



(12) 发明专利

(10) 授权公告号 CN 1871351 B

(45) 授权公告日 2010.06.23

(21) 申请号 200480031420.0

G01N 33/50(2006.01)

(22) 申请日 2004.08.25

C12R 1/645(2006.01)

(30) 优先权数据

60/498,318 2003.08.25 US

(56) 对比文件

CN 1261917 A, 2000.08.02, 说明书, 权利要求书.

(85) PCT申请进入国家阶段日

2006.04.24

WO 02/077223 A1, 2002.10.03, 说明书, 权利要求书.

(86) PCT申请的申请数据

PCT/IB2004/002963 2004.08.25

Monod michel 等人. Secreted proteases from pathogenic fungi. International journal of medical microbiology 292 5-6. 2002, 292(5-6), 405-419.

(87) PCT申请的公布数据

W02005/019251 EN 2005.03.03

Judith A Woodfolk 等人. Trichophyton antigens associated with IgE antibody mediated type hypersensitivity. Journal of biological chemistry 273 45. 1998, 273(45), 29489-29496.

(73) 专利权人 富诺齐梅生物技术股份有限公司  
地址 瑞士日内瓦

(72) 发明人 米歇尔·莫诺 雷托·斯托克林  
埃里克·格罗茨曼

(74) 专利代理机构 北京中誉威圣知识产权代理有限公司 11279

代理人 丛芳 彭晓玲

鲁衍强等人. 甲硫氨酰氨肽酶. 生命的化学 22 3. 2002, 22(3), 206-209.

审查员 邢云龙

(51) Int. Cl.

C12N 15/31(2006.01)

C12N 9/48(2006.01)

C12N 5/10(2006.01)

A61K 38/46(2006.01)

A61P 1/00(2006.01)

A61P 37/08(2006.01)

权利要求书 2 页 说明书 122 页 附图 10 页

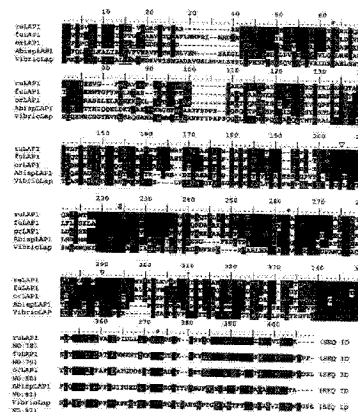
(54) 发明名称

一种新的真菌蛋白及其编码核酸

(57) 摘要

本发明涉及一种编码新的多肽酶的真菌核酸序列。同时也提供该核酸序列编码的多肽及其衍生物、变体或上述多肽的片段、多聚核苷或抗体。新的亮氨酸氨肽酶 (LAP) 和其他氨肽酶或羧肽酶多肽是指本发明所述的 EXOX 核酸和蛋白, 在医学、研究和商业应用上有着广泛的用途。

CN 1871351 B



1. 一种分离纯化的酶混合物,组成如下:
  - a) 一种分离纯化的二肽基肽酶 IV,由 SEQ ID NO :35 组成 ;以及
  - b) 一种分离纯化的氨肽酶,来自于一种皮肤真菌,该皮肤真菌选自于如下组 :絮状表皮癣菌,奥杜盎小孢子菌、铁锈色小孢子菌、同心性毛癣菌、卡那毛癣菌、麦格尼氏发癣菌、须癣发癣菌、罗比特斯克毛癣菌、红色毛癣菌、许兰氏毛癣菌、苏丹毛癣菌、断发毛癣菌、紫色毛癣菌、杨德氏发癣菌、犬小芽胞菌、马类小孢子菌、猪小孢子菌、桃色小孢子菌、马发癣菌、猴类毛癣菌、疣状毛癣菌、石膏样小孢子菌、阿耶罗毛癣菌和土生毛癣菌,该氨肽酶由成熟型的氨基酸序列组成,该氨基酸序列组成为 :
    - i) 一种分离纯化亮氨酸氨肽酶,由 SEQ ID NO :3 组成 ;或者是
    - ii) 一种分离纯化亮氨酸氨肽酶,由 SEQ ID NO :6 组成。
2. 一种分离纯化的酶混合物,组成如下 :
  - a) 一种分离纯化的二肽基肽酶 IV,由 SEQ ID NO :35 组成 ;以及
  - b) 一种分离纯化的氨肽酶,来自于一种皮肤真菌,该皮肤真菌选自于如下组 :絮状表皮癣菌,奥杜盎小孢子菌、铁锈色小孢子菌、同心性毛癣菌、卡那毛癣菌、麦格尼氏发癣菌、须癣发癣菌、罗比特斯克毛癣菌、红色毛癣菌、许兰氏毛癣菌、苏丹毛癣菌、断发毛癣菌、紫色毛癣菌、杨德氏发癣菌、犬小芽胞菌、马类小孢子菌、猪小孢子菌、桃色小孢子菌、马发癣菌、猴类毛癣菌、疣状毛癣菌、石膏样小孢子菌、阿耶罗毛癣菌和土生毛癣菌,该氨肽酶由成熟型的氨基酸序列组成,该氨基酸序列组成为 :
    - i) 一种分离纯化亮氨酸氨肽酶,由 SEQ ID NO :3 组成 ;和
    - ii) 一种分离纯化亮氨酸氨肽酶,由 SEQ ID NO :6 组成。
3. 如权利要求 1-2 中任一项所述的分离纯化的酶混合物,还含有一种或多种蛋白酶。
4. 如权利要求 3 所述的分离纯化的酶混合物,其中所述的一种或多种蛋白酶选自如下组 :胰蛋白酶、链霉菌蛋白酶、胰凝乳蛋白酶和蛋白酶 K。
5. 如权利要求 1-2 之任一项所述的分离纯化的酶混合物,其特征在于所述亮氨酸氨肽酶是红色毛癣菌亮氨酸氨肽酶 2。
6. 一种试剂盒,其中一个或多个容器内含有权利要求 1-5 之任一项所述的分离纯化的酶混合物。
7. 权利要求 1-5 之任一项所述的分离纯化的酶混合物的应用,用于蛋白水解 ;用于氨基酸分析时的多肽水解反应 ;用于去皮、脱毛、磨皮美容手术 ;用于假肢清洗和 / 或制备 ;用于织物柔软剂 ;用于香皂 ;用于对污染的水箱或其他含有要除掉蛋白的容器或要进行消毒的容器进行清洗或消毒。
8. 权利要求 1-5 之任一项所述的分离纯化的酶混合物在制备伤口清洗或伤口愈合的药物或治疗人类疾病综合症的药物中的应用,其中所述的疾病选自 :腹部疾病、消化道吸收不良、过敏反应、酶不足、或口炎性腹泻。
9. 如权利要求 8 所述的分离纯化的酶混合物的应用,其特征在于过敏反应是对麸质或其片段过敏。
10. 一种体外降解多肽底物的方法,所述方法包括将权利要求 1-5 之任一项所述的分

离纯化的酶混合物与多肽底物接触。

11. 如权利要求 10 所述的方法,其特征在于所述的分离纯化的酶混合物有序的酶解一个全长多肽底物或蛋白。

12. 如权利要求 10 或 11 所述的方法,其特征在于所述的多肽底物选自如下组:酪蛋白、麸朊、谷朊、小牛血清蛋白或其片段。

13. 如权利要求 10-11 之任一项所述的方法,其特征在于所述的多肽底物长度为 2 ~ 200 个氨基酸。

14. 一种依次从蛋白氨基 N 端去除一个或两个氨基酸的体外方法,包括将所述蛋白与权利要求 1-5 之任一项所述的分离纯化的酶混合物接触。

## 一种新的真菌蛋白及其编码核酸

### 技术领域

[0001] 本发明涉及一种新的具有独特催化性质的多肽酶及其编码核酸。具体的是，本发明涉及一种编码新的亮氨酸氨肽酶和其他氨基或羧基多肽酶的核酸，本发明中特指为 EXOX，还包括载体、宿主细胞、抗体，以及制造所述核酸和多肽酶的重组方法。所述基因可以从以下两种真菌中得到：红色毛癣菌 (*Trichophyton rubrum*)、烟曲霉菌 (*Aspergillus fumigatus*)。

### 背景技术

[0002] 细菌、酵母和丝状菌，还有些特殊的植物细胞，无脊椎动物和脊椎动物都会表达一种有利用吸收氨基酸、二肽或三肽的膜蛋白 (Lubkowitz, 等, *Microbiology* 143 : 387-396 (1997) ; Hauser 等, *Mol. Membr. Biol.* 18 (1) : 105-112 (2001) ; Stacey 等, *Trends Plant Sci.* 7 (6) : 257-263 (2002) ; Rubio-Aliaga & Daniel, *Trends Pharmacol. Sci.* 23 (9) : 434-440 (2002)。转运蛋白，如在酵母、丝状菌和植物细胞的，也能吸收大的寡肽 (4-5 个氨基酸)。在食品发酵工业中，早已证明微生物参与了蛋白被消化成氨基酸的过程。细菌属的乳酸菌 (*Lactobacillus*) (O'Cuinn 等, *Biochem. Soc. Trans.* 27 (4) : 730-734 (1999)) 和真菌属的曲霉菌 (*Aspergillus*) (Doumas 等, *Appl. Environ. Microbiol.* 64 : 4809-4815 (1998)) 分泌内部和外部蛋白酶，对蛋白消化很有用。

[0003] 氨基多肽酶的活性，也能在真菌被感染时发挥作用，早已证实通过 (De Bersaques & Dockx, *Arch. Belg. Dermatol. Syphiligr.* 29 : 135-140 (1973) ; Danew & Friedrich, *Mykosen* 23 : 502-511 (1980)) 菌丝体和培养基，然而，还没有氨基多肽酶或羧基多肽酶从皮肤真菌中分离并表征过。

### 发明内容

[0004] 本发明是基于一种分离的多肽的发明，所述多肽为含有选自 SEQ ID NOs : 3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的成熟多肽。本发明同时提供一种分离的含有选自 SEQ ID NOs : 3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的多肽，以及含有与上述序列具有 90% 同源性序列的分离多肽，所述多肽具有氨肽酶或羧肽酶活性。其中，所述多肽是亮氨酸氨肽酶中，如 ruLAP2。

[0005] 本发明另一方面提供含有一个或多个替换的保守氨基酸的分离多肽，所述多肽也具有氨肽酶活性。

[0006] 本发明同时包括自然发生的等位突变的选自 SEQ ID NOs : 3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的多肽。所述等位突变包括转录氨基酸序列的核酸序列有一个或多个核苷酸不同于选自 SEQ ID NOs : 3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列。在可选序列中任意的氨基酸变化的突变多肽是由保守替换提供的。

[0007] 本发明同时提供一种从多肽去除特定氨基酸的方法，例如从重组蛋白去除标签，其中具有去除氨基酸活性的多肽含有至少与选自 SEQ ID NOs : 3、6、9、12、15、18、21、24、27、

30、33 和 35 所示氨基酸序列 90% 同源性序列的多肽, 或具有类似功能的多肽。

[0008] 本发明所述的任一多肽均可天然产生。进一步, 所述多肽可以包含在含有试剂的组合物中, 所述组合物可用在含一个或多个试剂的试剂盒里。

[0009] 本发明同时提供一种含有所述多肽的皮肤真菌。例如, 适合的真菌包括: 絮状表皮癣菌 (*Epidermophyton floccosum*)、奥杜盎小孢子菌 (*Microsporum audouinii*)、铁锈色小孢子菌 (*Microsporum ferrugineum*)、同心性毛癣菌 (*Trichophyton concentricum*)、卡那毛癣菌 (*Trichophyton kanei*)、麦格尼氏发癣菌 (*Trichophyton megninii*)、须癣毛癣菌 (*Trichophyton mentagrophytes*)、罗比特斯克毛癣菌 (*Trichophyton raubitschekii*)、红色毛癣菌 (*Trichophyton rubrum*)、许兰氏毛癣菌 (*Trichophyton schoenleinii*)、苏丹毛癣菌 (*Trichophyton soudanense*)、断发毛癣菌 (*Trichophyton tonsurans*)、紫色毛癣菌 (*Trichophyton violaceum*)、杨德氏发癣菌 (*Trichophyton yaoundei*)、犬小孢子菌 (*Microsporum canis*)、马类小孢子菌 (*Microsporum equinum*)、猪小孢子菌 (*Microsporum nanum*)、桃色小孢子菌 (*Microsporum persicolor*)、马发癣菌 (*Trichophyton equinum*)、猴类毛癣菌 (*Trichophyton simii*)、疣状毛癣菌 (*Trichophyton verrucosum*)、石膏样小孢子菌 (*Microsporum gypseum*)、阿耶罗毛癣菌 (*Trichophyton ajelloi*) 和土生毛癣菌 (*Trichophyton terrestre*)。

[0010] 本发明同时提供一种含有本发明所述多肽的微生物培养上清液。

[0011] 本发明同时提供一种制备治疗相关病理的人类疾病药物中的应用, 所述的治疗包括本发明所述的多肽以及选择与本发明所述多肽相关的病理特征的疾病。

[0012] 本发明同时提供一种降解多肽底物的方法。所述方法包括将多肽底物与分离的一种或多种本发明多肽接触。其中, 多肽底物可以是一个全长的蛋白。进一步的, 其中一种或多种分离的多肽可以有序的分解多肽底物。所述的多肽底物可以是: 变性的酪蛋白、麸朊、谷朊、小牛血清蛋白或功能类似物。其中, 分离的多肽优选为亮氨酸氨肽酶中的 ruLAP2。

[0013] 本发明进一步提供一种验证具有潜在治疗真菌感染的治疗物质的方法, 其中所述真菌感染与本发明所述的多肽表达异常或与其异常的生理作用相关。所述方法包括提供一种能表达所述多肽且具有所述多肽特性或功能的细胞, 然后将包含候选物质的组合物与上述细胞接触, 再检测候选物质是否引起所述多肽功能特性的改变。其中, 当在有候选物质存在时观察到改变, 而在含有候选物质组合物的细胞接触时同样的改变没有出现, 那么该物质被证明具有潜在的治疗用途。其中, 所述功能性特征是多肽具有氨肽酶或羧肽酶活性的特征。

[0014] 本发明进一步提供一种治疗哺乳动物疾病的方法, 所述方法包括向动物给予一定量的多肽, 使其病证减轻。特别的是, 其中所述多肽为含有的氨基酸序列为选自 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示至少有 90% 同源的序列的多肽, 或其具有活性片段的多肽。所述的治疗病证包括真菌感染、腹部疾病、消化道吸收不良、口炎性腹泻、过敏反应和酶不足。优选的, 其中所述的过敏反应是对麸质过敏。

[0015] 本发明另外提供一种治疗哺乳动物疾病的方法, 所述方法包括向动物给予一定量的蛋白酶抑制剂, 使其病证减轻。其中所述蛋白酶抑制剂为含有的氨基酸序列为选自 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示至少有 90% 同源的序列的多肽, 或其具有活性片段的多肽。例如, 其中所述的疾病可以是真菌感染。

[0016] 本发明进一步提供一种分离的含有选自 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的多肽。所述多肽被一种细胞在诱导表达所述多肽的环境下产生。在一些具体实施方案中,所述细胞含有已重组了选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 和 34 所示核酸序列的核苷酸分子的载体。进一步的,所述细胞可以是真菌细胞、细菌细胞、昆虫细胞、植物细胞或哺乳动物细胞。

[0017] 本发明同时提供一种分离的含有选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 和 34 所示序列核酸序列的核苷酸分子。其中,所述核苷酸分子可以是自然存在的。

[0018] 本发明同时关于一种不同于一段核苷而是含有选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 和 34 所示序列,且编码成熟的具有选自 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的多肽的核酸序列。进一步,所述核苷酸分子可以是在严格的条件下与选自下组的核酸序列杂交,所述的组由选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32、34 或其互补序列物组成。在一些具体实施例中,所述核酸分子可以包含在载体内,进一步的可以包括一个与所述核酸分子功能相关的启动子。同时也提供含有所述载体的细胞。

[0019] 本发明同时提供制备本发明所述的多肽的方法。所述方法包括在诱导表达所述多肽的环境下培养一种细胞,其中所述细胞中含有已重组了选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 和 34 所示序列的核酸分子的载体。举某些例子,所述细胞可以是真菌细胞、细菌细胞、昆虫细胞、植物细胞或哺乳动物细胞。

[0020] 本发明同时提供一种制备蛋白的方法,包括在能够表达所述蛋白的环境下培养皮肤真菌,且从所述皮肤真菌中分离出来。其中,所述的蛋白是分泌型蛋白。同样的,所述的蛋白可以是氨肽酶或羧肽酶。特别地,所述氨肽酶是亮氨酸氨肽酶,如 ruLAP2。另外,所述的表皮真菌选自如下:絮状表皮癣菌 (*Epidermophyton floccosum*)、奥杜盎小孢子菌 (*Microsporum audouinii*)、铁锈色小孢子菌 (*Microsporum ferrugineum*)、同心性毛癣菌 (*Trichophyton concentricum*)、卡那毛癣菌 (*Trichophyton kanei*)、麦格尼氏发癣菌 (*Trichophyton megninii*)、须癣毛癣菌 (*Trichophyton mentagrophytes*)、罗比特斯克毛癣菌 (*Trichophyton raubitschekii*)、红色毛癣菌 (*Trichophyton rubrum*)、许兰氏毛癣菌 (*Trichophyton schoenleinii*)、苏丹毛癣菌 (*Trichophyton soudanense*)、断发毛癣菌 (*Trichophyton tonsurans*)、紫色毛癣菌 (*Trichophyton violaceum*)、杨德氏发癣菌 (*Trichophyton yaoundei*)、犬小孢子菌 (*Microsporum canis*)、马类小孢子菌 (*Microsporum equinum*)、猪小孢子菌 (*Microsporum nanum*)、桃色小孢子菌 (*Microsporum persicolor*)、马发癣菌 (*Trichophyton equinum*)、猴类毛癣菌 (*Trichophyton simii*)、疣状毛癣菌 (*Trichophyton verrucosum*)、石膏样小孢子菌 (*Microsporum gypseum*)、阿耶罗毛癣菌 (*Trichophyton ajelloi*) 和土生毛癣菌 (*Trichophyton terrestre*)。

[0021] 所述产物蛋白可以作用于多肽底物。例如,产物蛋白可以降解多肽或可以有顺序的分解一个全长的多肽底物。优选的,多肽底物可以是含有 2-200 个氨基酸的多肽。

[0022] 在一些实例中,所述产物蛋白添加一个或多个氨基酸到多肽底物上。另一些实例中,产物蛋白从多肽底物去除一个或多个氨基酸,使其成为一种被修饰多肽,然后所述的产物蛋白有序的将一个或多个氨基酸添加到被修饰多肽,这样形成一种含特殊氨基酸序列多

肽产物,与原多肽底物相比。

[0023] 本发明另外提供一种治疗真菌感染引起的疾病的方法,所述方法包括给予有效剂量的具有 EXOX 蛋白活性的一种抑制剂,其中抑制剂序列为选自 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列。例如,其中所述的 EXOX 蛋白含有 SEQ ID NO :2 所示序列。

[0024] 本发明进一步提供一种降解多肽底物的方法。所述方法包括将一个或多个本发明所述的分离的多肽与多肽底物接触。优选的,其中所述的多肽底物为全长的蛋白,其中所述的一个或多个分离的多肽有序的酶解多肽底物。所述的多肽底物选自如下组:变性的酪蛋白、麸朊、谷朊、小牛血清蛋白或功能类似物。进一步的在一些实例中,其中所述的分离的多肽是一种氨肽酶。所述的氨肽酶可以是亮氨酸氨肽酶,如 ruLAP2。

[0025] 另外,所述方法还可以用一个或多个蛋白酶与多肽底物接触。在一些实例中,所述的蛋白酶选自如下组:胰蛋白酶、链霉菌蛋白酶、胰凝乳蛋白酶和蛋白酶 K。

[0026] 本发明进一步提供一种去除蛋白氨基末端氨基酸的方法。所述方法包括将蛋白与本发明所述的分离的多肽接触。在一些实例中,所述的蛋白氨基末端带有组氨酸标签。在另一些实例中,所述的蛋白氨基末端带有 Xaa-Pro 标签。优选的,所述 Xaa 是一种含有至少两个相近亲核部位的氨基酸,如丝氨酸、苏氨酸或半胱氨酸。

[0027] 本发明进一步提供一种具有反蛋白水解活性的多肽。

[0028] 本发明进一步提供一种添加一个或多个氨基酸到底物多肽上的方法。方法包括将多肽底物与一种或多种本发明所述的分离的多肽接触。

[0029] 除非另有说明,本发明所述的所有技术和科学用语都和本发明所属一般技术人员理解的一样。尽管有类似或等同的方法和原料可以用来实现本发明,更适合的方法和原料将在以下描述。本发明引用所有的出版物、专利申请、专利及其他参考文献均参照其整体。在冲突的情况下,现在的说明书,包括定义将受到限制。另外,所述的原料、方法和实施例仅限于说明本发明而不是限制本发明。

## 附图说明

[0030] 图 1 是红色毛癣菌培养上清液用抗 -A.oryzue Alp(胶板 A,左边)和 Mep antisera(胶板 C,右边)Western 杂交结果图。胶板 B 显示用考马斯亮蓝染色的 10% SDS-PAGE 胶。泳道 1(lane 1)里,0.25ml 的红色毛癣菌培养上清液中的蛋白在加样到 SDS-PAGE 胶前预先用 TCA 处理。0.2g 纯化的 A.oryzue 重组蛋白 ALP 和 MEP 分别加样在泳道 2 和 3 里。标准的蛋白分子量对照在图左面。

[0031] 图 2 是显示 P.pastoris 产生的重组 ruLAP2(1,2)、fuLAP2(3,4)、ruLAP1(5,6)、fruLAP1(7,8)的 SDS-PAGE 蛋白特征图。每个纯化的重组 LAP 1g 加到 10% SDS-PAGE 胶。泳道 2、4、6 和 8 表示用 N-糖苷酶 F 处理后的去糖基蛋白。胶用考马斯亮蓝 R-250 染色。

[0032] 图 3 是红色毛癣菌培养上清液和重组 LAPs 作为对照探针,和抗 -ruLAP2(泳道 1-4)、抗 -ruLAP1 antisera(泳道 5-8)Western 杂交结果图。泳道 1、2、5 和 6 表示 0.25ml 的红色毛癣菌培养上清液中的蛋白在加样到 SDS-PAGE 胶前预先用 TCA 处理。0.1g 纯化的重组 ruLAP2(泳道 3、4)和 ruLAP1(泳道 7、8)作为对照加样。N-糖苷酶 F 用来去蛋白的糖基。标准的蛋白分子量对照在图左面。

[0033] 图4是红色毛癣菌 AMPP(氨基酶 P)在不同 pH 值下的酶活性图表。该图显示 AMPP 在很大的一个 pH 值范围内具有活性,范围为 pH6 ~ 11。

[0034] 图5是红色毛癣菌 AMPP(氨基酶 P)在不同温度下的酶活性图表。该酶在温度 25 ~ 60°C 间具有活性,优选的温度为 50°C。

[0035] 图6是显示在 37°C, 4h, (A) 不含 ruLP2 或 (B) 含 ruLAP2 以 E/S 比率 (w : w) 以 1/50 分解 gliadin 14mer 的结果图。

[0036] 图7是显示 (A) ru DPPIV 单独酶解或 (B) ruDPPIV/ruLAP2 的酶混和物分解 gliadin14mer 的结果图。

[0037] 图8是显示在 37°C, 4h, ruDPPIV1 以 E/S 比率 (w : w) 以 1/50 分解 gliadin 33mer 的结果图。

[0038] 图9是显示 DPPIV/ruLAP2 的酶混和物分解 gliadin 33mer 的结果图。

[0039] 图10A 和 10B 表示 ruLAP2 分解 Gly-Ser-proNPY (A) 前和 (B) 后的质谱图。

[0040] 图11A 和 11B 表示 ruLAP2 分解 Ala-proNPY (A) 前和 (B) 后的质谱图。

[0041] 图12A 和 12B 表示 ruLAP2 分解 TG47 (A) 前和 (B) 后的质谱图。

[0042] 图13A 和 13B 表示 DPPIV 分解 desMet-G-CSF (A) 前和 (B) 后的质谱图。

[0043] 图14是 M28E 亚家族氨基酶的推断的氨基酸序列比对图。

[0044] 图15是 M28E 亚家族氨基酶的推断的氨基酸序列比对图。

[0045] 本发明的详细描述

[0046] 本发明所述的蛋白酶与肽酶、蛋白水解酶和多肽水解酶同意。所述蛋白酶包括所有能够将 CO-NH 键断裂,将蛋白降解为多肽或氨基酸的酶。多肽外切酶作用于多肽链的氨基 (N) 或羧基 (C) 端邻近的基团。那些作用于 N 端释放一个氨基酸残基的酶命名为氨基酶。许多不同的高级蛋白酶参与了大量的生物生理反应。因此,现在选择新药的目标都放在控制相关多肽或酶的降解上。

[0047] 皮肤真菌是人类和动物的致病真菌,导致皮肤感染 (Vanbreuseghem 等, GUIDEPRATIQUE BE MYCOLOGUE MEDICALE ET VETERINAIRE. (1978); Kwong-Chong & Bennet, MEDICAL MYCOLOGY (1992); Weitzman & Summerbell, Clin. Microbiol. Rev. 8 : 240-259 (1995))。皮肤真菌包括但不限于: *T. ajelloi*, *A. uncinatum*, *K. ajelloi*, *T. asteroides*, *T. mentagrophytes*, *T. concentricum*, *T. cruris*, *E. floccosum*, *T. dankalienese*, *G. dankaliensis*, *T. equinum*, *T. equinum var. autotrophicum*, *T. equinum var. equinum*, *T. erinacei*, *T. fischeri*, *T. flavescens*, *T. floccosum*, *T. gloriae*, *T. gourvilii*, *T. granulare*, *T. granulatum*, *T. gypseum*, *T. inguinale*, *T. interdigitale*, *T. intertriginis*, *T. kanei*, *T. krajdennii*, *T. longifusum*, *T. megninii*, *A. quinckanum*, *A. benhamiae*, *A. vanbreuseghemii*, *T. pedis*, *T. proliferans*, *T. quickanum*, *T. radiolatum*, *T. mentagrophytes var. erinacei*, *T. mentagrophytes var. interdigitale*, *T. mentagrophytes var. mentagrophytes*, *T. mentagrophytes var. nodulare*, *T. mentagrophytes var. quinckeanum*, *T. niveum*, *T. nodulare*, *T. persicolor*, *M. persicolor*, *T. phaseoliforme*, *T. proliferans*, *T. purpureum*, *T. quinckeanum*, *T. radiolatum*, *T. raubitschekii*, *T. rubrum*, *S. ruber*, *T. schoenleinii*, *T. simii*, *A. simii*, *T. soudanense*, *T. sulphureum*, *T. tonsurans*, *A. insingulare*, *A. lenticularum*,



*A. quadrifidum*, *T. tonsurans*, *T. sulphureum*, *T. terrestre*, *T. tonsurans* var. *sulphureum*, *T. tonsurans* var. *tonsurans* subvar. *perforans*, *T. vanbreuseghemii*, *T. verrucosum*, *T. violaceum*, *T. yaoundei*, *E. floccosum*, *M. audouinii*, *M. ferrugineum*, *T. megninii*, *T. mentagrophytes*, *T. raubitschekii*, *T. schoenleinii*, *T. soudanese*, *T. violaceum*, *M. canis*, *M. equinum*, *M. nanum*, *M. persicolor*, *T. verrucosum*, 和 *M. gypseum*。在欧洲上述致病菌中,从医院和个人实例中分离得到的真菌,常见的有红色毛癣菌、须癣毛癣菌 (*T. mentagrophytes*)、犬小孢子菌 (*Microsporum canis*) (Monod 等, *Dermatology*, 205 :201-203, 2002)。实际上,皮肤真菌专门寄生在角质层、指甲或头发中,消化角质化细胞膜。至今,所有研究显示皮肤真菌产生具有体外活性的蛋白水解酶,许多研究者从个别菌种中发现一个或多个分泌型胞外蛋白酶。相关综述参见 Monod 等, *Int. J. Med. Microbiol.* 293 :405-419, 2002。特别地,犬小孢子菌拥有两个编码 S8 (枯草杆菌蛋白酶) 和 M36 (fungalsin) 家族的基因家族,参见 MEROPS 蛋白酶数据库 (<http://merops.sanger.ac.uk/>) (Brouta 等, *Infect. Immun.* 70 :5676-5683, 2002 ;Descamps 等, *J. Invest. Dermatol.* 70 :830-835, 2002)。分离的犬小孢子菌基因家族中其中之一编码从培养上清液中得到两个前述特征的胞内酶之一。(Mignon 等, *Med Mycol.* 36 :395-404, 1998 ;Brouta 等, *Med Mycol.* 39 :269-275, 2001)。两个酶都在感染的猫中发现,并表现出角质分解能力 (Mignon 等, *Med Mycol.* 36 :395-404, 1998)。所述蛋白水解酶的活性能使皮肤真菌专门在角质层、指甲或头发中生长,利用消化的角质化细胞膜,如单个氨基酸或小肽,作为细胞生长的营养来源。

[0048] 以下描述来自皮肤真菌红色毛癣菌的两个新的亮氨酸氨肽酶 :ruLAP1 和 ruLAP2。红色毛癣菌是发癣菌素属的一种,还包括 *T. ajelloi*, *T. asteroides*, *T. mentagrophytes*, *T. concentricum*, *T. cruris*, *T. dankalienese*, *T. equinum*, *T. equinum* var. *autotrophicum*, *T. equinum* var. *equinum*, *T. erinacei*, *T. fscheri*, *T. fla vesceus*, *T. floccosum*, *T. gloriae*, *T. gourvilii*, *T. granulare*, *T. granulorum*, *T. gypseum*, *T. inguinale*, *T. interdigitale*, *T. intertriginis*, *T. kanei*, *T. krajdennii*, *T. longfusum*, *T. megninii*, *T. pedis*, *T. proliferans*, *T. quickaneum*, *T. radiolatum*, *T. mentagrophytes* var. *erinacei*, *T. mentagrophytes* var. *interdigitale*, *T. mentagrophytes* var. *mentagrophytes*, *T. mentagrophytes* var. *nodulare*, *T. mentagrophytes* var. *quinckeanum*, *T. niveum*, *T. nodulare*, *T. persicolor*, *T. phaseoliforme*, *T. proliferans*, *T. purpureum*, *T. quinckeanum*, *T. radiolatum*, *T. raubitschekii*, *T. orthologues*。上述两个 LAPs 的特性与那些由机会型真菌烟曲霉的 fuLAP1 和 fuLAP2 直系同源性基因编码的分泌型酶、商业上可获得的从猪肾提取的微粒体 LAP (MEROPS > M1family) 相近。所有上述的酶具有亮氨酸氨肽酶活性。同时,烟曲霉的氨肽酶 fuLAP1 和 fuLAP2 表现出 70% 的氨基酸同源性和米曲霉的直系同源性基因,参考本发明引用的美国专利 US 6,127,161 和 5,994,113。进一步,ruLAP2 表现出独特的性质:(i) ruLAP1 和 ruLAP2 表现出 50% 的同源性,与烟曲霉直系同源性基因编码的 fuLAP1 和 fuLAP2、美国专利 US 6,127,161 和 5,994,113 中报道的米曲霉的 [xu1] 直系同源性基因;(ii) ruLAP2 和来源于毕氏酵母表达系统的类胰蛋白酶胞内酶的酶组合物能顺序酶解一个全长的多肽链,如变性的酪蛋白;(iii) ruLAP2 和 ruDPPIV (红色毛癣菌的另一个胞外酶) 的酶混合物降解一段麦醇溶蛋白,已知其能抵抗蛋

白酶水解,这样就提供了 ruLAP2 单独或与 ruDPPIV 组合用于治疗腹乳糜泻疾病或其他消化道疾病的证据,如吸收不良;(iv) ruLAP2 和其他蛋白酶组合(酶组合物)可应用于食品工业中,如降解苦味底物、降解 theves、处理肉类、香皂工业、分解朊病毒、分解病毒、降解有毒或副产物蛋白;(v) 和,既然 ruLAP2 和 / 或其他真菌分泌的蛋白酶是皮肤真菌在指甲角质化底物中生长必须的,那么 ruLAP2 和 / 或其他真菌分泌的蛋白酶的抑制剂将是治疗霉菌病的一种新方法。

[0049] 本发明提供一种新的真菌核酸和蛋白,所述蛋白具有亮氨酸氨肽酶活性。LAPs 在很多功能中都起到作用,包括但不限于血液凝固、控制细胞死亡、组织分化、肿瘤侵袭、一些致病微生物感染过程中等,所述作用使 LAPs 在新药领域中成为一种有价值的目标和有力的工具。除了在生理上的功能,氨肽酶还有商业应用价值,主要在清洁剂和食品领域。由于微生物有着丰富的生物多样性和对基因操作来说易于感染,微生物是蛋白酶的极佳来源,例如真菌。微生物降解蛋白并利用降解的产物作为生长的营养来源。因此,本发明所述的新的 LAPs 具有广泛的工业用途,包括但不限于用于:食品工业中水解蛋白;降解副反应产物(如羽毛);降解朊蛋白;蛋白质体学的降解蛋白;氨基酸分析时的多肽水解反应;伤口清洗(如攻击死组织);假肢清洗和 / 或准备;织物柔软剂;香皂;对污染的水箱或其他含有要除掉蛋白或消毒的容器(如保鲜桶、瓶子等)进行清洗或消毒;外科器械的清洗。

[0050] 本发明提供一种新的酶及其组合物,如相比与单个酶水解,多种酶的组合物容易分解不溶蛋白的结构,例如将角质细胞膜分解成小肽或游离氨基酸。事实上,除了 S8 和 M36 家族的胞外酶外,红色毛癣菌分泌的两个 LAPs 都各有不同的底物活性。ruLAP1 和 ruLAP2 都属于一样的 LAPs 家族(MEROPS > M28)。两个 LAPs 的特性与那些由机会型真菌烟曲霉的 fuLAP1 和 fuLAP2 直系同源性基因编码的分泌型酶、商业上可获得的从猪肾提取的微粒体 LAP(MEROPS > M1 family)相近。上述酶均具有亮氨酸氨肽酶活性。进一步的是,ruLAP2 具有独特的初级结构,同时具有一个特性:和 ruDPPIV 的酶混合物能有序的降解一段多肽链,如麦醇溶蛋白,已知其能抵抗其他蛋白酶的水解。更特别的是,纯化的 ruLAP2 和来源于毕氏酵母表达系统的类胰蛋白酶胞内酶的酶混合物物能顺序酶解一个全长的多肽链,如变性的酪蛋白。

[0051] 本发明部分是基于编码新的多肽的一段新的核酸序列。所述新的核酸序列和其编码的多肽指的是 ruLAP1、ruLAP2、fuLAP1 和 fuLAP2。所述核酸及其编码的多肽,本发明中统一简称为“EXOX”。

[0052] 本发明所述新的 EXOX 核酸包括表 1A、1B、2A、2B、3A、3B、4A、4B、5A、5B、6A、6B、7A、7B、8A、8B、9A、9B、10A、10B、11A、11B 和 12A 所示的序列或片段、衍生物、类似物或同源物。本发明所述新的 EXOX 蛋白包括表 1C、2C、3C、4C、5C、6C、7C、8C、9C、10C、11C 和 12B 所示的蛋白片段序列。下面将具体描述所述 EXOX 核酸和蛋白。

[0053] 同时,在本发明范围内还包括一种方法,所述方法是用蛋白酶抑制剂治疗或预防真菌感染或被真菌、酵母细胞和 / 或细菌感染的机会。

[0054] 利用反转录基因技术,红色毛癣菌分泌的两个亮氨酸氨肽酶的特性通过和真菌烟曲霉直系同源性基因编码的酶、从猪肾提取的微粒体氨酸氨肽酶 pkLAP 比较,得到了确定。所述四个真菌酶(ruLAP1、ruLAP2、fuLAP1 和 fuLAP2)被证实和 pkLAP 同样对底物 Leu-AMC 结构优先选择,同样具有亮氨酸氨肽功能。另外,亮氨酸氨肽酶 pkLAP 对 Leu-AMC

结构具有极高的选择性,因此也被称为丙胺酸氨肽酶 (MEROPS > M1.001)。

[0055] 本发明所述的编码 EXOX 蛋白的 EXOX 核酸,包括本发明提供的核酸序列及其片段。本 [xu1] 发明同时也包括相应的突变体或基于所示序列变换的仍可以编码具有 EXOX 类似活性或功能的蛋白的核酸,或核酸序列的片段。本发明进一步包括相应的与本发明所示序列互补的核酸,也包括核酸片段与本发明任意描述的序列互补。本发明同时包括核酸、核酸片段或其互补片段的化学修饰。所述修饰包括,但不限于如下所述,基团和核酸的糖磷酸骨架的修饰或衍生。运用上述修饰至少是为了提高被修饰核酸的化学稳定性,例如,修饰的核酸可以应用于相关治疗领域中的反义结合核酸技术。

[0056] 本发明所述的 EXOX 蛋白包括含有本发明所示序列的 EXO 蛋白。本发明同时还包括所示蛋白序列的突变体或相应残基的不同蛋白,其仍具有 EXO 类似活性或功能的蛋白,或蛋白的片段。本发明进一步还包括抗体或抗体片段,如  $F_{ab}$  或  $(F_{ab})_2$ ,所述抗体免疫特性的与本发明所述的蛋白结合。

[0057] EXOX 核酸和蛋白具有潜在的治疗应用,如治疗真菌感染。所述 EXOX 核酸、蛋白和抑制剂同时具有包括但不限于如下功能:(i) 改进蛋白产物的生物技术工具,如去除标签、含稀有氨基酸的产物;(ii) 某些领域疾病诊断的药物研发,如过敏反应(麸质过敏);(iii) 皮肤病药物研发,如抗真菌药物、疣的治疗、伤口愈合等;(iv) 美容术,如去皮、脱毛、磨皮和 dermaplaning;(v) 食品工业,如作为营养添加剂、糖精、通过预分解产生低变应原性的食品;(vi) 消毒工具,如净化蛋白污染物如朊蛋白或病毒;(vii) 清洁或回收废弃物,如羽毛、骨头、头发和毛皮;(viii) 清洗用品,如香波或液体清洁剂。

[0058] EXO 蛋白的抑制剂,特别是 ruLAP2 的抑制剂,可以作为抗真菌感染药物来治疗真菌感染病。LAPs 蛋白酶本身也可以作为治疗消化道疾病应用,如吸收不良或对麸朊引起过敏反应。麸 [xu2] 朊是小麦中发现的,谷蛋白和醇溶朊(醇溶谷蛋白)的蛋白混合物的专业名词。由于其内在的物理特性,如粘性和弹性,麸朊经常被用作食品添加剂。在食品质量控制和选择上检测麸朊的含量很重要,尤其是对那些麸朊过敏的个体,如麸朊过敏肠病、乳糜泻病、口炎性腹泻和相关的过敏反应,因此来自于小麦、裸麦,以及一些燕麦的食品中去除麸朊是很有必要的。

[0059] 胞外蛋白酶核酸及多肽

[0060] 本发明描述的红色毛癣菌亮氨酸氨肽酶活性及前人所研究的犬小孢子菌分泌的枯草杆菌蛋白酶和金属蛋白酶显示皮肤真菌分泌一系列蛋白酶,类似与那些在仅有碳和氮源的蛋白基质上的曲霉菌属。而且,ruDPPIV 和 ruDPPV;EMBL AF082514 的编码二肽-氨肽酶的 ruDPPV,与来自烟曲霉菌和米曲霉菌 DPPIV 和 DPPV 高度类似(Beauvais 等, J Biol. Chem. 272:6238-6244,1997;Beauvais 等, Immun. 65:3042-304,1997;Doumas 等, Appl. Environ. Microbiol. 64:4809-4815,1998;Doumas 等, J. Food Mycol. 2:271-279,1999),从基因组和红色毛癣菌的 cDNA 文库中分离得到。红色毛癣菌编码上述蛋白酶的基因中的内含子-外显子结构与从烟曲霉菌和米曲霉菌分离的同源基因类似。由于曲霉菌属与皮肤真菌属在分类上相近,出现上述结果就不奇怪了,同时它们属于同一个分类学上能产生闭囊壳原子囊(Eurotiomycete 分类)的子囊菌群。与编码枯草杆菌蛋白酶和 fungalysins 的基因相比,ruLAP1 和 ruLAP2 不是红色毛癣菌基因组中大基因家族的成员。

[0061] ruLAP1 显示 50% 的氨基酸同源性与 fuLAP1 和 / 或 LAP1 (参见表 19A 和图 14),

上述 3 个酶的结构都属于 M28E 亚家族,气单胞菌属和弧菌属的亮氨酸氨肽酶 (MEROPS > M28.002)。另外, ruLAP2 显示 50% 的氨基酸同源性与 fuLAP2 和 / 或 LAP2 (参见表 19B 和图 15), 上述 3 个酶的结构都属于 M28A 亚家族,有液泡的 *S. cerevisiae* 的蛋白酶 Y (MEROPS > M28.001), 灰色链霉菌分泌的氨肽酶 (MEROPS > M28.00X)。另外, M28A 和 M28E 亚家族的成员表现出很低的相似性。然而, 上述氨肽酶中的 2 个  $Zn^{++}$  结合位点的氨基酸很保守, 且在真菌 LAPs 被证明, 具体特征本发明详述 (表 20、21)。在灰色链霉菌和 *Aeromonas proteolytica* 分泌的氨肽酶中 His 和 Asp 残基结合一个  $Zn^{++}$  离子, 另两个 His 和 Glu 残基结合第二个  $Zn^{++}$  离子, 同时第二个 Asp 残基与两个  $Zn^{++}$  离子桥联 (Greenblat 等, *J. Mol. Biol.* 265 :620-636 (1997) ;Hasselgren 等, *J. Biol. Inorg. Chem.* 6 :120-127 (2001))。在灰色链霉菌氨肽酶中用不同的二价离子替换  $Zn^{++}$  受到  $Ca^{++}$  的影响, 且具有不同效果 (Ben-Meir 等, *Eur. J. Biochem* 212 :107-112 (1993) ;Lin 等, *J. Biol. Inorg. Chem.* 2 :744-749 (1997) ;Hasselgren 等, *J. Biol. Inorg. Chem.* 6 :120-127 (2001))。本发明所述的氨肽酶发现对不同的离子敏感性不同。与灰色链霉菌的氨肽酶一样, ruLAP2 与 fuLAP1 被  $Co^{++}$  高度激活。

[0062] 尽管有着很高的序列同一性, ruLAP2 与 fuLAP1 具有基本上不同的蛋白水解活性。特别的, ruLAP2 能够有效的水解 Asp- 和 Glu-7- 胺 -4- 甲基香豆素 (AMC), 同时 ruLAP2 到目前为止被证明是单一的 LAP, 能在 ruDPPIV 存在下水解能抵抗胃和胰腺的蛋白酶水解的麸朊, 或第二方面, 以来自 *P. pastoris* 表达系统的部分纯化的含有类胰酶胞外酶的形式消化一段全长多肽链, 如变性的酪蛋白。LAP 降解一段长的多肽的能力不能单独的从其分解氨酰 -AMC 残基基础上预测得到。皮肤真菌酶的部分特性已经在犬小孢子菌分泌的胞外酶上观察得到。与烟曲霉菌和米曲霉菌分泌的同源蛋白酶相比, 31.5kDa 犬小孢子菌枯草杆菌蛋白酶和 43.5kDa 犬小孢子菌金属蛋白酶都能够水解角朊。由于皮肤真菌从它们的在土壤中的习性进化后, 它们已发展出一套用特殊的蛋白酶分解角质组织进行感染的策略。ruLAP2 独特的特征能够表现在其高度特别的寄生的角质层和指甲的组织。

[0063] 除了上述 LAPs, 一系列新的从致病的真菌红色毛癣菌分离得到的蛋白酶将在下面详述。与 LAPs 一样, 上述蛋白酶都是胞外蛋白酶。它们包括: 两个羧肽酶、一个脯氨酸氨肽酶、一个氨肽酶 P、一个脯氨酸氨基酸酶和一个二肽基肽酶 IV。另外两个新的蛋白酶已经得到确认: 一个来自犬小孢子菌的亮氨酸氨肽酶 (caLAP1)、一个来自须癣毛癣菌的亮氨酸氨肽酶 (meLAP1)。

[0064] ruLAP2

[0065] ruLAP2 是红色毛癣菌的亮氨酸氨肽酶。1757 核苷酸 (SEQ ID NO :1) 的 ruLAP2 核酸如表 1A 所示。

[0066]

表 1A、ruLAP2 基因组核苷酸序列 (SEQIDNO :1)

```

ATGAAGTCGCAACTGTTGAGCCTGGCTGTGGCCGTCAACCATCTCCCAGGGCGTTGTTGGTCAAGAGCCCTTCGGATGGCCTTTCAA
GCCTATGGTCACTCAGGTGAGTTGCTCTCAACAGATCGATCGATCGATCTACCTTTGTCCCTGTACATCAAACCTCCAGCAGAGCCAAA
GAAACAGACACAATGTTCCCTGGGGAATTCTTATGGGCTAATGTAAATGTATAGGATGACCTGCAAAACAAGATAAAGCTCAAGGATATC
ATGGCAGGCGTCGAGAAGCTGCAAAGCTTTTCTGATGCTCATCCTGAAAAGAACCAGTGTGTTGGTGGTAATGGCCACAAGGACTGT
AGAGTGGATCTACAATGAGATCAAGGCCACTGGCTACTACGATGTGAAGAAGCAGGAGCAAGTACACCTGTGGTCTCATGCCAGGCTG
CTCTCAATGCCAATGGCAAGGACCTCAAGGCCAGCGCCATGTCTACAGCCCTCCTGCCAGCAAGATCATGGCTGAGCTTGTGTGTGCC
AAGAACAATGGCTGCAATGCTGTATGTGCCATACACTTTCTATACGTACATTCTCTCTAGAATGAAGAGCACGGGAGAACTAACTTTA
TGTATACAGACTGATTACCCAGCGAACACTCAGGGCAAGATCGTCCTCGTTGAGCGTGGTGTCTGCAGCTTCGGCGAGAAGTCTGCTCA
GGCTGGTGTGCAAAGGCTGTGGTGCCATTGTCTACAACAACGTCCCGGATCCCTTGTGGCACTCTTGGTGGCCTTGACAAGCGCC
ATGTCCAACCGCTGGTCTTTCCAGGAGGATGAAAAGAACCTTGCTACCCTCGTTGCTTCTGGTAAGATTGATGTCACCATGAACGTT
ATCAGTCTGTTTGAAGAACCAACCGTAAGTAGCTCAACGGCTGATCCAGCATCAATTGTCTCGAGTATATACTAAATCGATACCTCA
TAGCTGGAACGTCAATTGCTGAGACCAAGGGAGGAGACCACAACAACGTTATCATGCTCGGTGCTCACTCCGACTCCGTCGATGCCGGCC
CTGGTATTAACGACAACGGCTCGGGCTCCATTGGTATCATGACCGTTGCCAAAGCCCTCACCAACTTCAAGCTCAACAACGCCGTCGCG
TTTGCCCTGGTGGACCGCTGAGGAATTCGGTCTCCTTGAAGCACCTTCTACGTCAACAGCCTCGATGACCGTGAGCTGCACAAGGTCAA
GTTGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCCAACAGATCTACGACGGTGACGGTTCGGCCTACAACATGACCGGCC
CCGCTGGCTCTGCTGAAATCGAGTACCTGTTTCGAGAAGTCTTTGACGACCAGGGTATCCACACCAGCCCACTGCCTTCACTGGCCGA
TCCGACTACTCTGCTTTCATCAAGCGCAACGTGCCCGCTGGCGGCCTTCTACTGGAGCCGAGGTTGTCAAGACCCCGAGCAAGTCAA
GTTGTTTCGGTGGTGGGCTGGCGTTGCCATGACAAGAATAACATCGCAAGGGGACACCGTTGCCAACATCAACAAGGGAGCTATCT
TCCTTAACACTCGAGCCATCGCTACGCTATCGCCGAGTATGCCCGATCCCTCAAGGATTCCCAACCCGCCAAAGACCGGCAAGCGT
GACGTCAACCCCAAGTATTCTAAGATGCCTGGTGGTGGCTGCGGACACCACACTGTCTTCATGTAA

