如血、淋巴液或 CSF）可使用 ELISA 技术一方面检测子宫内膜异位基因编码的多肽的存在，另一方



面检测抗该多肽的自身抗体的存在。然后可使用特异抗体例如本发明的抗体在这些样品中检测子宫内膜异位基因编码的多肽或其片段。为检测自身抗体，优选可使用含有子宫内膜基因编码多肽的部分、域或甚至完整多肽、并与利于检测的蛋白域例如麦芽糖结合蛋白（MBP）相融合的重组融合蛋白。

诊断检测还可采用特异核酸探针在核酸水平上检测，例如在基因或转录本水平上进行。

本发明的核苷酸、氨基酸序列和抗体的提供进一步促进了该多肽/蛋白效应物的定向寻找。效应物是以抑制或激活方式作用于本发明的多肽并可选择性影响该多肽控制的细胞功能的试剂。然后这些试剂可用于适当病症例如基于侵袭过程的病症的治疗。因此本发明还涉及用于鉴定子宫内膜异位蛋白效应物的方法，其中将表达该蛋白的细胞与各种潜在效应物物质（如低分子量试剂）接触，然后分析细胞的改变，例如细胞激活、细胞抑制、细胞增殖和/或细胞遗传的改变。用该方法，还可鉴定子宫内膜异位蛋白的结合靶标。

由于许多肿瘤疾病伴有侵袭过程，根据本发明的基因的发现还提供了癌症诊断、预防和治疗的的可能性。

参与负责侵袭过程的基因的发现不仅开拓了治疗基于这种细胞改变的疾病的可能性，而且为使该过程有效，也可使用根据本发明的序列。这对于例如胚胎的植入可能是重要的。

因此本发明还涉及包含作为活性成分的如前所述的核酸、载体、细胞、多肽、肽和/或抗体的药物组合物。

根据本发明的药物组合物还可含有药理学常规载体、赋形剂和/或添加剂，而且，如果合适可含有其它的活性成分。该药物组合物可用于特别是侵袭过程相关的疾病的诊断、治疗或预防。而且本发明的组合物还可用于诊断这些疾病的易感性，特别是诊断子宫内膜异位危险。

本发明通过下列附图、序列和实施例来更详细说明。

图 1 图示了子宫内膜异位相关基因的 cDNA，其中仅显示了外显子 E1 至

E5.

SEQ ID NO. 1 代表包含编码子宫内膜异位相关基因的遗传信息的核苷酸序列，其中一个开放阅读框是第 3-911 位核苷酸，和

SEQ ID NO. 2 代表 SEQ ID NO.1 所示核苷酸的开放阅读框的氨基酸序列，其中该开放阅读框从第 1 位氨基酸延伸到第 302 位。

SEQ ID NO. 3 代表与 SEQ ID NO. 1 的核苷酸序列类似的核苷酸序列，但它在 5'端另外含有 84 个核苷酸，开放阅读框从第 3 位核苷酸延伸到第 995 位。

SEQ ID NO. 4 代表 SEQ ID NO.3 所示核苷酸序列开放阅读框的氨基酸序列，其中该氨基酸序列含有 320 个氨基酸，C 端 302 个氨基酸与 SEQ ID NO. 2 的氨基酸序列相同。

SEQ ID NO. 5 代表由所显示的 218 个核苷酸组成的可能存在的额外外显子 4a 的核苷酸序列，其中如果存在，外显子 4a 位于第 1054-1055 位核苷酸之间（相对于 SEQ ID NO. 3）。

实施例

实施例 1 细胞培养

为了鉴定子宫内膜异位相关基因，使用上皮子宫内膜异位细胞系 EEC145T⁺ 的侵袭和非侵袭细胞。将细胞在含有 10% 胎牛血清的 Dulbecco 培养基(DMEM)中培养，每周按 1:5 稀释 2 次（传代）。为通过 DDRT-PCR（见下）比较表达模式，使用传代 17 次的侵袭细胞和传代 33 次的非侵袭细胞。细胞用 SV40 转化并通过差异显示逆转录聚合酶链式反应（DDRT-PCR）进行分析。

实施例 2 DDRT-PCR

Liang 和 Pardee 开发的本方法是区分不同细胞类型的表达模式或一种细胞类型在不同生存条件下或处于不同发育阶段的表达模式变化的方法(Liang 和 Pardee (1992) , 科学 257, 967-971). DDRT-PCR 技术的基

础是每一个细胞中有约 15,000 个基因表达,而且原则上每一个单个 mRNA 分子均可通过逆转录和随机引物扩增来制备。

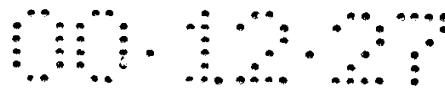
在本实施例中,首先采用几个不同 dT₁₁VX 引物(下游引物,锚定引物)将细胞 polyA⁺ RNA 转录成 cDNA。然后使用 Genhunter, Nashville (1994) RNA Map™ 试剂盒的 4 种下游引物和 20 种上游引物并加入放射性标记的核苷酸对获得的 cDNA 群体进行 PCR 扩增。扩增后将反应混合物真空浓缩,获得的 cDNA 片段在 6%天然 PAA(聚丙烯酰胺)凝胶中分级分离。用放射线自显影法进行 DNA 的检测。将待研究的两个细胞变体的条带模式显示显著不同的 PCR 混合物重复两次以检测重现性。如果前面发现的不同得到证实,则根据已知方法从凝胶中洗脱条带、再扩增、克隆和测序。