```

[0067] 1488 个核苷酸的 ruLAP2 开放阅读框 (ORF), 其中起始子 ATG 在第一个位置上 (表 1B 中下划线)。

[0068]

表 1B、ruLAP2 核苷酸序列 (SEQIDNO :2)

```

ATGAAGTCGCAACTGTTGAGCCTGGCTGTGGCCGTCAACCATCTCCCAGGGCGTTGTTGGTCAAGAGCCCTTCGGATGGCCTTTCAA
GCCTATGGTCACTCAGGATGACCTGCAAAACAAGATAAAGCTCAAGGATATCATGGCAGGCGTCGAGAAGCTGCAAAGCTTTTCTGATG
CTCATCCTGAAAAGAACCAGTGTGTTGGTGGTAATGGCCACAAGGACTGTAGAGTGGATCTACAATGAGATCAAGGCCACTGGCTAC
TACGATGTGAAGAAGCAGGAGCAAGTACACCTGTGGTCTCATGCCGAGGCTGCTCTCAAATGCCAATGGCAAGGACCTCAAGGCCAGCGC
CATGTCCTACAGCCCTCCTGCCAGCAAGATCATGGCTGAGCTTGTGTTGCCAAGAACAATGGCTGCAATGCTACTGATTACCCAGCGA
ACACTCAGGGCAAGATCGTCCTCGTTGAGCGTGGTGTCTGCAGCTTCGGCGAGAAGTCTGCTCAGGCTGGTGTGCAAAGGCTGTGGT
GCCATTGCTACAACAACGTCCCGGATCCCTTGTGGCACTCTTGGTGGCCTTGACAAAGCCCATGTCCAACCGCTGGTCTTTCCCA
GGAGGATGGAAGAACCTTGCTACCCCTCGTTGCTTCTGGTAAGATTGATGTCACCATGAACGTTATCAGTCTGTTTGAAGCCGAACCA
CCTGGAACGTCATTGCTGAGACCAAGGGAGGAGACCACAACAACGTTATCATGCTCGGTGCTCACTCCGACTCCGTCGATGCCGGCCCT
GGTATTAACGACAACGGCTCGGGCTCCATTGGTATCATGACCGTTGCCAAAGCCCTCACCAACTTCAAGCTCAACAACGCCGTCGCGCTT
TGCCCTGGTGGACCGCTGAGGAATTCGGTCTCCTTGAAGCACCTTCTACGTCAACAGCCTCGATGACCGTGAGCTGCACAAGGTCAAGT
TGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCCAACAGATCTACGACGGTGACGGTTCGGCCTACAACATGACCGGCCCC
GCTGGCTCTGCTGAAATCGAGTACCTGTTTCGAGAAGTCTTTGACGACCAGGGTATCCACACCAGCCCACTGCCTTCACTGGCCGATC
CGACTACTCTGCTTTCATCAAGCGCAACGTGCCCGCTGGCGGCCTTCTACTGGAGCCGAGGTTGTCAAGACCCCGAGCAAGTCAAGT
TGTTTCGGTGGTGGGCTGGCGTTGCCATGACAAGAATAACATCGCAAGGGGACACCGTTGCCAACATCAACAAGGGAGCTATCTTC
CTTAACACTCGAGCCATCGCTACGCTATCGCCGAGTATGCCCGATCCCTCAAGGATTCCCAACCCGCCAAAGACCGGCAAGCGTGA
CGTCAACCCCAAGTATTCTAAGATGCCTGGTGGTGGCTGCGGACACCACACTGTCTTCATGTAA

```

[0069] 上述 ruLAP2 核苷酸序列 (SEQ ID NO :2) 编码有 495 个氨基酸残基 (SEQ ID NO :3) 的蛋白, 在表 1C 中以一个字母的氨基酸代码表示。

[0070]

表 1C、ruLAP2 编码的蛋白序列 (SEQIDNO :3)

```
MKSQLLSLAVAVTTISQGVGQEPFGWPFKPMVTQDDLQNKIKLKDIMAGVEKLQSFSDAHPEKNRVFGGNGHKDTVEWYIYNEIKATGY
YDVKKQEQVHLWSHAEEAALNANGKDLKASAMSYSPASKIMAEELVAKNNGCNATDYPANTQGKIVLVERGVCSPFGEKSAQAGDAKAAG
AIVYNNVPGSLAGTLGGLDKRHVPTAGLSQEDGKNLATLVASGKIDVTMNVISLFENRTTWNVIAETKGGDHNNVIMLGAHSDSVDAGP
GINDNGSGSIGIMTVAKALTNFKLNNAVRFAWWTAEFGLLGSTFYVNSLDDRELHKVKLYLNFDMIGSPNFANQIYDGDGSAYNMTGP
AGSAEIEYLFKFFDDQGIHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTPEQVKLFGGEAGVAYDKNYHRKGDVTVANINKGAIF
LNTRAIAYAIAEYARSLKGFPTRPKTGKRVDVNPQYSKMPGGGCGHHTVFM
```

[0071] 上述 ruLAP2 具有与表 1D、1E 和 1F 中的 BLAST 氨基酸序列同源性。

[0072] 软件操作应用如下：

[0073] tblastn- 将蛋白“序列 1”和核酸“序列 2”比较, 其中核苷酸序列所有 6 阅读框转录。

[0074] blastx- 将核酸“序列 1”和蛋白“序列 2”比较。

[0075] blastp- 蛋白与蛋白比较。

[0076] 这里所述的 BLAST 比对结果中, “E 值”或“期待”值是一种可能性的数据显示, 表示比对序列可能在查询的数据库中与 BLAST 查询序列具有偶然的类似性。期待值 (E) 是一个描述找到可能“期待”看有机会与查询的数据库中的特殊位点吻合的参数。它被分值 (S) 减少其指数, S 是两个序列间配对情况。必要的, E 值描述了序列间配对存在的随机背景误差。

[0077] 表 1D. ruLAP2 TBLASTN 结果

[0078]

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi469363	Saccharomyces cerevisiae 氨肽酶 Y 基因	32421	170/477 (35%)	239/437 (55%)	8e-65
gi15839805	Mycobacterium tuberculosis CDC15551, 整个 280 个基因组中的 33 片段	18857	152/424 (35%)	225/424 (53%)	5e-57
gi9949032	Pseudomonas aruginosaPA01, 整个 529 个基因组中的 281 片段 氨肽酶 Y 基因	12547	129/317 (40%)	180/317 (56%)	1e-56

[0079] 表 1E. ruLAP2 BLASTX 结果

[0080]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi28918599	未知蛋白 / Neurosporacrassa	508	219/467 (46%)	287/467 (61%)	e-112
gi584764	APE3YEAST ; 氨肽酶前体 / Saccharomyces cerevisiae	537	170/477 (35%)	239/437 (53%)	1e-65
gi9949032	未知蛋白 / Thermobifidafusca	514	151/460 (32%)	237/460 (51%)	5e-61
gi15839805	水解酶 / Mycobacterium tuberculosis CDC15551	493	152/424 (35%)	225/424 (53%)	6e-58

[0081] 表 1F. ruLAP2 BLASTP 结果

[0082]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi28918599	未知蛋白 / Neurosporacrassa	508	219/467 (46%)	287/467 (61%)	e-105
gi584764	APE3YEAST ; 氨肽酶前体 / Saccharomyces cerevisiae	537	169/477 (35%)	237/477 (49%)	2e-64
gi15839805	水解酶 / Mycobacterium tuberculosis CDC15551	493	152/424 (35%)	225/424 (53%)	5e-57
gi123017467	未知蛋白 / Thermobifidafusca	514	150/460 (32%)	237/460 (51%)	1e-56

[0083] ruLAP1

[0084] ruLAP1 是红色毛癣菌的亮氨酸氨肽酶。1256 个核苷酸 (SEQ ID NO :4) 的 ruLAP1

核酸如表 2A 所示。

[0085]

表 2A、ruLAP1 基因组核苷酸序列 (SEQIDNO :4)

```
ATGAAGCTCCTCTCTGTTCTTGCCTGAGCGCTACCGCTACCTCCGTCCTCGGAGCTAGCATTCTGTTGATGCCCGGGCCGAGAAGTT
CCTCATCGAACTTGCCCTGGTGAGACTCGCTGGGTACCGAGGAGGAGAAGTGGGAGCTTAAGCGGGTATGTACCACATCCTACGCA
AAAGTTGATTTTCACTAGATAATATTGGTTATTAACACCCATTCTAGAAGGGTCAAGACTTCTTTGACATCACTGACGAGGAGGTTGG
ATTCACTGCTGCTGTTGCACAGCCAGCCATTGCCTACCCAACCTCCATCCGCCATGCTAATGCTGTAAACGCCATGATTGCTACCCTCT
CCAAGGAGAACATGCAGCGCATCTGACCAAGCTCAGCTCGTTCCAAACCGCTTACTATAAGGTTGACTTTGGCAAGCAGTCTGCCACC
TGGCTCCAGGAGCAAGTCCAGGCTGCCATCAATACCGCTGGTGCCAAATCGCTACGGAGCCAAGGTCGCCAGCTTCCGACACAACCTCGC
TCAGCACAGCATCATTGCCACTATTCGCGCCGCTCCCTGAAAGTCGTTGTCGTCGGTGTCAACCAAGACAGCATCAACCAACGCAGCC
CCATGACCGGCCGCGCTCCAGGTGCCATGACAACGGCAGTGGCTCCGTCACCATCCTTGAGGCCCTCCGTGGTGTCTCCGGGACCAG
ACCATCCTCCAGGGCAAGGCTGCCAACACCATTTGAGTTCCTGTTACGCGGTGAGGAAGCTGGTCTTCTGGGCTCCAGGCCATCTT
CGCCAATAACAACAGACCGGCAAGAAGGTCAGGGCATGCTCAACCAGGACATGACCGGTTACATCAAGGGAATGGTCGACAAGGGTC
TCAAGGTGTCCTTCGGTATCATCACCGACAACGTCACCGTAACTTGACCAAGTTCGTCGCCATGGTCATCACCAGGTAAGCTTCAAC
TCTTGATAAATATATTTTTTCATCGATGAAATGATGTCCTAATAATGCTTAAAGTACTGCTCAATCCCAACCATCGACACCCGCTGCGGCT
ATGCTTGCTGCTGACCACGCTCTGCCAACCGCAATGGCTACCCATCTGCCATGGTTGCCGAGTCTCCCATCGATCTCCTCGACCCCTCAC
CTCCACACTGACTCTGACAACATTAGCTACCTCGACTTCGACCACATGATCGAGCACGCTAAGCTCATTGTGCGGCTTCGTCAGCTGAGCT
CGCTAAGTAA
```

[0086] 1122 个核苷酸的 ruLAP1 开放阅读框 (ORF), 其中起始子 ATG 在第一个位置上 (表 2B 中下划线)。

[0087]

表 2B、ruLAP1 核苷酸序列 (SEQIDNO :5)

```
ATGAAGCTCCTCTCTGTTCTTGCCTGAGCGCTACCGCTACCTCCGTCCTCGGAGCTAGCATTCTGTTGATGCCCGGGCCGAGAAGTT
CCTCATCGAACTTGCCCTGGTGAGACTCGCTGGGTACCGAGGAGGAGAAGTGGGAGCTTAAGCGGAAGGGTCAAGACTTCTTTGACA
TCACTGACGAGGAGGTTGATTCACTGCTGCTGTTGCACAGCCAGCCATTGCCTACCCAACCTCCATCCGCCATGCTAATGCTGTAAAC
GCCATGATTGCTACCCTCTCCAAGGAGAACATGCAGCGCATCTGACCAAGCTCAGCTCGTTCCAAACCGCTTACTATAAGGTTGACTT
TGGCAAGCAGTCTGCCACCTGGCTCCAGGAGCAAGTCCAGGCTGCCATCAATACCGCTGGTGCCAAATCGCTACGGAGCCAAGGTCGCCA
GCTTCCGACACAACCTTCGCTCAGCACAGCATCATTGCCACTATTCGCGCCGCTCCCTGAAAGTCGTTGTCGTCGGTGTCAACCAAGAC
AGCATCAACCAACGCAGCCCATGACCGGCCGCGCTCCAGGTGCCATGACAACGGCAGTGGCTCCGTCACCATCCTTGAGGCCCTCCG
TGGTGTTCCTCCGGGACCAGACCATCCTCCAGGGCAAGGCTGCCAACACCATTTGAGTTCCTACTGGTACGCCGGTGAGGAAGCTGGTCTTC
TGGGCTCCAGGCCATCTTCGCAACTACAACAGACCGGCAAGAAGGTCAGGGCATGCTCAACCAGGACATGACCGGTTACATCAAG
GGAATGGTCGACAAGGGTCTCAAGGTGTCCTTCGGTATCATCACCGACAACGTCACCGTAACTTGACCAAGTTCGTCGCCATGGTCAT
CACCAAGTACTGCTCAATCCCAACCATCGACACCCGCTGCGGCTATGCTTGCTGCTGACCACGCTCTGCCAACCGCAATGGCTACCCAT
CTGCCATGGTTGCCGAGTCTCCCATCGATCTCCTCGACCCCTCACTCCACACTGACTCTGACAACATTAGCTACCTCGACTTCGACCAC
ATGATCGAGCACGCTAAGCTCATTGTGCGGCTTCGTCAGCTGAGCTCGCTAAGTAA
```

[0088] 上述 ruLAP1 核苷酸序列 (SEQ ID NO :5) 编码有 377 个氨基酸残基 (SEQ ID NO :6) 的蛋白, 在表 2C 中以一个字母的氨基酸代码表示。

[0089]



表 2C、ruLAP1 编码的蛋白序列 (SEQIDNO :6)

```

MKLLSVLALSATATSVLGASIPVDARAЕКFLIELAPGETRWVTEEEKWELKRRKGQDFFDITDEEVGFTAAVAQPAIAYPTSIRHANAVN
AMIATLSKENMQRDLTKLSSFQYAYKVDVFGKQSATWLQEQQVAAINTAGANRYGAKVASFRHNFAQHSI IATIPGRSPEVVVGAHQD
SINQRSPMTGRAPGADDNNGSGSVTILEALRGVLRDQTILQGKAANTIEFHWHYAGEEAGLLGSQATFANYKQTGKKVKGMLNQDMTG YIK
GMVDKGLKVSFGIITDNNANLTKFVRMVITKYCSIPTIDTRCGYACSDHASANRNGYPSAMVAESPIDLLDPLHHTSDNISYLDFFDH
MIEHAKLIVGFVTELAK

```

[0090] 上述 ruLAP1 具有与表 2D、2E 和 2F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0091] 表 2D. ruLAP1 TBLASTN 结果

[0092]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi1762234	Polyketide synthasePKSL2/	9894	131/247 (53%)	171/247 (69%)	1e-95
	Aspergillus parasiticus		40/76 (52%) 20/24 (83%)	57/76 (75%) 22/24 (91%)	
> gi23393798	亮氨酸氨肽酶 (Lap1) /Aspergillus sojiae	2547	77/159 (48%) 63/148 (42%) 14/30 (46%)	97/159 (61%) 89/148 (60%) 23/30 (76%)	4e-64
> gi927685	Saccharomyces cerevisiae 染色 IVlambda3641 和 cosmid9831, 和 9410	78500	137/350 (39%)	201/346 (57%)	3e-62
> gi7413486	Agricus 氨肽酶的部分 mRNA	1089	130/346 (37%)	189/346 (54%)	2e-55

[0093] 表 2E. ruLAP1 BLASTX 结果

[0094]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi23393799	亮氨酸氨肽酶 / Aspergillus sojiae	377	126/248 (50%) 37/78 (47%) 13/24 (83%)	162/248 (65%) 55/78 (70%) 20/24 (83%)	5e-87

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi28918132	未知蛋白 / Neurosporacrassa	402	115/247 (46%) 43/77 (55%) 18/24 (75%)	153/247 (61%) 58/77 (75%) 23/24 (95%)	8e-86
> gi6320623	未知 ORF ; ydr415cp/ Saccharomyces cerevisiae	374	96/254 (37%) 37/77 (46%)	143/254 (56%) 49/77 (63%)	7e-55
> gi28916832	未知蛋白 / Neurosporacrassa	409	96/226 (42%) 31/66 (46%)	135/226 (59%) 41/66 (62%)	4e-54

[0095] 表 2F. ruLAP1 BLASTP 结果

[0096]

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi23393799	亮氨酸氨肽酶 / Aspergillus sojiae	377	175/348 (50%)	234/348 (67%)	4e-99
> gi28918132	未知蛋白 / Neurosporacrassa	402	175/345 (50%)	230/345 (66%)	2e-97
> gi6320623	未知 ORF ; ydr415cp/ -Saccharomyces cerevisiae	374	140/351 (39%)	201/351 (57%)	7e-65
> gi28916832	未知蛋白 / Neurosporacrassa	409	129/296 (43%)	178/296 (60%)	3e-58

[0097] fuLAP2

[0098] fuLAP2 是烟曲霉菌的亮氨酸氨肽酶。1557 个核苷酸 (SEQ ID NO :7) 的 fuLAP2 核酸如表 3A 所示。

[0099]

表 3A、fuLAP2 基因组核苷酸序列 (SEQIDNO :7)

```

ATGAAGCTGCTCTACCTCACATCGTTTGCCTCTCTGGCCGTGGCCAAATGGCCCAGGATGGGACTGGAAGCCCCGAGTTCATCCGGTTAG
TGTTCCCTCTCGCCGGGTTGTCTGCTGTATGCTAACAGCATCCTGTCTATTACAGAAAAGTCCTGCCCAAATGATCCATTGTGGGATC
TTCTGCAGGGCGCTCAACAGCTGGAAGACTTCGCCATAGCCTACCCCGAGCGCAACCGCGTCTTTGGTGGACGGGCCACGAGGACACC
GTCAACTACCTCTACCGTGAGTTGAAGAAAACCGGCTACTACGACGTTTACAAGCAGCCCCAGGTTACCAGTGGACCCGAGCCGACCA
GGCTCTACCGTCGACGGCCAGTCCATAGACGCCACAACCATGACTTACAGCCCCAGCGTAAACGCCACGGCCCGCTGGCAGTGGTGA
AAACCTGGGCTGCGTCGAGGCTGACTATCCCGCCGATCTGACGGGCAAGATTGCTCTGATCTCGCGGGGCGAGTGCACCTTTGCGACC
AAATCCGTCTTGAGCGCAAGGCCGGGGCGGCGGCGCACTCGTGTACAACAATATCGAGGGTTCGATGGCGGGAACCTCTGGCGGCGC
GACCAGCGAGCTGGGTGCCTACGCTCCCATCGCCGGCATCAGCCTCGCGGACGGACAGGCGTGATCCAGATGATCCAGGCGGGCACGG
TGACAGCCAACTGTGGATCGACAGCCAGGTCGAGAACCGTACCACCTACAACGTGATCGCGCAGACCAAGGGCGGCGACCCCAACAAC
GTCTCGCGCTGGGTGGCCACACGGACTCGGTCGAGGCCGGCCCGGCATCAACGACGACGGCTCCGGCATCATCAGCAACCTCGTCGT
CGCCAAAGCGCTGACCCGCTTCTCGGTCAAGAACCGGTTGCGCTTCTGCTTCTGGACGGCGGAGGAGTTCCGGCTGTGGCAGCAAAT
ACTAGTCAACAGCCTCAATGCCACCGAGCAGGCCAAGATCCGCTGTATCTCAACTTCGACATGATCGCTCCCCCAACTACGCCCTG
ATGATCTATGACGGCGACGGCTCGGCTTCAACCTGACGGGGCGGCGGCTCGGCGCAGATCGAGCGGCTCTTCGAGGAGTACTACAC
GTGATCCGCAAGCCGTTCTGTCCGACCGAGTTCAACGGCCGCTCCGACTACCAGGCCTTTATTTCAACGGCATCCCCGCGGGAGGCC
TCTTACC GGCGGAGGCCATCAAGACCGAGGAACAGGCCAATTGTTTGGCGCCAGGCGCGCTGGCTCTGGACGCCAACTACCAC
GCCAAGGGTGACAATGACTAATCTCAACCGCGAGGCTTTCCTGATCAATTCAGGGCGACGGCCTTTGCCGTGGCGACGTACGCCAA
CAGCCTTGACTCGATCCCCCACGCAACATGACCACCGTGGTCAAGCGGTCGACGCTGGAGCAAGCCATGAAGAGGACCCCGCACACGC
ACACCGCGGGAACAGGATGCTACAAGGACCGGTTGAGCAGTAG

```

[0100] 1497 个核苷酸的 fuLAP2 开放阅读框 (ORF), 其中起始子 ATG 在第一个位置上 (表 3B 中下划线)。

[0101]

表 3B、fuLAP2 核苷酸序列 (SEQIDNO :8)

```

ATGAAGCTGCTCTACCTCACATCGTTTGCCTCTCTGGCCGTGGCCAAATGGCCCAGGATGGGACTGGAAGCCCCGAGTTCATCCGAAAAGT
CCTGCCCAAATGATCCATTTGTGGGATCTTCTGCAGGGCGCTCAACAGCTGGAAGACTTCGCCATAGCCTACCCCGAGCGCAACCGCG
TCTTTGGTGGACGGGCCACGAGGACACCGTCAACTACCTCTACCGTGAGTTGAAGAAAACCGGCTACTACGACGTTTACAAGCAGCCC
CAGGTTACCAGTGGACCCGAGCCGACCAGGCTCTACCGTCGACGGCCAGTCCATAGACGCCACAACCATGACTTACAGCCCCAGCGT
AAACGCCACGGCGCCGCTGGCAGTGGTGAACAACCTGGGCTGCGTCGAGGCTGACTATCCCGCGATCTGACGGGCAAGATTGCTCTGA
TCTCGCGGGGCGAGTGCACCTTTGCGACCAAATCCGTCTTGAGCGCAAGGCCGGGGCGGCGGCACTCGTGTACAACAATATCGAG
GGTTCGATGGCGGAACTCTGGGCGGCGGACCAGCGAGCTGGGTGCCTACGCTCCCATCGCCGGCATCAGCCTCGGGACGGACAGGC
GCTGATCCAGATGATCCAGCGGGCACGGTGACAGCCAACCTGTGGATCGACAGCCAGGTCGAGAACCGTACCACCTACAACGTGATCG
CGCAGACCAAGGGCGGCGACCCCAACAACGTCGTCGCGCTGGGTGGCCACACGGACTCGGTCGAGGCCGGGCCCGGCATCAACGACGAC
GGCTCCGGCATCATCAGCAACCTCGTCGTCGCAAGGCGCTGACCCGCTTCTCGGTCAAGAACCGGTCGCTTCTGCTTCTGGACGGC
GGAGGAGTTCGGCTGCTGGGCGAGCAACTACTACGTCAACAGCCTCAATGCCACCGAGCAGGCCAAGATCCGCTGTATCTCAACTTCG
ACATGATCGCTCCCCCAACTACGCCCTGATGATCTATGACGGCGACGGCTCGGCCTTCAACCTGACGGGGCCGGCGGCTCGGCGCAG
ATCGAGCGGCTCTTCGAGGACTACTACAGTCGATCCGCAAGCCGTTCTGTCGACCGAGTTCAACGGCCGCTCCGACTACCAGGCCTT
TATTTCAACGGCATCCCGCGGGAGGCCTTTCACCGCGCGGAGGCCATCAAGACCGAGGAACAGGCCAATTGTTTGGCGGCCAGG
CCGGCGTGGCTCTGGACGCCAACTACCACGCCAAGGGTGACAACATGACTAATCTCAACCGGAGGCTTTCCTGATCAATTCAGGGCG
ACGGCCTTTGCCGTGGCGAGCTACGCCAACAGCCTTGACTCGATCCCCCACGCAACATGACCACCGTGGTCAAGCGGTCGACGCTGGA
GCAAGCCATGAAGAGGACCCCGCACACGACACCGGCGGAACAGGATGCTACAAGGACCGGTTGAGCAGTAG

```

[0102] 上述 fuLAP2 核苷酸序列 (SEQ ID NO :8) 编码有 498 个氨基酸残基 (SEQ ID NO : 9) 的蛋白, 在表 3C 中以一个字母的氨基酸代码表示。

[0103]

表 3C、fuLAP2 编码的蛋白序列 (SEQIDNO :9)

```

MKLLYLTSFASLAVANGPGWDWKPRVHPKVLPMIHLWDLQGAQQLEDFAAYPERNRVFGGRAHEDTVNYLYRELKKTGYVDVYKQP
QVHQWTRADQALTVDGQSYDATTMTYSPSVNATAPLAVVNNLGCVEADYPADLTGKIALISRGECTFATKSVLSAKAGAAAALVYNNIE
GSMAGTLGGATSELGAYAPIAGISLADGQALIQMIQAGTVTANLWIDSQVENRTTYNVIAQTKGGDPNNVVALGGHTDSVEAGPGINDD
GSGIISNLVVAKALTRFSVKNVRFVCFWTAEEFGLLSNYVNSLNATEQAKIRLYLNFDMIASPNYALMIYDGDGSAFNLTGPAGSAQ
IERLFEDYYTSIRKPFVPTFEFNGRSDYQAFILNGIPAGGLFTGAEA IKTEEQAQLFGGQAGVALDANYHAKGDNMTNLNREAFLINSRA
TAFVATYANSLDISIPRNMTTVVKRSQLEQAMKRTPHHTGGTGCYKDRVEQ

```

[0104] 上述 fuLAP2 具有与表 3D、3E 和 3F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0105] 表 3D. fuLAP2T BLASTN 结果

[0106]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi469463	Saccharomyces Cerevisiae/ 氨肽酶 Y 基因	2272	184/464 (39%)	243/464 (52%)	7e-69
> gi9949032	Pseudomonas AruginosaPA01,, 整个 529 个基因组中的 281 片段 氨肽酶 Y 基因	12547	165/445 (37%)	231/445 (51%)	9e-67
> gi23017467	Mycobacterium Tuberculosis CDC15551, 整个 280 个基因组中的 33 片 段	18857	166/426 (38%)	218/426 (51%)	2e-62

[0107] 表 3E. fuLAP2 BLASTX 结果

[0108]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi28918599	未知蛋白 / Neurosporacrassa	508	250/479 (52%)	314/479 (65%)	e-131
> gi23017467	未知蛋白 / Thermobifidafusca	514	173/465 (37%)	251/465 (53%)	4e-74

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi584764	APE3YEAST ; 氨肽酶前体 / Saccharomyces cerevisiae	537	184/464 (39%)	243/464 (52%)	8e-70
> gi9949032	可能的氨肽酶 / Pseudomonas AeruginosaPA01	536	165/445 (37%)	231/445 (51%)	1e-67
> gi15839805	水解酶 / Mycobacterium Tuberculosis CDC15551	493	166/426 (38%)	218/426 (51%)	3e-63

[0109] 表 3F. fuLAP2 BLASTP 结果

[0110]

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi28918599	未知蛋白 / Neurosporacrassa	508	250/469 (52%)	314/479 (65%)	e-128
> gi123017467	未知蛋白 / Thermobifidafusca	514	173/465 (37%)	251/465 (53%)	3e-71
> gi584764	APE3YEAST ; 氨肽酶前体 / Saccharomyces cerevisiae	537	183/464 (39%)	243/464 (52%)	6e-70
> gi15598135	可能的氨肽酶 / Pseudomonas AeruginosaPA01	536	164/445 (36%)	230/445 (51%)	3e-65

[0111] fuLAP1

[0112] fuLAP1 是烟曲霉菌的亮氨酸氨肽酶。1298 个核苷酸 (SEQ ID NO:10) 的 ruLAP1 核酸如表 4A 所示。

[0113]

表 4A、fuLAP1 基因组核苷酸序列 (SEQIDNO :10)

```

ATGAAAGTTCTTACAGCTATTGCGCTGAGCGCAATAGCTTTTCACAGGGGCTGTAGCTGCAGTGATTACTCAGGAAGCATTCTTAAACAA
CCCCCGCATCCATCATGACCAGGAGAAGTACTTGATCGAACTGGCCCCCTTATCGAACACGATGGGTGACTGAAGAGGAGAAAATGGGCAT
TGAAATTGGTACCATACTTCCCCAAAATTTGGGTCTCCAAGTCCACGGGGGACTAAGTGCACGATTGCTTGAAGGACGGCGTGAATTTT
ATCGATATCACAGAAGAGCACAAACCCGGATTTTACCCGACTCTCCACAGCGCCAGCTATGTGAAATATCCACCGAAGATGCAGTATGC
AGAAGAAGTGGCTGCTCTTAAACAAGAAATTTATCGAAAGAAAACATGAAGGCCAACCTGGAACGATTACATCATTTCATACTCGCTATT
ACAAATCTCAGACGGGAATCCGATCGGCAACGTGGCTGTTTCGACCAAGTTCAGAGAGTTGTCTCTGAGTCTGGAGCCGCTGAGTATGGT
GCAACTGTTGAGCGATTCTCTCATCCATGGGGTTCAGTTCAGCATTATTGCCCGAATACCCGGCCGAACGAACAAGACTGTGGTGCTGGG
CGCCATCAGGACAGCATCAATTTGTTTCTCCCGTCAATCTTGGCTGCTCCCGGTGCTGATGACGATGGAAGTGGAACTGTCACCATTC
TTGAAGCGTTGCGCGTCTGCTGCAGTCAGACGCCATTGCCAAGGTAATGCATCCAATACTGTGCGATTCCACTGGTACTCTGCAGAA
GAAGGCGAATGCTGGGCTCCAGGCAATATTTTCCAATTACAAGCGGAATAGGCGGAAATCAAAGCCATGCTCCAGCAAGACATGAC
TGGCTACGTCCAGGGAGCTTTGAACGCCGTGTTGAGGAAGCCATAGGAATATGATGTCGATTATGTCGACCAGGGCCTCACACAGTTTC
TCAAGGACGTTGTTACAGCGGTAAGCCTCAGTTGTCCCCACGAAAAGCTGTTTAGTCGACAAAATGAAATGACGGCTGCATTAGTACT
GCTCTGTGGGTTACCTGGAGACGAAGTGGGATATGCCTGCTCCGACCACACCTCGGCCAGTAAATATGGTTATCCCGCGGCTATGGCG
ACAGAAGCAGAGATGGAAAATACCAATAAGAAGATACATACTACCGACGACAAGATCAAGTATTTGAGCTTCGATCATATGTTGGAGCA
TGCCAAGTTGAGTCTTGGCTTCGCTTTCGAATTGGCATTTCGCGCGTTTTAA

```

[0114] 1167 个核苷酸的 fuLAP1 开放阅读框 (ORF), 其中起始子 ATG 在第一个位置上 (表 4B 中下划线)。

[0115]

表 4B、fuLAP1 核苷酸序列 (SEQIDNO :11)

```

ATGAAGTTCTTACAGCTATTGCGCTGAGCGCAATAGCTTTTCACAGGGGCTGTAGCTGCAGTGATTACTCAGGAAGCATTCTTAAACAA
CCCCCGCATCCATCATGACCAGGAGAAGTACTTGATCGAACTGGCCCCCTTATCGAACACGATGGGTGACTGAAGAGGAGAAAATGGGCAT
TGAAATTGGACGGCGTGAATTTTATCGATATCACAGAAGAGCACAAACCCGGATTTTACCCGACTCTCCACAGCGCCAGCTATGTGAAA
TATCCACCGAAGATGCAGTATGCAGAAGAAGTGGCTGCTCTTAAACAAGAAATTTATCGAAAAGAAAACATGAAGGCCAACCTGGAACGATT
CACATCATTTCATACTCGCTATTACAAAATCTCAGACGGGAATCCGATCGGCAACGTGGCTGTTTCGACCAAGTTCAGAGAGTTGTCTCTG
AGTCTGGAGCCGCTGAGTATGGTGCAACTGTTGAGCGATTCTCTCATCCATGGGGTTCAGTTCAGCATTATTGCCCGAATACCCGGCCGA
ACGAACAAGACTGTGGTGCTGGGCGCCCATCAGGACAGCATCAATTTGTTTCTCCCGTCAATCTTGGCTGCTCCCGGTGCTGATGACGA
TGGAAGTGGAACTGTCACCATTTTGAAGCGTTGCGCGTCTGCTGCAGTCAGACGCCATTGCCAAGGTAATGCATCCAATACTGTGCG
AGTTCCTACTGGTACTCTGCAGAAGAAGCGGAATGTGGGCTCCAGGCAATATTTTCCAATTACAAGCGGAATAGGCGGAAATCAA
GCCATGCTCCAGCAAGACATGACTGGCTACGTCCAGGGAGCTTTGAACGCCGTGTTGAGGAAGCCATAGGAATATGGTCGATTATGT
CGACCAGGGCCTCACACAGTTTCTCAAGGACGTTGTTACAGCGTACTGCTCTGTGGGTTACCTGGAGACGAAGTGGGATATGCCTGCT
CCGACCACACCTCGGCCAGTAAATATGGTTATCCCGCGGCTATGGCGACAGAAGCAGAGATGGAAAATACCAATAAGAAGATACATACT
ACCGACGACAAGATCAAGTATTTGAGCTTCGATCATATGTTGGAGCATGCCAAGTTGAGTCTTGGCTTCGCTTTCGAATTGGCATTTCG
CGCGTTTTAA

```

[0116] 上述 fuLAP1 核苷酸序列 (SEQ ID NO :11) 编码有 388 个氨基酸残基 (SEQ ID NO : 12) 的蛋白, 在表 4C 中以一个字母的氨基酸代码表示。

[0117]

表 4C、fuLAP1 编码的蛋白序列 (SEQIDNO :12)

```

MKVLTAIALSATAFTGAVAAVITQEAFNNPRIHHDQEKYLIELAPYRTRWVTEEEKWALKLDGVNFIDITEEHNTGFYPTLHSASYVK
YPPKMQYAEVEVAALNKNLSKENMKANLERFTSFHTRYYSKQTGIRSATWLFQVQRRVSESGAAEYGGATVERFSPHWGQFSTIARIPGR
TNKTVVLGAHQDSINLFLPSILAAPGADDDGSGTVTILEALRGLLQSDAIAKGNASNTVEFHWYSAEEGMLGSQAIFSNYKRNRREIK
AMLQQDMTGYVQGALNAGVEEAIGIMVDYVDQGLTQFLKDVVTAYCSVGYLETCKGYACSDHTSASKYGYPAAMATEAMENTNKKIHT
TDDKIKYLSFDHMLEHAKLSLGFELAFAPF

```

[0118] 上述 fuLAP1 具有与表 4D、4E 和 4F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0119] 表 4D. fuLAP1 TBLASTN 结果

[0120]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi1762234	Polyketide synthasePKSL2/ Aspergillus parasiticus	9894	208/249(80%) 61/84(72%) 46/62(74%)	226/249(90%) 67/84(79%) 55/62(88%)	e-169
> gi23393798	亮氨酸氨肽酶 (Lap1)/ Aspergillus sojiae	2547	66/110(60%) 68/152(44%) 37/75(49%) 15/30(50%)	82/110(74%) 92/152(60%) 52/75(69%) 21/30(70%)	7e-82
> gi927685	Saccharomyces Cerevisiae 染色 IVlambda3641 和 cosmid9831,和 9410	78500	152/341 (44%)	207/341 (60%)	1e-71
> gi5832144	在无氮源情况下培养出的 Botrytis cinerea 菌株 T4cDNA 文库	780	89/134(66%) 27/53(50%)	106/134(79%) 33/53(62%)	7e-58

[0121] 表 4E. fuLAP1 BLASTX 结果

[0122]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi28918132	未知蛋白 / Neurosporacrassa	402	208/352(59%)	255/352(72%)	e-116

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi23393799	亮氨酸氨肽酶 / Aspergillus sojiae	377	183/355 (51%)	241/355 (67%)	3e-97
> gi6320623	未知 ORFYdr414cp/ Saccharomyces cerevisiae	374	152/341 (44%)	207/341 (60%)	2e-72
> gi18250467	氨肽酶 /Agaricus bisporus	384	139/352 (39%)	186/352 (52%)	1e-58

[0123] 表 4F. fuLAP1 BLASTP 结果

[0124]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
> gi28918132	未知 蛋白 / Neurosporacrassa	402	208/352 (59%)	255/352 (72%)	e-116
> gi23393799	亮氨酸氨肽酶 (LAP1) / Aspergillus sojiae	377	183/355 (51%)	241/355 (67%)	6e-98
> gi6320623	未知 ORFYdr414cp/ Saccharomyces cerevisiae	374	152/341 (44%)	207/341 (60%)	3e-73
> gi18250467	氨肽酶 /Agaricusbisporus	384	140/352 (39%)	190/352 (53%)	7e-59

[0125] ruCBPS1

[0126] ruCBPS1 是红色毛癣菌的羧肽酶。2106 个核苷酸 (SEQ ID NO:13) 的 ruCBPS1 核酸如表 5A 所示。

[0127]



表 5A、ruCBPS1 基因组核苷酸序列 (SEQIDNO :13)

```

ATGGTGTCAATTCTGCGGAGTGGCAGCCTGCCTGCTGACAGTTGCTGGCCATCTTGCGCAGGCTCAGTCCCACAAAACCGGAGGGAGT
CACTGTCCCTGGAGTCGAAATTCGGCAGCGGTGCTCGCATCACTTATAAGGAGGTCCGTTAGCTGCATAGAAAAGTCCACGTGAAGACGCT
GTAGCTAACAAATCCACTAGCCTGGCCTCTGTGAGACGACAGAGGGCGTCAAGTCGTACGCCGATATGTCATCTGCCTCCAGGCACGC
TCAGGGACTTCGGTGTGAGCAGGACTACCCATCAACACCTTTTTTTGGTTCTTTGAGGCAAGAAAAGGACCCTGAAAATGCCCTCTC
GGCATCTGGATGAACGGTGGCCCGGGTAGCTCGTCGATGTTTGAATGATGACTGAGAACGGGCCTTGCTTCGTCATGCAGACTCCAA
TTCTACTCGCCTGAACCCTCATTCTTGGAACAATGAAGGTATGCCATCAGCTTCTGATGGAAAATAAATATTGCTAACATTGTACTTT
CTGTGACTAGTCAATATGCTGTATATAGACCAGCCAGTGCAGGTCCGTCGTCTACGACACTTTGGCCAACCTCACCAGGAATCTAGT
CACGGATGAGATCACGAACTGAAACCCGGAGAACCTATCCGGAACAGAATGCCACTTCTCGGTAGGFACATATGCAAGCCGCAATA
TGAAACCACTGCACACGGAAGTACGATGCTGCCATGGCTCTCTGGCACTTCGCCAAAGTCTGGTTCCAAGAGTCCCAGGATATCAC
CCTAGGAACAACAAGATCAGCATTGCTACCGAATCCTACGGCGCCGTTATGGCCCGCCTTTACTGCCCTCTTTGAAGAGCAGAACCA
GAAGATCAAGAACGGCACATGGAAGGGACACGAGGGAACTATGCACGTGCTGCATCTCGACACCCTCATGATCGTCAACGGATGCATCG
ACCGTCTTGTCCAATGGCCGGCATATCCGCAATGGCGTACAACAACACATATAGCATCGAGGCAGTCAACGCCAGCATTTCATGCAGGA
ATGTGGATGCCCTTACC CGCAGCGTGGCTGTGAGACAAGATTAACCACTGCCGCTCCCTCTCTTCTGTGTTTCGATCTGAGAATCT
CGGCATCAACTCAACCGTCAATGATGTCTGCAAGGATGCCGAGACATTCTGCTCCAATGATGTTTCGCGATCCCTACCTCAAGTTCTCTG
GCCGCAACTACTATGACATCGACAGCTTGACCCAGCCATTCCAGCACCATTTCACATGGCTGGCTAAATCAGCCGCATGTGCAG
GCAGCACTGGGTGTGCCACTTAACCTGGACACAGTCAAACGATGTTGTGCTACCGCATTCCGTGCAATTTGGTACTACCTCGGCCAGG
GTGGCTGGAGAACCCTGGCTTATTTGCTGGAGAATGGCATCAAGG1TTCGCTTGTTCACGGTATCGGGACTACGCATGCAACTGGTTTCG
GTGGTGAGCTCTCAAGTCTGGGAATCAACTACACTGACACCCACGAATTCATAATGCCGGCTATGCAGGTATCCAGATCAATAGCAGC
TACATCGGTGGACAGGTGAGGCAGTACGGCAACCTCTCCTTCGCCCGCTTACGAGGCCGCCATGAGGTGCCATCGTACCAACCCGA
GACTGCACTGCAGATATCCACCCTCCCTGTTCAACAAGGATATCGTACTGGAACCAAGGACACGTATCGCGCATGGACGGAGGCA
AGTTTACGGCACCTCGGGCCCTGCGGACTCGTTTGGTTTCAAGAACAACCTCCACCGCAGCAGTCCACTTCTGTATATCTTAGAC
ACCAGCACCTGCACCAAGGAGCAGATCCAGTCAGTTGAGAACGGCACTGCCCGGTACGCAGCTGGATCATTGTCGACTCCAACCTCGAC
CTCTCTGTTCCCGAGGTAGTTGGCTCAGGGGAACCCACGCCAACCCCTATGCCTGGAGGGGCTACTACACTATCTGCTCACGGTTCT
TGTATGGCGTGACATTATGGGCTGTTATTGTTGTAGCTGTTATAGAGCTGGCAATGTAA

```

[0128] 1989 个核苷酸 (SEQ ID NO :14) 的 ruCBPS1 开放阅读框 (ORF) 见表 5B, 其中 ruCBPS1 开放阅读框 (ORF) 起始子 ATG 在第一个位置上 (表 5B 中下划线)。

[0129]

表 5B、ruCBPS1 核苷酸序列 (SEQ ID NO :14)

```

ATGGTGTCAATTCTGCGGAGTGGCAGCCTGCCTGCTGACAGTTGCTGGCCATCTTGCGCAGGCTCAGTTCACCAAAAACCGGAGGGA
GTCAGTGTCTGGAGTCGAAATTCGGCAGCGGTGCTCGCATCACTATAAGGAGCCTGGCCTCTGTGAGACGACAGAGGGCGTCAAG
TCGTACGCCGATATGTCCATCTGCCTCCAGGCACGCTCAGGGACTTCGGTGTGAGCAGGACTACCCTATCAACACCTTTTTTTGG
TTCTTTGAGGCAAGAAAGGACCTGAAAAATGCCCTCTCGGCATCTGGATGAACGGTGGCCCGGGTAGCTCGTCGATGTTTGGAATG
ATGACTGAGAACGGGCCTTGCTTCGTCGAATGCAGACTCCAATTCTACTGCGCTGAACCCTCATTCTTGAAACAATGAAGTCAATATG
CTGTATATAGACCAGCCAGTGCAGGTGCGTCTGTCTACGACACTTTGGCCAACTTCACCAGGAATCTAGTCACGGATGAGATCACG
AAACTGAAACCCGGAGAACCTATTCCGGAACAGAATGCCACTTTCCTGGTAGGTACATATGCAAGCCGAATATGAACACCCTGCA
CACGGAAGTGGCATGCTGCCATGGTCTCTGGCAGTTCGCCCAAGTCTGGTTCCAAGAGTTCACAGGATATCACCTAGGAACAAC
AAGATCAGCATTGCTACCGAATCCTACGGCGCCGTTATGGCCCGCCTTTACTGCCTTCTTTGAAGAGCAGAACCAGAAGATCAAG
AACGGCACATGGAAGGGACACGAGGAACTATGCACGTGCTGCATCTCGACACCCTCATGATCGTCAACGGATGCATCGACCGTCTT
GTCCAATGGCCGCATATCCGAAATGGCGTACAACAACACATATAGCATCGAGGCAGTCAACGCCAGCATTATGCAGGAATGCTG
GATGCCCTCTACCGCAGCGTGGCTGTCGAGACAAGATTAACCACTGCCGCTCCCTCTCTTCTGTGTTTCGATCCTGAGAAATCTCGGC
ATCAACTCAACCGTCAATGATGTCTGCAAGGATGCCGAGACATTCTGCTCCAATGATGTTTCGCGATCCCTACCTCAAGTTCTCTGGC
CGCAACTACTATGACATCGGACAGCTTGACCCAGCCCATTCACAGCACCATTTTACATGGCCTGGCTAAATCAGCCGCATGTGCAG
GCAGCACTGGGTGTGCCACTTAACTGGACACAGTCAAACGATGTTGTGTCTACCGCATTCCTGCAATTTGGTACTACCCTCGGCCA
GGGTGGCTGGAGAACCTGGCTTATTTGCTGGAGAATGGCATCAAGTTTCGCTGTTTACGGTATCGGGACTACGCATGCAACTGG
TTCGGTGGTAGCTCTCAAGTCTGGGAATCAACTACACTGACACCCACGAATTCATAATGCCGGCTATGCAGGTATCCAGATCAAT
AGCAGCTACATCGGTGGACAGGTAGGCAGTACGGCAACCTCTCCTTCGCCCGCTACGAGGCCGGCCATGAGGTGCCATCGTAC
CAACCCGAGACTGCACTGCAGATATTCACCGTTCCTGTTCACAAAGGATATCGCTACTGGAACCAAGGACACGTCATCGCGCATG
GACGGAGGCAAGTTTTACGGCACCTCGGGCCCTGCGGACTCGTTTGGTTTCAAGAACAACCTCCACCGCAGCAGTCCACTTCTGT
CATATCTTAGACACCAGCACCTGCACCAAGGAGCAGATCCAGTTCAGTTGAGAACGGCAGTCCCGCGTACGCAGCTGGATCATTGTC
GACTCCAACTCGACCTCTCTGTTCCCGAGGTAGTTGGCTCAGGGAAACCCACGCCAACCCCTATGCCTGGAGGGGCTACTACACTA
TCTGCTCACGGTTCTTGTATGGCGTGACATTATGGCTGTTATTGTTGTAGCTGTTATAGAGCTGGCAATGTAA