通过该方法发现一个侵袭细胞变体特异的 394 bp 片段(片段 1, SEQ ID NO. 1 所示核苷酸序列的第 1235-1628 位核苷酸,也见图 1)。该片段用作 Northern 杂交分析(见下)的探针。

实施例 3 在人 Northern 杂交中分析片段 1 的表达图谱

为了检测 DDRT-PCR 片段 1 的表达模式,进行 Northern 杂交。为此,将 20ug 总 RNA 或 4ug polyA⁺ RNA 在 1%变性琼脂糖凝胶中分级分离,并过夜转移至尼龙膜上。用 UV 光照射将 RNA 固定在膜上。在含甲醛的杂交液中用 ³²P 标记探针(用 Amersham 的 RPL 试剂盒标记)42 ℃ 杂交过夜。然后在递增严格的条件下洗膜直到斑点发射的放射性达到可测量强度。杂交模式通过放置在 X 光片(NEF-NEN, DuPont)上暴光数天来显现。为确定 DDRT-PCR 片段 1 的表达模式,用下列细胞或组织的 RNA 进行 Northern 印迹分析:

- 上皮子宫内膜异位细胞系 EEC145T⁺的侵袭细胞(传代 17 次)
- 上皮子宫内膜异位细胞系 EEC145T⁺的非侵袭细胞(传代 33 次)
- 腹膜细胞系 EEC143T⁺的细胞
- 子宫内膜组织
- 侵袭性人膀胱癌细胞系 EJ28 的细胞



- 非侵袭性人膀胱癌细胞系 RT112 的细胞

用 DDRT-PCR 片段 1 的探针杂交后检测到一个约 4kb 的 mRNA, 而且它仅可在子宫内膜异位细胞系 EEC145T⁺ 细胞的侵袭性变体中检测到。

还检测了其它人组织。在脾脏中发现明显与片段 1 杂交的 4kb mRNA, 在脑中分别发现长 4kb 和 > 9kb 的 mRNA 片段。

使用两个 Clontech 的人多组织 Northern (multiple tissue Northern, MTN) 按照厂家说明进行 Northern 印迹分析。在下列组织中检测表达: 结肠、小肠、心脏、脑、睾丸、肝、肺、脾、肾、卵巢、胰脏、外周血白细胞、胎盘、前列腺、骨骼肌、胸腺。用放射线标记 3' 探针 “DDRT-PCR 片段 1” 获得的表达模式表示如下:

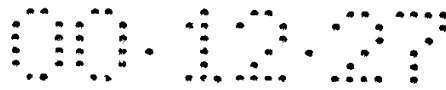
4 kb mRNA (期望大小): 脑、脾、胰脏

9.5 kb mRNA: 脑

在其它组织中没有检测到特异杂交。

原位杂交

为了阐述细胞表达模式, 对不同组织的 10 um 石蜡切片进行 mRNA 原位杂交。为此, 以 “DDRT-PCR 片段 1” 作为地高新配基标记的 RNA 探针。检测反应采用与碱性磷酸酶 (A) 偶联的地高新配基特异抗体进行。BM 紫作为 AP 的底物, 在去磷酸化之后它形成蓝色沉淀。结果列在下表中, 显示主要在侵袭/迁移细胞中表达。



强表达	弱、不很明显的表达
子宫内膜异位损伤的表皮细胞 癌 淋巴浸润物 胸腺 淋巴节生发中心(脾脏) 稍微弱些的: 子宫内膜上皮细胞 生血管性内皮细胞 迁移性神经细胞	骨骼肌 心脏 肉瘤

实施例 4 RT-PCR

RT-PCR (逆转录 PCR) 提供了一种测试表达模式的敏感方法。

为此,用 400U M-MLV 逆转录酶(Gibco-BRL)在 30ul 总体积中将 1ug 适当 polyA⁺ RNA 逆转录成 cDNA。取 1ul 该反应物用于随后用不同引物组合进行的 PCR。

所用 PCR 引物 P1 至 P7 见表 1 (也见图 1)。

表 1

序号	序列 (相对SEQ ID NO. 1的核苷酸位置)
P1	5'-CCAGCTGCTGCCAAATCC-3' (36-53)
P2	5'-CATCATGGTCATAGCTGC-3' (545-562)
P3	5'-AGCGTCTCATCGGTGTAC-3' (793-776, 反向引物)
P4	5'-AACAGAAGTGGTAGGTGC-3' (1080-1063, 反向引物)
P5	5'-AAAGGGACGGGAGGAAGC-3' (1243-1260)
P6	5'-CCAAAGTAGAAAACACTG-3' (1612-1595, 反向引物)
P7	5'-GCTTGTATGACACACACG-3' (2150-2133 反向引物)

RT-PCR 实验用不同细胞系和组织的 polyA⁺ RNA 和不同引物组合来

进行。结果见表 2。

表 2

PC	P17	P33	Per	EM	EJ28	RT112	E	EE	PEE
P1-P4	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P2-P6	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P5+P7	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P5+P6	+	-	-	+	-	-	+	+	+
P1+P3	+	-	-	+	-	-	+	+	+

PC = 引物组合

P17 = 子宫内膜异位细胞系 EEC145T, 传代 17 次, 侵袭

P33 = 子宫内膜异位细胞系 EEC145T, 传代 33 次, 非侵袭

Per = 腹膜细胞系 Perl43T

EM = 子宫内膜组织

EJ28 = 侵袭性膀胱癌细胞系

RT112 = 非侵袭性膀胱癌细胞系

E = 子宫内膜组织

EE = 子宫内膜异位患者的子宫内膜组织

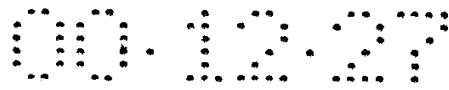
PEE = 腹膜子宫内膜异位活检

n.d. = 未测定

RT-PCR 结果证明在子宫内膜异位细胞系 EEC145T⁺ 的早期传代物 (17 代, 20 代) 中有片段 1 特异的表达。与 Northern 印迹分析不同, 在子宫内膜中显示可能也有弱表达。

用跨越内含子的引物进行 RT-PCR 分析

为了检测可能的其它外显子, 用跨越内含子的引物进行 RT-PCR 实验。可能在所述 mRNA 之外还存在至少一种 mRNA, 它在第 4 和第 5 外显子之间含有一个长 218bp 的外显子 (4a)。该外显子位于 3'-UTR (非



翻译区)，也就是说编码区之后。4a 外显子的序列如下。

```
gcggttgcc ggaatgccag tggctcctgg gcagatgtgc accccagatt
cagcctttgt gatagattcc aacacgttct ggcctcagac cacctttgtg
gtggggccag actgctctgg gcaaagtga gctggccttt atgctccaag
gaagggggcc tegagagcag gcctgcattg gctctcggac taattcgcga
tcattcttca tacagcag
```

可选择的外显子 4a 的核苷酸序列

实施例 5 cDNA 噬菌体文库 EEC14 的制备

按照 Short, J.M.等(1988)核酸研究(*Nucleic Acids Res.*)16: 7583-7600 的方法制备 cDNA 噬菌体文库 EEC14。

首先，进行上皮子宫内膜异位细胞系 EEC145T⁺侵袭性细胞（17 代）的 polyA⁺ RNA 的逆转录。所用引物包括 *XhoI* 切割位点和 18 个核苷酸长的 poly(dT)。将一含有 *EcoRI* 切割位点的接头与产生的 cDNA 片段连接。两个限制性位点允许 cDNA 片段直接插入 ZAP ExpressTM 载体。插入片段可以以抗卡那霉素的 pBK CMV 噬菌粒的形式从噬菌体中切割下来。

实施例 6 筛选噬菌体文库

按照厂家手册(Stratagene)，以 DDRT-PCR 片段 1（394bp）作为探针以筛选 10⁶ pfu（噬菌体形成单位）的 cDNA 噬菌体文库 EEC14。探针借助 PCR 用地高新配基(Boehringer Mannheim)标记。将感染细菌菌株 *XL 1blue MRF'*后形成的噬菌斑转移至尼龙膜上，并用上述探针在其上杂交。杂交的地高新配基标记探针按 Boehringer Mannheim 的化学发光手册进行检测。

选择阳性噬菌斑再次筛选。再次筛选的阳性克隆用于切割。通过 ExAssist 辅助噬菌体从噬菌体上切割载体部分获得了卡那霉素抗性的 pBK CMV 噬菌粒，该噬菌粒经过在细菌菌株 XL0LRTM 中扩增后可分

离和测序。分离的噬菌粒克隆 Q2A 含有 2.3 kb 大小的最长的插入片段，其序列已被测定并显示于 SEQ ID NO. 1。DDRT-PCR 片段 1 的序列是 SEQ ID NO. 1 第 1235-1628 位核苷酸。

实施例 7 Southern 印迹分析

用各种限制性内切酶切割 10ug 男女受检者的基因组 DNA。将片段在琼脂糖凝胶上分级分离并转移至尼龙膜上。用地高新配基标记的 DDRT-PCR 片段 1 对该膜杂交。

根据 Boehringer 手册用化学发光法检测杂交情况。在女性和男性 DNA 样品中使用各种限制酶均只检测到一个条带。该结果提示片段 1 所在基因是单一的非性别特异的基因。此后，用 DDRT-PCR 片段 1 分离了两个基因组克隆 PAC J1472 和 PAC N1977。

实施例 8 荧光原位杂交(FISH)

通过荧光原位杂交已将实施例 7 中获得的基因组克隆定位于 1 号染色体上 (1p36) (Lichter 等(1990), 科学 247:64-69)。

实施例 9 特异抗体的产生

通过适当限制性切割位点将上述 cDNA 序列的第 584-909 位核苷酸克隆至表达载体 pMAL cRI 中。为了表达该序列，将该结构转化至大肠杆菌 DH5 α 细胞中。翻译的蛋白质片段从 SDS 聚丙烯酰胺凝胶切割下来用于免疫兔。