```

[0130] 上述 ruCBPS1 核苷酸序列 (SEQ ID NO :14) 编码含 662 个氨基酸残基 (SEQ ID NO : 15) 的蛋白, 在表 5C 中以一个字母的氨基酸代码表示。

[0131]

表 5C、ruCBPS1 编码的蛋白序列 (SEQ ID NO :15)

```

MVSFCGVAACLLTVAGHLAQFPPKPEGVTVLESKFGSGARITYKEPGLCETTEGVKSYAGYVHLPPGTLRDFGVEQDYPINTFFW
FFEARKDPENAPLGIWMGGPGSSMFGMMTENGPCFVNADSNSTRLNPHSWNNEVNMLYIDQPVQVGLSYDTLANFTRNLVTDEIT
KLKPGPEIPEQNATFLVGYASRNMTTAHGTRHAAMALWHFAQVWFQEFPGYHPRNNKISIAATESYGGRYGPAFTAFFEEQNQKIK
NGTWWGHEGTMHVLHLDTLMIVNGCIDRLVQWPAYPQMAYNNTYSIEAVNASIHAGMLDALYRDGGCRDKINHCRSLSSVFDPENLG
INSTVNDVCKDAETFCSNDVRDYPYLFKFSGRNYDYGQLDPSPPFAPFYMAWLNQPHVQAALGVPLNWTQSNVVSSTAFRAIGDYPRP
GWLENLAYLLENGIKVSLVYGDYACNWFGGELSSLGINYTDTHEPHNAGYAGIQINSSYIGGQVRQYGNLSFARVYEAGHEVPSY
QPETALQIFHRSLFNKDIATGKDTSSRMDGGKFGYGTSGPADSFGFKNPPPQHVHVFCHILDTSTCTKEQIQSVENGTAAVRSWIIIV
DSNSTSLFPEVVGSGEPTPTPMPGGATLSAHGFLYGVTLWAVIVVAIVIELAM

```

[0132] 上述 ruCBPS1 具有与表 5D、5E 和 5F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0133] 表 5D. ruCBPS1 TBLASTN 结果

[0134]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   32410708	Neurosporacrassa 菌株 OR74A	1947	222/632 (35%)	321/632 (50%)	1e-90
gi   3046860	schizoaccharomyces prombecpy1 羧肽酶 Y 基 因	4308	137/481 (28%)	204/481 (42%)	6e-41
gi   18152938	Pichiaangusta 羧肽酶 Y (CPY) 基因	2214	141/520 (27%)	228/520 (43%)	4e-40
gi   4028157	Pichiaangusta 羧肽酶 Y 前体 (CPY) 基因	2509	140/520 (26%)	226/520 (43%)	7e-40
gi   170828	Candidaalbicans 羧肽 酶 Y 前体 (CPY1) 基因	1985	131/482 (27%)	205/482 (43%)	3e-36

[0135] 表 5E. ruCBPS1 BLASTX 结果

[0136]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   15004616	羧肽酶 S1/Aspergillusoryzae	555	209/535 (39%)	294/535 (54%)	1e-98
gi   435818	羧肽酶 S1CPD-S1/Penicillium janthinellum	423	159/498 (31%)	234/498 (46%)	6e-64
gi   995456	前 - 羧肽酶 Z/Absidiazychae	460	147/506 (29%)	219/506 (43%)	8e-48
gi   3046861	羧肽酶 Y/ schizoaccharomycesprombe	1002	137/481 (28%)	204/481 (42%)	7e-42
gi   18152939	羧肽酶 Y/Pichiaangusta	537	141/520 (27%)	228/520 (43%)	4e-41
gi   4028158	羧肽酶 Y 前体 ;vacuolar 羧肽酶 /Pichiaangusta	541	140/520 (26%)	226/520 (43%)	7e-41

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   7597001	羧肽酶 Y 前体 /Candidaalbicans	542	131/482 (27%)	206/482 (42%)	2e-37

[0137] 表 5F. ruCBPS1 BLASTP 结果

[0138]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   15004616	羧肽酶 S1/Aspergillus oryzae	555	210/537 (39%)	296/535 (55%)	2e-95
gi   435818	羧肽酶 S1 CPD-S1/Penicillium janthinellum	423	159/498 (31%)	234/498 (46%)	2e-60
gi   995456	前 - 羧肽酶 Z/Absidiazychae	460	146/500 (29%)	217/500 (43%)	6e-47
gi   19115337	羧肽酶 Y/ schizoaccharomycesprombe	1002	136/481 (28%)	204/481 (42%)	7e-41

[0139] ruCBPS1'

[0140] ruCBPS1' 是红色毛癣菌的羧肽酶。2030 个核苷酸 (SEQ ID NO :16) 的 ruCBPS1' 基因组 DNA 序列如表 6A 所示。

[0141]

表 6A、ruCBPS1' 基因组核苷酸序列 (SEQIDNO :16)

```

ATGCGCTTTGCTGCTAGCATTGCCGTGGCCCTGCCAGTCATTCACGGCGGAGTGCTCAAGGCTTCCCTCCACCCGTTAAGGGCGTC
ACCGTGGTCAAATCCAAGTTCGACGAAAACGTAAAGATCACATACAAGGAGGTATGTGTTTACATCATTTTCACATCCAGATCTTAT
ATCCTTACAATAAATCTGGCTAACTCACTGGATAGAATGACATATGTGAAACCACTCAAGGAGTTAGATCATTACCCGGTCATGTCC
ACCTTCTCCAGACAACGATGACTTTGGTGTCTACCGGAACTACTCCATCAACACATTCTTCTGGTTCTTGAAGCTCGTGAAGACC
CTAAGAATGCTCCTCTCTCCATCTGGCTGAACGGTGGTCCGGGATCGTCATCCATGATTGGACTCTTCCAGGAAAACGGTCCATGCT
GGGTCAATGAAGACTCTAAATCTACCACCAACAATTCATTTTCATGGAAACAATAAAGTAAATA TGCTCTACATTGATCAGCCAAACC
AAGTCGGTTTCAGTTATGACGTACCTACCAACATCACTTACTCTACCATCAATGATAACAATATCTGTTGCGGACTTCTCTAACGGTG
TCCCTGCGCAAAATCTTTCTACGTTGGTTGGAACCGGCAGCAGCCAGAACCCTTGGGCAACTGCCAATAACACTGTGAACGCTGCTC
GTTCTATCTGGCACTTTGCACAAGTGTGGTTCAGGAATTCCTGAAACACAAGCCTAACAAATAACAAGATCAGTATTTGGACAGAGT
CCTATGGAGGAAGATATGGTCCCTCATTGCGCTCTTACTTCCAGGAACAGAACGAAAAGATCAAAAACCATACCATTACTGAAGAAG
GAGAGATGCATATTCTGAACCTCGACACCCTCGGTATCATCAACGGCTGCATCGATCTTATGTTCCAAGCAGAAAAGTTATGCTGAAT
TCCCATACAACAACACCTATGGCATCAAAGCTTATAACCAAGGAGAAGCGTGACGCTATAATTACACGACATCCACCGTCTCGACGGCT
GCTTCGACAAGGTTACCAAGTCCGTGAGGCCGCGAAAAGAAGGAGACCCTCACTTCTACAGCAACAATGCAACCGTCAACACAATCT
GTGCGGATGCTAACTCTGCCTGCGACAAATATCTAATGGATCCTTTCCAAGAGACCAATCTTGGTTACTATGATATTGCTCATCCTC
TTCAGGATCCCTTCCCCCACCATTCTATAAGGGCTTCCTCAGCCAA TCCAGCGTTCTATCTGACATGGGATCGCCAGTCAACTTCT
CCCAATACGCCAAGCTGTGGGAAAATCATTCCATGGAGTTGGCGACTACGCTCGCCCTGATGTGCGCGGCTTACCAGGTGACATTG
CTTATCTTCTCGAGAGCGGAGTCAAGGTTGCTCTCGTCTATGGTGACAGAGACTACATCTGCAATTGGTTCGGTGGTGAAGCAGTCA
GTCTTGGCTTGAACACTACACTGGCACCCAAGACTTCCACAGGGCAAAATATGCCGATGTCAAGGTCAACTTTCATACGTCGGAGGCG
TAGTGCGTCAACATGGAACTTCTCTTTACCAGAGTTTTTCGAGGCCGGTCATGAAGTCCCTGGTTACCAACCCGAGACTGCCCTCA
AGATCTTTGAGCGCATCATGTTCAACAAGGATATTTCTACCGGTGAGATCGACATTGCTCAGAAAACCAGACTACGGTACCCTGGAA
CTGAGTCTACGTTCCATATCAAAAACGATATCCCTCCTTCGCTGAGCCGACCTGCTACCTCCTCAGTGTGACGGAACCTGTACCC
CGGAGCAGCTTAATGCTATTAAGGATGGAATGCAGTTGTTGAGA ACTACATTATTAAGAGCCCTGCTGCGTCAAGGGGAAACCCTC
CACCAACCACGACCTCATCTCCACAGCAGCCCTACCGCTGGAAGTGCCATGCTAAAAGGCTCCTGTGCAATGCTAGCAATATCAG
CTCTCACTGTCCTTGCTTTCTTCTGTAG

```

[0142] 1959 个核苷酸 (SEQ ID NO :17) 的 ruCBPS1' 见表 6B, 其中 ruCBPS1' 开放阅读框 (“ORF”) 起始子 ATG 在第一个位置上 (表 6B 中下划线)。

[0143]

表 6B、ruCBPS1' 核苷酸序列 (SEQIDNO :17)

```

ATGCGCTTTGCTGCTAGCATTGCCGTGGCCCTGCCAGTCATTACGCGGCGAGTGCTCAAGGCTTCCCTCCACCCGTTAAGGGCGTC
ACCGTGGTCAAATCCAAGTTCGACGAAAAACGTAAAGATCACATACAAGGAGAATGACATATGTGAAACCACTCAAGGAGTTAGATCA
TTCACCGGTCATGTCCACCTTCTCCAGACAACGATGACTTTGGTGTCTACCGGAACACTCCATCAACACATTCTTCTGGTTCTTT
GAAGCTCGTGAAGACCCTAAGAAATGCTCCTCTCTCCATCTGGTGAACGGTGGTCCGGGATCGTCATCCATGATTGGACTCTCCAG
GAAAACGGTCCATGCTGGGTCAATGAAGACTCTAAATCTACCACCAACAATTCATTTTCATGGAACAATAAAGTAAATATGCTCTAC
ATTGATCAGCCAAACCAAGTCGGTTTCAGTTATGACGTACTACCAACATCACTTACTCTACCATCAATGATACAATACTGTTGCG
GACTTCTCTAACGGTGTCCCTGCGCAAAAATCTTTCACGTTGGTTGGAACCGGCAGCAGCCAGAACCCTTGGGCAACTGCCAATAAC
ACTGTGAACGCTGCTCGTTCTATCTGGCACTTTGCACAAGTGTGGTCCAGGAATCCCTGAACACAAGCCTAACATAACAAGATC
AGTATTTGGACAGAGTCCATATGGAGGAAGATATGGTCCCTCATTCGCCCTTACTTCCAGGAACAGAACGAAAAGATCAAAAACCAT
ACCATTACTGAAGAAGGAGAGATGCATATTTGAACTCGACACCCCTCGGTATCATCAACGGCTGCATCGATCTTATGTTCCAAGCA
GAAAGTTATGCTGAATTCCTATACAACAACCTATGGCATCAAAGCTTATACCAAGGAGAAGCGTGACGCTATATTACACGACATC
CACCGTCTGACGGCTGCTTCGACAAGGTTACCAAGTGCCGTGAGGCCGCAAGAAGGAGACCCCTACTTCTACAGCAACAATGCA
ACCGTCAACACAATCTGTGCGGATGCTAACTCTGCCTGCGACAATACTAATGGATCCTTTCCAAGAGACCAATCTTGGTTACTAT
GATATTGCTCATCCTTTCAGGATCCCTTCCCCCACCATTCTATAAGGGCTTCTCAGCCAATCCAGCGTTCTATCTGACATGGGA
TCGCCAGTCAACTTCTCCAATACGCCAAGCTGTGGGAAAATCATTCATGGAGTTGGCGACTACGCTCGCCCTGATGTGCGCGGC
TTCACCGGTGACATTGCTTATCTTCTCGAGAGCGGAGTCAAGGTTGCTCTCGTCTATGGTGACAGAGACTACATCTGCAATTGGTTC
GGTGGTGAGCAGGTCAGTCTTGGCTTGAACACTGACCTGACCCCAAGACTTCCACAGGGCAAAAATGCGGATGTCAAGGTCAACTCT
TCATACGTCGGAGGCGTAGTGCCTAACATGAAACTTCTCTTACCAGAGTTTTTCAGGCGGTCATGAAAGTCCCTGGTTACCAA
CCCAGACTGCCCTCAAGATCTTTGAGCGCATCATGTTCAACAAGGATATTTCTACCGGTGAGATCGACATTGCTCAGAAAACAGAC
TACGGTACCACTGGAAGTGAAGTCTACGTTCCATATCAAAAACGATATCCCTCCTTCGCTGAGCCGACCTGCTACCTCCTCAGTGC
GACGGAACCTGTACCCCGAGCAGCTTAATGCTATTAAGGATGGAAGTGCAGTTGTTGAGAACTACATTATTAAGAGCCCTGCTGCG
TCGAAGGGAAACCTCCACCAACCACGACCTCATCTCCACAGCAGCCCTACCGCTGGAAGTGCATGCTAAAGGCTCCTGTGGCA
ATGCTAGCAATATCAGCTCTCACTGTCCTTGCTTTCTTCTGTAG

```

[0144] 上述 ruCBPS1' 核苷酸序列 (SEQ ID NO :17) 编码含 652 个氨基酸残基 (SEQ ID NO :18) 的蛋白,在表 6C 中以一个字母的氨基酸代码表示。

[0145]

表 6C、ruCBPS1' 编码的蛋白序列 (SEQIDNO :18)

```

MRFAASIAVALPVIHAASAQGFPPPVKGVTVVKSDFDENVKITYKENDICETTQGVRSFTGHVHLPDNDDFGVYRNYSINTFFWFF
EAREDPKNAPLSIWLNGGPGSSSMIGLFQENGPCWVNEDSKSTNNFSWNNKVNMLYIDQPNQVGFSDYVPTNITYSTINDTISVA
DFSNGVPAQNLSTLVGTGSSQNPWATANNTVNAARSIWHAQVWFQEFPEHKPNNKISIWTESYGGRYGPSFASYFQEQNEKIKNH
TITEEGEMHILNLDLGIINGCIDLMFQAESYAEFPYNNTYGIKAYTKEKRDAILHDIHRPDGCFDKVTKCREAAKEGDPHFYSNNA
TVNTICADANSACDKYLMDFQETNLGYDIAHPLQDPFPPPFYKGFLSQSSVLSDMGSPVNFQYQAVGKSFHGVGDYARPDVRG
FTGDIAYLLESQVGVKVALVYGDYICNWFQGEQVSLGLNYTGTQDFHRAKYADVKVNSYVGGVVRQHGNSFTRVFEAGHEVPGYQ
PETALKIFERIMFNKDISTGEIDIAQKPDYGTGTGTESTFHINKNDIPPSPEPTCYLLSADGTCTPEQLNAIKDGTAVVENYI IKSPAA
SKGNPPPTTTSPTAAPTAGSAMLKAPVAMLAISALTVLAFFL

```

[0146] 上述 ruCBPS1' 具有与表 6D、6E 和 6F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0147] 表 6D. ruCBPS1' TBLASTN 结果

[0148]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   32410708	Neurosporacrassa 菌株 OR74A	1947	246/632 (38%)	337/632 (53%)	e-104
gi   3046860	Schizoaccharomyces prombecpy1 羧肽酶 Y 基因	4308	137/480 (28%)	215/480 (44%)	1e-45
gi   18152938	Pichiaangusta 羧肽酶 Y (CPY) 基因	2214	139/508 (27%)	227/508 (44%)	2e-42

[0149] 表 6E. ruCBPS1' BLASTX 结果

[0150]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   15004616	羧肽酶 S1/ Aspergillusoryzae	555	221/567 (38%)	310/567 (54%)	e-102
gi   435818	羧肽酶 S1 CPD-S1/Penicillium janthinellum	423	174/499 (34%)	258/499 (51%)	4e-77
gi   995456	前 - 羧肽酶 Z/Absidia zychae	460	155/491 (31%)	243/491 (49%)	2e-58
gi   19115337	羧肽酶 Y/ Schizoaccharomyces prombe	1002	137/480 (28%)	215/480 (44%)	1e-46
gi   4028158	羧肽酶 Y 前体 ; vacuolar 羧肽酶 / Pichiaangusta	541	139/508 (27%)	226/508 (44%)	2e-43

[0151] 表 6F. ruCBPS1' BLASTP 结果

[0152]

基因目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   15004616	羧肽酶 S1/Aspergillus oryzae	555	222/567 (39%)	310/567 (54%)	7e-98
gi   435818	羧肽酶 S1 CPD-S1/Penicillium janthinellum	423	174/499 (34%)	259/499 (51%)	1e-71
gi   995456	前 - 羧肽酶 Z/Absidiazychae	460	156/491 (31%)	244/491 (49%)	2e-57
gi   19115337	羧肽酶 Y/ Schizoaccharomycesprombe	1002	137/480 (28%)	215/480 (44%)	4e-44

[0153] ruPAP

[0154] ruPAP 是红色毛癣菌的脯氨酰氨肽酶。1795 个核苷酸 (SEQ ID NO :19) 的 ruPAP 基因组 DNA 序列如表 7A 所示。

[0155]

表 7A、ruPAP 基因组核苷酸序列 (SEQIDNO :19)

```

ATGCGCTTTGCTGCTAGCATTGCCGTGGCCCTGCCAGTCATTCACGCGCGAGTGCTCAAGGCTTCCCTCCACCCGTTAAGGGCGTC
ACCGTGGTCAAATCCAAGTTCGACGAAAACGTAAAGATCACATACAAGGAGGTATGTGTTTACATCATTTTCACATCCAGATCTTAT
ATCCTTACAATAAATCTGGCTAACTACTGGATAGAATGACATATGTGAAACCACTCAAGGAGTTAGATCATTACCAGGTCATGTCC
ACCTTCCTCCAGACAACGATGACTTTGGTGTCTACCGGAACTACTCCATCAACACATTCTTCTGGTTCTTTGAAGCTCGTGAAGACC
CTAAGAATGCTCCTCTCTCCATCTGGCTGAACGGTGGTCCGGGATCGTCATCCATGATTGGACTCTTCCAGGAAAACGGTCCATGCT
GGGTCAATGAAGACTCTAAAATCTACCACCAACAATTCATTTTCATGGAACAATAAAGTAAATAATGCTCTACATTGATCAGCCAAACC
AAGTCGGTTTCAGTTATGACGTACCTACCAACATCACTTACTCTACCATCAATGATACAATATCTGTTGCGGACTTCTCTAACGGTG
TCCCTGCGCAAAATCTTTCTACGTTGGTTGGAACCGGCAGCAGCCAGAACCCTTGGGCAACTGCCAATAACACTGTGAACGCTGCTC
GTTCTATCTGGCACTTTGCACAAGTGTGGTTCCAGGAATTCCTGAAACACAAGCCTAACAAATAACAAGATCAGTATTTGGACAGAGT
CCTATGGAGGAAGATATGGTCCCTCATTGCTCTTACTTCCAGGAACAGAACGAAAAGATCAAAAACCATACCATTACTGAAGAAG
GAGAGATGCATATCTGAACCTCGACACCCTCGGTATCATCAACGGCTGCATCGATCTTATGTTCCAAGCAGAAAAGTTATGCTGAAT

```



表 7A、ruPAP 基因组核苷酸序列 (SEQIDNO :19)

```
TCCCATACAACAACACCTATGGCATCAAAGCTTATACCAAGGAGAAGCGTGACGCTATATTACACGACATCCACCGTCTGACGGCT
GCTTCGACAAGGTTACCAAGTGCCGTGAGGCCGCGAAAAGAAGGAGACCCTCACTTCTACAGCAACAATGCAACCGTCAACACAATCT
GTGCGGATGCTAACTCTGCCTGCGACAAATATCTAATGGATCCTTTCCAAGAGACCAATCTTGGTTACTATGATATTGCTCATCCTC
TTCAGGATCCCTTCCCCCACCATTCTATAAGGGCTTCCTCAGCCAATCCAGCGTTCTATCTGACATGGGATCGCCAGTCAACTTCT
CCCAATACGCCAAGCTGTGGGAAAATCATTCCATGGAGTTGGCGACTACGCTCGCCCTGATGTGCGCGGCTTACCAGGTGACATTG
CTTATCTTCTCGAGAGCGGAGTCAAGGTTGCTCTCGTCTATGGTGACAGAGACTACATCTGCAATTGGTTCGGTGGTGGAGAGTCA
GTCTTGGCTTGAACACTACACTGGCACCAAGACTTCCACAGGGCAAAATATGCCGATGTCAAGGCAACTTTCATACGTCGGAGGCG
TAGTGCCTCAACATGGAACTTCTCTTTCCAGAGTTTTCGAGGCCGGTCAATGAAGTCCCTGGTTACCAACCCGAGACTGCCCTCA
AGATCTTTGAGCGCATCATGTTCAACAAGGATATTTCTACCGTGAGATCGACATTGCTCAGAAAACCAGACTACGGTACCCTGGAA
CTGAGTCTACGTTCCATATCAAAAACGATATCCCTCCTTCGCTGAGCCGACCTGCTACCTCCTCAGTGTGACGGAACCTGTACCC
CGGAGCAGCTTAATGCTATTAAGGATGGAATGCAGTTGTTGAGAATACATTATTAAGAGCCCTGCTGCGTCAAGGGGAAACCCCTC
CACCAACCACGACCTCATCTCCACAGCAGCCCTACCGCTGGAAGTGCCATGCTAAAGGCTCCTGTGGCAATGCTAGCAATATCAG
CTCTCACTGTCCTTGCTTTCTTGTAG
```

[0156] 1326 个核苷酸 (SEQ ID NO :20) 的 ruPAP 见表 7B, 其中 ruPAP 开放阅读框 (“ORF”) 起始子 ATG 在第一个位置上 (表 7B 中下划线)。

[0157]

表 7B、ruPAP 核苷酸序列 (SEQIDNO :20)

```
ATGCAAGCAGCAAAATGTTGAGCCGGTACTGGCAAAATGTACCTGGAAGACTGAGGGTATCTGAGCTCCTTTTCGATGTCCCTTA
GACTACTCAAACCCGTCTTCCACTTCGCTCCGGTTGTCGCCAGGAGTGTGCAGCGGCGAATCCAGGGTCTCTCTCGATGATAAA
GACAGACAGCTACCCTGGATTGTTTTCTGCAGGGTGGACCAGGAGGAGCTTGCCACAACCTCAGGAGGTAGGCTGGGTGGGCCA
TTGCTGGATCGAGGATTCCAGATCCTTCTCCTTGACCAGCGAGGAACAGGGCTTTCAACCCCTATAACCGCTGCGACGCTTGCTCTT
CAGGGAAACGAGTAAAGCAAGCCGAATATCTTAGGCTATTCGCTGCCGATAATATCGTGCAGACTGTGAAGCAGTCCGTAAACTA
TTGACTGCTTATTACCCTCCAGATAAGCAGAAATGGAGCGTCTTGGCCAGAGTTTTGGAGGATTCTGTGCCGTACGATGTTTCT
AATCCTGAGGGACTTAAAGAAGTCTTCACAACTGGTGGATTACCCCTCTTGTGTCAAAGCCTGATCCTGTGTACGAGAGGACCTAC
GACAAGGTCCAGTCCCGAAATAAGTGTACTATTCCACTTTCCCCGAAGCAGAAAGATCGAGTGCAGGATTACTCAAGCATCTCCAA
ACCCAGATGTTAAGCTCCCGATGGCTCACCGTAACTCCGGAACGCTTTCTCCAGCTAGGAATTCATTTGGAATGAAAGGCATA
ATTTGAAGTGCATTAATGAACTGGAATACTTTGGCTTCCTCACACGACCTACTTTATCTCTGATTGAGAACGACACAGTGCAGAC
AACGGCATTCTATATGCCATAATGCATGAATCTATCTACTGCCAAGGGGAGGCTCAAACCTGGGCTGCCGAAAGACTACTACCAAAG
TTCTCTGGCTTCGAGGCGCTATAATCCTGATGGCATCTACTTACTGGGGAGATGGTATACAAACACTGGTTTGAGTCGTCCACA
GAACTCGGCCAGCTCAAAGAGGTAGCCGATATTCTTGCTTCTACAAATGACTGGCCGAGTTGTATGATAAGGAACAGCTCGCGCGC
AACGAGGTGCCAGTGATTTCCGCTACATATGTCGAGGATATGTACGTGCACCTCAGCTACGCCAACGAAACAGCTGCCACTATTAC
AATTGCAAACAGTTCATCACAACACGATGTACCACAACGGACTGCGTTCAGATTCCGCTGAACCTATTGCGCAGCTGTTTGCTCTT
CGTGATGATACGATTGACTAG
```

[0158] 上述 ruPAP 核苷酸序列 (SEQ ID NO :20) 编码含 441 个氨基酸残基 (SEQ ID NO : 21) 的蛋白, 在表 7C 中以一个字母的氨基酸代码表示。

[0159]

表 7C、ruPAP 编码的蛋白序列 (SEQIDNO :21)

```
MQAAKLLSRYWQNPVGRRLRVSELLFDVPLDYSNPSSTSLRRLFARSVQRRIPGSSLDKDRQLPWIVFLQGGPGGACQPQEVGWVGP
LLDRGFQILLLDQRGTGLSTPITAATLALQGNVAVKQAEYLRFRADNIVRDCEAVRKLTTAYYPPDKQKWSVLGGQSFGGFCAVTYVS
NPEGLKEVFTTGGLPPLVSKPDPVYERTYDKVQSRNKVYYSFPEDEDVRRIILKHLQTHDVKLPDGSPLTPERFLQLGIHFGMKG I
ILKCINELEYFGFLTRPTLSLIENDTSADNGILYA IMHESIYQCQEASNWAERLLPKFSGFRGAHNPDGIYFTGEMVYKHWFESST
ELGQLKEVADILASYNWPQLYDKEQLARNEVPVYSATYVEDMYVHFHSYANETAATIHCKQFITNTMYHNGLRSDSAELIAQLFAL
RDDTID
```

[0160] 上述 ruPAP 具有与表 7D、7E 和 7F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0161] 表 7D. ruPAP TBLASTN 结果

[0162]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   14329656	Aspergillus	3752	151/307	190/307 (61%)	e-118
	negerpapA 的前氨肽酶 A 基因		(49%)	61%)	
gi   32414442	Neurosporacrassa 菌株 OR74A	1449	212/477 (44%)	285/477 (59%)	e-100
gi   604877	Aeromonassobria 的前氨 肽酶基因	1740	175/420 (41%)	239/420 (56%)	4e-77

[0163] 表 7E. ruPAP BLASTX 结果

[0164]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   18307408	脯氨酰氨肽酶 A/Aspergillus neger	442	266/442 (60%)	334/442 (75%)	e-152
gi   4456054	推定的脯氨酰氨肽酶 A/Aspergillus nidulans	365	211/366 (57%)	263/366 (71%)	e-144
gi   22507295	脯氨酰氨肽酶 / Talaromycesemersonii	300	181/301 (60%)	226/301 (75%)	4e-99

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   236731	脯氨酰氨肽酶 A/ Aeromonassobria	425	175/420 (41%)	239/420 (56%)	4e-76

[0165] 表 7F. ruPAP BLASTP 结果

[0166]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   18307408	脯氨酰氨肽酶 A/Aspergillus neger	442	267/443 (60%)	336/442 (75%)	e-157
gi   4456054	推定的脯氨酰氨肽酶 /Aspergillus nidulans	365	211/366 (57%)	263/366 (71%)	e-116
gi   22507295	脯氨酰氨肽酶 / Talaromycesemersonii	300	181/301 (60%)	226/301 (75%)	e-102
gi   236731	脯氨酰氨肽酶 / Aeromonassobria	425	175/420 (41%)	239/420 (56%)	2e-78

[0167] ruAMPP

[0168] ruAMPP 是红色毛癣菌的氨肽酶 P。2418 个核苷酸 (SEQ ID NO:22) 的 ruAMPP 基因组

[0169] DNA 序列如表 8A 所示。

[0170]

表 8A、ruAMPP 基因组核苷酸序列 (SEQIDNO :22)

```

ATGCCGCCACCACCGTTGACACGACCCAGCGTCTCGCAAAGCTGCGAGAGCTGATGGCTCAGAACAAGGTCGATGTATATAGTATG
CAATTCAGATACACCATTAAGCTCCCTTGATAATAACAGTCGTATACTCATTTCTTCTTCTACTCCTCGCCTTAAAGTTGTGC
CTTCGGAAGACAGCCATCAGTCGGAGTACATTGCTCCATGTGATGGGCGTCGAGGTTAGACCTGTCCCTCCATAAAAGAATACCTAC
CCGTAATACCAGCCGGCAGACGCTCATACTGATCAGCTTTTCAATCCAGCTTCACTGGCTCGGCAGGATGTGCCATCGTCTC
TATGAGTAAAGCTGCTCTGTCTACAGACGGCAGATACTTCAGCCAAGCTGCAAAAACAGCTCGATGCCAACTGGATCCTGTGAAGCG
AGGTGTCGAGGGTGCCCAACCTGGGAAGAATGGTATATCTGCCCTGGTATCGACTTTTCCGGTATAATGGTTGACAGGCTGGATA
TAGGACCGCTGAGCAGGCCGAGACACGGCAAGGTTGTGGGTGTTGACCCGTCACTTATTACGGCAGGTGAGAATCTACAGTATGCCGT
CTCTTACAAGTGTCATCGTACTAAGTGTATGTTATAGCGGATGCACGAAAGCTTTCTCAGACGTTGAAGACCACCGGAGGCTCCTT
GGTTGGAATTGATCAGAACCCTGATTGATGCCGTCTGGGAGATGAACGTCCTGCACGGCTGCCAAACCAAATTACGGTACAGCCTGT
TGAGCGCGCGGAAAAGTCATTCGAGGAGAAAAGTGAAGACCTGCGAAAAGGAATTGACTGCGAAGAAGAGGCTGCTATGGTTATTG
TATGACGCTAGATCTATTTTGTATCAACATATACTAAACAAACGCAATATAGCCACCTTGGATGAGATTGCATGGCTCTTCAACCTC

```

表 8A、ruAMPP 基因组核苷酸序列 (SEQIDNO :22)

```

CGTGAAGCGAGTAAGTTTCTATATAAAATGGTATCTTTCACTTTATACAAAAAGCCATGCTGACTGGTGTAGTATTCCATATAACCC
CGTCTTTTTCTCGTACGCAATTGTGACGCCCTCAGTTGCGGAACTCTATGTGCGATGAGAGCAAGCTGTCTCCAGAAGCCAGAAAAACA
TCTCGAAGGCAAGGTCGTTCTCAAGCCATACGAGTCCATCTTCCAAGCTTCCAAAGTCCCTCGCCGAATCAAAGGCATCGGCTAGCAG
CGGTTCTCTGGGAAGTTCTTGTGTCTAACAAGGCTTCGTGGTCTTTGAGCCTCGCCCTCGGTGGGGAACAGAACGTCGTTGAGGT
TCGAAATCCCATCACTGACGCCAAAGCCATCAAGAACGAAGTTGAACTGGAAGGATTCAGAAAATGCCATATCCGAGACGGTGCAGC
TCTGATCGAGTACTTCGCCTGGCTTGAAAATGCATTGATCAAGAAGGTGCCAAGCTAGACGAAGTAGATGGAGCCGACAAAACCTTT
CGAGATCCGCAAGAAAATGACCTCTTCGTCGGCAACTCCTTCGACACCATCTCTTCTACCGGTGCTAACGGTGTACCATTCAATTA
CAAACCCGAGAAGTCAACTTGGCTATCATTGACCCGAAGGCTATGTACCTGTGTGACTCTGGTGGCCAAATACCTTGATGGTACTAC
TGATACTACCCGAAGTCTCCACTTTGGAGAGCCACGGAGTTCAGAGAAGGCTTATGCACTTGTCTAAAGGGACATATCAGCAT
TGACAAATGCCATTTTCCCAAAGGAACCACCGGATACGCCATTGACTCGTTTGCTCGACAGCATTTGTGGAAGGAGGGTCTGGATTA
CCTCCACGGCACCGGTCATGGTGTGGCTCATTTTTGGTACGGGGTTCCTTTTTCTTTTTTTTTTCTTTTTTTATTTTTATTATTA
CTTCTTAGGCTAACACATTCTCTCTAAGAACGTCCATGAGGGACCTATGGGCATAGGAAGCCGTGCTCAGTACGCTGAAGTTCCT
CTCTCTGCCAGCAATGTTCTTCCAACGGTAGGATTCTGCATCTCATCTTTCTTGAATCCTACTAATTGCAAAAATAGAGCCTGGAT
ATTATGAAGACGGCAACTTCGGCATTGCTCTCGAGAGTAAGTTCAATGACTGCGTATTCTAGTTTTTTCATACTGACGGCCTCTTTA
GACCTCGTAATCTGCAAGGAGGTCCAGACTGCACACAAAATTCGGCGACAAGCCCTTCCTCGGATTTGAGTCCATCACCTGGTACCT
TTCTGCCAAAAACTCCTTGATGCTTCTCTTGACCGAAGCTGAGAGAAAAGTGGGTGAATGATTACCATGCGAAAAGTCTGGGAGAAG
ACCAGTCCCTTCTTTGAGAAGGACGAGTTAACAACCGCCTGGCTAAAGCGCGAGACACAACCTATTTAA

```

[0171] 1878 个核苷酸 (SEQ ID NO :23) 的 ruAMPP 见表 8B, 其中 ruAMPP 开放阅读框 (“ORF”) 起始子 ATG 在第一个位置上 (表 8B 中下划线)。

[0172]

表 8B、ruAMPP 核苷酸序列 (SEQIDNO :23)

```

ATGCCGCCACCACCGTTGACACGACCCAGCGTCTCGCAAAGCTGCGAGAGCTGATGGCTCAGAACAAGGTCGATGTATATATTGTG
CCTTCGGAAGACAGCCATCAGTCGGAGTACATTGCTCCATGTGATGGCGCTCGAGCTTCATATCCAGCTTCACTGGCTCGGCAGGA
TGTGCCATCGTCTCTATGAGTAAAGCTGCTCTGTCTACAGACGGCAGATACTCAGCCAAGCTGCAAAACAGCTCGATGCCAACTGG
ATCCTGTTGAAGCGAGGTGTCGAGGGTGTCCCAACCTGGGAAGAATGGACCGCTGAGCAGGCCGAGACACGGCAAGGTTGTGGGTCG
GATGCACGAAAGCTTCTCAGACGTTGAAGACCACCGGAGGCTCCTTGGTTGGAATTGATCAGAACCTGATTGATGCCGTCTGGGGA
GATGAACGTCCTGCACGGCTGCCAACCAAAATTACGGTACAGCCTGTTGAGCGCGGGAAAGTCATTCGAGGAGAAAGTGAAGAC
CTGGCAAAGGAATTGACTGCGAAGAAGAGGTCTGCTATGGTTATTTTCGAGTAAGTTTCTATATAAAATGGTATCTTTCACCTTATACA
AAAAGCCAATGCTGACTGGTGTAGTATTCCATATAACCCCGTCTTTTTCTCGTACGCAATTGTGACGCCCTCAGTTGCCGAACTCTAT
GTGCATGAGAGCAAGCTGTCTCCAGAAGCCAGAAAACATCTCGAAGGCAAGGTCGTTCTCAAGCCATACGAGTCCATCTTCCAAGCT
TCCAAAGTCTCGCGAATCAAAGGCATCGGCTAGCAGCGTCTCTGCGAAGTTCCTGTTGTCTAACAAAGGCTTCGTGGTCTTTG
AGCCTCGCCCTCGGTGGGGAACAGAACGTCGTTGAGGTTGCAAGTCCCATCACTGACGCCAAAGCCATCAAGAACGAAGTGAAGTG
GAAGGATTCAGAAAATGCCATATCCGAGACGGTGCAGCTCTGATCGAGTACTTCGCCCTGGCTTAAAAATGCATTGATCAAAGAAGGT
GCCAAGCTAGACGAAGTAGATGGAGCCGACAAACTCTTCGAGATCCGCAAGAAATATGACCTCTTCGTCGGCAACTCCTTCGACACC
ATCTCTTCTACCGTGCTAACGGTGTCTACCATTCATTACAAACCCGAGAAGTCAACTTGCCTATCATTGACCCGAAGGCTATGTAC
CTGTGTGACTCTGGTGGCCAAATACCTTGATGGTACTACTGATACTACCCGAACCTCCACTTTGGAGAGCCCACGGAGTCCAGAAG
AAGGCTTATGCACTTGTCTTAAAGGGACATATCAGCATTGACAATGCCATTTCCCAAAGGAACCACCGGATACGCCATTGACTCG
TTTGTCTGACAGCATTGTGGAAGGAGGCTGATTACCTCCACGGCACCAGTTCATGGTGTGGCTCATTTTTGAACGTCCATGAG
GGACCTATGGGCATAGGAAGCCGTGCTCAGTACGCTGAAGTTCCTCTCTGCGCAGCAATAGCCTGGATAATTATGAAGACGGCAACT
TCGGCATTCTGTCGAGAGTAAGTCAATGACTGCGTATTCTAGTTTTTCTACTGACGGCTCTTTAGACCTCGTAATCTGCAAG
GAGGTCCAGACTGCACACAAAATTCGGCGACAAGCCCTTCTCGGATTTGAGTCCATCACCCCTGGTACCTTTTCGCCAAAAACTCCTT
GATGCTTCTCTTTCGACCAAGCTGAGAGAAAGTGGGTGAATGATTACCATGCGAAAGTCTGGGAGAAGACCAGTCCCTCTTTGAG
AAGGACGAGTTAACAAACCGCTGGCTAAAGCGCGAGACACAACCTATTTAA

```

[0173] 上述 ruAMPP 核苷酸序列 (SEQ ID NO :23) 编码含 625 个氨基酸残基 (SEQ ID NO : 24) 的蛋白, 在表 8C 中以一个字母的氨基酸代码表示。

[0174]

表 8C、ruAMPP 编码的蛋白序列 (SEQIDNO :24)

```

MPPPPVDTTQRLAKLRELMAQNKVDVYIVPSEDSHQSEYIAPCDGRRAFISSFTGSAGCAIVSMSKAALSTDGRYFSQAAKQLDANW
ILLKRGVEGVPTWEEWTAEQAETRQCGSDARKLSQTLKTTGGSLVGDQNLIDAVWDERPARPANQITVQPVVERAGKSFEEKVED
LRKELTAKKRSAMVISSKFLYKWYLSLYTKSHADWCSIPYNPVFFSYAIVTPSVAELYVDESKLSPEARKHLEGGKVVLPYESIFQA
SKVLAESKASASSGSSGKFLLSNKASWSLSLALGGEQNVVEVRSPITDAKA IKNEVELEGFRKCHIRDGAALIEYFAWLENALIKEG
AKLDEVGDADKLFERKKYDLFVGNFSFDTISSTGANGATIHYPKPEKSTCAI IDPKAMYLCDSGGQYLDGTTDTRTLHFGEPTFEQK
KAYALVLKGHISIDNAIFPKGTTGYAIDSFARQHLWKEGLDYLHGTGHVGSFLNVHEGPMGIGSRAQYAEVPLSASNSLDIMKTAT
SAFVSRVSSMTAYSFFILTASLDLVICKEVQTAHKFGDKPFLGFESITLVPFCQKLLDASLLTEAERKWNVDYHAKVWEKTSPPFFE
KDELTTAWLKRETQPI

```

[0175] 上述 ruAMPP 具有与表 8D、8E 和 8F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0176] 表 8D. ruAMPP TBLASTN 结果

[0177]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   32403169	Neurosporacrassa 菌株 OR74A	1845	339/630 (53%)	433/630 (68%)	0.0
gi   20453016	Drosophilamelanogaster 的氨肽 酶 P 基因	12647	268/638 (42%)	369/638 (57%)	e-127
gi   17571207	Drosophilamelanogaster (ApepP) 的染色体 2	12001	268/638 (42%)	369/638 (57%)	e-127
gi   4583560	Drosophilamelanogaster Daminopep-P 基因	2358	268/638 (42%)	369/638 (57%)	e-127

[0178] 表 8E. ruAMPP BLASTX 结果

[0179]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   25529603	X-Pro 氨肽酶细胞质形式 /Drosophila melanogaster	613	268/638 (42%)	369/638 (57%)	e-127
gi   4107172	氨肽酶 P/Drosophila melanogaster	613	258/638  (40%)	369/638  (57%)	e-124
gi   15384991	Xaa-Pro 氨肽酶 2/ Lycopersiconesculentum	654	268/674 (39%)	365/674 (54%)	e-120
gi   8489879	胞液的氨肽酶 P/ 人	623	254/646 (39%)	358/646 (55%)	e-119
gi   2584787	氨肽酶 P-like/ 人	623	254/646 (39%)	357/646 (55%)	e-119

[0180] 表 8F. ruAMPP BLASTP 结果

[0181]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   30923284	可能的多肽酶 C22G7.01C	598	291/629 (46%)	384/629 (61%)	e-156
gi   25529603	X-Pro 氨肽酶细胞质形式 /Drosophilamelanogaster	613	268/638 (42%)	369/638 (57%)	e-124
gi   15384991	Xaa-Pro 氨肽酶 2/ Lycopersicmesculentum	654	268/674 (39%)	365/674 (54%)	e-123
gi   8489879	胞液的氨肽酶 P/ 人	623	254/646 (39%)	358/646 (55%)	e-122
gi   2584787	氨肽酶 P-like/ 人	623	254/646 (39%)	357/646 (55%)	e-122
gi   4107172	氨肽酶 P/Drosophila melanogaster	613	258/638 (40%)	369/638 (57%)	e-121
gi   18777778	细胞质的氨肽酶 P/ Rattusnorvegicus	623	(39%)	(54%)	e-120
gi   18875372	胞液的氨肽酶 P/ Musmusculus	623	250/645 (38%)	354/645 (54%)	e-118
gi   15384989	Xaa-Pro 氨肽酶 1/ Lycopersicon	655	264/674 (39%)	361/674 (53%)	e-117

[0182] ruPLD

[0183] ruPLD 是红色毛癣菌的脯氨酰氨基酸酶。2344 个核苷酸 (SEQ ID NO :25) 的 ruPLD 基因组 DNA 序列如表 9A 所示。

[0184]

表 9A、ruPLD 基因组核苷酸序列 (SEQIDNO :25)

```

ATCAACCTCACCTCTTCACCGTCTCACGCCCTTCGTCCCGTCCAACCTTCATTTCCGCCCTCTCTATGATAACCAACAAACATCCGC
TGTTATGTAATCGAACCCGCCGTTAGCCATCCCTAGCCCGCGTTTTCTCCCAGCATCAATACGACCGAAATGAAGACAGACGGGGA
AGACGAGGCAAAACAATAACACATCAACAATTAACCCGTTGCCGTCTTCTACCCATCTTGTCTACGCATCGTCCAACCTTTTCTTG
CCCTATATCAGCCGAACTCGGCCATCATGGATATCCACGTCGACAAAATACCCGGCTAAGAGTCACGCCAGGCGCTCGCCGAGAAGC
TCAAGCCCGCGGGCAGCGCTCTACCGGCATCATCTTCGTGCAAGGCCAAAAGGAGCATAATTATCGATGATAGCGACGAGCCGTTTC
ACTTCCGGTGAGCCGTGGGAATACACTCGACTGGGCGGAATAAGCTAACAAAAGGGTGTGATAGTCAACGCCGAAACTTCCTCTATC
TGTCGGCTGTCTTGAGGCCGAGTGTCCGTTGCATACAACATCGAGAAAAGATGAGCTTACATTGTTTCATCCACCAGTCGACCCAG
CCTCGGTTATGTGTCCGGCTCCCTCTTGAGCCCGCCGAAAGCCTTGAAGCAGTTCGATGTTGATGCCGTGCTCCTCACAACCTGAGA
TAAACAACCTATCTCGCAAGTGTGGGGCGAGAAGGCTTCCACCATTGCAGACAGAGTTTGCCCGGAGGTCTCCTTCTCATCCTTCA
AGCACAACGACACCGATGCCCTGAAGCTTGCCATCGAGTCTGCGGTATAGTAAAAGACGAGTATGAAATGGTCTTCTCCGACGTG
CTAATGAGGTCTCCAGCCAAGCTCATAATTGAAGTATGAAAGCCGCAACCAAGTCAAAGAACGAGAGAGACTCTATGCTACTCTCA
ACTATGCTGCATGCTAATGGCTGCTCCGACCAGTCTTACCA TCCAATCTTGCATGTGGCCCAATGTGCCACTCTCCACTACA
CCAAGAACAACGGTGACCTAACTAACCCGGCTACCGGGATTAAGGACCAGCTCGTACTTATCGACGCTGGATGCCAGTACAAGGCGT
ACTGTGCAGATACACTCGTGCATTTCCCTTGTCCGGCAAATTCACCACGAGGGCCGCGCAGATCTATGATATTGCCCTTGGAGATGC
AGAAAGTCGCGTTTGGCATGATCAAACCTAATGTTTTGTTGACGACATGCATGCTGCGGTCCACCGGTTGCGATCAAGGGGCTGC
TCAAGATTGGCATTCTCACTGGCTCTGAGGATGAGATTTTCGATAAGGGAATCAGCACTGCCTTTTCCACATGGTCTAGGCCACC
ATCTCGGCATGGACACTCAGATGTTGGAGGAAACCCTAACCCGGCTGACCCGAA TCGCATGTTTAAATACTTGGCTCTGCGAGGCA
CTGTTCCAGAGGGATCCGTCATTACAATTGAGCCCGGTGTAAGTGTGAATCGAGTAGTTGCTCCGCCAATGTTTCACATACATTT
ACTAACCTTGTCTTAGGTCTACTTCTGCCGTTACATCATTTAGGCCATTCCTTACTAACCCCGAGACCAGCAAGTACATCAACTCCG
AAGTTCTAGACAAGTACTGGGCTGTTGGAGGTGTACGATCGAGGACAACGTCGTCGTCGCCGCAATGGCTTTGAGAACCTGACCA
CGGTGCCAAAGGAGCCCGAGGAGTTCGAACGCATTGTCCAGGAGGGTCTAAATAATTATGTTTTTATTTCAGTACACCGAGTGGTGC
GACACACGAGGAGCATGTACATATTTATGATCTACCCAGTTGATTTGCTACCAAAAAAAGACCGACCACAGCCCTATTTATTGATA
TTACATAGTAGGAATAAAGGCCACTTTGCCACCAGCAATAATAACAATAAGAAAAGCAACTACTCGTACAACCAGCCTAGAAAGCT
CTAGACCTCTTTCTCGCTGGGCCCTGAATGCCGGCTACTGGTGTATCACGCTCCCTGGCCCTCTTCTCCTCATGTCCAACACC
CGATTAAGCAAATCGAAACTGAACTGGGGATGCTCAAGACACAATGCCCTTGAAGTGTCTTTCAGCATCATGACGAGCAGCATCACTC
ATCTTAGCCAGAAGCGAGCAACCGTCTCTGATAGCAGTGTCTCCGGCGTGGTATGGCTGTACACGTATCTCGCATACTCGATC
TCACCCGTAGCACTACTCTCGATGCTACCAATCTTGTCTGAGCAAGCAGTTTGAAGTTTTTCGTTCCGAGCTTTTCGGCCA