实施例 10 RACE (cDNA 末端快速扩增)

因为 cDNA 克隆 Q2A (见实施例 6) 的长度与检测到的 mRNA 长度 (约 4 kb) 不同，所以进行 RACE 实验以获得更多的序列信息。借助该方法有可能从 mRNA 模板的已知内部序列和 5' 或 3' 末端的未知序列之间获得 cDNA 序列。克隆 Q2A 的 3' 端可通过从第 5 个外显子开始的 3' RACE 实验证实。

对于 5'RACE, 用与第 1 个外显子杂交的基因特异引物合成 cDNA 第一链, 然后用末端转移酶添加均聚核苷酸尾。添加的序列允许扩增位于基因特异引物和均聚核苷酸尾之间的序列区域。这使得可获得下列位于 Q2A 序列 5'端并属于第 1 个外显子的下列附加序列:

```

cc cgg ccg ccc cga gtg gag cgg atc cac ggg cag atg cag atg cct      47
  Arg Pro Pro Arg Val Glu Arg Ile His Gly Gln Met Gln Met Pro
    1           5           10           15

cga gcc aga cgg gcc cac agg ccc cgg gac cag gcg gcc gcc ctc gtg... 95
  Arg Ala Arg Arg Ala His Arg Pro Arg Asp Gln Ala Ala Ala Leu Val...
           20           25           30
  
```

加下划线的序列代表 Q2A 序列的头几个核苷酸, 它之前的序列是通过 5'RACE 获得的新序列。开放阅读框与已从片段推出的阅读框相符, 并含有两个推定的起始密码子(加下划线)。

前面获得序列的核苷酸序列见 SEQ ID NO. 1, 其 5'端额外的 84 个核苷酸见 SEQ ID NO. 3.

序列表

<110> Starzinski-Powitz, Anna 教授

<120>新的子宫内膜异位相关基因

<130>18505PWO 子宫内膜异位基因

<140>

<141>

<150>DE 198 24 230.1

<151>1998 - 05 - 29

<160>5

<170>PatentIn 2.1 版

<210>1

<211>2204

<212>DNA

<213>Homo sapiens

<220>

<221>CDS

<222> (3) .. (911)

cc	gcc	ctc	gtg	ccc	aag	gca	gga	ctg	gcc	aag	ccc	cca	gct	gct	gcc	47
	Ala	Leu	Val	Pro	Lys	Ala	Gly	Leu	Ala	Lys	Pro	Pro	Ala	Ala	Ala	
	1				5					10					15	
aaa	tcc	agc	cct	tcc	ctc	gcc	tct	tcg	tcc	tcg	tcc	tcg	tcc	tcc	gcg	95
Lys	Ser	Ser	Pro	Ser	Leu	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ala	
				20					25						30	
gtg	gcc	ggt	ggg	gcc	ccg	gag	cag	cag	gcc	ctc	ctg	agg	agg	ggc	aag	143
Val	Ala	Gly	Gly	Ala	Pro	Glu	Gln	Gln	Ala	Leu	Leu	Arg	Arg	Gly	Lys	
			35					40						45		
agg	cac	ctg	cag	ggg	gac	ggt	ctc	agc	agc	ttc	gac	tcc	aga	ggc	agc	191
Arg	His	Leu	Gln	Gly	Asp	Gly	Leu	Ser	Ser	Phe	Asp	Ser	Arg	Gly	Ser	
		50					55					60				
cgg	ccc	acc	aca	gag	act	gag	ttc	atc	gcc	tgg	ggg	ccc	acg	ggg	gac	239
Arg	Pro	Thr	Thr	Glu	Thr	Glu	Phe	Ile	Ala	Trp	Gly	Pro	Thr	Gly	Asp	
	65					70					75					
gag	gag	gcc	ctg	gag	tcc	aac	aca	ttt	ccg	ggc	ggt	tac	ggc	ccc	acc	287
Glu	Glu	Ala	Leu	Glu	Ser	Asn	Thr	Phe	Pro	Gly	Val	Tyr	Gly	Pro	Thr	
	80				85					90					95	
acg	gtc	tcc	atc	cta	caa	aca	cgg	aag	aca	act	gtg	gcc	gcc	acc	acc	335
Thr	Val	Ser	Ile	Leu	Gln	Thr	Arg	Lys	Thr	Thr	Val	Ala	Ala	Thr	Thr	
				100					105					110		
acc	acc	acc	acc	acg	gcc	acc	ccc	atg	acg	ctg	cag	act	aag	ggg	ttc	383
Thr	Thr	Thr	Thr	Thr	Ala	Thr	Pro	Met	Thr	Leu	Gln	Thr	Lys	Gly	Phe	
				115				120					125			

acc gag tcc ttg gat ccc cgg aga agg atc cca ggt ggg gtt agc aca 431
 Thr Glu Ser Leu Asp Pro Arg Arg Arg Ile Pro Gly Gly Val Ser Thr
 130 135 140