```

[0185] 1401 个核苷酸 (SEQ ID NO :26) 的 ruPLD 见表 9B, 其中获得的 ruPLD 部分开放阅读框 (“ORF”) 序列缺失在第一个位置上的起始子 ATG。

[0186]

表 9B、ruPLD 核苷酸序列 (SEQIDNO :26)

```

CCGAACTCGCCATCATGGATATCCACGTCGACAAAATACCCGGCTAAGAGTCACGCCAGGCGCTCGCCGAGAAGCTCAAGGCCGCG
GGGACGGCTCTACCGGCATCATCTTCGTGCAAGGCCAAAAGGAGCATAATTATCGATGATAGCGACGAGCCGTTTCACTTCCGTCAA
CGCCGAAACTTCCTCTATCTGTCCGGCTGTCTTGAGGCCGAGTGTCCGTTGCATACAACATCGAGAAAGATGAGCTTACATTGTTT
ATTCCACCAGTCGACCCAGCCTCGGTTATGTGGTCCGGCTCCCTCTTGAGCCCGCCGAAAGCCTTGAAGCAGTTCGATGTTGATGCC
GTGCTCCTCACAACCTGAGATAAAACAACCTATCTCGCAAGTGTGGGGCGAGAAGGCTTCCACCATTGCAGACAGAGTTTGCCCGAG
GTCTCCTTCTATCCTTCAAGCACAACGACACCGATGCCCTGAAGCTTGCCATCGAGTCTGCGGTATAGTAAAAGACGAGTATGAA
ATTGGTCTTCTCCGAGTGCTAATGAGGTCTCCAGCCAAGCTCATAATTGAAGTATGAAAGCCGCAACCAAGTCAAAGAACGAGAGA
GAGTCTATGCTACTCTCAACTATGTCTGCATGTCTAATGGCTGCTCCGACCAGTCTTACCATCCAATCTTGCATGTGGCCCAAT
GCTGCCACTCTCCACTACCAAGAACAACGGTGACCTAACTAACCCGGCTACCGGGATTAAGGACCAGCTCGTACTTATCGACGCT

```



表 9B、ruPLD 核苷酸序列 (SEQIDNO :26)

```
GGATGCCAGTACAAGGCGTACTGTGCAGATATCACTCGTGCATTCCCCTTGTCCGGCAAATTCACCACGGAGGGCCGCCAGATCTAT
GATATTGCCTTGGAGATGCAGAAAGTCGCGTTTGGCATGATCAAACCTAATGTTTTGTTTCGACGACATGCATGCTGCGGTCCACCGG
GTTGCGATCAAGGGGCTGCTCAAGATTGGCATTCTCACTGGCTCTGAGGATGAGATTTTCGATAAGGGAATCAGCACTGCCTTTTTTC
CCACATGGTCTAGGCCACCATCTCGGCATGGACACTCACGATGTTGGAGGAAACCCTAACCCGGCTGACCCGAATCGCATGTTTAAA
TACTTGGCTCTGCGAGGCACTGTTCCAGAGGGATCCGTCATTACAATTGAGCCCGGTGTCTACTTCTGCCGTTACATCATTGAGCCA
TTCTTACTAACCCCGAGACCAGCAAGTACATCAACTCCGAAGTTCTAGACAAGTACTGGGCTGTTGGAGGTGTACGTATCGAGGAC
AACGTCGTCGTCGCGCCAATGGCTTTGAGAACCTGACCACGGTGCCAAAGGAGCCCGAGGAGGTGCAACGCATTGTCCAGGAGGGT
GCTAAATAA
```

[0187] 上述部分 ruPLD 核苷酸序列 (SEQ ID NO :26) 编码含 466 个氨基酸残基 (SEQ ID NO :27) 部分的蛋白,在表 9C 中以一个字母的氨基酸代码表示。

[0188]

表 9C、ruPLD 编码的蛋白序列 (SEQIDNO :27)

```
PNSA IMDIHVDKYPKSHARRVAEKLKAAGHGSTGI IFVEGQKEHI IDSDPEPFHFRQRNFL YLSGCLAECSVAYNIEKDELTLF
IPPVDPASVMWSGLPLEPAEALKQFDVDAVLLTTE INNYLAKCGGEKVFTIADRVCPVEVSFSSFKHNDTDALKLAIESCRIKDEYE
IGLLRRANEVSSQAHIEVMKAATKSKNERELYATLNYVCMNSGSDQSYHP ILACGPNAA TLHYTKNNGDLTNPATG IKDQLVLIDA
GCQYKAYCADITRAFPLSGKFTTEGRQIYDIALEMQKVAFGMIKPNVLFDDMHAAVHRVAIKGLLKIGILTGSEDEIFDKGISTAFF
PHGLGHHLGMDTHDVGGNPNPADPNRMFKYLRLRGTVPEGSVITIEPGVYFCRYIIEPFLTNPETSKYINSEVLDKYWAVGGVRIED
NVVVRANGFENLTTVPKEPEEVERIVQEGAK
```

[0189] 上述 ruPLD 具有与表 9D、9E 和 9F 中的 BLAST 氨基酸序列同源性。由 parawise blast 程序分析的数据。

[0190] 表 9D. ruPLD TBLASTN 结果

[0191]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   14272360	脯氨酰氨基酸酶的 AspergillusnidulanspepP 基 因	2632	199/348 (57%)	249/348 (71%)	e-143
gi   32420910	Neurosporacrassa 菌株 OR74A	2562	235/457 (51%)	324/457 (70%)	e-136
gi   3114965	脯氨酰氨基酸酶的 Suberites domuncullamRNA, 形式 1	1688	157/464 (33%)	235/464 (50%)	4e-66
gi   22531161	ArabidopsisthalianaX-Pro 类二肽酶蛋白	1672	160/477 (33%)	242/477 (50%)	2e-64

[0192] 表 9E. ruPLD BLASTX 结果

[0193]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   14272361	脯氨酰氨基酸酶 /Emericellanidulans	496	267/463 (57%)	336/463 (72%)	e-153
gi   3114966	脯氨酰氨基酸酶 /Suberitesdomuncula	501	157/464 (33%)	235/464 (50%)	1e-66
gi   22531162	X-Pro 类二肽酶蛋白 /Arabidopsisthaliana	486	160/477 (33%)	242/477 (50%)	6e-65
gi   30582223	多肽酶 D/ 人	493	152/452 (33%)	231/452 (51%)	2e-63
gi   20271451	多肽酶 D/ 人	493	152/452 (33%)	230/452 (50%)	3e-63

[0194] 表 9F. ruPLD BLASTP 结果

[0195]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   14272361	脯氨酰氨基酸酶 /Emericella nidulans	496	267/463 (57%)	336/463 (72%)	e-158
gi   3114966	脯氨酰氨基酸酶 /Suberites domuncula	501	158/466 (33%)	235/466 (50%)	6e-67
gi   22531162	X-Pro 类二肽酶蛋白 /Arabidopsisthaliana	486	159/477 (33%)	241/477 (50%)	6e-64
gi   30584879	人 多肽酶 D	494	152/452 (33%)	231/452 (51%)	2e-63
gi   15929143	多肽酶 D/ 人	493	152/452 (33%)	231/452 (51%)	2e-63
gi   20271451	多肽酶 D/ 人	493	152/452 (33%)	230/452 (50%)	4e-63

[0196] caLAP2

[0197] caLAP2 是犬小孢子菌的亮氨酸氨肽酶。1730 个核苷酸 (SEQ ID NO :28) 的 caLAP2 基因组 DNA 序列如表 10A 所示。

[0198]

表 10A、caLAP2 基因组核苷酸序列 (SEQIDNO :28)

```
ATGAAGACACAGTTGTTGAGTCTGGGAGTTGCCCTCACGGCCATCTCTCAGGGCGTTATTGCTGAGGATGCCTTGAACCTGGCCAT
TCAAGCCGTTGGTTAATGCTGTGAGTATATACACAAGATCGATCGATCGTCTCTTGTCCCTGTCACTTATCGCTCTACAGTAAG
CAAAAATACTGGAGAATCATGTGCTGATGTAATGTATAGGATGACCTGCAAAAACAAGATTAAGCTCAAGGATCTTATGGCTGGC
GTACAGAAACTCCAAGACTTCGCCTACGCTCACCCCTGAGAAGAATCGAGTATTCGGTGGTGTGCGCCACAAGGATACCGTCGACT
GGATCTACAATGAGCTCAAGGCTACCGGCTACTACGATGTGAAGATGCAGCCACAAGTCCACCTGTGGTCTCATGCTGAGGCAGC
TGTCAATGCCAATGGCAAGGATCTCACTGCCAGTGCCAATGTCCTACAGCCCTCCAGCCGACAAGATCACTGCCGAGCTTGTCTG
GCCAAGAACATGGGATGCAATGCTGTATGTGCGCCCTTTTCCATTCTATATATCGACTGGTCGCTTGGAAATTCAGAAGAGCTG
ACAATTGCAAAACAGACTGATTACCCAGAGGGTACCAAGGGCAAGATTGTCTCATCGAGCGTGGTGTCTGCAGCTTTGGCGAGAA
GTCCGCTCAGGCTGGCGATGCAAAGGCTATTGGTGCCAATCGTCTACAACAACGTCCCTGGAAGCTTGGCCGGCACCCCTGGGTGGC
CTTGACAACCGCCATGCTCCAACCTGCTGGAATCTCTCAGGCTGATGGAAGAACCTCGTAGCCTTGTGCGCTCTGGCAAGGTTA
CCGTCACCATGAACGTTATCAGCAAGTTTGAGAACAGGACTACGTGAGTATTGTTCCATACTTTGGTCAACAATGATATATACAC
GTACTAACACTGCTCTATAGCTGGAACGTCATTGCCGAGACCAAGGGAGGAGACCACAACAACGTCATCATGCTCGGTTCTCACT
CTGACTCTGTGACGCCGCGCCCTGGTATCAACGACAACGGCTCCGGTACCATTGGTATCATGACCGTTGCCAAAAGCCCTCACCAA
CTTCAAGGTCAACAACGCCGTCGCTTCGGCTGGTGGACCGCCGAGGAGTTCGGCCTTCTCGGCAGCACTTCTACGTCGACAGC
CTTGACGACCGTGAACGTCACAAGGTCAAGCTGTACCTCAACTTCGACATGATTGGCTCCCCCAACTTCGCCAACAGATCTACG
ACGGAGACGGCTCCGCCTACAACATGACTGGCCCCGCCGATCTGCTGAAATCGAGTACCTGTTGAGAAGTCTTCGATGACCA
GGGAATCCCACACCAGCCCACCGCCTCACCGGCCGCTCCGACTACTCTGCCTTCAATCAAGCGCAACGTCCCTGCCGGAGGTCTG
TTTACTGGTGCTGAGGTCGTCAAGACCGCCGAGCAGGCTAAGCTATTTGGCGGCGAGGCTGGCGTTGCTTATGACAAGAACTACC
ACGGCAAGGGCGACTGTAGACAACATCAACAAGGGTGTATCTACCTCAACTCGAGGAAATCGCGTATGCCACTGCTCAGTA
TGCTAGTTCGCTGCGCGGATTCCCAACCCGCCAAAGACGGGTAAGCGTGACGTGAGCCCCGTGGCCAGTCTATGCCTGGTGGT
GGATGCGGACACCACAGCGTCTTCATGTAA
```

[0199] 1488 个核苷酸的 caLAP2 见表 10B, 其中起始子 ATG 在第一个位置上 (表 10B 中下划线)。

[0200]

表 10B、caLAP2 核苷酸序列 (SEQIDNO :29)

```
ATGAAGACACAGTTGTTGAGTCTGGGAGTTGCCCTCACGGCCATCTCTCAGGGCGTTATTGCTGAGGATGCCTTGAACCTGGCCAT
TCAAGCCGTTGGTTAATGCTGATGACCTGCAAAAACAAGATTAAGCTCAAGGATCTTATGGCTGGCGTACAGAAACTCCAAGACTT
CGCCTACGCTCACCCCTGAGAAGAATCGAGTATTCGGTGGTGTGCTGGCCACAAGGATACCGTCGACTGGATCTACAATGAGCTCAAG
GCTACCGGCTACTACGATGTGAAGATGCAGCCACAAGTCCACCTGTGGTCTCATGCTGAGGCAGCTGTCAATGCCAATGGCAAGG
ATCTCACTGCCAGTGCCAATGTCCTACAGCCCTCCAGCCGACAAGATCACTGCCGAGCTTGTCTGGCCAAGAACATGGGATGCAA
TGCTACTGATTACCCAGAGGGTACCAAGGGCAAGATTGTCCTCATCGAGCGTGGTGTCTGCAGCTTTGGCGAGAAGTCCGCTCAG
```

表 10B、caLAP2 核苷酸序列 (SEQIDNO :29)

```
GCTGGCGATGCAAAGGCTATTGGTGCCATCGTCTACAACAACGTCCCTGGAAGCTTGGCCGGCACCCCTGGGTGGCCTTGACAACC
GCCATGCTCCAAGTCTGGAATCTCTCAGGCTGATGGAAAGAACCTCGCTAGCCTTGTTCGCCTCTGGCAAGGTTACCGTCACCAT
GAACGTTATCAGCAAGTTTGAGAACAGGACTACCTGGAACGTCATTGCCGAGACCAAGGGAGGAGACCACAACAACGTCATCATG
CTCGGTTCTACTCTGACTCTGTCTGACGCCGGCCCTGGTATCAACGACAACGGCTCCGGTACCATTGGTATCATGACCGTTGCCA
AAGCCCTCACCAACTTCAAGGTCAACAACGCCGTCGGCTTCGGCTGGTGGACCGCCGAGGAGTTCCGGCCTTCTCGGCAGCACTTT
CTACGTCGACAGCCTTGACGACCGTGAAGTGCACAAGGTCAAGCTGTACCTCAACTTCGACATGATTGGCTCCCCAACTTCGCC
AACCAGATCTACGACGGAGACGGCTCCGCCTACAACATGACTGGCCCCGCCGATCTGCTGAAATCGAGTACCTGTTTCGAGAAGT
TCTTCGATGACCAGGGAATCCACACCAGCCACCGCCTTACCGGCCGCTCCGACTACTCTGCCTTCATCAAGCGCAACGTCCC
TGCCGGAGGTCTGTTTACTGGTGTGAGGTCGTAAGACCGCCGAGCAGGCTAAGCTATTTGGCGGCGAGGCTGGCGTTGCTTAT
GACAAGAACTACCACGGCAAGGGCGACACTGTAGACAACATCAACAAGGGTGCTATCTACCTCAACACTCGAGGAATCGCGTATG
CCACTGCTCAGTATGCTAGTTCGCTGCGCGGATTCCCAACCCGCCAAAGACGGGTAAGCGTGACGTGAGCCCCCGTGGCCAGTC
TATGCCTGGTGGTGGATGCGGACACCACAGCGTCTTCATGTAA
```

[0201] 上述 caLAP2 核苷酸序列 (SEQ ID NO :29) 编码含 495 个氨基酸残基 (SEQ ID NO :30) 的蛋白,在表 10C 中以一个字母的氨基酸代码表示。

[0202]

表 10C、caLAP2 编码的蛋白序列 (SEQIDNO :30)

```
MKTQLLSLGVALTAISQGVIAEDALNWPFKPLVNADDLQNKIKLKDLMAGVQKLQDFAYAHPEKNRVFGGAGHKDVTVDWIYNELK
ATGYVDVKMQPQVHLWSHAEEAVNANGKDLTASAMSYSPADKITAELVLAKNMGCNATDYPEGTKGKIVLIERGVCSFGEKSAQ
AGDAKAI GAIVYNNVPGSLAGTLGGLDNRHAPTAGISQADGKNLASLVASGKVTVMNVISKFENRTTWNVIAETKGGDHNNVIM
LGSHSDSVDAGPGINDNGSGTIGIMTVAKALTNFKVNNVRFGWWTAEFGLLGSTFYVDSLDDRELHKVKLYLNFDMIGSPNFA
NQIYDGDGSAYNMTGPAGSAEIEYLFEKFFDDQGIHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTAEQAKLFGGEAGVAY
DKNYHKGKDTVDNINKGAIYLNTRGIAAYATAQYASSLRGFPTRPKTKGRDVS PRGQSMPGGGCGHHSVFM
```

[0203] 上述 caLAP2 具有与表 10D、10E 和 10F 中的 BLAST 氨基酸序列同源性。由 parawise blast 程序分析的数据。

[0204] 表 10D. caLAP2 TBLASTN 结果

[0205]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   600025	Saccharomyces cerevisiae (s288c) RIF1, DPB3, YmL27 和 SNF5 基因	32421	182/477 (38%)	254/477 ( 53%)	8e-77
gi   469463	Saccharomyces cerevisiae 氨 肽酶 Y 基因	2272	182/477 (38%)	254/477 ( 53%)	8e-77

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi 16033407	Bacilluslicheniformis 亮氨酸氨肽酶前体, 基因	2054	132/474 (27%)	215/474( 45%)	3e-27

[0206] 表 10E. caLAP2BLASTX 结果

[0207]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi 6319763	氨肽酶 yscIII ;Ape3p/ Saccharomycescerevisiae	537	182/477 (38%)	254/477 (53%)	9e-78
gi 1077010	氨肽酶 Y 前体, 有液泡的 /Saccharomycescerevisiae	537	182/477 (38%)	254/477 (53%)	9e-78
gi 31791596	可能的脂蛋白氨肽酶 LPQL/ Mycobacteriumbovis	500	188/485 (38%)	269/485 (55%)	3e-77
gi 15839805	水解酶 / Mycobacteriumtuberculosis	493	187/481 (38%)	268/481 (55%)	6e-77

[0208] 表 10F. caLAP2BLASTP 结果

[0209]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi 6319763	氨肽酶 yscIII ;Ape3p/ Saccharomycescerevisiae	563	182/477 (38%)	254/477( 53%)	5e-78
gi 1077010	氨肽酶 Y 前体, 有液泡的 /Saccharomycescerevisiae	537	182/477 (38%)	254/477( 53%)	8e-78
gi 15839805	水解酶 / Mycobacterium tuberculosis	493	187/481 (38%)	268/481 (55%)	1e-71
gi 31617182	可能的脂蛋白氨肽酶 LPQL/ Mycobacteriumbovis	500	188/485 (38%)	269/485 (55%)	2e-71

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   15598135	可能的氨肽酶 / <i>Pseudomonasaeruginosa</i>	536	166/445 (37%)	242/445 (54%)	2e-65

[0210] meLAP2

[0211] meLAP2 是须癣毛癣菌的亮氨酸氨肽酶。1775 个核苷酸 (SEQ ID NO :31) 的 meLAP2 核酸如表 11A 所示。

[0212]

表 11A、meLAP2 基因组核苷酸序列 (SEQIDNO :31)

```

ATGAAGTCGCAACTGTTGAGCCTAGCCGTGGCCGTCACCACCATTTCCCAGGGCGTTGTTGGTCAAGAGCCCTTTGGATGGCCCT
TCAAGCCTATGGTCACTCAGGTGAGTTGCTGTCAACAGATCGATCGATCGATCTACCTTCGTCCCTGTCACCTATAACTCCACAG
CAGGACCAAGAAAACACAAGTTTTCCGGGGAATTCTTATGTGTGATGTAATGTATAGGATGACCTGCAAAACAAGATTAAGCT
CAAGGATATCATGGCAGGTGTCGAGAAGCTGCAAAGCTTTTCTGATGCTCATCTGAAAAGAACCAGTGTTCGGTGGTAATGGC
CACAAGGACACTGTCGAGTGGATCTACAATGAGCTCAAGGCCACCGGCTACTACAATGTGAAGAAGCAGGAGCAGGTACACCTGT
GGTCTCACGCTGAGGCCGCTCTCAGTGCCAATGGCAAGGACCTCAAGGCCAGCGCCATGTCGTACAGCCCTCTGCCAACAAGAT
CATGGCCGAGCTTGTGCTGCCAAGAACAATGGCTGCAATGCTGTAAGTGCCATACTTCCTATAACATCACATTCACCTTTAGAA
TGAAGAGCGCGGGAAGTGAATTTTTTTTTTTTTTTTTTTTTTTTGTAAACAGACCGATTACCCAGAGAACAACCTCAGGGAAAGAT
AGTCCATTCAGCGTGGTGTCTGCAGCTTCGGCGAGAAGTCTTCTCAGGCTGGTGATGCGAAGGCTATTGGTGCCGTTGTCTAC
AACAACGTCCCGGATCCCTTGCTGGCACTCTTGGTGGCCTTGACAAGCGCCATGTCCCAACCCTGGTCTTTCCAGGAGGATG
GAAAGAATCTTGCTAGCCTCGTTGCTTCTGGCAAGGTTGATGTCACCATGAACGTTGTCAGTCTGTTGAGAACCGAACCACGTA
AGTAACTCAACGTCATATCCAGCATTAATCTTCAGGAGTATATATACTAATTCGGTATCTCACAGCTGGAACGTCATTGCTGAGA
CCAAGGGAGGAGACCACAACAATGTTGTCATGCTTGGTGTCACTCCGACTCCGTCGATGCCGGCCCCGGTATCAACGACAACGG
CTCCGGCTCCATTGGTATCATGACCGTTGCCAAAGCCCTTACTAACTTCAAGCTCAACAACGCGTTCGCTTTGCCTGGTGGACC
GCTGAGGAATTCGGTCTCCTTGAAGCACCTTCTACGTCGACAGCCTTGATGACCGTGAGCTGCACAAGGTCAGCTGTACCTCA
ACTTCGACATGATCGGCTCTCCCAACTTCGCCAACCAGATCTACGACGGTGACGGTTCCGGCTACAACATGACTGGTCCCCTGG
CTCTGCTGAAATCGAGTACCTGTTCCGAGAAGTCTTTGACGACCAGGGTCTCCACACCAGCCACTGCCTTACCGGCCGATCC
GACTACTCTGCATTCATCAAGCGCAACGTCGCCGCTGGAGGCTTTTCACTGGTGCCGAGGTTGTCAAGACCCCGAGCAAGTTA
AGCTGTTCCGGTGGTGGGCTGGCGTTGCCTATGACAAGAACTACCATGGCAAGGGTGACACCGTTGCCAACATCAACAAGGGAGC
TATCTTCCCTAACACTCGAGCAATCGCCTACTCTGTGGCCGAGTATGCTCGATCCCTCAAGGGCTTCCCAACCCGCCAAAGACC
GGCAAGCGTGCCGTCACCCCTCAGTATGCTAAGATGCCTGGTGGTGGTTGCGGACACCACACTGTCTTCATGTAA

```

[0213] 1488 个核苷酸的 meLAP2 开放阅读框 (“ORF”), 其中起始子 ATG 在第一个位置上 (表 11B 中下划线)。

[0214]

表 11B、meLAP2 核苷酸序列 (SEQIDNO :32)

```

ATGAAGTCGCAACTGTTGAGCCTAGCCGTGCGTCACCACCATTTCCCAGGGCGTTGTTGGTCAAGAGCCCTTTGGATGGCCCT
TCAAGCCTATGGTCACTCAGGATGACCTGCAAAACAAGATTAAGCTCAAGGATATCATGGCAGGTGTCGAGAAGCTGCAAAGCTT
TTCTGATGCTCATCCTGAAAAGAACCGAGTGTTCGGTGGTAATGGCCACAAGGACACTGTCGAGTGGATCTACAATGAGCTCAAG
GCCACCGGCTACTACAATGTGAAGAAGCAGGAGCAGGTACACCTGTGGTCTCACGCTGAGGCCGCTCTCAGTGCCAATGGCAAGG
ACCTCAAGGCCAGCGCCATGTCGTACAGCCCTCTGCCAACAGATCATGGCCGAGCTTGTGCGTTGCCAAGAACAATGGCTGCAA
TGCTACCGATTACCCAGAGAACACTCAGGAAAGATAGTCCTCATTCAGCGTGGTGTCTGCAGCTTCGGCGAGAAGTCTTCTCAG
GCTGGTGTGCGAAGGCTATTGGTGCCGTTGTCTACAACAACGTCCCCGGATCCCTTGCTGGCACTCTTGGTGGCCTTGACAAGC
GCCATGTCCCAACCGCTGGTCTTTCCCAGGAGGATGGAAAGAATCTTGCTAGCCTCGTTGCTTCTGGCAAGGTTGATGTCACCAT
GAACGTTGTCAGTCTGTTTGAGAACCGAACCACTGGAACGTCAATTGCTGAGACCAAGGGAGGAGACCACAACAATGTTGTGATG
CTTGGTGTCACTCCGACTCCGTCGATGCCGGCCCCGGTATCAACGACAACGGCTCCGGCTCCATTGGTATCATGACCGTTGCCA
AAGCCCTTACTAACTTCAAGCTCAACAACGCCGTTTCGCTTTGCCTGGTGGACCGCTGAGGAATTCGGTCTCCTTGAAGCACCTT
CTACGTCGACAGCCTTGATGACCGTGAGCTGCACAAGGTCAAGCTGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCC
AACCAGATCTACGACGGTGACGGTTCGGCTACAACATGACTGGTCCCGCTGGCTCTGCTGAAATCGAGTACCTGTTTCGAGAAGT
TCTTTGACGACCAGGTCTCCACACCAGCCACTGCCTTACCGGCCGATCCGACTACTCTGCATTATCAAGCGCAACGTCCC
CGCTGGAGGTCTTTTACTGGTGCCGAGGTTGTCAAGACCCCCGAGCAAGTTAAGCTGTTCCGGTGGTGGGCTGGCCTTGCCTAT
GACAAGAACTACCATGGCAAGGGTGACACCGTTGCCAACATCAACAAGGGAGCTATCTTCCCTTAACACTCGAGCAATCGCCTACT
CTGTGGCCGAGTATGCTCGATCCCTCAAGGGCTTCCCAACCCGCCAAAGACCGGCAAGCGTGCCGTCAACCCTCAGTATGCTAA
GATGCCTGGTGGTGGTTGCGGACACCACACTGTCTTCATGTAA

```

[0215] 上述 meLAP2 核苷酸序列 (SEQ ID NO :32) 编码有 495 个氨基酸残基 (SEQ ID NO : 33) 的蛋白, 在表 11C 中以一个字母的氨基酸代码表示。

[0216]

表 11C、fuLAP2 编码的蛋白序列 (SEQIDNO :33)

```

MKSQLLSLAVAVTTISQGVVQEPFGWPFKPMVTQDDLQNKIKLKDIMAGVEKLQSFSDAHPEKNRVFGNGHKDIVEWIIYNE
LKATGYYNVKKQEQVHLWSHAEEAALSANGKDLKASAMSYSPANKIMAELVVAKNNGCNATDYPENTQGKIVLIQRGVCSFGE
KSSQAGDAKAI GAVVYNNVPGSLAGTLGGLDKRHVPTAGLSQEDGKNLASLVASGKVDVTMNVVSLFENRTTWNVIAETKGGD
HNNVVMLGAHSDSVDAGPGINDNGSGSIGIMTVAKALTNFKLNNAVRFWWWTAEEFGLLGSTFYVDSLDDRELHKVKLYLNF
MIGSPNFANQIYDGDGSAYNMTGPAGSAEIEYLFKFFDDQGLPHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTPPEQVK
LFGGEAGVAYDKNYHGKGDIVANINKGAI FLNTRAIAYSVAEYARSLKGFPTRPKTGKRAVNPQYAKMPGGGCGHHTVFM

```

[0217] 上述 meLAP2 具有与表 11D、11E 和 11F 中的 BLAST 氨基酸序列同源性。用 parawise blast 程序分析上述数据。

[0218] 表 11D. meLAP2 TBLASTN 结果

[0219]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   600025	Saccharomyces cerevisiae(s288c)RIF1, DPB3, YmL27 和 SNF5 基因	32421	180/479 (37%)	251/479 (52%)	2e-70
gi   469463	Saccharomyces cerevisiae 氨肽酶 Y 基因	2272	180/479 (37%)	251/479 (52%)	2e-70

[0220] 表 11E. meLAP2BLASTX 结果

[0221]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   1077010	氨肽酶 Y 前体, 有液泡的 /Saccharomyces cerevisiae	537	180/479 (37%)	251/479 (52%)	8e-71
gi   6319763	氨肽酶 yscIII ;Ape3p/ Saccharomycescerevisiae	563	180/479 (37%)	251/479 (52%)	8e-71
gi   15839805	水解酶 / Mycobacterium Tuberculosis	493	159/440 (36%)	236/440 (53%)	1e-63
gi   31791596	可能的脂蛋白氨肽酶 LPQL/ Mycobacteriumbovis	500	159/440 (36%)	236/440 (53%)	1e-63
gi   15598135	可能的 氨肽酶 /	536	158/445 (35%)	237/445 (53%)	1e-62
gi   15598135	可能的 氨肽酶 / Pseudomonas Aeruginosa	536	158/445 (35%)	237/445 (53%)	1e-62
gi   1045225	N- 乙酰化嘌呤霉素 N- 乙酰化水解酶 / Streptomycesanulatus	485	154/477 (32%)	218/477 (45%)	4e-48



基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   29831415	可能的氨肽酶 / Streptomyces avermitilis	315	95/244 (38%)	131/244 (53%)	2e-37

[0222] 表 11F. meLAP2BLASTP 结果

[0223]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   6319763	氨肽酶 yscIII ;Ape3p/ Saccharomyces cerevisiae	563	179/479 (37%)	248/479 (51%)	9e-71
gi   1077010	氨肽酶 Y 前体, 有液泡的 /Saccharomyces cerevisiae	537	179/479 (37%)	248/479 (51%)	9e-71
gi   31617182	可能的脂蛋白氨肽酶 LPQL/Mycobacterium bovis	500	159/440 (36%)	236/440 (53%)	2e-62
gi   15839805	水解酶 / Mycobacterium Tuberculosis	493	159/440 (36%)	236/440 (53%)	2e-62

[0224] ruDPPIV

[0225] ruDPPIV 是红色毛癣菌的二肽基多肽酶 IV。2326 个核苷酸序列 (SEQ ID NO :34) 的 ruDPPIV 核酸如表 12A 所示。所述的 ruDPPIV 开放阅读框 (“ORF”) 起始于 ATG 在第一个位置上 (表 12A 中下划线)。

[0226]

表 12A、ruDPPIV 核苷酸序列 (SEQIDNO :34)

ATGAAGCTCCTCTCGCTACTTATGCTGGCGGGCATCGCCCAAGCCATCGTTCCTCCTCGTGAGCCCCGTTACCAACTGGTG  
 GCGGCAACAAGCTGTTGACCTACAAGGAGTGTGTCCCTAGAGCTACTATCTCTCCAAGGTCGACGTCCCTTGCCCTGGATTAA  
 CAGTGAAGAAGATGGCCGGTACATCTCCCAGTCCGACGATGGAGCATTGATCCTCCAGAACATCGTCACGAACACCAACAAG  
 ACTCTCGTGGCCGACAGACAAGGTACCCAAGGGTACTATGACTACTGGTTCAAGCCAGACCTTTCTGCTGTCTTATGGGCAA  
 CCAATTACACCAAGCAGTACCGTCACTCTTACTTTGCCAACTACTTCATTCTAGACATCAAAAAGGGATCGTTGACCCCTCT  
 AGCCCAGGACCAGGCTGGTGACATCCAGTATGCTCAATGGAGCCCCATGAACAACCTCTATCGCCTATGTCCGTGGAAACGAC  
 CTGTATATCTGGAACAATGGCAAGACCAAGCGTATTACCGAAAAATGGCGGCCCGGATATCTTCAATGGTGTCCCTGACTGGG  
 TATACGAGGAAGAAATCTTCGGGGACCGGTTTCGCTCTTTGGTTCTCACCTGACGGTGAATACCTTGCGTACCTCCGCTTTAA  
 CGAGACTGGAGTCCCGACCTACACTATTCGGTACTACAAGAAACAAGCAAAAAGATTGCCCTGCCTACCCAAGGGAGCTGGAG  
 ATCCGTTACCCTAAAGTCTCTGCGAAGAACCAACCGTGCAGTTCCACCTGTAAACATTGCTTCATCCCAGGAGACAACATA  
 TCCCAGTTACTGCGTTCGCCGAAAACGATCTTGTGATCGGTGAGGTTGCTTGGCTCAGCAGTGGCCATGATAGTGTAGCATA  
 TCGTGCCTTCAACCGTGTCCAGGATAGAGAAAAGATTGTCAGCGTCAAGGTTGAGTCCAAGGAATCCAAGGTTATTCGCGAA  
 AGAGATGGCACCGACGGCTGGATCGACAACCTTCTCTCCATGTATATATCGGAAACGTTAACGGCAAGGAGTACTACGTCG  
 ATATATCTGATGCTTCTGGCTGGGCACATATCTACCTTACCCGGTTGATGGAGGAAAGGAGATTGACTAACAAAGGGAGA  
 ATGGGAAGTCGTTGCCATTCTCAAGGTTGACACGAAGAAGAAGCTGATCTACTTCACCTTACCAAAATATCACAGCACCCT  
 CGACACGTCTACTCTGTCTCGTATGACACAAAAGGTCATGACCCCTCTCGTCAACGATAAGGAGGCTGCGTACTACTGCAT  
 CCTTCTCGGCCAAGGGTGGTTACTATATCTTGTCTACCAAGGTCCAAATGTTCCATACCAAGAACTTACTCCACCAAGGA  
 CAGTAAGAAGCCTCTCAAGACAATCACTAGCAATGATGCAATGCTCGAGAAGCTGAAGGAGTACAAGCTCCCAAGGTTAGC  
 TTCTTTGAGATCAAGCTTCCATCTGGTGAACCCCTTAATGTTAAGCAACGCCCTACCACCTAACTTCAACCCACACAAGAAGT  
 ACCCCGTCCTCTTACTCCGATGGTGGCCCTGGTGCCAAGAGGTAAGCCAGGCATGGAATTCATTGGACTTCAAGTCCTA  
 CATTACATCTGACCCTGAGCTTGAATACGTTACCTGGACTGTTGACAACCGTGAACCCGGCTACAAGGGCCGCAAGTCCGC  
 AGCGCCGTAGCTAAGCGTCTCGGTTTCCCTCGAAGCCCAGGACCAGGTCTTTGCTGCTAAGGAGGTGCTGAAAAACCGTTGGG  
 CTGATAAGGACCATATGGAACTCTGGGGCTGGAGCTATGGCGGCTTCCTGACCGCTAAGACCCCTCGAGACCGACAGTGGTGT  
 ATTCACTTTTGGTATCAGTACTGCTCCTGTCTCTGATTTACAGACTCTACGACAGCATGTACTGAGCGTTACATGAAGACC  
 GTTGAACATAACCGCTGACGGCTACAGTGAGACCGCCGTGCACAAGGTTGATGGCTTTAAGAACCTCAAAGGTCATTACTTCA  
 TCCAGCATGGAACCGGTGACGACAACGTCCACTTCCAAAACGCCGCTGTCTTTCCAACACCCTGATGAACGGCGGTGTAAC  
 TGCAGACAAGTTGACTACTCAGTGGTTTACTGACTCGGACCAGGCATCAGATACGATATGGACTCCACTTACCAGTACAAG  
 CAGCTTCTAAGATGGTCTACGACCAGAAGCAACGAAGGCCAGAAAGCCACCAATGCACCAATGGAGCAAGAGAGTTTGG  
 CTGCCCTGTTGGTGAGAGGGCAGAGGAATGA

[0227] 上述 ruDPPIV 核苷酸序列 (SEQ ID NO :34) 编码含 775 个氨基酸残基 (SEQ ID NO : 35) 的蛋白, 在表 12B 中以一个字母的氨基酸代码表示。

[0228]

表 12B、ruDPPIV 编码的蛋白序列 (SEQIDNO :35)

```

MKLLSLLMLAGIAQAIVPPREPRSPTGGGNKLLTYKECVPRATISPRSTSLAWINSEEDGRYISQSDDGALILQNIVTNTNKT
LVAADKVPKGYDYWFKPDL SAVLWATNYTKQYRHSYFANYFILDIKKGSLTPLAQDQAGDIQYAQWSPMNSIAYVRXNDLY
IWNNGKTKRITENGGPDIFNGVDPWVYEEEIFGDRFALWFSPDGEYLA YLRFNETGVPTYTIPYYKKNQKIAPAYPRELEIRY
PKVSAKNPTVQFHLLNIASSQETTIPVTAFFPENDLVIGEVAWLSSGHDSVAYRAFNRVQDREKIVSVKVESKESKIVIRERDGT
DGWIDNLLSMSYIGNVNGKEYVVDISDASGWAHIYLYPVDGGKEIALTKGEVVA I LKVDTKKKLIYFTSTKYHSTTRHVYS
VSYDTKVMTPLVNDKEAAYTASFSAKGGYYILSYQGNVPYQELYSTKDSKPLKTTITNDALLEKLKEYKLPKVSFFEIKL
PSGETLNVKQRLPPNPNPHKKYPVLFPTPYGGPGAQEVSAWNSLDFKSYITSDPELEYVTWTVDNRGTGYKGRKFRSAVAKRL
GFLEAQDQVFAAKEVLKNRWADKDHIGIWGXSYGGFLTAKTLETDSGVFTFGISTAPVSDFRLYDSMYTERYMKTVELNADGY
SETAVHKVDGFKNLKGHYLIQHGTGDDNVHFQNAAVLSNTLMNGGVTADKLTQWFTSDHGI RYDMDSTYQYKQLSKMVYDQ
KQRRPESPPMHQWSKRVLAAALFGERAEE

```

[0229] 上述 ruDPPIV 具有与表 10C、10D 和 10E 中的 BLAST 氨基酸序列同源性。用 parawiseblast 程序分析上述数据。

[0230] 表 12C. ruDPPIV TBLASTN 结果

[0231]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi 2351699	Aspergillus fumigatus 二肽基多肽酶 IV (Dpp4) 基因	2352	469/761 (61%)	585/761 (76%)	0.0
gi 2924304	Aspergillus oryzaeDpp4 基因	4771	448/769 (58%)	568/769 (73%)	0.0
gi 32422540	Neurosporacrassa 菌株 OR74A	2688	256/720 (35%)	374/420 (51%)	e-114
gi 14330262	二肽基氨肽酶 IV 型的 Aspergillus nigerdapB 基因, 外显子 1-3	3989	224/637 (35%)	333/637 (52%)	e-111
gi 1621278	二肽基多肽酶的 Xenopus laevismRNA	3337	244/752 (32%)	375/752 (49%)	e-100
gi 6978772	Rattusnorvegicus 二肽基多肽 酶 4 (Dpp4)	4835	246/742 (33%)	373/742 (50%)	8e-98

[0232] 表 12D. ruDPPIV BLASTX 结果

[0233]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   2351700	二肽基多肽酶 IV/Aspergillus fumigatus	765	218/341 (63%)	270/341 (79%)	0.0
gi   2924305	脯氨酰二肽基多肽酶 /Aspergillus oryzae	771	213/344 (61%)	270/340 (78%)	0.0
gi   1621279	二肽基多肽酶 IV/XenopuslaevismRNA	748	118/349 (33%)	186/349 (53%)	8e-93
gi   535588	二肽基多肽酶 IV/ 人	766	125/375 (33%)	191/375 (50%)	3e-90

[0234] 表 12E. ruDPPIV BLASTP 结果

[0235]

基因 目录 / 索引	蛋白 / 组织	长度 (aa)	一致性 (%)	得分值 (%)	E 值
gi   2351700	二肽基多肽酶 IV/Aspergillus fumigatus	765	468/761 (61%)	585/761 (76%)	0.0
gi   2924305	脯氨酰二肽基多肽酶 /Aspergillus oryzae	771	448/769 (58%)	568/769 (73%)	0.0
gi   14330263	二肽基氨肽酶 IV 型 /Aspergillus  niger	901	261/733 (35%)	387/733 (52%)	e-114
gi   19114882	二肽基氨肽酶 / Schizoaccharomycesprombe	793	258/742  (34%)	396/742  (53%)	e-106
gi   3660	二肽基氨肽酶 /Saccharomycescerevisiae	841	254/750 (33%)	370/750 (49%)	2e-95

[0236] 本发明另一方面是关于分离的编码 EXOX 多肽或具有其生物活性部分的核酸分子。同时本发明也包括能充分应用的作为杂交探针证明 EXOX 编码核酸（如 EXOX mRNAs）的核酸片段，作为扩增或突变 EXOX 核酸分子的 PCR 引物的核酸片段。所述应用，所述术语“核酸分子”指的是包括 DNA 分子（如 cDNA 或基因组 DNA），RNA 分子（如 mRNA），利用核酸同源性产生的上述 DNA 或 RNA 的类似物、衍生物、片段或功能类似物等。所述核酸分子可以是单链或双链。

[0237] 一个 EXOX 核酸能编码一个成熟的 EXOX 多肽。本发明所指，“成熟”形式的多肽或蛋白表示自然产生的多肽、前体形式或蛋白前体。所述自然产生的多肽、前体形式或蛋白前体包括但不限于，全长的编码相关蛋白的基因产物。同样的，上述自然产生的多肽、前体形式或蛋白前体也可以被定义为由一段本发明所述的 ORF 编码。所述的产物的“成熟”形式产生于，包括但不限于，一个或多个自然发生的基因表达过程，如发生在细胞、宿主细胞里出现基因表达。上述产生步骤导致“成熟”形式的多肽或蛋白的例子包括去除由 ORF 开始密码子编码的 N 末端甲硫氨酸残基，或水解去除一段信号肽或引导序列。上述从具有残基 1 到 N 的多肽或蛋白前体产生的成熟形式，所述残基 1 是指 N 末端甲硫氨酸，去除 N 末端甲硫氨酸后具有残基 2 到 N 的残基。另外的，上述从具有残基 1 到 N 的多肽或蛋白前体产生的成熟形式，其中 N 末端从残基 1 到残基 M 的信号序列被除去，保留残基 M+1 到残基 N 的残基。这里进一步引述，多肽或蛋白的“成熟”形式可以从后转录修饰中产生，而不用蛋白水解去除现象。上述另外的过程包括，但不限于如，糖基化（N-、O- 和 W 型）、十四酰化、磷酸化、硫化、N 末端环化或 C 末端酰胺化。一般的，成熟的多肽或蛋白可以通过上述一种方法产生，或上述方法的任意组合产生。

[0238] 所述的术语“探针”，在本发明中指的是具有不同长度的核酸序列，可能至少是 10 个核苷酸（nt），或到 100 个，又或到最大的 6000 个 nt，这取决于其相关作用。探针被用于证明核酸序列的一致性、相关性或互补性。较长的探针一般来自于天然或重组的，具有高度特异性，同时比起短的探针具有更长的杂交时间。探针可以是单链或双链，可以设计用于 PCR 反应，膜杂交技术，或象 ELISA 技术 [xu4]。

[0239] 所述术语“分离的”核酸分子，在本发明中指的是从自然存在的核酸中分离得到的一种核酸分子。优选的，“分离的”核酸不含有该核酸来源生物的基因组 DNA 中天然位于该核酸侧翼的序列（如位于该核酸的 5' 和 3' 端序列）。例如，在不同的实施方案中，分离的 EXOX 核酸分子可以含有至少约 5kb、4kb、3kb、2kb、1kb、0.5kb 或 0.1kb 的核苷酸序列，该核苷酸序列天然位于上述核酸分子来源的细胞 / 组织 / 物种基因组 DNA 中的核酸分子的侧翼。此外，“分离的”核酸分子，例如 cDNA 分子，在用重组技术产生时，可以基本上不含其他细胞物质或培养基；或在用化学法合成时，基本上不含化学前体或其他化学物质。特别的，其中核酸分子或蛋白的纯度至少约 50%，进一步优选为至少约 85%，最优选为至少约 99%。

[0240] 本发明所述的术语“重组体”指的是参考一种细胞，显示该细胞复制异源核酸或表达由异源核酸编码的多肽或蛋白。重组体细胞可以含有不是天然的（非重组体）细胞的基因。重组体细胞也可以含有天然的细胞里经过修饰过的和通过人工手段引入到细胞里的基因。所述术语还包括含有内生的被修饰过，但没有去除核酸的细胞里的核酸；所述修饰包括可以获得的基因替换、定点突变和其他相关技术。本领域技术人员熟知的是所述细胞可以被用作单细胞或多细胞转基因组织，例如产生 EXOX 的转基因真菌。

[0241] 本发明提供的一种核酸分子,如含有 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示核酸序列的核酸分子或其互补系列,可以用标准的分子生物学技术分离得到,所述序列信息如本发明所示。用全部或部分 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示核酸序列作为杂交探针, EXOX 分子用标准的杂交和克隆技术被分离得到(如以下文献描述:Sambrook 等,(eds),分子克隆试验操作手册第二版,冷泉港实验室出版,冷泉港, NY, 1989; Ausubel 等,(eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993)。

[0242] 本发明所述的核酸可以通过以 cDNA、mRNA 或等位基因组 DNA 作为模板,以合适的寡核苷酸片段为引物,进行一定的 PCR 反应得到。所述的核酸扩增后可以克隆到合适的载体中,并通过序列分析得到确认。进一步的, EXOX 核苷酸序列对应的寡核苷酸可以通过标准的合成技术制备,如 DNA 自动合成仪。