5 acg gag cct tcc acc agt ccc agc aac aac ggg gaa gtc acc cag ccc 479
 Thr Glu Pro Ser Thr Ser Pro Ser Asn Asn Gly Glu Val Thr Gln Pro
 145 150 155

cca agg att ctg ggg gag gcc tca ggt ctg gct gtc cat cag atc atc 527
 10 Pro Arg Ile Leu Gly Glu Ala Ser Gly Leu Ala Val His Gln Ile Ile
 160 165 170 175

acc atc acc gtc tcc ctc atc atg gtc ata gct gct ctc atc aca act 575
 15 Thr Ile Thr Val Ser Leu Ile Met Val Ile Ala Ala Leu Ile Thr Thr
 180 185 190

ctt gtc tta aaa aat tgc tgt gcc caa agc ggg aac act cgt cgg aac 623
 Leu Val Leu Lys Asn Cys Cys Ala Gln Ser Gly Asn Thr Arg Arg Asn
 195 200 205

20 agc cac cag cgg aag acc aac cag cag gag gag agc tgc cag aac ctc 671
 Ser His Gln Arg Lys Thr Asn Gln Gln Glu Glu Ser Cys Gln Asn Leu
 210 215 220

25 acg gac ttc ccc tcg gcc cgg gtg ccc agc agc ctg gac ata ttc acg 719
 Thr Asp Phe Pro Ser Ala Arg Val Pro Ser Ser Leu Asp Ile Phe Thr
 225 230 235

gcc tat aac gag acc ctg cag tgt tct cac gag tgc gtc agg gca tct 767
 30 Ala Tyr Asn Glu Thr Leu Gln Cys Ser His Glu Cys Val Arg Ala Ser
 240 245 250 255

gtg ccc gtg tac acc gat gag acg ctg cac tcg acg acg ggg gag tac 815
 35 Val Pro Val Tyr Thr Asp Glu Thr Leu His Ser Thr Thr Gly Glu Tyr
 260 265 270

aaa tcc aca ttt aat gga aac cga ccc tcc tct tct gat cgg cat ctt 863
 Lys Ser Thr Phe Asn Gly Asn Arg Pro Ser Ser Ser Asp Arg His Leu
 275 280 285

40 att cct gtg gcc ttc gtg tct gag aaa tgg ttt gaa atc tcc tgc tga 911
 Ile Pro Val Ala Phe Val Ser Glu Lys Trp Phe Glu Ile Ser Cys
 290 295 300

45 ctggccgaag tcttttttac ctctctggggg cagggcagac gccgtgtgtc tgtttcacgg 971

attccgttgg tgaacctgta aaaacaaaac aaacaaaaca aaacaaaaaa gacaaaacct 1031

50 aaaactgagc tatctaaggg ggagggtccc cgcacctacc acttctgttt gccggtggga 1091

aactcacaga gcaggacgct ctaggccaaa tctatttttg taaaaatgct catgcctatg 1151

gggtgactgcc ttctcccaga gttttctttg gagaacagaa agaagaaagg aaagaaagga 1211

55 accagaggca gagagacgag gataccacgc gaaagggacg ggaggaagca tccgaaacct 1271

aggattcgtc ctacgattct gaacctgtgc caataatacc attatgtgcc atgtactgac 1331

ccgaaaggct cggccacaga gccggggccc agcgaatcac gcagagaaat cttacagaaa 1391

60 acaggggtgg gaatctcttc cgatagagtc gctatttctg gttaatatac atatataaat 1451

atataaatac aaacacacac acacactttt tttgtactgt agcaattttt gaagatctta 1511

65 aatgttcctt tttaaaaaaa agaattgtgt tataggttac aaaatctgat ttatttaaca 1571

tgcttagtat gagcagaata aaccagtgtt ttctactttg gcaactcacg tcacacacat 1631

attacacaca tgtgcgcata cacacacaca atacacatat atgcatatag acgcatctat 1691
 tggaaatgca gttccacagg tgagcatggt ctttctgggtg acctgggtatt ccatcaccat 1751
 tcaccccagg ggacagcctc gaccgagaca aggaggccct taaatgacag cctgcatttg 1811
 ctagacgggtt ggtgagtggc atcaaatgtg tgacttacta tcttgggcca gaactaagaa 1871
 tgccaagggtt ttatatatgt gtgtatatat atatatatat atatatatat atgtttgtgt 1931
 gtgtatatat atatatatat atatatgttt gtgtgtgtat atatatgttt gtgtatatat 1991
 atacacatat gcatacatat gatttttttt ttttcattta agtggtggaa gatgctacct 2051
 aacagccacg ttcacattta cgtagctggt tgcttacaana cgggctgag cccctgggtg 2111
 ggtgggtggt ggattcttgg acgtgtgtgt catacaagca tagactggat taaagaagtt 2171
 ttccagttcc aaaaattaaa ggaatatatc ctt 2204

<210> 2

<211> 302

<212> PRT

<213> Homo sapiens

<400> 2
 Ala Leu Val Pro Lys Ala Gly Leu Ala Lys Pro Pro Ala Ala Ala Lys
 1 5 10 15
 30 Ser Ser Pro Ser Leu Ala Ser Ser Ser Ser Ser Ser Ser Ser Ala Val
 20 25 30
 Ala Gly Gly Ala Pro Glu Gln Gln Ala Leu Leu Arg Arg Gly Lys Arg
 35 35 40 45
 His Leu Gln Gly Asp Gly Leu Ser Ser Phe Asp Ser Arg Gly Ser Arg
 50 55 60
 40 Pro Thr Thr Glu Thr Glu Phe Ile Ala Trp Gly Pro Thr Gly Asp Glu
 65 70 75 80
 Glu Ala Leu Glu Ser Asn Thr Phe Pro Gly Val Tyr Gly Pro Thr Thr
 85 90 95
 45 Val Ser Ile Leu Gln Thr Arg Lys Thr Thr Val Ala Ala Thr Thr Thr
 100 105 110
 Thr Thr Thr Thr Ala Thr Pro Met Thr Leu Gln Thr Lys Gly Phe Thr
 115 120 125
 Glu Ser Leu Asp Pro Arg Arg Arg Ile Pro Gly Gly Val Ser Thr Thr
 130 135 140
 55 Glu Pro Ser Thr Ser Pro Ser Asn Asn Gly Glu Val Thr Gln Pro Pro
 145 150 155 160
 Arg Ile Leu Gly Glu Ala Ser Gly Leu Ala Val His Gln Ile Ile Thr
 165 170 175
 60 Ile Thr Val Ser Leu Ile Met Val Ile Ala Ala Leu Ile Thr Thr Leu
 180 185 190
 Val Leu Lys Asn Cys Cys Ala Gln Ser Gly Asn Thr Arg Arg Asn Ser
 195 200 205

	His	Gln	Arg	Lys	Thr	Asn	Gln	Gln	Glu	Glu	Ser	Cys	Gln	Asn	Leu	Thr
	210						215					220				
5	Asp	Phe	Pro	Ser	Ala	Arg	Val	Pro	Ser	Ser	Leu	Asp	Ile	Phe	Thr	Ala
225						230					235					240
	Tyr	Asn	Glu	Thr	Leu	Gln	Cys	Ser	His	Glu	Cys	Val	Arg	Ala	Ser	Val
					245					250					255	
10	Pro	Val	Tyr	Thr	Asp	Glu	Thr	Leu	His	Ser	Thr	Thr	Gly	Glu	Tyr	Lys
				260					265					270		
	Ser	Thr	Phe	Asn	Gly	Asn	Arg	Pro	Ser	Ser	Ser	Asp	Arg	His	Leu	Ile
			275					280					285			
15	Pro	Val	Ala	Phe	Val	Ser	Glu	Lys	Trp	Phe	Glu	Ile	Ser	Cys		
	290						295					300				

- <210> 3
- <211> 2288
- <212> DNA
- <212> Homo sapiens

- <220>
- <221> CDS
- <222> (3) .. (995)

<400> 3

30 cc cgg ccg ccc cga gtg gag cgg atc cac ggg cag atg cag atg cct 47
 Arg Pro Pro Arg Val Glu Arg Ile His Gly Gln Met Gln Met Pro
 1 5 10 15

cga gcc aga cgg gcc cac agg ccc cgg gac cag gcg gcc gcc ctc gtg 95
 35 Arg Ala Arg Arg Ala His Arg Pro Arg Asp Gln Ala Ala Ala Leu Val
 20 25 30

ccc aag gca gga ctg gcc aag ccc cca gct gct gcc aaa tcc agc cct 143
 40 Pro Lys Ala Gly Leu Ala Lys Pro Pro Ala Ala Ala Lys Ser Ser Pro
 35 40 45

tcc ctc gcc tct tcg tcc tcg tcc tcg tcc tcc gcg gtg gcc ggt ggg 191
 Ser Leu Ala Ser Ser Ser Ser Ser Ser Ser Ser Ala Val Ala Gly Gly
 50 55 60

45 gcc ccg gag cag cag gcc ctc ctg agg agg ggc aag agg cac ctg cag 239
 Ala Pro Glu Gln Gln Ala Leu Leu Arg Arg Gly Lys Arg His Leu Gln
 65 70 75

50 ggg gac ggt ctc agc agc ttc gac tcc aga ggc agc cgg ccc acc aca 287
 Gly Asp Gly Leu Ser Ser Phe Asp Ser Arg Gly Ser Arg Pro Thr Thr
 80 85 90 95

gag act gag ttc atc gcc tgg ggg ccc acg ggg gac gag gag gcc ctg 335
 55 Glu Thr Glu Phe Ile Ala Trp Gly Pro Thr Gly Asp Glu Glu Ala Leu
 100 105 110

gag tcc aac aca ttt ccg ggc gtt tac ggc ccc acc acg gtc tcc atc 383
 60 Glu Ser Asn Thr Phe Pro Gly Val Tyr Gly Pro Thr Thr Val Ser Ile
 115 120 125

cta caa aca cgg aag aca act gtg gcc gcc acc acc acc acc acc acc 431
 Leu Gln Thr Arg Lys Thr Thr Val Ala Ala Thr Thr Thr Thr Thr Thr
 130 135 140