[0243] 本发明所述的术语“寡核苷酸”指的是一系列连接在一起的核苷酸残基,所述寡核苷酸具有足够多的碱基来保证应用于 PCR 反应。一段短的寡核苷酸序列可以基于或来自于一个基因组 DNA 或 cDNA,然后被应用于扩增、确认或揭示在特定细胞或组织中存在的同一种的、类似的或互补的 DNA 或 RNA。寡核苷酸含有部分核酸序列,长度可为 10nt、50nt 或 100nt,优选为 15nt ~ 30nt。在本发明一个具体实施方案中,一段长度小于 100nt 的寡核苷酸进一步的包括了 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 或其互补序列中至少 6 个邻近核苷酸。寡核苷酸可以通过化学合成,作为探针应用。

[0244] 在另一个实施方案中,本发明所述的分离的核酸分子包括了一种含有与选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列互补的核酸分子(例如,一段作为探针或引物的核酸片段、一段编码具有 EXOX 多肽活性部分的核酸片段)。上述与选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列互补的核酸分子具有与选自 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列很少或没有错配的氢键连接,从而形成稳定的双螺旋结构。

[0245] 在这里,术语“互补体”指的是核酸分子的核苷酸间为 Watson-Crick 或 Hoogsteen 碱基配对。术语“连接”指的是两个多肽、化合物、多肽相关物质、组合物或其功能类似物之间的物理或化学相互作用力。连接包括离子键、非离子键、范德华引力、氢键引力等。物理作用力可以是直接的,也可以是间接的。间接的作用力可以通过或由另一个多肽或化合物的影响。直接连接指的是没有空间的阻隔,或者是由于另一个多肽或化合物作用,但不是指其他具体的化学反应作用。

[0246] 本发明提供的片段定义为具有至少 6 个(相邻的)核酸或 4 个(相邻的)氨基酸,其长度足够允许严格的核酸杂交或与特殊的抗体结合,相对的,部分片段的反应小于全长反应。片段可以从所选择的核酸或氨基酸序列中相邻的部分得到。由核酸序列或氨基酸序列形成的衍生物可以是直接从自然获得的化合物的或被修饰过的或部分替换的。同功物指的是具有类似核酸或氨基酸序列结构,但不限于的是,具有不同的组分或侧链。同源或异源是指核酸序列或氨基酸序列从不同物种分离得到的一类特定基因。

[0247] 如果衍生物或同功物含有被修饰过的本发明所述核酸或蛋白,所述衍生物和同功物可以是全长或超过全长。本发明所述的核酸或蛋白的衍生物或同功物包括但不限于:在不同的重组体中含有与本发明所述的核酸或蛋白具有至少约 70%、80% 或 95% 同源性的

的分子（优选为具有 80-95%同源性的分子）；或是按照本领域技术人员熟知的经过计算机同源程序配对过的线性序列；或是能够在严格的、温和的、低度严格的条件下与编码上述的蛋白的互补序列杂交的编码的核酸。参见如 Ausubel 等，CURRENT PROTOCOLS IN MOLECULARBIOLOGY, John Wiley & Sons, New York, NY, 1993 和以下所述。

[0248] 本发明中，一种“同源核酸序列”或“同源氨基酸序列”或是其它不同名称指的是在核苷酸或氨基酸水平上具有同源性特征的序列。同源核酸序列包括那些编码 EXOX 多肽的重组异构体的序列。重组异构体能够在同样的组织中表达，如选择性剪切的 RNA。另一方面，重组异构体可以被不同的基因编码。在本发明中，同源核酸序列可以是编码除真菌外的其他物种的 EXOX 多肽的核酸序列。同源核酸序列也包括但不限于自然发生的本发明提供的核酸等位替换或突变。同源核酸序列还包括那些编码如 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的保守氨基酸亚单位（参见下面），一种多肽具有 EXOX 生物活性。EXOX 蛋白的不同生物活性将在下面详述。

[0249] EXOX 多肽由 EXOX 核酸的开放阅读框 (ORF) 编码。ORF 被终止子结束其延伸。一段开放阅读框代表了一种全长蛋白的编码序列，由 ATG “起始子”开始，到 TAA、TAG 或 TGA 三个终止子之一终止。为了本发明目的，所述 ORF 可以是编码序列的任何部分，含有或不含有起始子和终止子之一或两个。作为一个好的编码真正的细胞质蛋白的 ORF，其最小的长度经常是编码 50 个氨基酸或更多氨基酸蛋白的 DNA 结构。

[0250] 经来自真菌 EXOX 基因克隆确认的核苷酸序列可以作为探针和引物应用，所述探针和引物设计用来证明和 / 或克隆在其他物种的 EXOX 同源基因，也包括来自其他物种的 EXOX 同源基因。所述引物 / 探针一般包括一段基本上纯化的寡核苷酸。所述寡核苷酸一般包括核苷酸序列的一个区，所述核苷酸序列在严格条件下与至少约 12、25、50、100、150、200、250、300、350 或 400 连续的核苷酸序列的核酸杂交，其中杂交的核酸序列为 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 的正义核苷酸序列；或 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 的反义核苷酸序列；或 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 的天然存在的突变体序列。

[0251] “具有 EXOX 多肽部分活性的多肽”指的是多肽具有活性类似，但不限于，本发明所述多肽活性，包括成熟形式，经过特殊的生物检测，有或没有剂量依赖。一种编码“EXOX 生物活性部分”的核酸片段可以从 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示的编码具有 EXOX 生物活性的多肽 (EXOX 蛋白生物活性在下面详述) 的部分序列分离得到，表达 EXOX 蛋白被编码部分（如体内表达的重组子），评价 EXOX 蛋白被编码部分活性。

[0252] EXOX 核酸和多肽的变异体

[0253] 本发明进一步包括不同于 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的核酸分子，其导致基因编码的兼并，这样同样的编码如 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列编码的 EXOX 蛋白。在另一实例中，一种本发明所述分离的核酸分子具有编码含有 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列多肽的核酸序列。除了 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示的真菌 EXOX 核酸序列外，核酸序列 EXOX 多肽的氨基酸序列存在于不同的种群中。上述 EXOX 基因的多样性可以存在于不同的由自然等位变异导致的真菌种类中。本发明中，术语“基因”和“重组基因”指的是含有编码 EXOX 蛋白开放阅读框 (ORF) 的核酸分子，优选的是一种真

菌 EXOX 蛋白。上述自然等位变异可以是典型的在 EXOX 基因的核酸序列中发生 1 ~ 5% 的变异。所述任一和所有 EXOX 多肽的核酸变异和引起的氨基酸多样性都属于本发明范围,所述变异和多样性是自然等位变异的结果,不会改变 EXOX 多肽的功能、活性。

[0254] 更进一步的是,其他种类中编码 EXOX 多肽的核酸分子,但具有与来自于真菌的 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列不同的核酸序列,也属于本发明范围。与本发明 EXOX cDNAs 同源和相对应的自然等位变异的核酸分子可以通过其与本发明所述真菌 EXOX 核酸序列同源,用真菌 cDNAs 或类似部分作为杂交探针通过标准的杂交技术在严格的杂交条件下分离得到。

[0255] 同样的,在另一个实施方案中,一种本发明所述的分离核酸分子具有至少 6 个核苷酸,能够在严格条件与含有 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的核酸分子杂交。

[0256] 在另一个实施方案中,上述核酸至少具有 10、25、50、100、250、500、750、1000、1500 或 2000 核苷酸。在另一个实施方案中,本发明所述的分离核酸分子与编码区杂交。在本发明中,术语“在严格条件下杂交”描述的是在核酸序列彼此间至少有 60% 同源情况下进行杂交和清洗,核酸之间仍保持杂交状态。

[0257] 同源或其他相关序列(如直系同源、侧系同源)可以通过用所有或部分所述真菌的特殊序列作为探针,用本领域常用的核酸杂交和克隆方法,在低的、温和或高的严格杂交条件下获得。

[0258] 本发明所述术语“严格杂交条件”指的是探针、引物或寡核苷酸只与目标序列杂交,而不是其他序列。严格条件是序列依赖型,根据不同反应而不同。特别的,长序列的杂交温度比短序列杂交温度高。一般的,对于特别序列,其严格条件选择比热熔点 ( $T_m$ ) 低 5°C,在一定的离子强度和 pH 下。所述  $T_m$  是指(在一定的离子强度、pH 和核酸浓度下)探针 50% 的与目标序列互补的序列和目标序列达到平衡的杂交状态时的温度。既然目标序列一般存在过量的,在  $T_m$  时,探针的 50% 序列在平衡状态时都被反应了。一般的,严格条件中的盐浓度小于约 1.0M 钠离子,一般为约 0.01 到 1.0M 钠离子(或其他盐),pH7.0 ~ 8.3,对于短的探针、引物或寡核苷酸(如 10nt 到 50nt)温度至少约 30°C,对于长的探针、引物或寡核苷酸温度至少约 60°C。严格条件中还可以添加破坏 DNA 稳定的试剂,例如甲酰胺。

[0259] 严格条件是本领域的技术人员熟知的,而且能在 Ausubel 等, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 查到。优选的,在所述条件下,核酸序列相互之间至少约 65%、70%、75%、85%、90%、98% 或 99% 同源性的序列仍能相互杂交。非限定的严格杂交条件例子是在高的盐缓冲液,包括: 6×SSC, 50mM Tris-HCl (pH7.5), 1mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, 和 500mg/ml 的变性鲑鱼精 DNA, 在 65°C 下杂交,随后在 50°C 下用 0.2×SSC 洗一次或多次。本发明所述的一种分离的核酸分子可以在严格条件下与 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列对应的自然-存在的核酸分子杂交。在本发明中,“自然-存在”核酸分子指的是自然存在的一种 RNA 或 DNA 分子(如编码一种天然的蛋白)。

[0260] 在第二个实施方案中,本发明还提高一种在中度严格条件下与含有 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的核酸分子或其片段、同功物或衍生物杂交的核酸分子。非限定的中度严格杂交条件例子是在 6×SSC, 5×denhardt' 溶液, 0.5%



SDS 和 100mg/ml 的变性鲑鱼精 DNA, 在 55°C 下杂交, 随后在 37°C 下用 1×SSC、0.1% SDS 洗一次或多次。其他的中度严格杂交条件是本领域技术人员熟知的, 如 Ausubel 等, 1993, *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y.; Kriegler, 1990, *Gene transfer and Expression, A Laboratory Manual*, Stockton 出版, NY。

[0261] 在第三个实施方案中, 本发明还提高一种在低度严格条件下与含有 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的核酸分子或其片段、同功物或衍生物杂交的核酸分子。非限定的中度严格杂交条件例子是在 35% 甲酰胺、5×SSC、50mM Tris-HCl (pH7.5)、5mM EDTA、0.02% PVP、0.02% Ficoll、0.02% BSA、100mg/ml 的变性鲑鱼精 DNA 和 10% (w/v) 葡聚糖硫酸盐, 在 40°C 下杂交, 随后在 50°C 下用 2×SSC、25mM Tris-HCl (pH7.4)、5mM EDTA 和 0.1% SDS 洗一次或多次。其他的低度严格杂交条件是本领域技术人员熟知的 (如种间杂交应用), 参见如 Ausubel 等, 1993, *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y.; Kriegler, 1990, *Gene transfer and Expression, A Laboratory Manual*, Stockton 出版, NY; Shilo & Weinberg, *Proc Natl Acad Sci USA* 78 :6789-6792 (1981)。

#### [0262] 保守的突变

[0263] 除了自然存在的 EXOX 序列的等位突变可以存在上述生物体中, 本领域技术人员可以得出如下的改变, 所述改变由 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列的核酸突变而来, 从而导致编码的 EXOX 蛋白的氨基酸序列发生变化, 但上述变化不会改变所述 EXOX 蛋白的功能活性。例如, 导致在“非必需”氨基酸残基上替换氨基酸的核苷酸替换可以发生在 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示序列中。所述“非必需”氨基酸残基是指一种在 EXOX 蛋白序列中野生型序列中的不改变 EXOX 蛋白生物活性的氨基酸残基改变, 而“必需”的氨基酸是指生物活性需要的基团。

[0264] 本发明中, 术语“生物活性”或“功能活性”指的是自然或一般的 EXO 蛋白具有的功能, 例如, 降解其他蛋白的能力。本发明所述的 EXOX 蛋白中的氨基酸残基可以通过特殊的 non-amenable 改变来预测。保守性的氨基酸替换是本领域技术人员熟知的。其中一项本领域技术人员熟知的技术是通过标准的技术, 每一个核酸编码子 (除了 AUG, 一般只对应甲硫氨酸) 可以被修饰成功能一致的分子。进一步的, 在一段编码序列中改变、增加或删除一个氨基酸或一小部分氨基酸 (通常小于 5%, 更特别的是小于 1%) 的单独的替换、删除或增加的改变是“保守突变”, 所述改变是化学性类似的氨基酸替换的结果。

[0265] 本发明另一方面还涉及编码含有对于活性不是必需的氨基酸残基改变的 EXOX 蛋白的核酸分子。所述 EXOX 蛋白的氨基酸序列不同于 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列, 但其仍有生物活性。在一个实施方案中, 所述分离的核酸分子包括一种编码蛋白的核苷酸序列, 所述蛋白的氨基酸序列至少与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列有 45% 的同源性。优选的, 所述核酸分子编码的蛋白的氨基酸序列至少与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列有 60% 的同源性; 更优选的是至少与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列有 70% 的同源性; 更进一步优选的是至少与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列有 80% 的同源性; 再优选的是至少与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列有 90% 的同源性; 最 [xu5] 优选的是至少与 SEQ ID NOs :3、6、

9、12、15、18、21、24、27、30、33 和 35 所示序列有 95% 的同源性。

[0266] 编码与含有 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列的蛋白同源的 EXOX 蛋白的分离的核酸分子可以在 SEQ ID NOs :2、5、8、11、14、17、20、23、26、29、32 或 34 所示核苷酸序列中通过一个或多个核苷酸替换、增加或删除得到, 这样一个或多个的氨基酸替换、增加或删除被引入到所述的被编码的蛋白。

[0267] 通过标准的技术可以在 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列引入突变, 如点突变、PCR 介导的突变和 DNA 移位。优选的, 保守的氨基酸替换发生在一个或多个可预测的、非必需的氨基酸残基上。在人类 DNA 中单个碱基的替换是参见的。如果替换是发生在编码区, 那么替换可能是保守或非保守替换。所述“保守的氨基酸替换”是具有类似特性的、与被替换的氨基酸具有类似的侧链结构新的氨基酸。非必需替换指的是具有不同特性的新的氨基酸。现有技术对具有类似侧链的氨基酸家族已有定义。氨基酸家族包括碱性侧链的(如赖氨酸、精氨酸、组氨酸)、酸性侧链的(如天冬氨酸、谷氨酸)、不带电的极性侧链的(如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、非极性侧链的(如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、蛋氨酸、色氨酸)、 $\beta$ -分支侧链的(如苏氨酸、缬氨酸、异亮氨酸)和芳香族侧链的(如酪氨酸、苯丙氨酸、色氨酸、组氨酸)。因此, 对于保守替换, 在 EXOX 蛋白中预测为非必需的氨基酸残基可以被同一个侧链家族的另一个氨基酸替换。可选择的是, 在另一个实施方案中, 突变也可以被随机的引入到所有或部分 EXOX 编码序列中, 如饱和突变, 然后对产生的突变体筛选其 EXOX 生物活性, 以鉴定出保持活性的突变体。SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列发生突变后, 可对其编码蛋白进行重组表达和蛋白活性的测定。

[0268] 氨基酸 [xu6] 家族的相关性也可以被侧链相互作用的基础决定。替换的氨基酸可以是保守“强”的残基或完全保守“弱”的残基。所述保守氨基酸残基的“强”的组包括以下组 :STA、NEQK、NHQK、NDEQ、QHRK、MILV、MILF、HY、FYW, 其中单个字母的氨基酸代号代表那些可以互相替换的氨基酸。另外, 保守氨基酸残基的“强”的组包括以下组 :CSA、ATV、SAG、STNK、STPA、SGND、SNDEQK、NDEQHK、NEQHRK、HFY, 其中每组中的字母代表单个字母的氨基酸代号。

[0269] 在一个实施方案中, 突变的 EXOX 蛋白可以按以下方法检验 : (i) 形成蛋白的能力 ; 和其他 EXOX 蛋白、其他细胞表面蛋白或生物活性类似物部分的相互作用 ; (ii) 突变的 EXOX 蛋白与 EXOX 配体形成联合体形式 ; 或 (iii) 突变的 EXOX 蛋白与细胞内的目标蛋白或生物活性类似物结合的能力 (如抗生物素蛋白)。

[0270] 在另一个实施方案中, 突变的 EXOX 蛋白可以通过调控特殊的生物功能 (如分解蛋白的能力) 被检验。

[0271] EXOX 多肽

[0272] 本发明对应的多肽包括含有 EXOX 多肽具有的如 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示氨基酸序列的多肽。本发明还包括从 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示序列对应的残基发生变化的突变或变异的蛋白, 所述蛋白的还具有 EXOX 蛋白的活性或生理功能, 或其功能类似物片段。

[0273] 一般地, 保存了 EXOX 类似功能的 EXOX 变异体, 包括序列中特殊位置的残基被其他氨基酸替换的任意的变异体, 进一步还包括在原蛋白的两个基团中插入一个或多个额外的

氨基酸残基或在原蛋白中删除一个或多个氨基酸残基的可能性。任何的氨基酸替换、插入或删除都属于本发明范围。在有利的条件下,所述替换如本文上面定义。

[0274] 本发明另一发明还涉及分离的 EXOX 蛋白、其类似生物活性部分或衍生物、片段、异源体或类似同源体。生物活性 [xu7] 部分指的是 EXOX 蛋白中对于一般功能必需的部分,如氨肽酶活性。同时也提供多肽片段适合于用于培养抗 EXOX 抗体的免疫原。在一个实施方案中,天然的 EXOX 蛋白可以从细胞、组织或培养基用适当的蛋白提纯方法分离得到。在另一个实施方案中,EXOX 蛋白可以通过重组 DNA 技术得到。EXOX 蛋白或多肽还可以通过标准的肽合成技术,由化学法合成,而代替重组表达。

[0275] 所述“分离的”或“纯化的”多肽、蛋白或其类似生物活性部分基本上不含来自 EXOX 蛋白来源的细胞或组织的细胞物质或其他杂质蛋白,或者用化学法合成时,基本上不含化学前体或其他化学物质。术语“基本上不含细胞物质”包括 EXOX 蛋白与细胞组分分开的 EXOX 蛋白的前体,其中细胞是用来分离或重组产生改蛋白的细胞。在一个实施方案中,术语“基本上不含细胞物质”包括含有少于约 30% (按干重量) 的非 EXOX 蛋白 (这里也称为“杂质蛋白”) 的 EXOX 蛋白前体,优选少于约 20% 的非 EXOX 蛋白的 EXOX 蛋白前体,更优选为少于约 10% 的非 EXOX 蛋白的 EXOX 蛋白前体,最优选为少于约 5% 的非 EXOX 蛋白的 EXOX 蛋白前体。当该 EXOX 蛋白或其类似生物活性部分是由重组方法产生时,还优选基本上不含培养基,即培养基的含量少于约 20%,优选为少于约 10%,更优选为少于约 5% 的 EXOX 蛋白前体。

[0276] 术语“基本上不含化学前体或其他化学物质”是指蛋白与化学前体或其他化学物质分离的 EXOX 蛋白前体,其中的化学前体或其他化学物质参与了蛋白的合成。在一个实施方案中,术语“基本上不含化学前体或其他化学物质”包括含有少于约 30% (按干重量) 的化学前体或非 EXOX 化学物质的 EXOX 蛋白前体,优选少于约 20% 的学前体或非 EXOX 化学物质的 EXOX 蛋白前体,更优选为少于约 10% 的学前体或非 EXOX 化学物质的 EXOX 蛋白前体,最优选为少于约 5% 的学前体或非 EXOX 化学物质的 EXOX 蛋白前体。进一步的,“基本上不含化学前体或其他化学物质”还包括氧化副产物。本 [xu8] 领域技术人员都知道如何防止氧化反应,如将化学物质置于一个无氧的环境中。

[0277] EXOX 蛋白的生物活性部分所包含的肽含有的氨基酸序列与 EXOX 蛋白的氨基酸序列有足够的同源性,或来源于 EXOX 蛋白的氨基酸序列 (如 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示的氨基酸序列),它包含的氨基酸序列比全长的 EXOX 蛋白的少,并表现出 EXOX 蛋白的至少一种活性。通常,生物活性部分包含至少一种 EXOX 蛋白的活性结构域或片段。EXOX 蛋白的生物活性部分可以是多肽,例如,它的长度为 10、25、50、100 或更多个氨基酸。

[0278] 此外,其他蛋白区域缺失的其他生物活性部分可以通过重组技术得到,并用来评价 EXOX 蛋白的一种或多种功能活性。

[0279] 在一个实施方案中,EXOX 蛋白含有 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示的氨基酸序列。在其他实施方案中,EXOX 蛋白基本上与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 同源,并保留了 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 蛋白的变异功能活性,但其氨基酸序列有差别,其原因如下所详述的,是由于天然的等位基因变异或突变造成的。因此,在另一个实施方案中,EXOX 蛋白是如以下所述的

蛋白:含有与 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 所示的氨基酸序列至少约 90%同源的氨基酸序列,并保留了 SEQ ID NOs :3、6、9、12、15、18、21、24、27、30、33 和 35 的 EXOX 蛋白的功能活性。本发明中,术语“生物活性”或“功能活性”指的是自然或一般的 EXO 蛋白具有的功能,例如,降解其他蛋白的能力。

[0280] 两个或多个序列同源性的测定

[0281] 为了测定两种氨基酸序列或两种核酸序列的同源性百分数,可将序列进行比对,用于最优比较目的(例如,在第一和第二氨基酸或核苷酸序列中之一,或两者中引入空位,用于最优比对,非相同的序列可考虑进行比较)。氨基酸残基、对应氨基酸位置的核苷酸或核苷酸位置可以比较。当第一序列的某个位置与第二序列相应相同位置具有相同氨基酸残基或核苷酸时,则此两个分子在该位置上具有同源性(这里所用的氨基酸或核酸“同源性”与氨基酸或核酸的“同一性”的意义是相同的)。

[0282] 核酸序列的同源性可用两个序列间的同源程度衡量。本领域技术人员可以用计算机程序计算同源性,如 GCG 软件包的 GAP 程序,参见 Needleman & Wunsch, J. Mol. Biol. 48 : 443-453, 1970。通过以下操作,利用 GCG GAP 程序来对比核酸序列:空位罚分 5.0 和空位拓展罚分 0.3,对比的核酸序列的编码区按照上述参数,表现一定的至少 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%同源可能性,与 SEQ ID NOs :2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 和 34 所示的序列的 CDS(编码)部分相比。

[0283] 术语“序列同源”是指两个多聚核苷酸或多肽在对比的特定区内具有的残基对残基相同性的程度。术语“序列同源性的百分比”是通过以下方法计算得出:在整个对比区比较两个最佳配比的序列,在两个序列中具有同样的核酸碱基(如 A、T、C、G、U 或 I,核酸中的碱基)数目产生匹配位置的数目,除以匹配位置整个对比区的数目(如整个屏幕),然后将上述结果乘以 100 产生序列同源的百分比。这里所述术语“基本同源”表示多聚核苷酸序列的一种特征,其中多聚核苷酸包括与参考序列的对比区相比至少有 80%序列同源,优选至少 85%同源,更优选 90 ~ 95%同源,最优选为有至少 99%序列同源。

[0284] 嵌合和融合蛋白

[0285] 本发明还提供 EXOX 嵌合或融合蛋白。这里所用的 EXOX “嵌合蛋白”或“融合蛋白”包括与非 EXOX 多肽选择性连接的 EXOX 多肽。“EXOX 多肽”是指含有相当于 EXOX 蛋白(SEQID NOs :3、6、9、12、15、18、21、24、27、30、33 或 35)氨基酸序列的多肽,而“非 EXOX 多肽”是指该多肽含有的氨基酸序列相当于与 EXOX 蛋白基本上不同源的多肽,例如与 EXOX 蛋白不同的蛋白,和来源于相同或不同组织的蛋白。在 EXOX 融合蛋白中,EXOX 多肽可以相当于全部或部分的 [xu9]EXOX 蛋白。在一个实施方案中,EXOX 融合蛋白含有 EXOX 蛋白的至少一个生物活性部分。在另一个实施方案中,EXOX 融合蛋白含有 EXOX 蛋白的至少二个生物活性部分。在另一个实施方案中,EXOX 融合蛋白含有 EXOX 蛋白的至少三个生物活性部分。在融合蛋白中,术语“选择性连接”是要表明 EXOX 蛋白和非 EXOX 蛋白彼此融合在同一开放读码序列中。非 EXOX 多肽可以融合到 EXOX 多肽的 N 端或 C 端。

[0286] 在一个实施方案中,该融合蛋白是 GST-EXOX 融合蛋白,其中 EXOX 序列融合到 GST 序列的 C 端。这种融合蛋白可促进重组 EXOX 蛋白的纯化。

[0287] 在另一个实施方案中,该融合蛋白是在 N 端含有异源信号序列的 EXOX 蛋白。在某些宿主细胞中(例如哺乳动物宿主细胞),通过使用异源信号序列可以提高 EXOX 的表达和

/或分泌水平。

[0288] 在另一个实施方案中,融合蛋白是一种 EXOX- 免疫球蛋白融合蛋白,其中 EXOX 基因与来自免疫球蛋白家族的一个基因融合。本发明所述的 EXOX- 免疫球蛋白融合蛋白可以作为药物组合物治疗抑制 EXOX 配体和 EXOX 蛋白在细胞表面上的结合的药物中的应用,这样抑制生物体内 EXOX- 介导的信号传导。EXOX- 免疫球蛋白融合蛋白还可以用于影响 EXOX 相关的配体的生物利用度。对 EXOX 配体 /EXOX 相互作用的抑制在医疗上有用,可以针对治疗增殖和分化的紊乱,调节(如促进或抑制)细胞生长。进一步,本发明所述的 EXOX- 免疫球蛋白融合蛋白可以作为免疫原,在生物体内产生抗 -EXOX 抗体,用来纯化 EXOX 配体,和扫描分析验证抑制 EXOX 和 EXOX 配体相互作用的分子。

[0289] 本发明所述的 EXOX 嵌合体或融合蛋白可以通过标准的重组 DNA 技术获得。例如,编码不同多肽序列的 DNA 片段通过常规的技术连接在一起,如为连接提供平头末端或粘性末端,用限制酶切提供合适的末端,添加合适的粘性末端,用碱磷酸化处理防止错配,最后酶连接。在另一个实施方案中,融合蛋白可以通过常规的技术合成获得,包括 DNA 自动合成。可供选择的,基因片段的 PCR 扩增运用锚引物,可以在两个连续的基因片段上产生互补的突出序列,这样可以顺序的将两个片段融合和扩增,产生一个嵌合体基因序列(参见 Ausubel 等(eds.)CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992)。进一步的,许多商业上可获得的表达系统已编码了融合的一部分(如, GST 多肽)。EXOX- 编码的核酸可以克隆到上述表达载体中,这样 EXOX 蛋白与融合的一部分连接在一起。EXOX 激动剂和拮抗剂

[0290] 本发明还包括 EXOX 蛋白的变体,它可以起 EXOX 激动剂(模拟物)的功能,或起 EXOX 拮抗剂的功能。EXOX 蛋白的变体可通过诱变产生(例如, EXOX 蛋白的不连续点突变或截短)。EXOX 蛋白的激动剂可使其基本上保留与天然存在形式的 EXOX 蛋白相同、或一部分的生物学活性。EXOX 的拮抗剂可抑制天然存在的 EXOX 蛋白形式的一种或多种活性,例如,通过竞争性抑制细胞信号通道的上游或下游,其中包括 EXOX 蛋白酶。因此,用限定功能的变体治疗,可引起特定的生物学效应。在一个实施方案中,与用天然存在形式的 EXOX 蛋白治疗相比,用具有天然存在形式的 EXOX 蛋白的生物活性中的一部分的变体治疗,对受试者的副作用较少。

[0291] 作为 EXOX 蛋白激动剂(模拟物),或 EXOX 蛋白拮抗剂的 EXOX 蛋白的变体,可用下述方法鉴定:对 EXOX 蛋白的突变体(例如截短突变体的)组合文库,筛选其 EXOX 蛋白激动剂或拮抗剂活性。在一个实施方案中,多样化的 EXOX 突变体文库可通过在核酸水平上的组合诱变来产生,并由多样化的基因文库所编码。多样化的 EXOX 变体文库可以通过下述的方法产生,例如,将合成的寡核苷酸混合物用酶法连接成各种基因序列,从而使一组简并的潜在 EXOX 序列可以按单个多肽的形式表达,或者以包含一组 EXOX 序列的一组大融合蛋白(如噬菌体表达)的形式表达。由简并的寡核苷酸序列生产潜在 EXOX 变体的方法有很多种。可以在自动 DNA 合成仪中进行简并基因序列的化学合成,然后将合成的基因连接到适合的表达载体中。使用一组简并的基因,可以在一种混合物中提供全部编码所需要的一组潜在 EXOX 序列。合成简并的寡核苷酸的方法在本领域是已知的,参见例如, Narang, Tetrahedron(1983); Itakura 等, Annu. Rev. Biochem. 53 :323(1984); Itakura 等, Science 198 :1056(1984); Ike 等, Nucleic Acid Res. 11 :477(1983)。

### [0292] 多肽文库

[0293] 此外, EXOX 蛋白编码序列片段的文库可以用来产生多样化的一组 EXOX 片段, 用于筛选和随后选择 EXOX 蛋白的变体。在一个实施方案中, 产生编码序列片段的文库的方法如下: 用核酸酶处理 EXOX 编码序列的双链 PCR 片段, 处理的条件是使每个分子仅发生约一次切口, 然后使双链 DNA 变性, 再使 DNA 复性形成双链 DNA, 该 DNA 可以包括不同切口的产物的正义 / 反义对, 用 S1 核酸酶处理重新形成的双链体, 以除去单链部分, 再将得到的片段文库连接到表达载体中。用这种方法可获得编码各种大小的 EXOX 蛋白的 N 端、C 端和内部片段的表达文库。

[0294] 已知在本领域有几种用于筛选通过点突变或截短制备的组合文库的基因产物, 以及用于对 cDNA 文库筛选具有选择特性的基因产物的技术。这些技术适用于对 EXOX 蛋白组合诱变所产生的基因文库进行快速筛选。能进行高通量分析、使用最广的用于筛选大基因文库的技术通常包括: 将基因文库克隆到可复制的表达载体中, 用产生的载体文库转化适合的细胞, 以及在某种条件下使组合的基因表达, 在该表达条件下, 通过对所需活性的检测, 促进了分离出编码该被测产物基因的载体。Recursive ensemble 诱变 (REM), 是一种新的技术, 它可提高文库中功能突变体的频度, 可以结合筛选测定用于鉴别 EXOX 变体, 参见例如, Arkin & Yourvan Proc. Natl. Acad. Sci. USA 89 :7811-7815 (1992) ; Delgrave 等, 蛋白 Engineering 6 :327-331 (1993)。

[0295] 文库也可以通过 DNA 重排产生。DNA 重排采用来自不同种属或基因的相关基因, 其在功能、片段上相关, 通过重组使其重新组合。可以确认其中含有有用或潜在令人感兴趣的产物的重组基因。所有重组的基因相当于片段来说具有用途, 可以重新组合形成新的重组基因。来自不同种属或基因的不同片段经过复性和延伸, 在文库中形成多样性。筛选出一个令人感兴趣的蛋白的过程是合适的。利用 DNA 重排产生重组基因的重要因素包括复性时的温度、基因的相似性和 DNA 片段的大小。

[0296] Stemmer 等 (Nature 370 :389-391 (1994))、Stemmer (Proc. Natl. Acad. USA 91 :10747-10751 (1994))、美国专利 No. 5, 603, 793、No. 5, 830, 721 和 No. 5, 811, 238, 在这里作为具体的参考文献, 描述了如生物体外的蛋白重排方法, 如诱变的重复循环、重排和选择, 如同一系列的产生 displayed 多肽和抗体的方法, 如同一系列通过 DNA 片段进行的 DNA 重新组配技术, 以及它们生物体内外的诱变复制。进一步的, DNA 重排技术的不同应用是本领域技术人员熟知的。除了上述引用的公开文献, 美国专利 No. 5, 837, 458 也提供了在递归的重排技术里新的代谢途径和生物产率的提高的进化, 和 Cramer 等 (Nature Medicine 2(1) :100-103 (1996)) 描述了抗体重排应用于抗体噬菌体文库。同时也可参考 W095/22625, W097/20078, W096/33207, W097/33957, W098/27230, W097/35966, W098/31837, W098/13487, W098/13485 和 W0989/42832。

### [0297] 表达载体

[0298] 本发明的另一个方面是关于载体, 优选表达载体, 它含有编码 EXOX 蛋白、或衍生物、片段、同一物或其同源物的核酸。这里所用的术语“载体”是指一种核酸分子, 它能转运与其连接的另一核酸。载体的一种类型是“质粒”, 它是指可将另外的 DNA 片段连接在其中的环状双链 DNA。另一种类型的载体是病毒载体, 在该载体中, 另外的 DNA 片段可与病毒的基因组连接。一些载体能够在其导入的宿主细胞中自主复制 (例如具有细菌复制起点的细

菌载体和游离型的哺乳动物的载体)。其他的载体(例如,非游离型的哺乳动物载体)是通过导入宿主细胞而整合到宿主细胞的基因组中,从而与宿主的基因组一起复制。此外,某些载体能指导与其选择性连接的基因的表达,这种载体在此称为“表达载体”。一般,重组 DNA 技术中使用的表达载体常常为质粒的形式。本说明书中,“质粒”和“载体”可以互换使用,因为质粒是最常用的一种载体形式。但是,本发明还包括其他形式的有同等功能的表达载体,例如,病毒载体(例如复制缺陷逆转录病毒、腺病毒和腺伴随病毒)。

[0299] 功能蛋白的产生与组织产生蛋白的细胞机器紧密相关。在选择表达系统时,大肠杆菌是经典的“工厂”,因为它的基因组全部测得;组织容易处理;生长快;只需便宜容易制备的基质;分泌蛋白到基质中,从而有利于收集蛋白。但是,大肠杆菌是原核细胞,缺少细胞器,如真核具有的内质网和高尔基体,它们能够修饰产生的蛋白。许多真核蛋白可以在大肠杆菌中表达得到,但许多是没有功能的、无用的形式,因为糖基化或后转录修饰没有发生。

[0300] 这样,研究者最近转向了真核酵母、哺乳动物核植物蛋白表达系统。例如,在过去几年里嗜甲醇酵母 *P. pastoris* 成为了不同蛋白表达系统中的有力的宿主,同时建立了一种选择性表达人类蛋白的高产量的真核宿主技术。

[0301] 另一个例子,植物也被大量的蛋白表达系统用作表达宿主,其提供了潜在的比经典的表达系统更具有经济价值、可控和安全的优势 [xu10]。最近,有系列的植物不同表达系统出现,包括短期表达、植物细胞悬浮液培养、重组植物病毒和叶绿体转基因系统。当在植物里表达的蛋白与哺乳动物有一些差异时,目前没有证据表明这些差异导致在病人身上产生不良反应。

[0302] 另一个适合不同表达系统利用了昆虫细胞,经常与杆状病毒表达载体结合。杆状病毒载体适用于在培养的昆虫细胞表达蛋白,例如, SF9 细胞,包括 pAc 系列 (Smith 等, *Mol. Cell. Biol.* 3 :2156-2165(1983)) 和 pVL 系列 (Lucklow 和 Summers, *Virology* 170 : 31-39(1989))。

[0303] 本发明的宿主细胞也可以用于生产除人以外的转基因动物,所述动物的基因组中已经转入了外源的序列。转基因动物是除人以外的动物,优选哺乳动物,更优选啮齿动物,如大鼠或小鼠,其中该动物的一个或多个细胞包含转移基因。其他转基因动物的例子包括除人以外的灵长类动物、绵羊、狗、牛、山羊、鸡、两栖类等。通过胚胎的操作和显微注射而产生转基因动物的方法,特别是对小鼠之类的动物,在本领域中已成为常规的方法,并在以下文献中有描述,例如:美国专利 NOs. 4, 736, 886、4, 870, 009 和 4, 873, 191 ;Hogan, 1986. In : *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.)。类似的方法也用来生产其他的转基因动物。Pichia pastoris 表达系统

[0304] 真核酵母之一是嗜甲醇 *Pichia pastoris*。P. pastoris 已经被开发成一个产生外源蛋白合理的宿主,因它的乙醇氧化启动子被分离和克隆了。Pichia pastoris 转录在 1985 年就已经有报道了。Pichia pastoris 异源蛋白表达系统由 Phillips Petroleum 发展,参见美国专利 NOs. 4, 855, 231, 4, 857, 4674, 879, 231 和 4, 929, 555, 本发明中全面引述上述专利。该系统最近被体外标记。与其他真核表达系统相比,Pichia 具有很多优势,因为它没有细菌相关的内毒素问题,也没有动物细胞培养产生的蛋白中含有滤过性毒菌的问题。进一步的,P. pastoris 可以在缺乏葡萄糖的情况下利用甲醇作为碳源。P. pastois 表达系统利用甲醇介导的乙醇氧化酶 (AOX1) 启动子,该启动子控制乙醇氧化酶的表达基因,该酶在甲

醇代谢过程中第一个启动。上述启动子已研究清楚并整合到了一系列 *P. pastoris* 表达系统中。由于 *P. pastoris* 产生的蛋白具有正常的结构并能分泌到胞外, 基因工程 *P. pastoris* 发酵提供了一种很好的代替大肠杆菌表达系统的选择。进一步的, *P. pastoris* 还有超过大肠杆菌的优势: 自发的对产生的蛋白糖基化。该表达系统已制得了很多的蛋白, 包括破伤风菌疫苗片段、*Bordetella pertussis pertactin*、人血清蛋白和溶菌酶。

[0305] EXOX 蛋白标签去除

[0306] 已经开发了不少系统用来快速有效的纯化在细菌中表达的重组蛋白。其中大部分依赖与谷胱甘肽转移酶 (GST) 区域、钙结合多肽 (CBP) 组氨酸标签融合的蛋白表达。例如, 结合谷胱甘肽转移酶 (GST) 的片段的多肽表达可以从细菌粗提取物中在不变性的条件下用谷胱甘肽凝胶亲和层析色谱法纯化。

[0307] 进一步的, 上述载体表达系统一般引入特殊的蛋白酶水解位点, 方便蛋白酶对细菌融合蛋白进行水解, 载体引入的水解位点有凝血酶、肠激酶、或 Xa 蛋白因子。凝血酶专门对包含 Leu-Val-Pro-Arg ↓ Gly-Ser (SEQ ID NO:44) 序列的目标蛋白进行酶解。肠激酶的酶切位点是 Asp-Asp-Asp-Asp-Lys ↓ (SEQ ID NO:45)。象肠激酶, Xa 蛋白因子酶解位点在其结合序列 Ile-Glu-Gly-Arg ↓ (SEQ ID NO:46) 的 C 末端, 这样就可以用来从适当设计结构中去除所有载体编码的序列。上述所有蛋白水解酶现在都可以在商业上获得高纯度的, 以避免其包含的其他酶的二次降解。每个试剂盒都提供了所有的适合酶保存、利用在酶解基质中去除该酶的生物试剂等工具。最近 Qiagen 公司开发了一种 TAGZyme 系统, 能有效的去除从蛋白 N 末端的组氨酸标签, 该蛋白包括能有序地从 N 末端到“停止点”氨基酸结构域分解双肽的外肽酶, 包括 ↓ Lys-Xaa、↓ Arg-Xaa、↓ Xaa-Xaa-Pro-Xaa-、↓ Xaa-Pro-Xaa-Xaa- 或 ↓ Gln-Xaa。

[0308] 尽管没有太大的必要去除一个纯化后的重组蛋白中短的组氨酸相似标签 (不管组氨酸残基为多少), 但在一些应用中需要去除组氨酸, 如 X 射线晶体结构分析或 NMR。同样的对于凝血酶酶切位点 Gly-Ser 残基或其他多余的 N 末端氨基酸残基也需要上述处理, 其存在和所用表达系统相关。

[0309] 最近利用在巯基基团和具有两个相邻亲核基团的分子之间浓缩反应方法来纯化。例如具有两个相邻亲核基团的氨基酸包括丝氨酸、苏氨酸和半胱氨酸。纯化蛋白或多肽包括由如 N 末端半胱氨酸、苏氨酸或丝氨酸残基和树脂间形成可逆的共价的键, 参见 Villian 等, *Chem. & Biol.* 8:673-679(2001)。在重组蛋白的 N 末端增加一对残基, 如 Thr-Pro、Cys-Pro 或 Ser-Pro, 或化学合成的蛋白或多肽, 通过以下两个步骤纯化: (1) 通过共价连接纯化; (2) 去除二肽标签。该方法有效的回收重组蛋白的成熟形式, 而没有二肽标签序列。

[0310] EXOX 蛋白的逆蛋白水解活性

[0311] 本发明的另一个方面是关于将一个或多个氨基酸接到另一个氨基酸、多肽、寡多肽、多聚肽或任意的可连接的二级胺化合物的方法, 该方法是利用 EXOX 蛋白的逆蛋白水解活性。本发明中, 术语“逆蛋白水解活性”是指能促进将一个或多个氨基酸接到另一个氨基酸、多肽、寡多肽、多聚肽或任意的可连接的二级胺化合物的酶活性。本技术领域一个公知的技术是在合适的热力学的条件下, 蛋白水解酶具有逆蛋白水解活性。

[0312] 其中一个蛋白水解酶具有逆蛋白水解活性的酶是胰蛋白酶, 该酶是胰腺的丝氨酸酶, 其专门针对带正电的赖氨酸和精氨酸侧链的底物。胰蛋白酶被广泛的应用于从猪胰岛



素制备人胰岛素,其中猪胰岛素除了在 B 链最后一个氨基酸丙氨酸代替了人胰岛素的苏氨酸,其余类似。在胰蛋白酶存在下,将猪胰岛素与苏氨酸酯反应,通过去除其末端的丙氨酸后加上苏氨酸酯,形成人胰岛素苏氨酸酯。随后将人胰岛素苏氨酸酯用三氟醋酸水解该酯,从而形成人胰岛素。

[0313] 在一些实施方案中,EXO 蛋白被用来促进逆蛋白水解活性。举些例子,将 EXO 蛋白和一个多肽或一个或多个氨基酸在增加一个或多个氨基酸到多肽的环境里培育。

[0314] 有很多利用本发明 EXO 蛋白的逆蛋白水解活性的途径。例如,利用 EXO 蛋白的逆蛋白水解活性合成多肽链。EXO 蛋白也可以作为一种连接工具,将一个或多个氨基酸接到另一个氨基酸、多肽或任意的可连接的二级胺化合物。

[0315] 药物组合物

[0316] 本发明的 EXO 核酸分子、EXO 蛋白、抗 EXO 抗体(这里亦称为活性化合物)和衍生物、片段、同一物和其同源物可以掺入适合于给药的药物组合物中。这种组合物通常含有核酸分子、蛋白、抗体,或调节化合物和可药用载体。这里所用的术语“可药用载体”是指包括一些和全部的与药物的给药相容的溶剂、分散介质、包衣、抗细菌和抗真菌剂、等渗剂和吸收延迟剂等适合于给药的载体。适合的载体在最新版的 Remington's Pharmaceutical Sciences 一书中有描述,该书是本领域标准的参考,也是本发明全面引述的。优选的载体或稀释剂包括但不限于水、盐溶液、Ringer's 溶液、葡萄糖溶液和 5% 人血清清蛋白。脂质体和无水赋形剂,如不挥发的油,也可以使用。这种介质和试剂用于药物活性物质中在本领域来说是众所周知的。迄今为止除了任何常规使用的介质或试剂与活性化合物不相容以外,上述介质和试剂在组合物中的使用是经过仔细考虑的。补充的活性化合物也可掺入组合物中。

[0317] 封装技术已广泛应用于许多工业领域中。例子包括控释药剂、食品和饮料中的色料、食品中抗氧化剂、农业中控释外激素。胶囊、微胶囊和微球是小球体微粒,活性成分包含在微粒混合物中或附着在微粒表面。例如,封装在生物可降解的藻酸盐微粒里。生物封装技术趋向于封装细胞、酶和生物活性材料。

[0318] 本发明的药物组合物配制成与其给药途径相容的形式。给药途径的例子包括非肠道的给药,例如静脉内、真皮内、皮下的、口服的(例如吸入)、经皮的(局部的)、经粘膜的,和直肠的给药。用于非肠、真皮内、或皮下应用的溶液或悬浮液可以包括以下组分:无菌的稀释剂,例如注射用水、盐溶液、不挥发油类、聚乙二醇、甘油、丙二醇或其他合成的溶剂;抗细菌剂,例如苯甲醇或对羟基苯甲酸甲酯;抗氧化剂,例如抗坏血酸或二亚硫酸钠;螯合剂,如乙二胺四乙酸(EDTA);缓冲液,例如乙酸盐、柠檬酸盐或磷酸盐缓冲液和调节渗透压的试剂,例如氯化钠或葡萄糖。pH 可用酸或碱调节,例如盐酸或氢氧化钠。非肠道制剂可以密封入安瓿、一次性注射器或由玻璃或塑料制成的多剂量的小瓶中。

[0319] 适用于可注射用的药物组合物包括无菌的含水溶液(水可溶的情况)或分散液,或者,对临时配制的无菌可注射溶液或分散液的制剂来说,包括无菌的粉末。对于静脉内给药,合适的载体包括生理盐水、抑菌水、Cremophor EL™(BASF, Parsippany, N. J.) 或磷酸盐缓冲液(PBS)。在所有场合,该组合物必须是无菌的,并应是流动的,其流动性应达到易于注射的程度。在生产和储存的条件下,该组合物必须是稳定的,并且必须能抗微生物,如细菌和真菌的污染作用而能保存。该载体可以是溶剂或含有,举例来说,水、乙醇、多元醇(例

如,甘油、丙二醇和液体聚乙烯醇等),及其它合适的混合物的分散介质。适当的流动性可通过以下方法保持:例如,使用包衣,如卵磷脂;在分散体的场合,保持所要求的颗粒大小;以及使用表面活性剂。微生物作用的预防可通过各种抗菌剂和抗真菌剂而达到,例如对羟基苯甲酸酯、氯丁醇、酚、抗坏血酸、汞硫代水杨酸钠等。在很多场合,组合物中最好包括等渗剂,例如蔗糖、多元醇,如甘露醇、山梨醇、氯化钠。在组合物中包括延迟吸收的试剂,例如单硬脂酸酯铝和明胶,可达到使注射组合物延长吸收的目的。

[0320] 无菌的注射溶液的制备方法如下:将活性化合物(例如, EXOX 蛋白或抗 mGluR5M 抗体)按所要求的量掺入适当溶剂中,该溶剂含有以上列举的组分中的一种或其组合,然后,按照要求通过无菌过滤。一般,分散剂的制备方法是,将活性化合物掺入无菌载体中,该载体含有基本的分散介质和选自以上列举组分中的其他必须组分。对于用于制备无菌注射溶液的无菌粉末,优选的制备方法是真空干燥和冷冻干燥,该方法由预先经过无菌过滤的溶液生成含活性组分和其它所需组分的粉末。

[0321] 从红色毛癣菌或产生 EXOX 或从红色毛癣菌分离的 EXOX 的转基因真菌的培养基质中分离得到的粗提物可以用于口服给药,因为蛋白酶是胞外的。口服组合物一般包括惰性稀释剂或可食的载体。口服组合物可以封装在明胶胶囊内,或压成片。用于口服治疗给药时,活性化合物中可掺入赋形剂,并以片剂、锭剂或胶囊的形式使用。口服组合物也可以用流体载体来制备,以使用作嗽口剂,其中在流体载体中的化合物,被经口施用并漱口和吐出,或吞下。药物相容的粘结剂,和/或佐剂物质也可作为组合物的一部分包括在内。片剂、丸剂、胶囊、锭剂等可以含有下列的任一种组分,或类似性质的化合物:粘结剂,例如微晶纤维素、黄蓍胶或明胶;赋形剂,例如淀粉或乳糖;崩解剂,例如藻朊酸、Primogel 或玉米淀粉;润滑剂,例如硬脂酸镁或 Sterotes;滑动剂,例如胶体的二氧化硅;甜味剂,例如蔗糖或糖精;或调味剂,例如薄荷、水杨酸甲酯或橙味调料。

[0322] 用于吸入给药时,组合物可以用气溶胶喷雾的形式由压力容器或分配器,或者喷雾器来递送,压力容器或分配器含有合适的推进剂,例如,二氧化碳气体。

[0323] 全身的给药也可通过经粘膜或经皮给药的方法。对于经粘膜或经皮给药,在该制剂使用适合于需要透过屏障的渗透剂。这种渗透剂在本领域中通常是熟知的,例如,对于经粘膜给药,包括清洁剂、胆盐,和梭链孢酸衍生物。经粘膜给药也可以通过使用鼻的喷雾剂或栓剂来完成。对于经皮给药,如本领域通常所知,可将活性化合物配制在软膏、油膏、凝胶或霜剂中。

[0324] 该化合物也可以制成栓剂(例如,用常规的栓剂基,如可可酯和其他甘油酯)或保留灌肠的形式,用于直肠递送。

[0325] 在一个实施方案中,活性化合物采用能防止该化合物从人体迅速消除的载体来制备,例如控释制剂,包括植入物和微胶囊化传递系统。生物可降解的、生物相容的聚合物也可以使用,例如乙酸乙烯酯、聚酞、聚乙醇酸、胶原、聚原酸酯和聚乳酸。制备这种制剂的方法对本领域的技术人员来说是显而易见的。上述的材料也可从市场上由 Alza 公司和 Nova 制药公司获得。脂质体悬浮液(包括用病毒抗原单克隆抗体导向受感染细胞的脂质体)也可以用作可药用载体。这些载体可以按照本领域技术人员熟知的方法制备,举例来说,如美国专利 NO. 4,522,811 中所述。

[0326] 尤其有利的是,按剂量单位形式来配制口服或非肠道给药的组合物,以便于给药

和有利于剂量的统一。这里所用的剂量单位形式是指以单元剂量适用于治疗受试者的、物理形式上分开的单位；每单位含有预定量的活性化合物，该量计算为能与所需药用载体结合产生所需疗效的量。本发明的剂量单位形式的规格根据以下方面而定，或直接依赖于以下方面：活性化合物的特性和要达到的疗效，以及在合成这种用于治疗个体的活性化合物的技术本身存在的限制。

[0327] 本发明的核酸分子也可以插入载体中并用作基因治疗载体。基因治疗载体可以递送给治疗受试者，例如，通过静脉内注射、局部给药（参见美国专利 5,328,470）或通过立体定位注射（参见，例如 chen 等（1994）PNAS 91:3054-3057）。基因治疗载体的药物制剂可包括可接受的稀释剂中的基因治疗载体，或可包括包埋了基因递送载体的缓释基质。或者，在完整的基因递送载体可以由重组细胞，例如逆转录病毒载体完整地产生的场合，该药物制剂可以包括一种或多种产生基因递送系统的细胞。

[0328] 上述药物组合物可以连同给药说明一起包含在容器、小盒或喷雾器内。

## 具体实施方式

### [0329] 实施例 1：材料与方法

#### [0330] 菌株和质粒

[0331] 本研究使用的是红色毛癣菌的一个临床分离株 CHUV 862-00。用大肠杆菌 (*E. coli*) LE392 来繁殖噬菌体  $\lambda$  EMBL3 (Promega, Wallisellen, Switzerland)。所有的质粒-亚克隆实验使用质粒 pMTL2I 在大肠杆菌 DH5 $\alpha$  中进行 (Chambers 等, Gene 68:139-149(1988))。使用 *P. pastoris* GSI15 和表达载体 pKJ 113 (Borg-Von Zepelin et al, Mol. Microbiol 28:543-554(1998)) 来表达重组肽酶。众所周知，在本领域 *P. pastoris* 可以被用来表达大量的重组蛋白。

#### [0332] 红色毛癣菌生长培养基

[0333] 红色毛癣菌在 Sabouraud 琼脂糖或液体培养基 (Bio-Rad, Munchen, Germany) 上生长，或为了提高产物的蛋白水解活性，在含有 0.2% 大豆蛋白 (Supro 1711, Protein Technology International, St. Louise, MO) 作为唯一氮源和碳源的液体培养基中培养。这种培养基中不添加盐。那些本领域的技术人员熟知的盐也可以添加到生长培养基。100ml 的液体培养基和少量新鲜生长菌丝一起在 800ml 组织培养瓶中孵育。培养物在 30°C 下无须震荡培养 10 天。

#### [0334] 基因组和 cDNA 文库

[0335] 从新鲜的生长旺盛的菌丝中分离 DNA 用于构建红色毛癣菌的基因组 DNA 文库 (Yelton et al. PNAS 81:1470-1474(1984))。用 Sau3A 部分消化 DNA，从低熔点琼脂糖凝胶中回收 12 ~ 20kb 的 DNA 片段。使用一个合适的克隆体系 (Promega)，将这些 DNA 片段插入到噬菌体 XEMBL3 中。

[0336] 使用一种微量 mRNA 系统和 500ug 总 RNA，一个红色毛癣菌的 cDNA 文库被制备，cDNA 被插入进 pSPORT6 质粒中 (Invitrogen Life Technologies; Rockville, Maryland, USA)。从大豆蛋白液体培养基的 10 天培养物 (10 $\times$ 100 毫升) 中制备 RNA。用研钵在液氮中将菌丝研磨成细的粉末，用专供植物和真菌的 RNeasy 总 RNA 纯化试剂盒 (Qiagen, Basel, Switzerland) 分离总 RNA。

[0337] 以前已经用 CHUVI 92-88 菌株构建了一个烟曲霉菌的 cDNA 文库, 该烟曲霉菌菌株在含 0.2% 胶原蛋白作为唯一氮源和碳源的液体培养基中, 在 30°C 的条件下生长 40 小时 (Monod 等, 1991)。总 RNA 依照 Applegate 和 Monod 描述的方法提取, 依据标准程序 (Sambrook 等, 1989) 用 oligo(dT) 纤维素膜 (Sigma, Buchs, Switzerland) 纯化获得 mRNA。按照制造商提供的程序, 用  $\lambda$  噬菌体 gt11 (Promega) 和这些 mRNA 制备一个文库。

[0338] 表 13. 编码氨肽酶的红色毛癣菌和烟曲霉菌的基因

[0339]

基因	基因组 DNA (bp. 从 ATG 到终止密码子)	cDNA :ORF 长度 (bp.) 自 ATG 密码子开始	自 ATG 密码子开始编码的氨基酸残基 (aa)	内含子 (bp. 自 ATG 密码子开始的基因组 DNA)
ruLAP2	1757	1488	495	3 个内含子 (bp106-231 ;556-632 ;917-982) 4 个内含子, 编码 35, 108, 95, 257aa
fuLAP2	1557	1497	498	1 个内含子 (bp85-144) 2 个内含子编码 28, 470aa
ruLAP1	1256	1122	373	2 个内含子 (bp157-226 ;968-1031) 3 个内含子编码 52, 247, 74aa
fuLAP1	1298	1167	388	2 个内含子 (bp187-252 ;1000-1064) 3 个内含子编码 62, 249, 77aa

[0340] LAP 基因的克隆

[0341] 基因组文库的重组噬菌斑 ( $10^4$ ) 被转移到 GeneScreen 尼龙膜上 (NEN Life Science Products, Boston, MA)。在低严格条件下, 用  $^{32}\text{P}$  标记的探针对这些膜进行杂交 (Monod 等, Mol. Microbiol. 13 :357-368 (1994))。采用 Grossberger 描述的方法, 所有的阳性噬菌斑被纯化并提取相应的噬菌体 DNA (Grossberger, Nucleic Acid Res. 15 : 6737 (1987))。按照标准程序, 源自 EMBL3 噬菌体的杂交筛选到的片段被亚克隆进 pMTL 2I 中。用 Microsynth (Balgach, Switzerland) 进行核酸测序。

[0342] 用标准 PCR 方法分离 cDNA

[0343] 制备 cDNA 文库中 106 个克隆的 DNA, 用 PCR 方法获得红色毛癣菌和烟曲霉菌的 cDNA。

[0344] 根据源自不同肽酶基因 (表 13) 的 DNA 序列设计的同源引物, 在标准条件下进行 PCR。200ngDNA、10u1 正义和反义寡核苷酸引物 (浓度为 42mM)、8u1 的脱氧核苷酸混合物 (每种 dNTP 为 10mM), 溶于 100u1 PCR 缓冲液中 (10mM Tris-HCl pH8.3 ;50mm KCl ;1.5Mm MgCl<sub>2</sub>)。每个反应体系加入 2.5 单位 AmpliTaq DNA 聚合酶 (PE., Zurich, Switzerland)。PCR 反应程序为 :反应混合物先在 94°C 温育 5 分钟, 接着 25 个循环 (94°C 0.5 分钟, 55°C 0.5 分钟, 72°C 0.5 分钟), 最后 72°C 保温 10 分钟。

[0345] 重组 LAPs 的制备

[0346] 通过将克隆 cDNA 的 PCR 产物连进 *E. coli*-*P. pastoris* 穿梭载体 pKJ 113 的多克隆位点, 构建表达质粒。用 PCR 纯化试剂盒 (ROCHE Diagnostics) 纯化 PCR 产物, 用先前设计的 5'-引物 (表 14) 上的限制性内切酶来消化。用 10pg 经 EcoR I 或 Sma I 线性化的质粒 DNA 电激转化 *P. pastoris* GSI 15 (Invitrogen) 在组氨酸缺陷型培养基 (1M 山梨糖, 1% (W/V) 葡萄糖, 1.34% (W/V) 不含氨基酸酵母 YNB,  $4 \times 10^{-5}$ % (W/V) 生物素,  $5 \times 10^{-3}$ % 氨基酸 (W/V) (例如 L-谷氨酸, L-甲硫氨酸, L-亮氨酸, L-异亮氨酸各  $5 \times 10^{-3}$ %), 2% (W/V) 琼脂糖) 选择转化株, 用最小的甲醇平板 (1.34% (W/V) 不含氨基酸的 YNB,  $4 \times 10^{-5}$ % (W/V) 生物素、0.5% (V/V) 甲醇、2% (W/V) 琼脂糖) 筛选插入到 AOX1 位点的结构中。转化株不能在含有甲醇作为唯一氮源的培养基上生长, 经过在 AOX1 位点取代 AOX1 编码区域的整合事件, 保证了在正确的酵母基因组位置包含上述结构。这些转化株在 10ml 甘油酵母培养基 (0.1M 磷酸钾缓冲液, pH6.0, 包括 1% (W/V) 酵母提取物、2% (W/V) 蛋白胨、1.34% (W/V) 不含氨基酸的 YNB, 1% (V/V) 甘油和  $4 \times 10^{-5}$ % (W/V) 生物素) 中, 在 30°C 下, 生长至接近饱和 (600nm 下 OD 20)。收集细胞、将其重新悬浮于 2ml 培养基中, (除了用 0.5% (V/V) 甲醇取代甘油外, 其他都相同) 孵育 2 天。两天后, 收集上清液, 用 SDS-PAGE 胶检测蛋白产物。从 400ml 细胞培养物的上清液中获得大量重组肽酶。

[0347] 表 14. 描述了在 *P. pastoris* 中表达不同 LAP 所用的材料

[0348]

基因	寡核苷酸	方向	编码的氨基酸序列*	PCR 产物 (含克隆位点)	载体
<i>ruLAP2</i>	GT TG/T CGA CTT GTT GGT CAA GAG CCC TTC GGA TGG (SEQ ID NO : 47) CT TGC/GGC CGC TTA CAT GAA GAC AGT GTG GTG TCC (SEQ ID NO : 48)	正向 反向	(R) (L) VGQEPFGW (SEQ ID NO: 63) GHHTVEMSTOP (SEQ ID NO: 64)	<i>ruLAP2</i> (58-1486) SalI---NotI	pKJ113 XhoI---NotI
<i>fulAP2</i> <sup>†</sup>	GT TC/T CGA GGC CCA GCA TGG GAC TGG AAG (SEQ ID NO : 49) CGC AAA GG/T GCA CTC GCC CCG CGA (SEQ ID NO : 50) TCG CGG GGC GAG/TGC ACC TTT GCG (SEQ ID NO : 51) CTT A/GA TCT CTA CTG CTC AAC CCG GTC CTT (SEQ ID NO : 52)	正向 反向 正向	(R) GPGWDWK (SEQ ID NO: 65) SRGECTFA (SEQ ID NO: 66) SRGECTFA (SEQ ID NO: 67)	<i>fulAP2a</i> (49-460) XhoI---ApaLI <i>fulAP2b</i> (461-1494)	pKJ113 XhoI---BamHI
<i>ruLAPI</i>	GT TC/T CGA GGC ATT CCT GTT GAT GCC CGG GCC G (SEQ ID NO : 53) CTT A/GA TCT TTA CTT AGC AAG CTC AGT GAC GAA GCC GAC (SEQ ID NO : 54)	反向 正向 反向	KDRVEQSTOP (SEQ ID NO: 68) (R) (G) IPVDARA (SEQ ID NO: 69) VGFVTELAkSToP (SEQ ID NO: 70)	ApaLI---BgIII <i>ruLAPI</i> (61-1119) X7toI---BgIII	pKJ113 XhoI---BamHI
<i>fulAPI</i>	GT TC/T CGA GGG GCT GTA GCT GCA GTG ATT (SEQ ID NO : 55) CTT A/GA TCT TTA AAA CGG CGC AAA TGC CAA (SEQ ID NO : 56)	正向 反向	(R) GAVAAVI (SEQ ID NO: 71) LAFAPfSToP (SEQ ID NO: 72)	<i>fulAPI</i> (46-1164) XhoI---BgIII	pKJ113 XhoI---BamHI
<i>RuDPPIV</i> <sup>§</sup>	CT TC/T CGA GTC GTT CCT CCT CGT GAG CCC CG (SEQ ID NO : 57) G TTC CAT GGT/CAT GAC CTT TGT GTC ATA CGA GAC AG (SEQ ID NO : 58) GT TCC ATG GT/C ATG ACC CCT CTC GTC AAC GAT AAG G (SEQ ID NO : 59) CTT G/GA TCC TCA TTC CTC TGC CCT CTC ACC (SEQ ID NO : 60)	正向 反向 正向 反向	(R) (V) VPPREPR (SEQ ID NO: 73) VSYDIKVM (SEQ ID NO: 74) VMTPLVNDK (SEQ ID NO: 75) GPRAEESSTOP (SEQ ID NO: 76)	<i>ruDPPIVa</i> (49-1266) XhoI---RcaI <i>ruDPPIZb</i> (1267-2325)	PKJ111 XhoI---BamHI
<i>ruDPPV</i>	CCG G/AA TTC TTT ACC CCA GAG GAC TTC (SEQ ID NO : 61) GAG T/CT AGA CTA GTA GTC GAA GTA AGA GTG (SEQ ID NO : 62)	正向 反向	(E) (F) FTPEDF (SEQ ID NO: 77) HSYFDYSTOP (SEQ ID NO: 78)	RcaI---BamHI <i>ruDPPV</i> (58-2178) EcoRI---XbaI	pPICZaA EcoRI---XbaI

\*插入语表示限制位点序列编码的氨基酸，并被重组酶加上了 N 末端点。

<sup>†</sup>插入语的数字表示在 LAP 和 DPPcDNAs 中存在的核苷酸位置。

<sup>§</sup>*fulAP2* 和 *ruDPPIV* PCR 片段与 *E. coli-P. pastoris* 穿梭载体首尾相连。

[0349] 重组 LAP 的纯化

[0350] 使用 Amicon cell 和 Ultracel Amicon YM30 膜 (30kDa 为分离点) (Millipore,

Volketswil, Switzerland) 用超滤的方法浓缩 400ml *P. pastoris* 培养物上清液中的分泌蛋白。先用 50mM pH7.5 的 Tris-HCl 洗浓缩物,用相同的缓冲液平衡 Mono Q-Sepharose 柱 (Amersham Pharmacia,Switzerland)。用 50mM pH7.5 的 Tris-HCl 洗脱后,用 0-0.5M 线性梯度的 NaCl 以 1ml/min 的流速洗脱。从 Mono Q-Sepharose 柱洗下的不同组分被用于酶活性筛选。用亮氨酸-7-氨基-4-甲基香豆素 (Leu-AMC) 作为底物。含有 LAP 的组分被汇聚在一起。用含有 Ultracel Amincon YM30 膜的超滤 Amicon cell 浓缩后,用 20ml pH6.0 的 Tris-HCl 洗脱,LAP 提取物上样到排阻的 superose 6 FPLP (Amersham Pharmacia),用 20ml Tris-HClpH6.0 的作为洗脱液以 0.2ml/min 的流速淋洗。收集洗脱下来的活性组分。在做进一步的功能确认之前,在 4℃ 条件下,用 30kDa 的分离点 (Millipore) 的 Centricon 浓缩器浓缩 LAP 酶至终体积 0.4-1.0ml。

[0351] 在一个替代的纯化方案中,每一步纯化都在 4℃ 下操作。将 400ml *P. pastoris* 培养物上清液中的分泌蛋白用 Amicon cell 和 Ultracel Amicon YM30 膜 (30kDa 的分离点) (Millipore, Volketswil, Switzerland) 超滤的方法浓缩。用 100ml 20mM pH6.0 的醋酸钾洗浓缩物,用相同的缓冲液平衡 Mono Q-Sepharose 柱 (Amersham Pharmacia, Switzerland)。用 20mM pH6.0 的 Tris-HCl 缓冲液洗脱柱子后,用 0-0.2M 线性梯度的 NaCl 以 1ml/min 流速,以超过 142 分钟时间洗脱酶。从 Mono Q-Sepharose 柱上洗脱下的不同组分用 Leu-AMC 作为底物来筛选酶活性,含有 LAP 的组分被汇集在了一起。在带有 UltracelAmicon YM30 膜的 Amicon cell 超滤浓缩后,用 PBS 缓冲液洗脱。LAP 提取物被上样到一个排除大小的 superdex 200 FPLC 柱 (Amersham Pharmacia)。用 20mM pH6.0 的醋酸钠缓冲液以 0.2ml/min 的流速洗脱。洗脱得到的有活性的片段被收集到一起。在 4℃ 下在 30kDa 分离点的 Centricon 浓缩器中将 LAP 浓缩至 0.4-1.0ml 以供进一步的特征分析。

[0352] 在 Mono Q 中含有 ruLAP2 活性的组分洗脱在 30-40 分钟 (大约 50mM NaCl),在 Superdax200 中是在 65-70 分钟 (峰 3)。然而,在用 1M NaCl 洗脱时,大量 LAP2 活性没有保留下来。因此,这个组分在用 20mM 醋酸钠脱盐后样品被加到相同的 MonoQ 柱,用更宽的浓度范围 (0-1M) 的 NaCl 以 0.5/min 的速度洗脱超过 142 分钟。第一个活性峰出现在洗脱 7-15 分钟时和 70-140mM NaCl 对应,第二个峰出现在 150-250mM NaCl 洗脱时 (具有更强的酶活性)。在 Superdex 上,洗脱 78-80 分钟 70-140mM NaCl 的组分和它上面提到的峰一起被收集。在 Superdex 上,150-250mM NaCl 的组分提供了两个有活性的洗脱组分,分别在 44-49 分钟 (峰 1) 和 50-63 分钟 (峰 2)。

[0353] 蛋白质提取分析

[0354] 使用 12% 聚丙烯酰胺分离胶,采用 SDS-PAGE 方法分析蛋白质提取物。用考马斯亮蓝 R-250 (Bio-Rad) 对胶进行染色。N-糖苷酶 F 的消化按先前描述的方法进行 (Doumas et al., Appl. Environ. Microbiol. 64 :4809-4815 (1998))

[0355] Western 杂交

[0356] 膜先用红-丽春红染色,用针标记出主要蛋白带的位置。用兔抗血清和碱性磷酸酶耦联的山羊抗-兔 IgG (Bio-Rad) 或结合有过氧化物酶的山羊抗-兔 IgG (Amersham Pharmacia, Switzerland) 作为二次标记抗体。针对 ruLAP1、ruLAP2、fungallysin 家族的 *A. Oryzae* 分泌的碱性磷酸酶 (ALP) 和 *A. oryzae* 分泌的中性蛋白酶 (NPI) 的兔抗血清由 Eurogentec (Liege, Belgium) 用纯化的重组酶制造。

## [0357] 氨肽酶活性检测

[0358] 用不同多肽的 aminoacyl-4-methylcoumaryl-7-amide 荧光衍生物来确定氨肽酶活性,在内部淬灭 Lys (Abz)-pro-pro-Pna 荧光底物特异性确定氨肽酶 P 的活性 (Stockel et al, Adv Exp. Med. Biol. 421 :31-35(1997)。所有的底物都购自 Bachem(Bubendorf, Switzerland)。根据制造商的建议,底物被制备为 0.1M 贮存液,在 -20℃ 保存。反应混合物包括:在 25ul 50mM 缓冲液中包含浓度为 5mM 的底物和酶制备物(根据每种酶对底物的水解活性,每次检测的用量在 56-2662ng),调节 pH 值到每种 LAP 的最适值。37℃ 温育 60 分钟后,加入 5ul 冰醋酸中止反应,用 3.5ml 稀释反应混合物。用荧光光度计 (Spectrofluorophotometer) (PE,LS-5 fluorometer) 检测释放的 AMC,激发光波长 370nm,发射波长 460nm。根据用合成的 AMC 作的标准曲线被用来估计释放的 AMC 的量。在 310nm 激发光和 410nm 发射光下,测定释放的 diprolyl-p-nitroanilide。LA 活性以释放的 AMC 的 nmol 数或 pNA/分钟 /ug 蛋白来表示。

[0359] 表 15

[0360]

底物	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
Leu-AMC	100.0	100.0	100.0	100.0	100.0
IIE-AMC	6.4	1.8	7.4	13.2	6.3
Val-AMC	4.8	0.8	4.9	27.6	4.0
Ala-AMC	33.3	11.7	5.2	4.7	584.7
Gly-AMC	3.3	2.2	5.1	0.8	74.8
Ser-AMC	26.1	10.3	5.9	10.3	24.6
Thr-AMC	0.9	0.1	1.7	5.1	4.4
Cys-AMC	14.9	2.1	18.5	5.0	35.5
Met-AMC	119.7	89.5	41.3	116.9	46.1
Asn-AMC	114.6	73.5	6.8	29.4	33.9
Gln-AMC	49.9	37.0	2.3	44.9	50.7
Asp-AMC	3.8	0.3	0.0	0.8	0.9
Glu-AMC	3.7	1.1	0.0	0.0	4.7
Lys-AMC	4.6	2.3	9.1	7.7	70.1



底物	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
Arg-AMC	1.9	2.3	12.3	53.9	174.8
His-AMC	0.6	1.9	0.1	0.8	17.6
Phe-AMC	17.1	8.9	4.6	163.7	184.4
Pro-AMC	21.4	7.4	1.4	12.0	7.9
Hyp-AMC	14.2	13.3	0.3	3.9	1.7
Gly-Pro-AMC	7.2	74.1	0.0	5.4	16.7
Pyr-AMC	0.0	0.0	0.0	0.0	0.0
Lys (Abz) Pro-PropNA	0.0	0.0	0.0	0.0	0.0

[0361] 各种化学试剂对 LAP 的影响

[0362] 抑制剂和金属阳离子在 37℃ 下和酶提前温育 15 分钟, 然后, 加入终浓度为 5mM 的 Leu-AMC。再温育 60 分钟后, 依前面描述的方法测定酶活性。试验纯化的 LAP 的抑制剂和它们的浓度分别是: 500ul amastatin (Bachem)、40um 苯甲脒 (Sigma)、500ul bestatin (Bachem)、100um E-64 (L-trans-epoxysuccinyl-leu-4-guanidinobutylamide) (Bachem)、100um leupeptin (Sigma)、40um PMSF (Sigma)、20um TLCK (RocheDiagnostics) 和 20um TPCK (Roche Diagnostics)。0.5mM 和 1mM 的 CaCl<sub>2</sub>、MgCl<sub>2</sub>、CoCl<sub>2</sub>、ZnCl<sub>2</sub>、NiCl<sub>2</sub> 和 CuCl<sub>2</sub> 被用于试验。

[0363] 表 16 详细的描述了不同蛋白酶抑制剂存在下不同的 EXOXs 水解活性, 以 Leu-MCA 作为 LAP 的底物。活性以其与没有抑制剂时的活性百分比来表示。

[0364] 表 16

[0365]

抑制剂	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
乙二醇四乙酸 5mM	5	50	0	16	99
乙二醇四乙酸 1mM	7	77	7	19	68
邻二氮杂菲 5mM	0	0	0	0	0
邻二氮杂菲 1mM	0	0	0	0	0
抑氨肽酶 B500Pu	55	88	0	11	24

抑制剂	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
抑氨肽酶肽 500M	0	0	0	17	0
对氯汞苯甲酸 500M	21	96	32	90	59
E64100uM	34	71	103	190	93
亮抑酶肽 100M	113	61	233	149	86
胃酶抑素 100M	45	73	160	14	64
苯甲基磺酰氟 40uM	79	84	78	156	58
苯甲脒 40M	89	91	85	77	75
甲苯磺酰 - 左旋赖氨酸 - 氯 甲基酮盐酸盐 20M	96	120	68	80	113
甲苯磺酰苯丙氨酰氯甲 酮 20M	79	87	68	95	108

[0366] 表 17 详细的描述了不同阳离子存在下不同的 EXOXs 水解活性, 以 Leu-MCA 作为 LAP 的底物。活性以其与没有抑制剂时的活性百分比来表示。

[0367] 表 17

[0368]

	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
CaCl <sub>2</sub> 0.5mM	126.6	110.0	151.7	54.9	177.4
CaCl <sub>2</sub> 1mM	141.9	165.4	175.6	43.3	161.8
MgCl <sub>2</sub> 0.5mM	121.2	97.6	129.9	68.5	130.1
MgCl <sub>2</sub> 1mM	110.2	108.0	132.6	72.6	146.1
MnCl <sub>2</sub> 0.5mM	77.5	84.3	120.7	25.9	157.6
MnCl <sub>2</sub> 1mM	86.8	140.2	105.2	28.4	165.8
CoCl <sub>2</sub> 0.5mM	591.2	378.0	210.2	104.3	876.1

	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pkLAP
CoCl <sub>2</sub> 1mM	789.7	662.7	202.1	96.5	899.8
ZnCl <sub>2</sub> 0.5mM	77.9	51.4	43.0	60.7	437.6
ZnCl <sub>2</sub> 1mM	88.9	119.5	68.9	53.2	297.9
NiCl <sub>2</sub> 0.5mM	130.5	98.4	74.8	51.7	1187.7
NiCl <sub>2</sub> 1mM	147.9	149.3	58.1	37.2	1158.7
CuCl <sub>2</sub> 0.5mM	50.9	68.9	40.1	25.8	1422.0
CuCl <sub>2</sub> 1mM	34.7	73.6	13.7	17.0	1092.4

[0369] EXOX 活性的最适 pH 值

[0370] 采用 Ellis 和 Morrison 缓冲系统来确定酶活性的最适 pH 值 (Ellis & Morrison, Methods Enzymol. 87 :405-426(1982))。缓冲液中含有具有不同 pKa 值的三种成分,在试验的整个 pH 值范围内,缓冲液的离子强度保持恒定。缓冲液的 pH 值可以用 1M HCl 或 1M NaOH 以半个 pH 值单位在 6 到 11 的范围内进行调节。基于 Leu-AMC 底物的活性检测条件,除了 Tris/HCl 缓冲液被一定 pH 值的 Ellis 和 Morrison 缓冲液取代以外,其他条件和前面描述的一样。

[0371] 表 18. 详述了内源和重组的红色毛癣菌和烟曲霉菌氨肽酶的特征

[0372]

表 18

基因	基因长度 (nt)	内含子数	蛋白前体 (aa)	信号 (aa)	成熟域 (aa)	成熟多肽链的分子重量 (kDa)	内源/重组酶化后的分子重量 (kDa)	去糖基化后酶分子重量 (kDa)	可能的糖基化位点的数目	经计算的 pI 值	重组蛋白的产量 (ug/ml)*	GenBank 序列号
<i>ruLAP1</i>	1256	2	373	19	354	38,804	31-33/38-40	38-40	3	6.39(6.23)	40	AY496930
<i>fuLAP1</i>	1298	2	388	17	371	41,465	§NI/40	40	3	5.67(5.67)	80	AY436356
<i>ruLAP2</i>	1757	3	495	18	477	51,487	58/58-65	52	4	7.32(6.94)	40	AY496929
<i>fuLAP2</i>	1557	1	498	15	383	52,270	§NI/75-100	52	6	5.57(5.46)	100	AY436357
<i>ruDPPIV</i>	2326	0	775	15	760	86,610	90/90	84	4	(8.05)	10	AY497021

§NI: 表示未决定

\*括号中数字对应不含前序列的全长多肽

[0373] EXOX 活性的最适温度

[0374] 每份 LAP 和 Leu-AMC (5mM) 在 20、30、40、50、60、70 和 80℃ 的条件下分别孵育 10、

30 和 60 分钟,通过测试酶活性和它们最适的 pH 来确定最适的温度条件。

[0375] 蛋白水解检验

[0376] 使用 resorufin 标记的酪蛋白在磷酸缓冲液 (20mM, pH7.4) 中测定蛋白水解活性。反应混合物总体积约 0.5ml, 其中包含 0.002% 的底物。37°C 温育后, 加入三氯乙酸 (终浓度为 4%) 沉淀未被消化的底物, 通过离心分离上清液。加入 500ul Tris 缓冲液 (500mM, pH9.4) 碱化后, 测定上清液在 574nm 下的吸光度。为了实用方便, 一个蛋白水解活性单位 (U) 被定义为每分钟产生 0.001 吸光度值的酶量。

[0377] 实施例 2: 红色毛癣菌分泌蛋白的水解酶活性

[0378] 在含有 0.2% 大豆蛋白作为唯一氮源和碳源的培养基上在 30°C 的条件下培养红色毛癣菌。生长 14 天后, 记录到培养基伴有澄清现象, 用 resorufin 标记的酪蛋白作为底物检测真实的蛋白水解酶活性 (400U/ml)。其蛋白水解活性分别被 PMSF 和邻-二氮杂菲抑制 15% 和 85%, 证明红色毛癣菌分泌丝氨酸和金属蛋白酶。对培养物上清液的 Western 杂交分析揭示红色毛癣菌和犬小孢子菌相像, 分泌枯草杆菌蛋白酶 (subtilisin) 家族 (MEROPS > S8) 和 fungalyisin 家族 (MEROPS > M36) 的内源蛋白酶, 类似于 *A. oryzae* 分泌的碱性磷酸酶 ALP 和中性的金属蛋白酶 NPI (参见图 1)。另外, 在红色毛癣菌培养物的上清液中检测到对象 Leu-AMC 和 Leu-pNA 这样底物的高活性。

[0379] 实施例 3: 红色毛癣菌分泌物的氨肽酶活性

[0380] 犬小孢子菌内蛋白酶基因的核苷酸序列和同源基因的相似性为 50-70%。这些基因编码 *A. oryzae* 和 *A. fumigatus* 分泌的枯草杆菌蛋白酶和 fungalyisin。另外, 犬小孢子菌和 *Aspergillus* 的基因具有内含子-外显子结构的线性对应关系。因此, 来自 *A. oryzae* 和酵母编码氨肽酶, 基因的 DNA 序列被用来设计探针, 用于筛选红色毛癣菌的基因组 DNA 文库。使用重组蛋白, 通过与机会致病菌烟曲霉菌分泌的蛋白进行比较。确定红色毛癣菌分泌的氨肽酶的特性。

[0381] 实施例 4: 克隆编码红色毛癣菌和烟曲霉菌氨肽酶的基因

[0382] 表 19A 和 19B 详述了各种 LAP 的两两比较结果

[0383] 表 19A

[0384]

M28E 酶	%相似性或一致性 <sup>a</sup>			
	ruLAP1	fuLAP1	或 LAP1	Vibrio LAP
ruLAP1		72	72	41
fuLAP1	50		70	39
或 LAP1	48	49		42
Vibrio LAP	22	21	23	

[0385] 表 19B

[0386]

M28E 酶	%相似性或一致性 <sup>a</sup>			
	ruLAP2	fuLAP2	或 LAP2	<i>S. cer. aaY</i>
ruLAP2		69	71	53
fuLAP2	51		85	52

[0387]

或 LAP2	49	72		53
<i>S. cer. aaY</i>	32	33	34	

[0388] <sup>a</sup> 相似性百分比 (上边右手处) 和一致性百分比 (下边左手处) 的比值可以通过 Gapimplemented in the GCG package of the Genetics Computer Group, University of Wisconsin, Madison 程序获得。

[0389] 图 14 是 M28E 亚家族氨肽酶推导氨基酸序列的对比结果。用下划线表示推定的信号序列加工位点。实心三角形显示为 ruLAP1 推定的 KR 加工位点。Open 箭头显示的是 *S. griseus* 的氨肽酶以及在其他 LAP 中的保守的两个 Zn<sup>++</sup> 结合位点。使用威斯康星大学 GCG 软件包中的经 Boxshade 3.2 改良的对比。AbispLAP1 是指 *Agaricus bisporus* 的 LAP。

[0390] 图 15 是 M28A 亚家族氨肽酶推导氨基酸序列的对比结果。下划线表示推定的信号序列加工位点。用空心三角表示。两个氨基酸残基, His 和 Asp, 在真菌的 LAP 中保守, 在 *S. griseus* 氨肽酶中结合第一个 Zn<sup>++</sup>。实心菱形表示结合第二个 Zn<sup>++</sup> 离子的另外两个残基——His 和 Glu, 空心箭头表示连接两个 Zn<sup>++</sup> 离子的 ASP 残基。\* 代表仅在 ruLAP 中发现的甲硫氨酸残基。

[0391] 氨基酸序列 GPGINDDGSG (SEQ ID NO :36) 和 DM(Q/M)ASPN (SEQ ID NO :37) 被发现存在于 *A. oryzae* 分泌的 52kDa 氨肽酶 (美国专利 :6127161) 和酵母氨肽酶中。(Nishizawa, et. al. 1994) 来自这些数据, 两段大致上一致的寡核苷酸 (GGXATXAAYGAYGGXTCXGG, SEQ ID NO :38) 和 (TTXGGXGAXGCXATCATRTC, SEQ ID NO :39) 被分别用来作为正义和反义引物来扩增红色毛癣菌的 DNA。获得一段 220bp 的 PCR 片段被用来作为探针筛选  $\lambda$  噬菌体 EMBL 3 基因组 DNA 文库, 找到一条编码可能的氨肽酶 (ruLAP2) 的核苷酸序列。从烟曲霉菌基因组序列中找到一条编码相似分泌型氨肽酶 (fuLAP2) 的核苷酸序列 (参见网站 www. TIGR. com)。

[0392] 使用寡核苷酸 GCATTCCTGUGATGCCCGGGCCG (正义) (SEQ ID NO :40) 和 TTACTIONGCAAGCTCAGTGACGAAGCCGAC (反义) (SEQ ID NO :41), 通过 PCR 扩增 *A. oryzae* 基因组 DNA, 得到一段 1200bp 的片段, 该片段包含编码 *A. oryzae* 31kDa 氨肽酶 (美国专利 5994113) 基因的核苷酸序列。这个片段被用作探针二次筛选红色毛癣菌基因组 DNA 文库。在红色毛癣菌基因组文库的  $\lambda$  噬菌体 EMBL3DNA 中, 发现了一段核苷酸序列 (EMBL), 它和那些编码 *A. oryzae* 30kDa 氨肽酶和来自烟曲霉菌基因组序列的其他分泌型氨肽酶 (网站 : www. TIGR. com) 相似。这些红色毛癣菌和烟曲霉菌推定的氨肽酶被分别称为 ruLAP1 和 fuLAP1。

[0393] 鉴定的 ruLAP1、ruLAP2、fuLAP1 和 fuLAP2 每个都包含一个 17 ~ 20 氨基酸的信号序列。用 5' - 正向和 3' - 反向引物 PCR 扩增基因组 DNA (参见表 14) 和包含 10<sup>8</sup> 个克隆的红色毛癣菌或烟曲霉菌 -cDNA 文库的总 DNA, 对 PCR 产物进行测序, 确定红色毛癣菌和烟曲霉菌基因的内含子 - 外显子结构。RuLAP2 的 3 个内含子的第一个在位置上和 fuLAP2 唯一的内含子在位置上相似 (参见表 13)。基因 ruLAP1 和 fuLAP1 具有两个内含子和三个外显子, 有相似的共线性结构。

[0394] 实施例 5 : 重组红色毛癣菌和烟曲霉菌氨肽酶的产生

[0395] 由 RT-PCR 获得的红色毛癣菌和烟曲霉菌的 cDNA 被克隆进 pKJ 113, 在 *P. pastoris* 中表达。取决于肽酶的产量, 在 Leu-AMC 上得到大约 10-80ug/ml 有活性的酶 (参见表 18)。在相同的培养条件下, 野生型的 *P. pastoris* 不向培养基中分泌任何有亮氨酸氨肽酶活性的物质。SDS-PAGE 分析 *P. pastoris* 转化株分泌的重组 ruLAP2、fuLAP1 和 fuLAP2 呈现弥散的带 (图 2)。N-糖苷酶处理后, 在胶上只有快迁移率的主带出现, 说明和 ruLAP1 相反, 这三个 LAP 是糖蛋白 (图 2) 去糖基化的重组 LAP 的表现分子量和根据编码蛋白酶基因的核苷酸序列推导出的多肽链计算出的分子量接近。表 18 提供了每个重组酶推导出的初级结构 (氨基酸序列)。

[0396] 实施例 6:检测红色毛癣菌培养物上清液中的 ruLAP1 和 ruLAP2

[0397] 使用抗 ruLAP1 的抗血清, 在红色毛癣菌培养物上清液中检测到 LAP1 产物的电泳迁移率高于重组 ruLAP1 的。

[0398] 使用抗 ruLAP2 抗血清, 对红色毛癣菌培养物上清液的 Western 杂交分析结果显示红色毛癣菌分泌的糖基化的 LAP2 和来自 *P. pastoris* (的重组酶有相同的电泳迁移率)。

[0399] 实施例 7:重组 LAP 的特性

[0400] 氨肽酶 ruLAP1、ruLAP2、fuLAP1 以及猪肾微粒体氨肽酶 (pkLAP) 都能有效地水解 Leu-AMC。这种底物被用来确定活性最适的温度和 pH 值, 通过测定 (i) 各种已知的肽酶抑制剂 (参见表 16) 和 (ii) 不同的二价离子 (参见表 17) 对酶活性的影响来进一步了解它们的特性。每种 LAP 在 20°C 时都有水解 Leu-AMC 的能力, 最适的温度范围从 40°C ~ 50°C。最适的 pH 在 7.0 到 8.5 之间 (参见表 18)。80°C 预处理 10 分钟可使酶不可逆地完全失活。

[0401] 供试的氨肽酶能被浓度 500uM 的 amastatin (参见表 16) 强烈或完全抑制。RuLAP1、fuLAP1 和 pkLAP 也被 bestatin 抑制, 但这种抑制剂仅对 ruLAP2 和 fuLAP2 有部分抑制效应。在螯合剂试验中, 邻二氮杂菲在浓度为 1mM 和 5mM 时完全抑制这 5 种酶。fuLAP1、ruLAP2 和 ruLAP1 比其他 LAP 对 EDTA 更敏感。E64 和 Chloromercuribenzoate (半胱氨酸蛋白酶抑制剂) 使 ruLAP2 的活性减弱。暗示在这种酶的氨基酸序列上存在对于活性关键的巯基残基。对于供试的所有 LAP, Leupeptin (丝氨酸 / 半胱氨酸蛋白酶抑制剂)、PMSF (丝氨酸蛋白酶抑制剂)、苯甲脒、TLCK 和 TPCK 没有明显的抑制效果。令人吃惊的是, fuLAP1 和 ruLAP1 对 0.1mM pepstatin (天冬氨酸蛋白酶抑制剂) 表现有一些灵敏性。

[0402] 除了 fuLAP1 表现对二价离子的一般性敏感, 当浓度达到 1mM 时,  $\text{Co}^{++}$  可以使 LAP 活性增加 200% 到 900%, 四种真菌的 LAP 对二价阳离子表现不同的灵敏性。例如 fuLAP2 被  $\text{Mn}^{++}$  和  $\text{Ca}^{++}$  激活, 而 fuLAP1 被相同的离子抑制。不同 M28 家族的 4 种真菌 LAP1, 微粒体的 pkLAP 被 Zn、Ni 和  $\text{Cu}^{++}$  高度活化。

[0403] 用 Leu-AMC 作为参照, 对这些酶对于不同的 aminoacyl-AMC 的水解活性进行了比较 (参见表 15)。根据供试的氨肽酶, 检测不同氨酰残基的不同参数选择。例如, 氨肽酶 pkLAP 和四种真菌的 LAP 不同之处在于它对 Ala-AMC 和 Arg-AMC 有特别高的效率。RuLAP1 明显对 Leu-AMC 有最强的选择性。然而, 观察到和 ruLAP2、fuLAP1 和 fuLAP2 有关的其他一些优先的水解活性。例如, ser- 和 pro-AMC 被 ruLAP2 更有效水解, 而 fuLAP1 喜爱 Asp- 和 Glu-AMC。因为它们不能切开 Lys (Abz)-pro-pro-pNA, 这些酶都不表现氨肽酶 P 的活性。

[0404] 实施例 8:ruLAP2 和 ruDPPIV 在消化醇溶蛋白多肽中的应用

[0405] Celiac 病 (CD) 是一种消化疾病, 它能破坏小肠, 干扰从食物中吸收营养。患有

celiac 病的人无法忍受一种叫做麸质 (gluten) 的蛋白质,这种蛋白质存在于小麦、黑麦和大麦中。当患 CD 的病人摄入含有麸质的食物时,他们的免疫系统产生反应,对小肠造成破坏。在大部分世界人群中,这种疾病的发病率大约是 1/200,治疗 CD 的唯一方法是终生都绝不食用含有麸质的膳食。对于大部分患者而言,遵循这样的饮食将会使症状消失,已有的肠破坏得到恢复,并且防止进一步恶化。

[0406] 小麦麸质的主要毒性成分是一种叫做醇溶蛋白 (gliadin) 的富含脯氨酸和 Gln 的蛋白家族,它们在消化道中不被降解并且饱含许多 T- 细胞刺激抗原决定部位。因为不同的体外系统已被用来进行这些研究,关于有效诱导 HLA-DQ2 阳性 gut- 来源和外周 T 细胞和免疫活性的抗原决定部位仍然存在一些争议 (Vader et al., Gastroenterology 122 : 1729-1737 (2002))。在一种 CD 器官培养模板中,醇溶蛋白产生毒性的能力和刺激 T- 细胞并不一致,反之亦然。(McAdam & Sollid, Gut 47 :743-745 (2000))。而且,许多麸质抗原决定部位同 HLA-DQ2 和 HLA-DQ8 结合,但这并不是所有的被脱酰胺作用增强。通过小肠组织转谷氨酰胺酶的作用,某些谷氨酰胺残基转化为谷氨酸。因为它们有可能的刺激 T- 细胞的能力。Molberg et al., Nat. Med. 4 :713-717 (1998)。但是,对于 T- 细胞激活,脱酰胺作用并不是绝对必要的。(Arentz-Hansen et al., Gastroenterology 123 :803-809 (2002))。

[0407] 其他治疗或预防 CD 的措施,最好的方式找到一种替代的“不含麸质”食品,在过去几年里已提出建议,包括用化合物抑制 T- 细胞激活来阻止多肽和 HLA-DQ2 的结合,抑制组织转谷氨酰胺酶来阻止麸质脱酰胺 (Sollid, Nat. Rev. Immunol. 2 :647-655 (2002)) 和口腔肽酶的补充。后面的方法被认为对于免疫抑制多肽消化有根本性的作用,具体是含有细菌脯氨酰胞内肽酶,该酶对于含有脯氨酸的多肽有很宽范围的耐受性 (Shan 等, Science 297 : 2275-2279 (2002) ;Hausch et al., Am. J. Physiol. Gastrointest Liver Physiol. 283 : G996-G1003 (2002))。相对大片段的醇溶蛋白能抵抗消化酶的降解,而且,这种肽是来自被麸质刺激的 CD 病人肠组织不同的 HLA-DQ2- 限制性 T- 细胞克隆的潜在的刺激物。每个这样的 T- 细胞克隆识别不同的 33mer 抗原决定部位。脯氨酰胞内肽酶,偏爱 Pro-Xaa-Pro 结构域,能够切开 33mer 的醇溶蛋白肽,和刷状缘的氨肽酶的协调效应,迅速减少这些肽的 T- 细胞刺激潜力。