65

```

acg gcc acc ccc atg acg ctg cag act aag ggg ttc acc gag tcc ttg 479
Thr Ala Thr Pro Met Thr Leu Gln Thr Lys Gly Phe Thr Glu Ser Leu
145 150 155

5 gat ccc cgg aga agg atc cca ggt ggg gtt agc aca acg gag cct tcc 527
Asp Pro Arg Arg Arg Ile Pro Gly Gly Val Ser Thr Thr Glu Pro Ser
160 165 170 175

acc agt ccc agc aac aac ggg gaa gtc acc cag ccc cca agg att ctg 575
10 Thr Ser Pro Ser Asn Asn Gly Glu Val Thr Gln Pro Pro Arg Ile Leu
180 185 190

ggg gag gcc tca ggt ctg gct gtc cat cag atc atc acc atc acc gtc 623
15 Gly Glu Ala Ser Gly Leu Ala Val His Gln Ile Ile Thr Ile Thr Val
195 200 205

tcc ctc atc atg gtc ata gct gct ctc atc aca act ctt gtc tta aaa 671
Ser Leu Ile Met Val Ile Ala Ala Leu Ile Thr Thr Leu Val Leu Lys
210 215 220

20 aat tgc tgt gcc caa agc ggg aac act cgt cgg aac agc cac cag cgg 719
Asn Cys Cys Ala Gln Ser Gly Asn Thr Arg Arg Asn Ser His Gln Arg
225 230 235

25 aag acc aac cag cag gag gag agc tgc cag aac ctc acg gac ttc ccc 767
Lys Thr Asn Gln Gln Glu Glu Ser Cys Gln Asn Leu Thr Asp Phe Pro
240 245 250 255

tcg gcc cgg gtg ccc agc agc ctg gac ata ttc acg gcc tat aac gag 815
30 Ser Ala Arg Val Pro Ser Ser Leu Asp Ile Phe Thr Ala Tyr Asn Glu
260 265 270

acc ctg cag tgt tct cac gag tgc gtc agg gca tct gtg ccc gtg tac 863
35 Thr Leu Gln Cys Ser His Glu Cys Val Arg Ala Ser Val Pro Val Tyr
275 280 285

acc gat gag acg ctg cac tcg acg acg ggg gag tac aaa tcc aca ttt 911
Thr Asp Glu Thr Leu His Ser Thr Thr Gly Glu Tyr Lys Ser Thr Phe
290 295 300

40 aat gga aac cga ccc tcc tct tct gat cgg cat ctt att cct gtg gcc 959
Asn Gly Asn Arg Pro Ser Ser Ser Asp Arg His Leu Ile Pro Val Ala
305 310 315

45 ttc gtg tct gag aaa tgg ttt gaa atc tcc tgc tga ctggccgaag 1005
Phe Val Ser Glu Lys Trp Phe Glu Ile Ser Cys
320 325 330

tcttttttac ctctctggggg cagggcagac gccgtgtgtc tgtttcacgg attccgttgg 1065
50 tgaacctgta aaaacaaaac aaacaaaaca aaacaaaaaa gacaaaacct aaaactgagc 1125

tatctaaggg ggaggggtccc cgcacctacc acttctgttt gccgggtggga aactcacaga 1185

55 gcaggacgct ctaggccaaa tctatttttg taaaaatgct catgcctatg ggtgactgcc 1245

ttctcccaga gttttctttg gagaacagaa agaagaaagg aaagaaagga accagaggca 1305

gagagacgag gatacccagc gaaagggacg ggaggaagca tccgaaacct aggattcgtc 1365
60 ctacgattct gaacctgtgc caataatacc attatgtgcc atgtactgac ccgaaaggct 1425

cggccacaga gccggggccc agcgaatcac gcagagaaat cttacagaaa acaggggtgg 1485

65 gaatctcttc cgatagagtc gctattttctg gttaatatac atatataaat atataaatac 1545

aaacacacac acacactttt tttgtactgt agcaattttt gaagatctta aatgttccct 1605

```

tttaaaaaaa agaattgtgt tataggttac aaaatctgat ttatttaaca tgcttagtat 1665
 gagcagaata aaccagtgtt ttctactttg gcaactcacg tcacacacat attacacaca 1725
 5 tgtgcgcata cacacacaca atacacatat atgcatatag acgcatctat tggaaatgca 1785
 gttccacagg tgagcatggt ctttctgggt acctggatt ccatcaccat tcaccccagg 1845
 ggacagcctc gaccgagaca aggaggccct taaatgacag cctgcatttg ctagacgggt 1905
 10 ggtgagtggc atcaaagtgt tgacttacta tcttgggcca gaactaagaa tgccaagggt 1965
 ttatatatgt gtgtatatat atatatatat atatatatat atgtttgtgt gtgtatatat 2025
 15 atatatatat atatatgttt gtgtgtgtat atatatgttt gtgtatatat atacacatat 2085
 gcatacatat gatttttttt ttttcattta agtgttggaa gatgctacct aacagccacg 2145
 ttcacattta cgtagctggt tgcttaca aa cgggctgag cccctgggtg ggtgggtggt 2205
 20 ggattcttgg acgtgtgtgt catacaagca tagactggat taaagaagtt ttccagttcc 2265
 aaaaattaa ggaatatatc ctt 2288