[0408] 尽管在大麦和黑麦有稳定的同源物,这些已经描述的抵抗消化道降解的 gluten 肽结构域在我们的案例中作为不同 LAP 的模式底物,单独或同 ruDPPIV 联合一起消化 :对应于  $\alpha/\beta$  醇溶蛋白 AIV (P04724) 82-95 片段 PQPQLPYPQPQLPY (SEQ ID NO :42) (14mer) 或对应于醇溶蛋白 MM1 (P18573) 57-89 片段的 LQLQFPQPQLPYPQPQLPYPQPQLPYPQPQP (SEQ IDNO : 43) (33mer)。

[0409] 33mer 的 N- 端乙酰化的形式 (Ac-33mer) 也被合成作为酶切试验的对照。通过一种杂质酶用外肽酶去干扰内肽酶的酶切。

[0410] 被用于评估的酶有 :ruLAP1 (Trichophyton rubrum 的氨肽酶 I)、ruLAP2 (Trichophyton rubrum 的氨肽酶 II)、或 LAP2 (Aspergillus orizae 的氨肽酶 II)、fuLAP2 (Aspergillus fumigatus 的氨肽酶 II)、MicpkLAP (来自猪肾的微粒体亮氨酸氨肽酶, Sigma)、CytphkLAP (来自猪肾的胞质亮氨酸氨肽酶, Sigma) 和 ruDPPIV。

[0411] 多肽的合成 :

[0412] 采用固相合成的方法。在 Applied Biosystems 的 custom-modified 430A



型态合成仪上进行,用原位 neutralization/2-(1H-benzotriazol-1-yl)-1,1,1,3,3-tetramethyluronium hexa fluoro-phosphate(HBTU)activation 试剂盒,在 standard -O-CH<sub>2</sub>-phenylacetamidomethyl resin 逐步的进行 Boc 化学链延长 (Schulzer 等, Int. J. Peptide Protein Res. 40 :180-193(1992))。

[0413] 在合成结束时,用无水 HF 在 0℃处理 1 小时,用 5%对 - 甲酚作为净化剂,肽被去保护并被从树脂上切下。切下后,用冰冻的乙醚沉淀合成的肽,溶解在乙腈水溶液中,冻干。用 RP-HPLC 的方法纯化肽,使用来自 Waters 的 C18 柱,线性梯度的缓冲 B(99%乙腈/10% H<sub>2</sub>O/0.1%三氟乙酸),在缓冲液 A(H<sub>2</sub>O/0.1%三氟乙酸),214nm 下紫外检测。使用 PlatformII 设备 (Micromass, Manchester, England) 用 electrospray 质谱对样品进行分析。

[0414] 降解反应的条件:

[0415] 温育在 37℃, PH7.2 的 50mM Tris-HCl 中进行,1mM CoCl<sub>2</sub>,底物浓度 1mg/ml, E/S 比例为 1 : 20。用乙酸酸化中止反应,在 C8z 柱上用 RP-HPLC 进行中点分析,在 0.1% TFA 中用 2% / 分钟乙腈梯度。所有的峰由 ESI-MS 确认。

[0416] 14mer 的消化:

[0417] 如图 6 所示,在 4 小时内 14mer 没有被 ruLAP2 消化。和对照相比,HPLC 谱没有变化。实际上,消化仅仅切除了 N- 末端的脯氨酸。另一方面,补充 ruDPPIV 导致完全降解成氨基酸和二肽,ruDPPIV 自身不能水解肽。(图 7)

[0418] 33mer 的消化:

[0419] 单独用 ruLAP2 消化 33mer 导致在 4 小时内肽的部分降解(少于 50%)。这个肽不是 ruDPPIV 的底物(图 8)。然而,当两种酶被混合在一起,33mer 被完全消化(图 9)成氨基酸和二肽,其中的一些能够被 ESI-MS 鉴定(Y、L、F、P、PY 和 PF)。

[0420] 当 ruDPPIV 和 ruLAP2 或 fuLAP2 混合时,得到相同的 HPLC 波形。然而,和 ruLAP1 混和时,一些高分子量的成分仍然存在,少于 10%是初始的底物。

[0421] 另一方面,和猪肾微粒体氨肽酶一起温育导致仅仅部分 N- 末端 Leu 和 C- 末端 phe 被部分除去(由于 Carboxypeptidasic 污染),添加 DPPIV 并不改变 HPLC 图谱。胞质的猪肾氨肽酶对 33mer 完全没有活性。

[0422] 在 LAP 或 DPPIV 单独的或混合在一起的消化实验中,Ac- 醇溶蛋白 33mer 的稳定性进一步证实自由的氨基基团是醇溶蛋白 33mer 被这些外肽酶完全降解所必须的。

[0423] 用其他酶消化:

[0424] 用链霉菌蛋白酶 (E/S = 1/25) 消化超过 20 小时仅部分(少于 40%)水解,加入 ruLAP2(两种酶都以 E/S = 1/50 的比率)没有提高水解性。另一方面,在相同条件下加入 DPPIV 导致肽的完全降解,因为氨肽酶和二肽肽酶的协同作用。Chtmotrypsin 单独或和 ruLAP 或 DPPIV 混和不能降解肽。

[0425] 实施例 9 :ruLAP2 在表达融合其他蛋白的重组蛋白的过程中或添加 N- 末端标签中的应用

[0426] LAP2 被鉴定酶切 proNPY N 末端的 Gly-Ser 以及同一个肽的 N- 末端另外一个 Ala。为了增加 LAP2 单独或 / 和其他外肽酶结合时在加工较大重组蛋白过程中应用,一个有 N- 末端 Met-Thr-Pro 序列的 G-CSF 重组蛋白 (Cys<sup>17</sup> → Ser, Lys<sup>16,23,34</sup>, 40 → Arg) 被依次

和 ruLAP2 和 ruDPPIV 温育,从有 175 个残基的蛋白上依次除去 met 和 Thr-Pro 二肽。

[0427] ruLAP2 消化 Gly-Ser-ProNPY :

[0428] 1mg/ml 肽加入 50mM Tris-HCl, 1mM CoCl<sub>2</sub> 缓冲液中和 ruLAP2 以 E/S 比率以 1 : 20 和 1 : 100 (W : W) 混合, 37°C 温育过夜。用 RP-HPLC 分离酶切材料, 用 ESI-MS 鉴定。如图 10 所示, 和 ruLAP2 温育结果是切去 N- 端的两个残基, 理论上损失 144. 1amu (发现的 144. 2) E/S 比率为 1 : 100 得到相同的结果。当酶到达 Xaa-Pro 结构域时酶切停止, 在 proNPY 的结构域是 Tyr-Pro。

[0429] ruLAP2 消化 Ala-ProNPY :

[0430] 消化条件和 Gly-Ser-ProNPY 的相同。图 11B 显示 N- 末端的丙氨酸差不多被全部从 ProNPY 中除去 (分子量损失 71amu)。

[0431] 从 C-CSF 的 N- 末端依次切去甲硫氨酸和 Thr-Pro :

[0432] 被用于这些试验的是名为 TG47 的一种 G-CSF 的突变类似物, 它是 methinoyl-[C17S, K16, 23, 34, 40R]G-CSF, 其重新折叠的蛋白质理论上的质量是 18894. 90。

[0433] 用 ruLAP2 消化 :

[0434] 将 G-CSF 的贮备液 (在含 0. 1% Sarcosyl PBS 中浓度为 1. 9mg/ml) 稀释 4 倍至 pH7. 2 的 50mM Tris-HCl 中, 含 1mM CoCl<sub>2</sub>。和 ruLAP2 (E/S = 1 : 20 和 1 : 100, W : W) 在 37°C 温育 15 小时。用 30% (V : V) 乙腈稀释溶液, 乙酸酸化, RP-HPLC 分离蛋白质, 进行 MS 分析。如图 12A 和 B 所示, 过夜温育导致 N- 末端甲硫氨酸完全切除。理论上质量损失 131. 2amu。在 E/S 比率为 1 : 100 时, 过夜温育后, 仍有痕量未被切开的物质存在。

[0435] 为了从半-制备的 RP-HPLC 柱上分离截短的物质, 该实验在 2mg 尺度上被重复。以 E/S 比率 1 : 25 (W : W) 在 37°C 条件下消化 15 小时以上。分离到的物质 (0. 8mg) 通过 ESI-MS 鉴定 (图 12B, desMet-G-CSF, 计算分子量为 18763. 7amu, 实测分子质量为 18762. 5)。

[0436] DPPIV 消化 desMet-G-CSF :

[0437] 冻干的材料重新溶于 pH7. 5 的含 0. 1% Sarcosyl 的 50mM Tris-HCl 中, 浓度为 1mg/ml。和 DPPIV 以 E/S 比率 1/20 (W : W) 在 37°C 温育过夜。像以前一样用 RP-HPLC 分离蛋白质, 用 ESI-MS 鉴定 (图 13A 和 13B)。DPPIV 消化 (图 13B) 结果是切除 N- 末端二肽 Thr-Pro (计算的分子量为 18, 564. 8amu, 实测分子质量为 18, 563)。在反应介质中, 痕量的未消化物质仍然存在。

[0438] 于是, 依次应用 LAP2 和 DPPIV 的结果是有效地移去重组蛋白 N- 末端序列。当酶到达一个“停止点结构域”, 如 Xaa-Pro-Xaa, 或 Xaa-Pro 结构域, 它们可能是 LAP2 特异的停止点, RuLAP2 的消化被终止, 但可被 DPPIV 依次切去。

[0439] 然而, 对 N- 末端残基切除的起始高度依赖于序列, 因为在与 LAP 和 DPPIV 一起温育时, Met (His)<sub>6</sub> 标记没有从 Met (His)<sub>6</sub>-proPNY 中除去。

[0440] 其他实施方案

[0441] 尽管以上实施例详细的描述了本发明, 但仅用来说明本发明目的, 而不是用来限制本发明的权利要求范围。特别的是, 可预测到不偏离本发明精神和范围的不同的替换、改变和修饰均属于本发明权利要求范围。本领域普通技术人员知晓的有关重组的核酸起始原料、涉及的克隆或文库类型等方面的选择均认为是常规选择。其他的方面, 如优势和修饰也被认为属于以下权利要求范围内。

[0442] seq. ST25  
 [0443] 序列表  
 [0444] <110> 富诺齐梅生物技术股份有限公司  
 [0445] <120> 一种新的真菌蛋白及其编码核酸  
 [0446] <130>PFD1060017F-CF  
 [0447] <150>US 60/498, 398  
 [0448] <151>2003-08-25  
 [0449] <160>35  
 [0450] <170>PatentIn version 3.1  
 [0451] <210>1  
 [0452] <211>1757  
 [0453] <212>DNA  
 [0454] <213> 红色毛癣菌 (*T. rubrum*)  
 [0455] <400>1  
 [0456] atgaagtcgc aactgttgag cctggctgtg gccgtcacia ccatctccca gggcgttggt 60  
 [0457] ggtcaagagc ccttcgatg gcctttcaag cctatggtca ctcaggtgag ttgctctcaa 120  
 [0458] cagatcgatc gatcgatcta cttttgtccc tgtcacatca aactccagca gagccaaaga 180  
 [0459] aacagacaca atgttcctgg ggaattctta tgggctaata taaatgtata ggatgacctg 240  
 [0460] caaaacaaga taaagctcaa ggatatcatg gcaggcgtcg agaagctgca aagcttttct 300  
 [0461] gatgctcatc ctgaaaagaa ccgagtgttt ggttgtaata gccacaagga cactgtagag 360  
 [0462] tggatctaca atgagatcaa ggccactggc tactacgatg tgaagaagca ggagcaagta 420  
 [0463] cacctgtggt ctcatgccga ggctgctctc aatgccaatg gcaaggacct caaggccagc 480  
 [0464] gccatgtcct acagccctcc tgccagcaag atcatggctg agcttgttgt tgccaagaac 540  
 [0465] aatggctgca atgctgtatg tgccatacac tttctatacg tcacattctc tctagaatga 600  
 [0466] agagcacggg agaactaact ttatgtatac agactgatta cccagcgaac actcagggca 660  
 [0467] agatcgtcct cgttgagcgt ggtgtctgca gcttcggcga gaagtctgct caggctggtg 720  
 [0468] atgcaaaggc tgctggtgcc attgtctaca acaacgtccc cggatccctt gctggcactc 780  
 [0469] ttggtggcct tgacaagegc catgtcccaa ccgctggtct ttcccaggag gatggaaaga 840  
 [0470] accttgctac cctcgttgct tctggtaaga ttgatgtcac catgaacgtt atcagtctgt 900  
 [0471] ttgagaaccg aaccacgtaa gtagctcaac ggctgatcca gcatcaattg tctcagatg 960  
 [0472] atactaaatc gatacctcat agctggaacg teattgctga gaccaaggga ggagaccaca 1020  
 [0473] acaacgttat catgctcggg gctcactccg actccgtcga tgccggccct ggtattaacg 1080  
 [0474] acaacggctc gggtccatt ggtatcatga ccgttgccaa agccctcacc aacttcaagc 1140  
 [0475] tcaacaacgc cgtccgcttt gcctggtgga ccgtgagga attcggctctc cttggaagca 1200  
 [0476] ctttctacgt caacagcctc gatgaccgtg agctgcacia ggtcaagtgg tacctcaact 1260  
 [0477] tcgacatgat cggtctctcc aacttcgcca accagatcta cgacgggtgac ggttcggcct 1320  
 [0478] acaacatgac cggccccgct ggctctgctg aaatcgagta cctgttcgag aagttctttg 1380  
 [0479] acgaccaggg tatccacac cageccactg cttcaactgg ccgatccgac tactctgctt 1440  
 [0480] tcatcaagcg caacgtgccc gctggcggcc tttcaactgg agccgaggtt gteaagacc 1500

[0481]	ccgagcaagt caagttgttc ggtggtgagg ctggcgttgc ctatgacaag aactaccatc	1560
[0482]	gcaagggcga caccgttgcc aacatcaaca agggagctat cttccttaac actcgagcca	1620
[0483]	tcgcctacgc tatcgccgag tatgcccgat ccctcaaggg attcccaacc cgcccaaaga	1680
[0484]	ccggcaagcg tgacgtcaac ccccagtatt ctaagatgcc tgggtgtggc tgcggacacc	1740
[0485]	acactgtctt catgtaa	1757
[0486]	<210>2	
[0487]	<211>1488	
[0488]	<212>DNA	
[0489]	<213> 红色毛癣菌 ( <i>T. rubrum</i> )	
[0490]	<400>2	
[0491]	atgaagtcgc aactgttgag cctggctgtg gccgtcaca ccatctccca gggcgttgtt	60
[0492]	ggtcaagagc ctttcggatg gcctttcaag cctatggtca ctcaggatga cctgcaaac	120
[0493]	aagataaagc tcaaggatat catggcaggc gtcgagaagc tgcaaagctt ttctgatgct	180
[0494]	catcctgaaa agaaccgagt gtttggtggt aatggccaca aggacactgt agagtggatc	240
[0495]	tacaatgaga tcaaggccac tggctactac gatgtgaaga agcaggagca agtacacctg	300
[0496]	tggctcatg ccgaggctgc tctcaatgcc aatggcaagg acctcaaggc cagcgccatg	360
[0497]	tcctacagcc ctcctgccag caagatcatg gctgagcttg ttgttgcaa gaacaatggc	420
[0498]	tgcaatgcta ctgattacc agcgaacact cagggcaaga tcgtcctcgt tgagcgtggt	480
[0499]	gtctgcagct tcggcgagaa gtctgctcag gctggtgatg caaaggctgc tggtgccatt	540
[0500]	gtctacaaca acgtccccg atcccttgct ggcactcttg gtggccttga caagcgccat	600
[0501]	gtcccaaccg ctggtctttc ccaggaggat ggaaagaacc ttgctaccct cgttgcttct	660
[0502]	ggtaagattg atgtcacat gaacgttate agtctgtttg agaaccgaac cacctggaac	720
[0503]	gtcattgctg agaccaaggg aggagaccac aacaacgtta tcatgctcgg tgctcactcc	780
[0504]	gactccgtcg atgccggccc tggattaaac gacaacggct cgggctccat tggatatcatg	840
[0505]	accgttgcca aagccctcac caacttcaag ctcaacaacg ccgtccgctt tgccctggtgg	900
[0506]	accgctgagg aattcggctc ctttgggaagc accttctacg tcaacagcct cgatgacctg	960
[0507]	gagctgcaca aggtcaagtt gtacctcaac ttgacatga teggctctcc caacttcgcc	1020
[0508]	aaccagatct acgacggtga cggttcggcc tacaacatga ccggccccgc tggetctgct	1080
[0509]	gaaatcgagt acctgttcga gaagttcttt gacgaccagg gtatcccaca ccageccact	1140
[0510]	gccttcaactg gccgatccga ctactctgct tcatcaagc gcaacgtgcc cgetggcggc	1200
[0511]	ctcttcaactg gagccgaggt tgtcaagacc cccgagcaag tcaagttggt cgggtgtgag	1260
[0512]	gctggcggtg cctatgacaa gaactacat cgcaagggcg acaccgttgc caacatcaac	1320
[0513]	aaggageta tcttccttaa cactcgagcc atcgctacg ctatcgccga gtatgccga	1380
[0514]	tcctcaagg gattcccac ccgcccgaag accggcaagc gtgacgtcaa ccccagtat	1440
[0515]	tctaagatgc ctggtgtgg ctgcggacac cacactgtct tcatgtaa	1488
[0516]	<210>3	
[0517]	<211>495	
[0518]	<212>PRT	
[0519]	<213> 红色毛癣菌 ( <i>T. rubrum</i> )	

[0520] <400>3  
 [0521] Met Lys Ser Gln Leu Leu Ser Leu Ala Val Ala Val Thr Thr Ile Ser  
 [0522] 1 5 10 15  
 [0523] Gln Gly Val Val Gly Gln Glu Pro Phe Gly Trp Pro Phe Lys Pro Met  
 [0524] 20 25 30  
 [0525] Val Thr Gln Asp Asp Leu Gln Asn Lys Ile Lys Leu Lys Asp Ile Met  
 [0526] 35 40 45  
 [0527] Ala Gly Val Glu Lys Leu Gln Ser Phe Ser Asp Ala His Pro Glu Lys  
 [0528] 50 55 60  
 [0529] Asn Arg Val Phe Gly Gly Asn Gly His Lys Asp Thr Val Glu Trp Ile  
 [0530] 65 70 75 80  
 [0531] Tyr Asn Glu Ile Lys Ala Thr Gly Tyr Tyr Asp Val Lys Lys Gln Glu  
 [0532] 85 90 95  
 [0533] Gln Val His Leu Trp Ser His Ala Glu Ala Ala Leu Asn Ala Asn Gly  
 [0534] 100 105 110  
 [0535] Lys Asp Leu Lys Ala Ser Ala Met Ser Tyr Ser Pro Pro Ala Ser Lys  
 [0536] 115 120 125  
 [0537] Ile Met Ala Glu Leu Val Val Ala Lys Asn Asn Gly Cys Asn Ala Thr  
 [0538] 130 135 140  
 [0539] Asp Tyr Pro Ala Asn Thr Gln Gly Lys Ile Val Leu Val Glu Arg Gly  
 [0540] 145 150 155 160  
 [0541] Val Cys Ser Phe Gly Glu Lys Ser Ala Gln Ala Gly Asp Ala Lys Ala  
 [0542] 165 170 175  
 [0543] Ala Gly Ala Ile Val Tyr Asn Asn Val Pro Gly Ser Leu Ala Gly Thr  
 [0544] 180 185 190  
 [0545] Leu Gly Gly Leu Asp Lys Arg His Val Pro Thr Ala Gly Leu Ser Gln  
 [0546] 195 200 205  
 [0547] Glu Asp Gly Lys Asn Leu Ala Thr Leu Val Ala Ser Gly Lys Ile Asp  
 [0548] 210 215 220  
 [0549] Val Thr Met Asn Val Ile Ser Leu Phe Glu Asn Arg Thr Thr Trp Asn  
 [0550] 225 230 235 240  
 [0551] Val Ile Ala Glu Thr Lys Gly Gly Asp His Asn Asn Val Ile Met Leu  
 [0552] 245 250 255  
 [0553] Gly Ala His Ser Asp Ser Val Asp Ala Gly Pro Gly Ile Asn Asp Asn  
 [0554] 260 265 270  
 [0555] Gly Ser Gly Ser Ile Gly Ile Met Thr Val Ala Lys Ala Leu Thr Asn  
 [0556] 275 280 285  
 [0557] Phe Lys Leu Asn Asn Ala Val Arg Phe Ala Trp Trp Thr Ala Glu Glu  
 [0558] 290 295 300

[0559]	Phe Gly Leu Leu Gly Ser Thr Phe Tyr Val Asn Ser Leu Asp Asp Arg			
[0560]	305	310	315	320
[0561]	Glu Leu His Lys Val Lys Leu Tyr Leu Asn Phe Asp Met Ile Gly Ser			
[0562]		325	330	335
[0563]	Pro Asn Phe Ala Asn Gln Ile Tyr Asp Gly Asp Gly Ser Ala Tyr Asn			
[0564]		340	345	350
[0565]	Met Thr Gly Pro Ala Gly Ser Ala Glu Ile Glu Tyr Leu Phe Glu Lys			
[0566]		355	360	365
[0567]	Phe Phe Asp Asp Gln Gly Ile Pro His Gln Pro Thr Ala Phe Thr Gly			
[0568]		370	375	380
[0569]	Arg Ser Asp Tyr Ser Ala Phe Ile Lys Arg Asn Val Pro Ala Gly Gly			
[0570]		385	390	395
[0571]	Leu Phe Thr Gly Ala Glu Val Val Lys Thr Pro Glu Gln Val Lys Leu			
[0572]		405	410	415
[0573]	Phe Gly Gly Glu Ala Gly Val Ala Tyr Asp Lys Asn Tyr His Arg Lys			
[0574]		420	425	430
[0575]	Gly Asp Thr Val Ala Asn Ile Asn Lys Gly Ala Ile Phe Leu Asn Thr			
[0576]		435	440	445
[0577]	Arg Ala Ile Ala Tyr Ala Ile Ala Glu Tyr Ala Arg Ser Leu Lys Gly			
[0578]		450	455	460
[0579]	Phe Pro Thr Arg Pro Lys Thr Gly Lys Arg Asp Val Asn Pro Gln Tyr			
[0580]		465	470	475
[0581]	Ser Lys Met Pro Gly Gly Gly Cys Gly His His Thr Val Phe Met			
[0582]		485	490	495
[0583]	<210>4			
[0584]	<211>1256			
[0585]	<212>DNA			
[0586]	<213> 红色毛癣菌 (T. rubrum)			
[0587]	<400>4			
[0588]	atgaagctcc tctctgttct tgcgctgagc gctaccgcta cctccgtcct cggagctagc	60		
[0589]	attcctgttg atgcccgggc cgagaagttc ctcatcgaac ttgccctgg tgagactcgc	120		
[0590]	tgggttaccg aggaggagaa gtgggagctt aagcgggtat gtaccactat cctacgcaaa	180		
[0591]	agttgtatatt ttagtagata atattggtta ttaacacca ttctagaagg gtcaagactt	240		
[0592]	ctttgacatc actgacgagg aggttggatt cactgctgct gttgcacagc cagccattgc	300		
[0593]	ctaccaacc tccatccgcc atgctaagtc tgtaacgcc atgattgcta ccctctccaa	360		
[0594]	ggagaacatg cagecgcgac tgaccaagct cagctcgttc caaacgctt actataaggt	420		
[0595]	tgactttggc aagcagtctg ccacctggct ccaggagcaa gtccaggctg ccatcaatac	480		
[0596]	cgctggtgcc aatcgtacg gageccaaggt cgccagcttc cgacacaact tegetcagca	540		
[0597]	cagcatcatt gccactatc cggcgcgctc ccctgaagtc gttgtcgtcg gtgetcacca	600		

[0598]	agacagcatc aaccaacgca gcccacatgac cggccgcgct ccaggtgccg atgacaacgg	660
[0599]	cagtggctcc gtcaccatcc ttgaggccct ccgtgggtgtt ctccgggacc agaccatcct	720
[0600]	ccagggcaag gctgccaaca ccattgagtt ccaactggtac gccggtgagg aagctggtct	780
[0601]	tctgggctcc caggccatct tcgccaacta caaacagacc ggcaagaagg tcaagggcat	840
[0602]	gctcaaccag gacatgaccg gttacatcaa gggaatggtc gacaagggtc tcaaggtgtc	900
[0603]	cttcgggtatc atcaccgaca acgtcaacgc taacttgacc aagtctgtec gcattggtcat	960
[0604]	caccaaggta agcttcaact cttgataaat atatttttca tcgatgaaat gatgtcctaa	1020
[0605]	taatgcttaa gtactgetca atcccaacca tcgacaccgc ctgceggctat gcttgetctg	1080
[0606]	accacgcctc tgccaaccgc aatggctacc catctgccaat ggttgccgag tctcccatcg	1140
[0607]	atctcctega cctcacctc cacactgact ctgacaacat tagctacctc gacttcgacc	1200
[0608]	acatgatcga gcacgctaag ctcatgtcgc gcttcgtcac tgagctcgtc aagtaa	1256
[0609]	<210>5	
[0610]	<211>1122	
[0611]	<212>DNA	
[0612]	<213> 红色毛癣菌 ( <i>T. rubrum</i> )	
[0613]	<400>5	
[0614]	atgaagctcc tctctgttct tgcgctgagc gctaccgcta cctccgtcct cggagctagc	60
[0615]	attcctgttg atgccgggc cgagaagttc ctcatcgaa ttgccctgg tgagactcgc	120
[0616]	tgggttaccg aggaggagaa gtgggagctt aagcggagg gtcaagactt ctttgacatc	180
[0617]	actgacgagg aggttgatt cactgctgct gttgcacagc cagccattgc ctaccaacc	240
[0618]	tccatccgcc atgctaatgc tgtaacgcc atgattgcta cctctccaa ggagaacatg	300
[0619]	cagcgcgac tgaccaagct cagctcgttc caaacgctt actataaggt tgactttggc	360
[0620]	aagcagtctg ccacctggct ccaggagcaa gtccaggctg ccatcaatac cgctggtgcc	420
[0621]	aatcgctacg gagccaaggt cgcagcttc cgacacaact tcgctcagca cagcatcatt	480
[0622]	gccactatc cggccgctc cctgaagtc gttgtcgtc gtgctacca agacagcatc	540
[0623]	aaccaacgca gcccacatgac cggccgcgct ccaggtgccg atgacaacgg cagtggctcc	600
[0624]	gtcaccatcc ttgaggccct ccgtgggtgtt ctccgggacc agaccatcct ccagggcaag	660
[0625]	gctgccaaca ccattgagtt ccaactggtac gccggtgagg aagctggtct tctgggctcc	720
[0626]	caggccatct tcgccaacta caaacagacc ggcaagaagg tcaagggcat gctcaaccag	780
[0627]	gacatgaccg gttacatcaa gggaatggtc gacaagggtc tcaaggtgtc cttcgggtatc	840
[0628]	atcaccgaca acgtcaacgc taacttgacc aagtctgtec gcattggtcat caccaagtac	900
[0629]	tgctcaatcc caaccatcga caccgcgtgc ggctatgctt gctctgacca cgcctctgcc	960
[0630]	aaccgcaatg gtaccatc tgccatggtt gccagctc ccatgatct cctcgacct	1020
[0631]	cacctcaca ctgactctga caacattagc tacctcgact tcgaccacat gatcgagcac	1080
[0632]	gctaagctea ttgtcggctt cgtcactgag ctgctaaagt aa	1122
[0633]	<210>6	
[0634]	<211>373	
[0635]	<212>PRT	
[0636]	<213> 红色毛癣菌 ( <i>T. rubrum</i> )	

[0637] <400>6  
 [0638] Met Lys Leu Leu Ser Val Leu Ala Leu Ser Ala Thr Ala Thr Ser Val  
 [0639] 1 5 10 15  
 [0640] Leu Gly Ala Ser Ile Pro Val Asp Ala Arg Ala Glu Lys Phe Leu Ile  
 [0641] 20 25 30  
 [0642] Glu Leu Ala Pro Gly Glu Thr Arg Trp Val Thr Glu Glu Glu Lys Trp  
 [0643] 35 40 45  
 [0644] Glu Leu Lys Arg Lys Gly Gln Asp Phe Phe Asp Ile Thr Asp Glu Glu  
 [0645] 50 55 60  
 [0646] Val Gly Phe Thr Ala Ala Val Ala Gln Pro Ala Ile Ala Tyr Pro Thr  
 [0647] 65 70 75 80  
 [0648] Ser Ile Arg His Ala Asn Ala Val Asn Ala Met Ile Ala Thr Leu Ser  
 [0649] 85 90 95  
 [0650] Lys Glu Asn Met Gln Arg Asp Leu Thr Lys Leu Ser Ser Phe Gln Thr  
 [0651] 100 105 110  
 [0652] Ala Tyr Tyr Lys Val Asp Phe Gly Lys Gln Ser Ala Thr Trp Leu Gln  
 [0653] 115 120 125  
 [0654] Glu Gln Val Gln Ala Ala Ile Asn Thr Ala Gly Ala Asn Arg Tyr Gly  
 [0655] 130 135 140  
 [0656] Ala Lys Val Ala Ser Phe Arg His Asn Phe Ala Gln His Ser Ile Ile  
 [0657] 145 150 155 160  
 [0658] Ala Thr Ile Pro Gly Arg Ser Pro Glu Val Val Val Val Gly Ala His  
 [0659] 165 170 175  
 [0660] Gln Asp Ser Ile Asn Gln Arg Ser Pro Met Thr Gly Arg Ala Pro Gly  
 [0661] 180 185 190  
 [0662] Ala Asp Asp Asn Gly Ser Gly Ser Val Thr Ile Leu Glu Ala Leu Arg  
 [0663] 195 200 205  
 [0664] Gly Val Leu Arg Asp Gln Thr Ile Leu Gln Gly Lys Ala Ala Asn Thr  
 [0665] 210 215 220  
 [0666] Ile Glu Phe His Trp Tyr Ala Gly Glu Glu Ala Gly Leu Leu Gly Ser  
 [0667] 225 230 235 240  
 [0668] Gln Ala Ile Phe Ala Asn Tyr Lys Gln Thr Gly Lys Lys Val Lys Gly  
 [0669] 245 250 255  
 [0670] Met Leu Asn Gln Asp Met Thr Gly Tyr Ile Lys Gly Met Val Asp Lys  
 [0671] 260 265 270  
 [0672] Gly Leu Lys Val Ser Phe Gly Ile Ile Thr Asp Asn Val Asn Ala Asn  
 [0673] 275 280 285  
 [0674] Leu Thr Lys Phe Val Arg Met Val Ile Thr Lys Tyr Cys Ser Ile Pro  
 [0675] 290 295 300



[0676]	Thr Ile Asp Thr Arg Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Ala	
[0677]	305	310 315 320
[0678]	Asn Arg Asn Gly Tyr Pro Ser Ala Met Val Ala Glu Ser Pro Ile Asp	
[0679]		325 330 335
[0680]	Leu Leu Asp Pro His Leu His Thr Asp Ser Asp Asn Ile Ser Tyr Leu	
[0681]		340 345 350
[0682]	Asp Phe Asp His Met Ile Glu His Ala Lys Leu Ile Val Gly Phe Val	
[0683]		355 360 365
[0684]	Thr Glu Leu Ala Lys	
[0685]	370	
[0686]	<210>7	
[0687]	<211>1557	
[0688]	<212>DNA	
[0689]	<213> 烟曲霉菌 ( <i>A. fumigatus</i> )	
[0690]	<400>7	
[0691]	atgaagctgc tctacctcac atcgtttgcc tctctggccg tggccaatgg cccaggatgg	60
[0692]	gactggaagc cccgagtcca tccggttagt gttcctctcg ccgggtttgt ctgctgtatg	120
[0693]	ctaacagcat cctgtctatt acagaaagtc ctgccccaaa tgatccattt gtgggatctt	180
[0694]	ctgcagggcg ctcaacagct ggaagacttc gcctatgcct accccgagcg caaccgcgtc	240
[0695]	tttggtggac gggcccacga ggacaccgtc aactacctct accgtgagtt gaagaaaacc	300
[0696]	ggctactacg acgtttacaa gcagccccag gttcaccagt ggacccgagc cgaccaggct	360
[0697]	ctcaccgteg acggccagtc ctatgacgcc acaacatga cttacagccc cagcgtaaac	420
[0698]	gccacggcgc cgctggcagt ggtgaacaac ctgggctgcg tcgaggctga ctateccgcc	480
[0699]	gatctgacgg gcaagattgc tctgatctcg cggggcgagt gcacctttgc gaccaaatcc	540
[0700]	gtcttgagcg ccaaggccgg ggcggcggcg gcaactcgtgt acaacaatat cgagggttcg	600
[0701]	atggcgggaa ctctgggagg cgcgaccagc gagctgggtg cctacgctcc catcgccggc	660
[0702]	atcagcctcg cggacggaca ggcgctgata cagatgatcc aggcgggcac ggtgacagcc	720
[0703]	aacctgtgga tcgacagcca ggtcgagaac cgtaccacct acaacgtgat cgcgcagacc	780
[0704]	aagggcggcg accccaacaa cgtcgtcgcg ctgggtggcc acacggactc ggtcgaggcc	840
[0705]	gggcccggca tcaacgacga cggtccggc atcatcagea acctcgtcgt cgccaaggcg	900
[0706]	ctgaccgctt tctcggtaaa gaacgcggtg cgcttctgct tctggacggc ggaggagttc	960
[0707]	ggcctgctgg gcagcaacta ctacgtcaac agcctcaatg ccaccgagca ggccaagatc	1020
[0708]	cgctgtatc tcaacttcga catgatgcc tcccccaact acgccctgat gatctatgac	1080
[0709]	ggcgacggct cggccttcaa cctgacgggg cggcggct cggcgcagat cgagcggctc	1140
[0710]	ttcaggact actacacgtc gatccgcaag ccgttcgtgc cgaccgagtt caacggccgc	1200
[0711]	tccgactacc aggcctttat tctcaacggc atccccggg gaggcctctt caccggcgcg	1260
[0712]	gaggcgatea agaccgagga acaggcccaa ttgtttggcg gccaggccgg cgtggctctg	1320
[0713]	gacgccaact accacgcca gggtgacaac atgactaatc tcaaccgca ggctttcctg	1380
[0714]	atcaattcea ggcgacggc ctttgccgtg ggcagctacg ccaacagcct tgactcgatc	1440

[0715]	ccccacgca acatgaccac cgtggtcaag cggtcgcagc tggagcaagc catgaagagg	1500
[0716]	accccgca ca cgcacaccgg cggaacagga tgctacaagg accgggttga gcagtag	1557
[0717]	<210>8	
[0718]	<211>1497	
[0719]	<212>DNA	
[0720]	<213>烟曲霉菌 (A. fumigatus)	
[0721]	<400>8	
[0722]	atgaagctgc tctacctcac atcgtttgcc tctctggccg tggccaatgg cccaggatgg	60
[0723]	gactggaagc cccgagtca tccgaaagtc ctgccccaaa tgatccattt gtgggatctt	120
[0724]	ctgcaggggcg ctcaacagct ggaagacttc gcctatgcct accccgagcg caaccgctc	180
[0725]	tttggtggac gggccccacga ggacaccgtc aactacctct accgtgagtt gaagaaaacc	240
[0726]	ggctactacg acgtttacaa gcagccccag gttcaccagt ggaccgagc cgaccaggct	300
[0727]	ctcaccgtcg acggccagtc ctatgacgcc acaacatga cttacagccc cagcgtaaac	360
[0728]	gccacggcgc cgctggcagt ggtgaacaac ctgggctgcg tcgaggctga ctatcccgcc	420
[0729]	gatctgacgg gcaagattgc tctgatctcg cggggcgagt gcacctttgc gaccaaatcc	480
[0730]	gtcttgagcg ccaaggccgg ggccggcggc gcactcgtgt acaacaatat cgagggttcg	540
[0731]	atggcgggaa ctctgggcgg cgcgaccagc gagctgggtg cctacgtcc catcgccggc	600
[0732]	atcagcctcg cggacggaca ggcgctgata cagatgatcc aggcgggcac ggtgacagcc	660
[0733]	aacctgtgga tcgacagcca ggtcgagaac cgtaccacct acaacgtgat cgcgcagacc	720
[0734]	aagggcggcg accccaacaa cgtcgtcgcg ctgggtggcc acacggactc ggtcgaggcc	780
[0735]	gggcccggca tcaacgacga cggctccggc atcatcagca acctcgtcgt cgccaaggcg	840
[0736]	ctgaccgct tctcggtaaa gaacgcggtg cgcttctgct tctggacggc ggaggagttc	900
[0737]	ggcctgctgg gcagcaacta ctacgtcaac agcctcaatg ccaccgagca ggccaagatc	960
[0738]	cgcctgtatc tcaacttega catgatgcc tccccaaact acgccctgat gatctatgac	1020
[0739]	ggcgacggct cggccttcaa cctgacgggg ccggccggct cggcgagat cgagcggctc	1080
[0740]	ttcaggact actacacgtc gatccgcaag ccgttcgtgc cgaccgagtt caacggccgc	1140
[0741]	tccgactacc aggctttat tctcaacggc atcccccgcg gaggcctctt caccggcgcg	1200
[0742]	gaggcgatca agaccgagga acaggcccaa ttgtttggcg gccaggccgg cgtggctctg	1260
[0743]	gacgccaact accacgcaaa gggtgacaac atgactaatc tcaaccgca ggctttcctg	1320
[0744]	atcaattcca gggcgacggc ctttgccgtg gcgacgtacg ccaacagcct tgactcgatc	1380
[0745]	ccccacgca acatgaccac cgtggtcaag cggtcgcagc tggagcaagc catgaagagg	1440
[0746]	accccgca ca cgcacaccgg cggaacagga tgctacaagg accgggttga gcagtag	1497
[0747]	<210>9	
[0748]	<211>498	
[0749]	<212>PRT	
[0750]	<213>烟曲霉菌 (A. fumigatus)	
[0751]	<400>9	
[0752]	Met Lys Leu Leu Tyr Leu Thr Ser Phe Ala Ser Leu Ala Val Ala Asn	
[0753]	1 5 10 15	

[0754]	Gly Pro Gly Trp Asp Trp Lys Pro Arg Val His Pro Lys Val Leu Pro
[0755]	20 25 30
[0756]	Gln Met Ile His Leu Trp Asp Leu Leu Gln Gly Ala Gln Gln Leu Glu
[0757]	35 40 45
[0758]	Asp Phe Ala Tyr Ala Tyr Pro Glu Arg Asn Arg Val Phe Gly Gly Arg
[0759]	50 55 60
[0760]	Ala His Glu Asp Thr Val Asn Tyr Leu Tyr Arg Glu Leu Lys Lys Thr
[0761]	65 70 75 80
[0762]	Gly Tyr Tyr Asp Val Tyr Lys Gln Pro Gln Val His Gln Trp Thr Arg
[0763]	85 90 95
[0764]	Ala Asp Gln Ala Leu Thr Val Asp Gly Gln Ser Tyr Asp Ala Thr Thr
[0765]	100 105 110
[0766]	Met Thr Tyr Ser Pro Ser Val Asn Ala Thr Ala Pro Leu Ala Val Val
[0767]	115 120 125
[0768]	Asn Asn Leu Gly Cys Val Glu Ala Asp Tyr Pro Ala Asp Leu Thr Gly
[0769]	130 135 140
[0770]	Lys Ile Ala Leu Ile Ser Arg Gly Glu Cys Thr Phe Ala Thr Lys Ser
[0771]	145 150 155 160
[0772]	Val Leu Ser Ala Lys Ala Gly Ala Ala Ala Ala Leu Val Tyr Asn Asn
[0773]	165 170 175
[0774]	Ile Glu Gly Ser Met Ala Gly Thr Leu Gly Gly Ala Thr Ser Glu Leu
[0775]	180 185 190
[0776]	Gly Ala Tyr Ala Pro Ile Ala Gly Ile Ser Leu Ala Asp Gly Gln Ala
[0777]	195 200 205
[0778]	Leu Ile Gln Met Ile Gln Ala Gly Thr Val Thr Ala Asn Leu Trp Ile
[0779]	210 215 220
[0780]	Asp Ser Gln Val Glu Asn Arg Thr Thr Tyr Asn Val Ile Ala Gln Thr
[0781]	225 230 235 240
[0782]	Lys Gly Gly Asp Pro Asn Asn Val Val Ala Leu Gly Gly His Thr Asp
[0783]	245 250 255
[0784]	Ser Val Glu Ala Gly Pro Gly Ile Asn Asp Asp Gly Ser Gly Ile Ile
[0785]	260 265 270
[0786]	Ser Asn Leu Val Val Ala Lys Ala Leu Thr Arg Phe Ser Val Lys Asn
[0787]	275 280 285
[0788]	Ala Val Arg Phe Cys Phe Trp Thr Ala Glu Glu Phe Gly Leu Leu Gly
[0789]	290 295 300
[0790]	Ser Asn Tyr Tyr Val Asn Ser Leu Asn Ala Thr Glu Gln Ala Lys Ile
[0791]	305 310 315 320
[0792]	Arg Leu Tyr Leu Asn Phe Asp Met Ile Ala Ser Pro Asn Tyr Ala Leu

[0793]		325		330		335
[0794]	Mat	Ila Tyr Asp Gly Asp Gly Ser Ala Phe Asn Leu Thr Gly Pro Ala				
[0795]		340		345		350
[0796]	Gly Ser Ala Gln Ile Glu Arg Leu Phe Glu Asp Tyr Tyr Thr Ser Ile					
[0797]		355		360		365
[0798]	Arg Lys Pro Phe Val Pro Thr Glu Phe Asn Gly Arg Ser Asp Tyr Gln					
[0799]		370		375		380
[0800]	Ala Phe Ile Leu Asn Gly Ile Pro Ala Gly Gly Leu Phe Thr Gly Ala					
[0801]		385		390		400
[0802]	Glu Ala Ile Lys Thr Glu Glu Gln Ala Gln Leu Phe Gly Gly Gln Ala					
[0803]		405		410		415
[0804]	Gly Val Ala Leu Asp Ala Asn Tyr His Ala Lys Gly Asp Asn Met Thr					
[0805]		420		425		430
[0806]	Asn Leu Asn Arg Glu Ala Phe Leu Ile Asn Ser Arg Ala Thr Ala Phe					
[0807]		435		440		445
[0808]	Ala Val Ala Thr Tyr Ala Asn Ser Leu Asp Ser Ile Pro Pro Arg Asn					
[0809]		450		455		460
[0810]	Met Thr Thr Val Val Lys Arg Ser Gln Leu Glu Gln Ala Met Lys Arg					
[0811]		465		470		480
[0812]	Thr Pro His Thr His Thr Gly Gly Thr Gly Cys Tyr Lys Asp Arg Val					
[0813]		485		490		495
[0814]	Glu Gln					
[0815]	<210>10					
[0816]	<211>1298					
[0817]	<212>DNA					
[0818]	<213> 烟曲霉菌 (A. fumigatus)					
[0819]	atgaaagttc ttacagctat tgcgctgagc gcaatagctt tcacaggggc tgtagctgca	60				
[0820]	gtgattactc aggaagcatt cttaaacaac ccccgcattc atcatgacca ggagaagtac	120				
[0821]	ttgatcgaac tggcccctta tcgaacacga tgggtgactg aagaggagaa atgggcattg	180				
[0822]	aaattggtac catacttccc caaaatttgg gtctccaagt ccacgggcga ctaactgcac	240				
[0823]	gattgcttga aggacggcgt gaattttatc gatatcacag aagagcacia caccgattt	300				
[0824]	taccgcactc tccacagcgc cagctatgtg aaatatccac cgaagatgca gtatgcagaa	360				
[0825]	gaagtggctg ctcttaacaa gaatttatcg aaagaaaaca tgaaggccaa cctggaacga	420				
[0826]	ttcacatcat ttcatactcg ctattacaaa tctcagacgg gaatccgate ggcaacgtgg	480				
[0827]	ctgttcgacc aagttcagag agttgtctct gagtctggag ccgctgagta tgggtgcaact	540				
[0828]	gttgagcgat tctctcatcc atggggctcag ttcagcatta ttgcccgaat acccggcccga	600				
[0829]	acgaacaaga ctgtggtgct gggegcccat caggacagca tcaatttgtt tctcccgtca	660				
[0830]	atcttggtg ctcccgtgct tgatgacgat ggaagtggaa ctgtcaccat tcttgaageg	720				
[0831]	ttgcgcggtc tgetgeagtc agacgccatt gccaaaggta atgcatccaa tactgtcgag	780				

[0832]	ttccactggt actctgcaga agaaggcgga atgctgggct cccaggcaat attttccaat	840
[0833]	tacaagcgga ataggcgga aatcaaagcc atgctccagc aagacatgac tggctacgtc	900
[0834]	caggagctt tgaacgccgg tgttgaggaa gccataggaa ttatggtcga ttatgtcgac	960
[0835]	cagggcctca cacagtttct caaggacggt gttacagcgg taagcctcag ttgtcccca	1020
[0836]	cgaaaagctg ttagtcgac aatgaaatt gacggctgca ttagtactgc tctgtgggtt	1080
[0837]	acctggagac gaagtgcgga tatgcctgct ccgaccacac ctcggccagt aatatggtt	1140
[0838]	atcccgcggc tatggcgaca gaagcagaga tggaaaatac caataagaag atacatacta	1200
[0839]	ccgacgaca gatcaagtat ttgagcttcg atcatatggt ggagcatgcc aagttgagtc	1260
[0840]	ttggcttcgc ttcgaattg gcatttgcgc gttttaa	1298
[0841]	<210>11	
[0842]	<211>1167	
[0843]	<212>DNA	
[0844]	<213> 烟曲霉菌 (A. fumigatus)	
[0845]	<400>11	
[0846]	atgaaagttc ttacagctat tgcgctgagc gcaatagctt tcacaggggc tgtagctgca	60
[0847]	gtgattactc aggaagcatt cttaaacaac ccccgcattc atcatgacca ggagaagtac	120
[0848]	ttgatcgaac tggcccctta tcgaacacga tgggtgactg aagaggagaa atgggcattg	180
[0849]	aaattggacg gcggtgaattt tatcgatatac acagaagagc acaacaccgg attttaccg	240
[0850]	actctccaca gcgccagcta tgtgaaatat ccaccgaaga tgcagtatgc agaagaagtg	300
[0851]	gctgctctta acaagaattt atcgaaagaa aacatgaagg ccaacctgga acgattcaca	360
[0852]	tcatttcata ctcgctatta caaatctcag acgggaatcc gatcggcaac gtggctgttc	420
[0853]	gaccaagttc agagagttgt ctctgagtct ggagccgctg agtatggtgc aactgttgag	480
[0854]	cgattctctc atccatgggg tcagttcagc attattgccc gaataccgg ccgaacgaac	540
[0855]	aagactgtgg tgctgggcgc ccatcaggac agcatcaatt tgtttctccc gtcaatcttg	600
[0856]	gctgctcccg gtgctgatga cgatggaagt ggaactgtca ccattcttga agcgttgcgc	660
[0857]	ggtctgctgc agtcagacgc cattgccaag ggtaatgcat ccaatactgt cgagttccac	720
[0858]	tggtactctg cagaagaagg cggaatgctg ggctcccagg caatatttcc caattacaag	780
[0859]	cggaatagggc gggaaatcaa agccatgctc cagcaagaca tgactggcta cgtccaggga	840
[0860]	gctttgaacg ccggtgttga ggaagccata ggaattatgg tcgattatgt cgaccagggc	900
[0861]	ctcacacagt ttctcaagga cgttgttaca gcgctactgct ctgtgggtta cctggagacg	960
[0862]	aagtgcggat atgctgctc cgaccacacc teggccagta aatatggtta tcccgggct	1020
[0863]	atggcgacag aagcagagat ggaaaatacc aataagaaga tacatactac cgacgacaag	1080
[0864]	atcaagtatt tgagcttcga tcatatgttg gagcatgcca agttgagtct tggttcgct	1140
[0865]	ttcgaattgg catttgcgcc gttttaa	1167
[0866]	<210>12	
[0867]	<211>388	
[0868]	<212>PRT	
[0869]	<213> 烟曲霉菌 (A. fumigatus)	
[0870]	<400>12	