<210> 4

<211> 330

<212> PRT

<213> Homo sapiens

<400> 4

Arg	Pro	Pro	Arg	Val	Glu	Arg	Ile	His	Gly	Gln	Met	Gln	Met	Pro	Arg
1				5					10					15	
Ala	Arg	Arg	Ala	His	Arg	Pro	Arg	Asp	Gln	Ala	Ala	Ala	Leu	Val	Pro
			20					25					30		
Lys	Ala	Gly	Leu	Ala	Lys	Pro	Pro	Ala	Ala	Ala	Lys	Ser	Ser	Pro	Ser
		35					40					45			
Leu	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ala	Val	Ala	Gly	Gly	Ala
	50					55					60				
Pro	Glu	Gln	Gln	Ala	Leu	Leu	Arg	Arg	Gly	Lys	Arg	His	Leu	Gln	Gly
	65				70					75				80	
Asp	Gly	Leu	Ser	Ser	Phe	Asp	Ser	Arg	Gly	Ser	Arg	Pro	Thr	Thr	Glu
				85					90					95	
Thr	Glu	Phe	Ile	Ala	Trp	Gly	Pro	Thr	Gly	Asp	Glu	Glu	Ala	Leu	Glu
			100					105					110		
Ser	Asn	Thr	Phe	Pro	Gly	Val	Tyr	Gly	Pro	Thr	Thr	Val	Ser	Ile	Leu
		115					120					125			
Gln	Thr	Arg	Lys	Thr	Thr	Val	Ala	Ala	Thr	Thr	Thr	Thr	Thr	Thr	Thr
	130					135					140				
Ala	Thr	Pro	Met	Thr	Leu	Gln	Thr	Lys	Gly	Phe	Thr	Glu	Ser	Leu	Asp
	145				150					155					160
Pro	Arg	Arg	Arg	Ile	Pro	Gly	Gly	Val	Ser	Thr	Thr	Glu	Pro	Ser	Thr
				165					170					175	
Ser	Pro	Ser	Asn	Asn	Gly	Glu	Val	Thr	Gln	Pro	Pro	Arg	Ile	Leu	Gly
			180					185					190		

Glu Ala Ser Gly Leu Ala Val His Gln Ile Ile Thr Ile Thr Val Ser
 195 200 205
 Leu Ile Met Val Ile Ala Ala Leu Ile Thr Thr Leu Val Leu Lys Asn
 5 210 215 220
 Cys Cys Ala Gln Ser Gly Asn Thr Arg Arg Asn Ser His Gln Arg Lys
 225 230 235 240
 10 Thr Asn Gln Gln Glu Glu Ser Cys Gln Asn Leu Thr Asp Phe Pro Ser
 245 250 255
 Ala Arg Val Pro Ser Ser Leu Asp Ile Phe Thr Ala Tyr Asn Glu Thr
 260 265 270
 15 Leu Gln Cys Ser His Glu Cys Val Arg Ala Ser Val Pro Val Tyr Thr
 275 280 285
 Asp Glu Thr Leu His Ser Thr Thr Gly Glu Tyr Lys Ser Thr Phe Asn
 20 290 295 300
 Gly Asn Arg Pro Ser Ser Ser Asp Arg His Leu Ile Pro Val Ala Phe
 305 310 315 320
 25 Val Ser Glu Lys Trp Phe Glu Ile Ser Cys
 325 330

<210> 5

<211> 218

<212> DNA

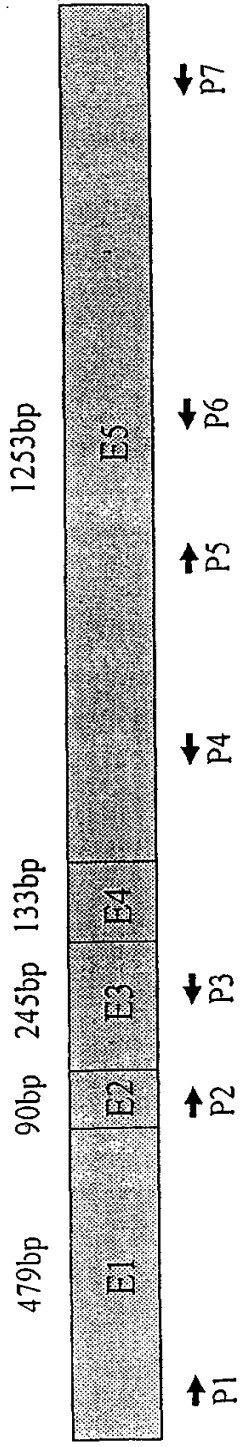
<213> Homo sapiens

<400> 5

gcgggttgcc ggaatgccag tggctcctgg gcagatgtgc accccagatt cagcctttgt 60
 gatagattcc aacacgttct ggcctcagac cacctttgtg gtggggccag actgctctgg 120
 gcaaagtgaa gctggccttt atgctccaag gaagggggcc tcgagagcag gcctgcattg 180
 gctctcggac taattcgca tcatctttca tacagcag 218

说明书附图

图 1
DDRT-PCR-
片段-1 (394bp)



E1-E5: 外显子1到5

图上方分别标出了各外显子的大小

P1-P7: PCR引物

表1中给出了这些引物的序列和确切位置

DDRT-PCR片段1序列的位置用黑色条表示。