[0871] Met Lys Val Leu Thr Ala Ile Ala Leu Ser Ala Ile Ala Phe Thr Gly  
 [0872] 1 5 10 15  
 [0873] Ala Val Ala Ala Val Ile Thr Gln Glu Ala Phe Leu Asn Asn Pro Arg  
 [0874] 20 25 30  
 [0875] Ile His His Asp Gln Glu Lys Tyr Leu Ile Glu Leu Ala Pro Tyr Arg  
 [0876] 35 40 45  
 [0877] Thr Arg Trp Val Thr Glu Glu Glu Lys Trp Ala Leu Lys Leu Asp Gly  
 [0878] 50 55 60  
 [0879] Val Asn Phe Ile Asp Ile Thr Glu Glu His Asn Thr Gly Phe Tyr Pro  
 [0880] 65 70 75 80  
 [0881] Thr Leu His Ser Ala Ser Tyr Val Lys Tyr Pro Pro Lys Met Gln Tyr  
 [0882] 85 90 95  
 [0883] Ala Glu Glu Val Ala Ala Leu Asn Lys Asn Leu Ser Lys Glu Asn Met  
 [0884] 100 105 110  
 [0885] Lys Ala Asn Leu Glu Arg Phe Thr Ser Phe His Thr Arg Tyr Tyr Lys  
 [0886] 115 120 125  
 [0887] Ser Gln Thr Gly Ile Arg Ser Ala Thr Trp Leu Phe Asp Gln Val Gln  
 [0888] 130 135 140  
 [0889] Arg Val Val Ser Glu Ser Gly Ala Ala Glu Tyr Gly Ala Thr Val Glu  
 [0890] 145 150 155 160  
 [0891] Arg Phe Ser His Pro Trp Gly Gln Phe Ser Ile Ile Ala Arg Ile Pro  
 [0892] 165 170 175  
 [0893] Gly Arg Thr Asn Lys Thr Val Val Leu Gly Ala His Gln Asp Ser Ile  
 [0894] 180 185 190  
 [0895] Asn Leu Phe Leu Pro Ser Ile Leu Ala Ala Pro Gly Ala Asp Asp Asp  
 [0896] 195 200 205  
 [0897] Gly Ser Gly Thr Val Thr Ile Leu Glu Ala Leu Arg Gly Leu Leu Gln  
 [0898] 210 215 220  
 [0899] Ser Asp Ala Ile Ala Lys Gly Asn Ala Ser Asn Thr Val Glu Phe His  
 [0900] 225 230 235 240  
 [0901] Trp Tyr Ser Ala Glu Glu Gly Gly Met Leu Gly Ser Gln Ala Ile Phe  
 [0902] 245 250 255  
 [0903] Ser Asn Tyr Lys Arg Asn Arg Arg Glu Ile Lys Ala Met Leu Gln Gln  
 [0904] 260 265 270  
 [0905] Asp Met Thr Gly Tyr Val Gln Gly Ala Leu Asn Ala Gly Val Glu Glu  
 [0906] 275 280 285  
 [0907] Ala Ile Gly Ile Met Val Asp Tyr Val Asp Gln Gly Leu Thr Gln Phe  
 [0908] 290 295 300  
 [0909] Leu Lys Asp Val Val Thr Ala Tyr Cys Ser Val Gly Tyr Leu Glu Thr

[0910]	305	310	315	320
[0911]	Lys Cys Gly Tyr Ala Cys Ser Asp His Thr Ser Ala Ser Lys Tyr Gly			
[0912]		325	330	335
[0913]	Tyr Pro Ala Ala Met Ala Thr Glu Ala Glu Met Glu Asn Thr Asn Lys			
[0914]		340	345	350
[0915]	Lys Ile His Thr Thr Asp Asp Lys Ile Lys Tyr Leu Ser Phe Asp His			
[0916]		355	360	365
[0917]	Met Leu Glu His Ala Lys Leu Ser Leu Gly Phe Ala Phe Glu Leu Ala			
[0918]		370	375	380
[0919]	Phe Ala Pro Phe			
[0920]	385			
[0921]	<210>13			
[0922]	<211>2106			
[0923]	<212>DNA			
[0924]	<213> 红色毛癣菌 (T. rubrum)			
[0925]	<400>13			
[0926]	atggtgtcat tctgcggagt ggcagcctgc ctgctgacag ttgctggcca tcttgcgcag			60
[0927]	gctcagttcc caccaaaacc ggaggagtc actgtcctgg agtcgaaatt cggcagcggc			120
[0928]	gctcgcata cttataagga ggtccgtag ctgcatagaa agtccacgtg aagacgctgt			180
[0929]	agctaacaat ccactagcct ggcctctgtg agacgacaga gggcgtcaag tegtacgccg			240
[0930]	gatatgtcca tctgcctcca ggcagctca gggacttcgg tgctgagcag gactacccta			300
[0931]	tcaacacctt tttttggttc tttgaggcaa gaaaggacce tgaaaatgcc cctctcggca			360
[0932]	tctggatgaa cgggtggccc ggtagctcgt cgatgtttgg aatgatgact gagaacgggc			420
[0933]	cttgcttcgt caatgcagac tccaattcta ctgcctgaa ccctcattct tggaacaatg			480
[0934]	aaggtatgcc atcagcttct gatggaaaac taaatattgc taacattgta ctttctgtga			540
[0935]	ctagtcaata tgctgtatat agaccagcca gtgcaggctg gtctgtccta cgacactttg			600
[0936]	gccaacttca ccaggaatct agtcacggat gagatcacga aactgaaacc cggagaacct			660
[0937]	attccggaac agaatgccac tttcctggta ggtacatatg caagccgcaa tatgaacacc			720
[0938]	actgcacacg gaactaggca tgctgccatg gctctctggc acttegccca agtctggttc			780
[0939]	caagagttcc caggatatca ccctaggaac aacaagatca gcattgctac cgaatcctac			840
[0940]	ggcggccggt atggcccggc ctttactgcc ttctttgaag agcagaacca gaagatcaag			900
[0941]	aacggccat ggaagggaca cgagggaact atgcacgtgc tgcattctga caccctcatg			960
[0942]	atcgtcaacg gatgcatcga cegtcttctc caatggccgg catatccga aatggcgtac			1020
[0943]	aacaacacat atagcatcga ggcagtcaac gccagcattc atgcaggaat gctggatgcc			1080
[0944]	ctctaccgg acggtggctg tegagacaag attaaccact gccgtccct ctctctgtg			1140
[0945]	ttcgatcctg agaatctcgg catcaactca accgtcaatg atgtctgcaa ggatgccag			1200
[0946]	acattctget ccaatgatgt teggatccc tactcaagt tctctggccg caactactat			1260
[0947]	gacatcggac agettgacce cageccattc ccagcaccat tttacatggc ctggetaaat			1320
[0948]	cagccgatg tgcaggcagc actgggtgtg ceacttaact ggacacagtc aaacgatgtt			1380

[0949]	gtgtctaccg cattccgtgc aattggtgac taccctcggc cagggtggtt ggagaacctg	1440
[0950]	gcttatttgc tggagaatgg catcaagggtt tcgcttgttt acggtgatcg ggactacgca	1500
[0951]	tgcaactggt tcggtggtga gctctcaagt ctgggaatca actacactga cacccacgaa	1560
[0952]	ttccataatg ccggctatgc aggtatccag atcaatagca gctacatcgg tggacaggtg	1620
[0953]	aggcagtacg gcaacctctc cttegccgc gtctacgagg ccggccatga ggtgccatcg	1680
[0954]	taccaacctg agactgcact gcagatattc caccgttccc tgttcaaaa ggatatcgct	1740
[0955]	actggaacca aggacacgtc atcgcgcatg gacggaggca agttttacgg cacctcgggc	1800
[0956]	cctgcggact cgtttggtt caagaacaaa cctccaccgc agcacgtcca cttctgtcat	1860
[0957]	atcttagaca ccagcacctg caccaaggag cagatccagt cagttgagaa cggcactgcc	1920
[0958]	gccgtacgca gctggatcat tctcactcc aactcgacct cctctgtccc cgaggtagtt	1980
[0959]	ggctcagggg aaccacgcc aaccctatg cctggagggg ctactacact atctgtcac	2040
[0960]	gggttcttgt atggcgtgac attatgggct gttattgtt tagctgttat agagctggca	2100
[0961]	atgtaa	2106
[0962]	<210>14	
[0963]	<211>1989	
[0964]	<212>DNA	
[0965]	<213> 红色毛癣菌 ( <i>T. rubrum</i> )	
[0966]	<400>14	
[0967]	atggtgtcat tctgcggagt ggcagcctgc ctgctgacag ttgctggcca tcttgcgcag	60
[0968]	gctcagttcc caccaaaacc ggagggagtc actgtcctgg agtcgaaatt cggcagcggc	120
[0969]	gctcgcatac cttataagga gcctggcctc tctgagacga cagagggcgt caagtcgtac	180
[0970]	gccgatatg tccatctgcc tccaggcagc ctgaggact tcggtgtcga gcaggactac	240
[0971]	cctatcaaca ctttttttg gttctttgag gcaagaaagg accctgaaaa tgcccctctc	300
[0972]	ggcatctgga tgaacggtgg cccgggtgac tcgtcgtatg ttggaatgat gactgagaac	360
[0973]	gggccttgct tcgtcaatgc agactccaat tctactcgc tgaacctca ttcttggaac	420
[0974]	aatgaagtca atatgctgta tatagaccag ccagtgcagg tcggtctgtc ctacgacact	480
[0975]	ttggccaact tcaccaggaa tctagtcacg gatgagatca cgaactgaa acccggagaa	540
[0976]	cctattccgg aacagaatgc cactttcctg gtaggtacat atgcaagccg caatatgaac	600
[0977]	accactgcac acggaactag gcatgctgcc atggctctct ggcaactcgc ccaagtctgg	660
[0978]	ttccaagagt tcccaggata tcaccctagg aacaacaaga tcagcattgc taccgaatcc	720
[0979]	tacggcggcc gttatggccc ggcccttact gccttctttg aagagcagaa ccagaagatc	780
[0980]	aagaacggca catggaaggg acacgaggga actatgcacg tctgcatct cgacaccctc	840
[0981]	atgatcgtca acggatgcat cgaccgtctt gtccaatggc cggcatalcc gcaaatggcg	900
[0982]	tacaacaaca catatagcat cgaggcagtc aacgccagca ttcattgcagg aatgctggat	960
[0983]	gccctctacc gcgacggtgg ctgtcgagac aagattaacc actgccgctc cctctctct	1020
[0984]	gtgttcgatc ctgagaatct cggcatcaac tcaaccgtca atgatgtctg caaggatgcc	1080
[0985]	gagacattct gtccaatga tgttcgcat cctacctca agttctctgg ccgcaactac	1140
[0986]	tatgacatcg gacagcttga cccagccca ttcccagcac cattttacat ggctggcta	1200
[0987]	aatcagccgc atgtgcaggc agcactgggt gtgccactta actggacaca gtcaaacgat	1260



[0988]	gttgtgtcta ccgattccg tgcaattggt gactaccctc ggccagggtg gctggagaac	1320
[0989]	ctggcttatt tgctggagaa tggcatcaag gtttcgcttg tttacggtga tcgggactac	1380
[0990]	gcatgcaact ggttcggtgg tgagctctca agtctgggaa tcaactacac tgacaccac	1440
[0991]	gaattccata atgccggcta tgcaggtatc cagatcaata gcagctacat cgggtggacag	1500
[0992]	gtgaggcagt acggcaacct ctcccttcgcc cgcgtctacg aggccggcca tgaggtgcca	1560
[0993]	tcgtaccaac ccgagactgc actgcagata ttcaccggtt ccctgttcaa caaggatata	1620
[0994]	gctactggaa ccaaggacac gtcatcgcgc atggacggag gcaagtttta cggcacctcg	1680
[0995]	ggccctgcgg actcgtttgg tttcaagaac aaacctccac cgcagcacgt ccacttctgt	1740
[0996]	catatcttag acaccagcac ctgcaccaag gagcagatcc agtcagttga gaacggcact	1800
[0997]	gccgccgtac gcagctggat cattgtcgcac tccaactcga cctctctggt ccccaggta	1860
[0998]	gttggctcag ggaaccac gccaaccct atgcctggag gggctactac actatctgt	1920
[0999]	cacgggttct tgtatggcgt gacattatgg gctgttattg ttgtagctgt tatagagctg	1980
[1000]	gcaatgtaa	1989
[1001]	<210>15	
[1002]	<211>662	
[1003]	<212>PRT	
[1004]	<213> 红色毛癣菌 (T. rubrum)	
[1005]	<400>15	
[1006]	Met Val Ser Phe Cys Gly Val Ala Ala Cys Leu Leu Thr Val Ala Gly	
[1007]	1                      5                      10                      15	
[1008]	His Leu Ala Gln Ala Gln Phe Pro Pro Lys Pro Glu Gly Val Thr Val	
[1009]	20                      25                      30	
[1010]	Leu Glu Ser Lys Phe Gly Ser Gly Ala Arg Ile Thr Tyr Lys Glu Pro	
[1011]	35                      40                      45	
[1012]	Gly Leu Cys Glu Thr Thr Glu Gly Val Lys Ser Tyr Ala Gly Tyr Val	
[1013]	50                      55                      60	
[1014]	His Leu Pro Pro Gly Thr Leu Arg Asp Phe Gly Val Glu Gln Asp Tyr	
[1015]	65                      70                      75                      80	
[1016]	Pro Ile Asn Thr Phe Phe Trp Phe Phe Glu Ala Arg Lys Asp Pro Glu	
[1017]	85                      90                      95	
[1018]	Asn Ala Pro Leu Gly Ile Trp Met Asn Gly Gly Pro Gly Ser Ser Ser	
[1019]	100                      105                      110	
[1020]	Met Phe Gly Met Met Thr Glu Asn Gly Pro Cys Phe Val Asn Ala Asp	
[1021]	115                      120                      125	
[1022]	Ser Asn Ser Thr Arg Leu Asn Pro His Ser Trp Asn Asn Glu Val Asn	
[1023]	130                      135                      140	
[1024]	Met Leu Tyr Ile Asp Gln Pro Val Gln Val Gly Leu Ser Tyr Asp Thr	
[1025]	145                      150                      155                      160	
[1026]	Leu Ala Asn Phe Thr Arg Asn Leu Val Thr Asp Glu Ile Thr Lys Leu	

[1027]		165		170		175
[1028]	Lys Pro Gly Glu Pro Ile Pro Glu Gln Asn Ala Thr Phe Leu Val Gly					
[1029]		180		185		190
[1030]	Thr Tyr Ala Ser Arg Asn Met Asn Thr Thr Ala His Gly Thr Arg His					
[1031]		195		200		205
[1032]	Ala Ala Met Ala Leu Trp His Phe Ala Gln Val Trp Phe Gln Glu Phe					
[1033]		210		215		220
[1034]	Pro Gly Tyr His Pro Arg Asn Asn Lys Ile Ser Ile Ala Thr Glu Ser					
[1035]		225		230		235
[1036]	Tyr Gly Gly Arg Tyr Gly Pro Ala Phe Thr Ala Phe Phe Glu Glu Gln					
[1037]		245		250		255
[1038]	Asn Gln Lys Ile Lys Asn Gly Thr Trp Lys Gly His Glu Gly Thr Met					
[1039]		260		265		270
[1040]	His Val Leu His Leu Asp Thr Leu Met Ile Val Asn Gly Cys Ile Asp					
[1041]		275		280		285
[1042]	Arg Leu Val Gln Trp Pro Ala Tyr Pro Gln Met Ala Tyr Asn Asn Thr					
[1043]		290		295		300
[1044]	Tyr Ser Ile Glu Ala Val Asn Ala Ser Ile His Ala Gly Met Leu Asp					
[1045]		305		310		315
[1046]	Ala Leu Tyr Arg Asp Gly Gly Cys Arg Asp Lys Ile Asn His Cys Arg					
[1047]		325		330		335
[1048]	Ser Leu Ser Ser Val Phe Asp Pro Glu Asn Leu Gly Ile Asn Ser Thr					
[1049]		340		345		350
[1050]	Val Asn Asp Val Cys Lys Asp Ala Glu Thr Phe Cys Ser Asn Asp Val					
[1051]		355		360		365
[1052]	Arg Asp Pro Tyr Leu Lys Phe Ser Gly Arg Asn Tyr Tyr Asp Ile Gly					
[1053]		370		375		380
[1054]	Gln Leu Asp Pro Ser Pro Phe Pro Ala Pro Phe Tyr Met Ala Trp Leu					
[1055]		385		390		395
[1056]	Asn Gln Pro His Val Gln Ala Ala Leu Gly Val Pro Leu Asn Trp Thr					
[1057]		405		410		415
[1058]	Gln Ser Asn Asp Val Val Ser Thr Ala Phe Arg Ala Ile Gly Asp Tyr					
[1059]		420		425		430
[1060]	Pro Arg Pro Gly Trp Leu Glu Asn Leu Ala Tyr Leu Leu Glu Asn Gly					
[1061]		435		440		445
[1062]	Ile Lys Val Ser Leu Val Tyr Gly Asp Arg Asp Tyr Ala Cys Asn Trp					
[1063]		450		455		460
[1064]	Phe Gly Gly Glu Leu Ser Ser Leu Gly Ile Asn Tyr Thr Asp Thr His					
[1065]		465		470		475
						480

[1066]	Glu Phe His Asn Ala Gly Tyr Ala Gly Ile Gln Ile Asn Ser Ser Tyr		
[1067]		485	490 495
[1068]	Ile Gly Gly Gln Val Arg Gln Tyr Gly Asn Leu Ser Phe Ala Arg Val		
[1069]		500	505 510
[1070]	Tyr Glu Ala Gly His Glu Val Pro Ser Tyr Gln Pro Glu Thr Ala Leu		
[1071]		515	520 525
[1072]	Gln Ile Phe His Arg Ser Leu Phe Asn Lys Asp Ile Ala Thr Gly Thr		
[1073]		530	535 540
[1074]	Lys Asp Thr Ser Ser Arg Met Asp Gly Gly Lys Phe Tyr Gly Thr Ser		
[1075]		545	550 555 560
[1076]	Gly Pro Ala Asp Ser Phe Gly Phe Lys Asn Lys Pro Pro Pro Gln His		
[1077]		565	570 575
[1078]	Val His Phe Cys His Ile Leu Asp Thr Ser Thr Cys Thr Lys Glu Gln		
[1079]		580	585 590
[1080]	Ile Gln Ser Val Glu Asn Gly Thr Ala Ala Val Arg Ser Trp Ile Ile		
[1081]		595	600 605
[1082]	Val Asp Ser Asn Ser Thr Ser Leu Phe Pro Glu Val Val Gly Ser Gly		
[1083]		610	615 620
[1084]	Glu Pro Thr Pro Thr Pro Met Pro Gly Gly Ala Thr Thr Leu Ser Ala		
[1085]		625	630 635 640
[1086]	His Gly Phe Leu Tyr Gly Val Thr Leu Trp Ala Val Ile Val Val Ala		
[1087]		645	650 655
[1088]	Val Ile Glu Leu Ala Met		
[1089]		660	
[1090]	<210>16		
[1091]	<211>2030		
[1092]	<212>DNA		
[1093]	<213> 红色毛癣菌 (T. rubrum)		
[1094]	<400>16		
[1095]	atgcgctttg ctgctagcat tgccgtggcc ctgccagtca ttcacgcggc gagtgetcaa	60	
[1096]	ggcttccctc cacccgtaa gggcgtcacc gtggtcaaat ccaagttcga cgaaaacgta	120	
[1097]	aagatcacat acaaggaggt atgtgtttac atcattttca catccagate ttatatectt	180	
[1098]	acaataaate tggctaacte actggataga atgacatatg tgaaccact caaggagtta	240	
[1099]	gatcattcac cggatcatgtc caccttccctc cagacaacga tgactttggg gtctaccgga	300	
[1100]	actactccat caacacatte ttctggttct ttgaagctcg tgaagaccct aagaatgctc	360	
[1101]	ctctctccat ctggctgaac ggtgggtccgg gatcgtcacc catgattgga ctcttccagg	420	
[1102]	aaaacggtcc atgctggggtc aatgaagact ctaaattctac caccaacaat tcattttcat	480	
[1103]	ggaacaataa agtaaataatg ctctacattg atcageccaaa ccaagtcggg ttcagttatg	540	
[1104]	acgtacctac caacataact tactetacca tcaatgatac aatatctggt geggacttct	600	

[1105]	ctaacggtgt ccctgcgcaa aatctttcta cgttggttgg aaccggcagc agccagaacc	660
[1106]	cttgggcaac tgccaataac actgtgaacg ctgctcgttc tatctggcac tttgcacaag	720
[1107]	tgtggttcca ggaattccct gaacacaagc ctaacaataa caagatcagt atttggacag	780
[1108]	agtcctatgg aggaagatat ggtccctcat tcgcctctta cttccaggaa cagaacgaaa	840
[1109]	agatcaaaaa ccataccatt actgaagaag gagagatgca tattctgaac ctcgacaccc	900
[1110]	tcggtatcat caacggctgc atcgatctta tgttccaagc agaaagttat gctgaattcc	960
[1111]	catacaaaa cacctatggc atcaaagctt ataccaagga gaagcgtgac gctatattac	1020
[1112]	acgacatcca ccgtcctgac ggctgcttcg acaaggttac caagtgccgt gaggccgca	1080
[1113]	aagaaggaga ccctcacttc tacagcaaca atgcaaccgt caacacaate tgtgcggatg	1140
[1114]	ctaactctgc ctgcgacaaa tatctaattg atcctttcca agagaccaat cttggttact	1200
[1115]	atgatattgc tcctctctt caggatccct tccccacc attctataag ggcttctca	1260
[1116]	gccaatccag cgttctatct gacatgggat cgcagtcaa cttctccaa tacgccaag	1320
[1117]	ctgtgggaaa atcattccat ggagtggcg actacgctcg ccctgatgtg cgcggctca	1380
[1118]	ccggtgacat tgcttatctt ctcgagagcg gactcaaggt tgctctcgtc tatggtgaca	1440
[1119]	gagactacat ctgcaattgg ttcggtggtg agcaggtcag tcttggttg aactacactg	1500
[1120]	gcaccaaga cttccacagg gcaaaatatg ccgatgcaa ggtcaactct tcatacgtc	1560
[1121]	gaggcgtagt gcgtcaacat ggaaacttct cttcaccag agttttcgag gccggtcatg	1620
[1122]	aagtccctgg ttaccaacce gagactgcc tcaagatctt tgagcgcac atgttcaaca	1680
[1123]	aggatatttc taccggtgag atcgacattg ctcagaaacc agactacggt accactggaa	1740
[1124]	ctgagtctac gttccatate aaaaacgata tccctcttc gcctgagccg acctgctacc	1800
[1125]	tcctcagtgc tgacggaacc tgtaccccg agcagcttaa tgctattaag gatggaactg	1860
[1126]	cagttgttga gaactacatt attaagagcc ctgctcgtc gaaggggaac cctccacaa	1920
[1127]	ccacgacctc atctcccaca gcagccccta ccgtggaag tgccatgcta aaggtcctg	1980
[1128]	tggcaatgct agcaatatca gctctcactg tcttgcctt cttctttag	2030
[1129]	<210>17	
[1130]	<211>1959	
[1131]	<212>DNA	
[1132]	<213> 红色毛癣菌 (T. rubrum)	
[1133]	<400>17	
[1134]	atgcgctttg ctgctagcat tgccgtggcc ctgccagtca ttcacgcggc gactgctcaa	60
[1135]	ggcttccctc caccggttaa ggccgtcacc gtggtcaaat ccaagttcga cgaaaacgta	120
[1136]	aagatcacat acaaggagaa tgacatatgt gaaaccactc aaggagttag atcattcacc	180
[1137]	ggctcatgtc accttctctc agacaacgat gactttggtg tctaccggaa ctactccatc	240
[1138]	aacacattct tctggttctt tgaagctcgt gaagacccta agaattgctc tctctccatc	300
[1139]	tggctgaacg gtggtccggg atcgctatcc atgattggac tcttccagga aaacggtcca	360
[1140]	tgctgggtca atgaagactc taaatctacc accaacaatt cattttcatg gaacaataaa	420
[1141]	gtaaatatgc tctacattga tcagccaaac caagtcggtt tcagttatga cgtacctacc	480
[1142]	aacatcactt actctaccat caatgataca atatctgttg cggacttctc taacggtgctc	540
[1143]	cctgcgcaaa atctttctac gttggttggga accggcagca gccagaacce ttgggcaact	600

[1144]	gccaataaca ctgtgaacgc tgctcgttct atctggcact ttgcacaagt gtggttccag	660
[1145]	gaattccctg aacacaagcc taacaataac aagatcagta tttggacaga gtcctatgga	720
[1146]	ggaagatatg gtccttcatt cgcctcttac ttccaggaac agaacgaaaa gatcaaaaac	780
[1147]	cataccatta ctgaagaagg agagatgcat attctgaacc tcgacaccct cggtatcatc	840
[1148]	aacggctgca tcgatcttat gttccaagca gaaagttatg ctgaattccc atacaacaac	900
[1149]	acctatggca tcaaagctta taccaaggag aagcgtgacg ctatattaca cgacatccac	960
[1150]	cgctctgacg gctgcttcga caaggttacc aagtgccgtg aggccgcgaa agaaggagac	1020
[1151]	cctcacttct acagcaacaa tgcaaccgtc aacacaatct gtgcggatgc taactctgcc	1080
[1152]	tgcgacaaat atctaattga tcctttccaa gagaccaatc ttggttacta tgatattgct	1140
[1153]	catcctcttc aggatccctt cccccacca ttctataagg gcttcctcag ccaatccagc	1200
[1154]	gttctatctg acatgggate gccagtcaac ttctccaat acgcccgaagc tgtgggaaaa	1260
[1155]	tcattccatg gagttggcga ctacgctcgc cctgatgtgc geggcttcac cggtgacatt	1320
[1156]	gcttatcttc tcgagagcgg agtcaaggtt gctctcgtct atggtgacag agactacatc	1380
[1157]	tgcaattgggt tcggtggtga gcaggtcagt cttggcttga actacactgg cacccaagac	1440
[1158]	ttccacaggg caaaatatgc cgatgtcaag gtcaactctt catacgtcgg aggcgtagtg	1500
[1159]	cgtaacatg gaaacttctc tttcaccaga gtttccgagg ccggtcatga agtccttgg	1560
[1160]	taccaaccg agactgcct caagatcttt gagcgcacatca tgttcaacaa ggatattct	1620
[1161]	accggtgaga tcgacattgc tcagaaacca gactacggta ccaactggaac tgagtctacg	1680
[1162]	ttccatatca aaaacgatat cctcctctcg cctgagccga cctgctacct cctcagtgt	1740
[1163]	gacggaacct gtaccccgga gcagcttaat gctattaagg atggaactgc agttgttgag	1800
[1164]	aactacatta ttaagagccc tgctgcgtcg aaggggaacc ctccaccaac cacgacctca	1860
[1165]	tctcccacag cagcccctac cgctggaagt gccatgctaa aggctcctgt ggcaatgcta	1920
[1166]	gcaatateag ctctcactgt ccttgcttcc ttctttag	1959
[1167]	<210>18	
[1168]	<211>652	
[1169]	<212>PRT	
[1170]	<213> 红色毛癣菌 (T. rubrum)	
[1171]	<400>18	
[1172]	Met Arg Phe Ala Ala Ser Ile Ala Val Ala Leu Pro Val Ile His Ala	
[1173]	1 5 10 15	
[1174]	Ala Ser Ala Gln Gly Phe Pro Pro Pro Val Lys Gly Val Thr Val Val	
[1175]	20 25 30	
[1176]	Lys Ser Lys Phe Asp Glu Asn Val Lys Ile Thr Tyr Lys Glu Asn Asp	
[1177]	35 40 45	
[1178]	Ile Cys Glu Thr Thr Gln Gly Val Arg Ser Phe Thr Gly His Val His	
[1179]	50 55 60	
[1180]	Leu Pro Pro Asp Asn Asp Asp Phe Gly Val Tyr Arg Asn Tyr Ser Ile	
[1181]	65 70 75 80	
[1182]	Asn Thr Phe Phe Trp Phe Phe Glu Ala Arg Glu Asp Pro Lys Asn Ala	

[1183]		85		90		95										
[1184]	Pro	Leu	Ser	Ile	Trp	Leu	Asn	Gly	Gly	Pro	Gly	Ser	Ser	Ser	Met	Ile
[1185]		100		105		110										
[1186]	Gly	Leu	Phe	Gln	Glu	Asn	Gly	Pro	Cys	Trp	Val	Asn	Glu	Asp	Ser	Lys
[1187]		115		120		125										
[1188]	Ser	Thr	Thr	Asn	Asn	Ser	Phe	Ser	Trp	Asn	Asn	Lys	Val	Asn	Met	Leu
[1189]		130		135		140										
[1190]	Tyr	Ile	Asp	Gln	Pro	Asn	Gln	Val	Gly	Phe	Ser	Tyr	Asp	Val	Pro	Thr
[1191]		145		150		155		160								
[1192]	Asn	Ile	Thr	Tyr	Ser	Thr	Ile	Asn	Asp	Thr	Ile	Ser	Val	Ala	Asp	Phe
[1193]		165		170		175										
[1194]	Ser	Asn	Gly	Val	Pro	Ala	Gln	Asn	Leu	Ser	Thr	Leu	Val	Gly	Thr	Gly
[1195]		180		185		190										
[1196]	Ser	Ser	Gln	Asn	Pro	Trp	Ala	Thr	Ala	Asn	Asn	Thr	Val	Asn	Ala	Ala
[1197]		195		200		205										
[1198]	Arg	Ser	Ile	Trp	His	Phe	Ala	Gln	Val	Trp	Phe	Gln	Glu	Phe	Pro	Glu
[1199]		210		215		220										
[1200]	His	Lys	Pro	Asn	Asn	Asn	Lys	Ile	Ser	Ile	Trp	Thr	Glu	Ser	Tyr	Gly
[1201]		225		230		235		240								
[1202]	Gly	Arg	Tyr	Gly	Pro	Ser	Phe	Ala	Ser	Tyr	Phe	Gln	Glu	Gln	Asn	Glu
[1203]		245		250		255										
[1204]	Lys	Ile	Lys	Asn	His	Thr	Ile	Thr	Glu	Glu	Gly	Glu	Met	His	Ile	Leu
[1205]		260		265		270										
[1206]	Asn	Leu	Asp	Thr	Leu	Gly	Ile	Ile	Asn	Gly	Cys	Ile	Asp	Leu	Met	Phe
[1207]		275		280		285										
[1208]	Gln	Ala	Glu	Ser	Tyr	Ala	Glu	Phe	Pro	Tyr	Asn	Asn	Thr	Tyr	Gly	Ile
[1209]		290		295		300										
[1210]	Lys	Ala	Tyr	Thr	Lys	Glu	Lys	Arg	Asp	Ala	Ile	Leu	His	Asp	Ile	His
[1211]		305		310		315		320								
[1212]	Arg	Pro	Asp	Gly	Cys	Phe	Asp	Lys	Val	Thr	Lys	Cys	Arg	Glu	Ala	Ala
[1213]		325		330		335										
[1214]	Lys	Glu	Gly	Asp	Pro	His	Phe	Tyr	Ser	Asn	Asn	Ala	Thr	Val	Asn	Thr
[1215]		340		345		350										
[1216]	Ile	Cys	Ala	Asp	Ala	Asn	Ser	Ala	Cys	Asp	Lys	Tyr	Leu	Met	Asp	Pro
[1217]		355		360		365										
[1218]	Phe	Gln	Glu	Thr	Asn	Leu	Gly	Tyr	Tyr	Asp	Ile	Ala	His	Pro	Leu	Gln
[1219]		370		375		380										
[1220]	Asp	Pro	Phe	Pro	Pro	Pro	Phe	Tyr	Lys	Gly	Phe	Leu	Ser	Gln	Ser	Ser
[1221]		385		390		395		400								

[1222] Val Leu Ser Asp Met Gly Ser Pro Val Asn Phe Ser Gln Tyr Ala Gln  
 [1223] 405 410 415  
 [1224] Ala Val Gly Lys Ser Phe His Gly Val Gly Asp Tyr Ala Arg Pro Asp  
 [1225] 420 425 430  
 [1226] Val Arg Gly Phe Thr Gly Asp Ile Ala Tyr Leu Leu Glu Ser Gly Val  
 [1227] 435 440 445  
 [1228] Lys Val Ala Leu Val Tyr Gly Asp Arg Asp Tyr Ile Cys Asn Trp Phe  
 [1229] 450 455 460  
 [1230] Gly Gly Glu Gln Val Ser Leu Gly Leu Asn Tyr Thr Gly Thr Gln Asp  
 [1231] 465 470 475 480  
 [1232] Phe His Arg Ala Lys Tyr Ala Asp Val Lys Val Asn Ser Ser Tyr Val  
 [1233] 485 490 495  
 [1234] Gly Gly Val Val Arg Gln His Gly Asn Phe Ser Phe Thr Arg Val Phe  
 [1235] 500 505 510  
 [1236] Glu Ala Gly His Glu Val Pro Gly Tyr Gln Pro Glu Thr Ala Leu Lys  
 [1237] 515 520 525  
 [1238] Ile Phe Glu Arg Ile Met Phe Asn Lys Asp Ile Ser Thr Gly Glu Ile  
 [1239] 530 535 540  
 [1240] Asp Ile Ala Gln Lys Pro Asp Tyr Gly Thr Thr Gly Thr Glu Ser Thr  
 [1241] 545 550 555 560  
 [1242] Phe HisIle Lys Asn Asp Ile Pro Pro Ser Pro Glu Pro Thr Cys Tyr  
 [1243] 565 570 575  
 [1244] Leu Leu Ser Ala Asp Gly Thr Cys Thr Pro Glu Gln Leu Asn Ala Ile  
 [1245] 580 585 590  
 [1246] Lys Asp Gly Thr Ala Val Val Glu Asn Tyr Ile Ile Lys Ser Pro Ala  
 [1247] 595 600 605  
 [1248] Ala Ser Lys Gly Asn Pro Pro Pro Thr Thr Thr Ser Ser Pro Thr Ala  
 [1249] 610 615 620  
 [1250] Ala Pro Thr Ala Gly Ser Ala Met Leu Lys Ala Pro Val Ala Met Leu  
 [1251] 625 630 635 640  
 [1252] Ala Ile Ser Ala Leu Thr Val Leu Ala Phe Phe Leu  
 [1253] 645 650  
 [1254] <210>19  
 [1255] <211>1795  
 [1256] <212>DNA  
 [1257] <213> 红色毛癣菌 (T. rubrum)  
 [1258] <220>  
 [1259] <221>misc feature  
 [1260] <222>(283).. (283)

[1261]	<223>n	
[1262]	<400>19	
[1263]	atgcaagcag caaaattggt gagccgggtac tggcaaaatg tacctgggta gtgcagctaa	60
[1264]	tcttgagtca catcatgcat agttaaccga gtatcacaac acaatctact attgcgtttt	120
[1265]	tgctaattggc taccatagga agactgaggg tatctgagct ccttttcgat gtccttttag	180
[1266]	actactcaaa cccgtcttcc acttegetcc ggttggtcgc caggagtgtg cageggcgaa	240
[1267]	ttccagggtc ctctctcgat gataaagaca gacagctacc ctnggattgt tttcctgcag	300
[1268]	ggtggaccag gaggagcttg cccacaacct caggaggtag gctgggttg gccattgctg	360
[1269]	gatcgaggat tccagggtgag tctccagaat cgggatgagt aactgtagaa caccttggtg	420
[1270]	aatttcttga ttagatcctt ctctctgacc agcgaggaac agggctttca acccctataa	480
[1271]	ccgctgcgac gcttctctt cagggaacg cagtaaagca agccgaatat cttaggctat	540
[1272]	tccgtgccga taatctctg cgagactgtg aagcagtgcg taaactattg actgcttatt	600
[1273]	accctccaga taagcagaaa tggagcgtcc ttggccagag ttttgagga ttctgtgccg	660
[1274]	tcacgtatgt ttctaagtag tgagtaacta ctcttcaaa tccacctgct atagattgtc	720
[1275]	gtgcaaatct aaccttcac atctagtcct gagggactta aagaagtctt cacaactggt	780
[1276]	ggattacccc ctcttggtc aaagcctgat cctgtgtacg agaggaccta cggtaagtgtg	840
[1277]	ggatagattg ggctattttt agtttaatat acagctgaca tctacagaca aggtccagtc	900
[1278]	ccggaataaa gtgtactatt ccactttccc cgaagacgaa gatcgagtgc ggattatact	960
[1279]	caagcatctc caaacccacg atgttaagct ccccgatggc tcaccgttaa ctccggaacg	1020
[1280]	ctttctccag ctaggaattc attttggaaat gaaaggtacg ccatacttcg caggtgactt	1080
[1281]	ctcgtaacca atgactaaca tatgcatata gggggcatcg gcttagttca tagtatgata	1140
[1282]	ccatcaataa cttacattat acttattcac tgactaacaa tgtecgaaata tcaggcataa	1200
[1283]	ttttgaagtg cattaatgaa ctggaatact ttggcttctt cacacgacct actttatctc	1260
[1284]	tgattgagaa cgacacgagt gcagacaacg gcattctata tgccataatg catgaatcta	1320
[1285]	tctactgcca agggtaaaac gtctctcttg atcgagtcaa taccagaatc taactgata	1380
[1286]	ccgtagggag gctcaaaact gggctgccga aagactacta ccaaagttct ctggcttccg	1440
[1287]	aggcgctcat aatcctgatg gcacttactt cactggggag atggtataca aacctggtt	1500
[1288]	tgagtcgtcc acagaactcg gccagctcaa agaggtagcc gatattcttg ctctctacaa	1560
[1289]	tgactggccg cagttgtatg ataaggaaca gctcgcgcgc aacgaggtgc cagtgtattc	1620
[1290]	cgctacatat gtcgaggata tgtacgtgca ctccagctac gccaacgaaa cagctgccac	1680
[1291]	tattcacaat tgcaaacagt tcatcaccia cagatgtac cacaacggac tgcgttcaga	1740
[1292]	ttccgctgaa cttattgcgc agctgtttgc tcttcgtgat gatacagattg actag	1795
[1293]	<210>20	
[1294]	<211>1326	
[1295]	<212>DNA	
[1296]	<213> 红色毛癣菌 (T. rubrum)	
[1297]	<400>20	
[1298]	atgcaagcag caaaattggt gagccgggtac tggcaaaatg tacctggaag actgagggta	60
[1299]	tctgagctcc ttttcgatgt ccttttagac tactcaaaacc cgtcttccac ttcgctccgg	120



[1300]	ttgttcgcca ggagtgtgca gcggcgaatt ccagggtcct ctctcgatga taaagacaga	180
[1301]	cagctaccct ggattgtttt cctgcagggt ggaccaggag gagcttgccc acaacctcag	240
[1302]	gaggtaggct gggttgggcc attgctggat cgaggattcc agatccttct ccttgaccag	300
[1303]	cgaggaacag ggctttcaac ccctataacc gctgcgacgc ttgctcttca gggaaacgca	360
[1304]	gtaaagcaag ccgaatatct taggctattc cgtgccgata atatcgtgcg agactgtgaa	420
[1305]	gcagtgcgta aactattgac tgcttattac cctccagata agcagaaatg gagcgtcctt	480
[1306]	ggccagagtt ttggaggatt ctgtgccgtc acgtatgttt ctaatcctga gggacttaaa	540
[1307]	gaagtcttca caactggtgg attaccacct cttgtgtcaa agcctgatcc tgtgtacgag	600
[1308]	aggacctacg acaaggcca gtcccgaat aaagtgtact attccacttt ccccgaagac	660
[1309]	gaagatcgag tgcgattat actcaagcat ctccaaacct acgatgttaa gctccccgat	720
[1310]	ggctcacctg taactccgga acgctttctc cagctaggaa ttcattttgg aatgaaaggc	780
[1311]	ataattttga agtgcattaa tgaactggaa tactttggct tcctcacacg acctacttta	840
[1312]	tctctgattg agaacgacac gagtgcagac aacggcattc tatatgcat aatgcatgaa	900
[1313]	tctatctact gccaagggga ggcctcaaac tgggctgccg aaagactact accaaagttc	960
[1314]	tctggcttcc gaggcgctca taatcctgat ggcactact tcaactggga gatggtatac	1020
[1315]	aaacactggt ttgagtcgtc cacagaactc ggccagctca aagaggtagc cgatattctt	1080
[1316]	gcttctaca atgactggcc gcagttgtat gataaggaac agctcgcgcg caacgaggtg	1140
[1317]	ccagtgtatt ccgctacata tgctcaggat atgtacgtgc acttcagcta cgccaacgaa	1200
[1318]	acagctgcca ctattcaciaa ttgcaaacag ttcatacaca acacgatgta ccacaacgga	1260
[1319]	ctgcgttcag attccgctga acttattgcg cagctgtttg ctcttcgtga tgatacgatt	1320
[1320]	gactag	1326
[1321]	<210>21	
[1322]	<211>441	
[1323]	<212>PRT	
[1324]	<213> 红色毛癣菌 (T. rubrum)	
[1325]	<220>	
[1326]	<221>UNSURE	
[1327]	<222>(283).. (283)	
[1328]	<223>I	
[1329]	<400>21	
[1330]	Met Gln Ala Ala Lys Leu Leu Ser Arg Tyr Trp Gln Asn Val Pro Gly	
[1331]	1                    5                    10                    15	
[1332]	Arg Leu Arg Val Ser Glu Leu Leu Phe Asp Val Pro Leu Asp Tyr Ser	
[1333]	20                    25                    30	
[1334]	Asn Pro Ser Ser Thr Ser Leu Arg Leu Phe Ala Arg Ser Val Gln Arg	
[1335]	35                    40                    45	
[1336]	Arg Ile Pro Gly Ser Ser Leu Asp Asp Lys Asp Arg Gln Leu Pro Trp	
[1337]	50                    55                    60	
[1338]	Ile Val Phe Leu Gln Gly Gly Pro Gly Gly Ala Cys Pro Gln Pro Gln	

[1339]	65	70	75	80
[1340]	Glu Val Gly Trp Val Gly Pro Leu Leu Asp Arg Gly Phe Gln Ile Leu			
[1341]		85	90	95
[1342]	Leu Leu Asp Gln Arg Gly Thr Gly Leu Ser Thr Pro Ile Thr Ala Ala			
[1343]		100	105	110
[1344]	Thr Leu Ala Leu Gln Gly Asn Ala Val Lys Gln Ala Glu Tyr Leu Arg			
[1345]		115	120	125
[1346]	Leu Phe Arg Ala Asp Asn Ile Val Arg Asp Cys Glu Ala Val Arg Lys			
[1347]		130	135	140
[1348]	Leu Leu Thr Ala Tyr Tyr Pro Pro Asp Lys Gln Lys Trp Ser Val Leu			
[1349]		145	150	155
[1350]	Gly Gln Ser Phe Gly Gly Phe Cys Ala Val Thr Tyr Val Ser Asn Pro			
[1351]		165	170	175
[1352]	Glu Gly Leu Lys Glu Val Phe Thr Thr Gly Gly Leu Pro Pro Leu Val			
[1353]		180	185	190
[1354]	Ser Lys Pro Asp Pro Val Tyr Glu Arg Thr Tyr Asp Lys Val Gln Ser			
[1355]		195	200	205
[1356]	Arg Asn Lys Val Tyr Tyr Ser Thr Phe Pro Glu Asp Glu Asp Arg Val			
[1357]		210	215	220
[1358]	Arg Ile Ile Leu Lys His Leu Gln Thr His Asp Val Lys Leu Pro Asp			
[1359]		225	230	235
[1360]	Gly Ser Pro Leu Thr Pro Glu Arg Phe Leu Gln Leu Gly Ile His Phe			
[1361]		245	250	255
[1362]	Gly Met Lys Gly Ile Ile Leu Lys Cys Ile Asn Glu Leu Glu Tyr Phe			
[1363]		260	265	270
[1364]	Gly Phe Leu Thr Arg Pro Thr Leu Ser Leu Ile Glu Asn Asp Thr Ser			
[1365]		275	280	285
[1366]	Ala Asp Asn Gly Ile Leu Tyr Ala Ile Met His Glu Ser Ile Tyr Cys			
[1367]		290	295	300
[1368]	Gln Gly Glu Ala Ser Asn Trp Ala Ala Glu Arg Leu Leu Pro Lys Phe			
[1369]		305	310	315
[1370]	Ser Gly Phe Arg Gly Ala His Asn Pro Asp Gly Ile Tyr Phe Thr Gly			
[1371]		325	330	335
[1372]	Glu Met Val Tyr Lys His Trp Phe Glu Ser Ser Thr Glu Leu Gly Gln			
[1373]		340	345	350
[1374]	Leu Lys Glu Val Ala Asp Ile Leu Ala Ser Tyr Asn Asp Trp Pro Gln			
[1375]		355	360	365
[1376]	Leu Tyr Asp Lys Glu Gln Leu Ala Arg Asn Glu Val Pro Val Tyr Ser			
[1377]		370	375	380

[1378]	Ala Thr Tyr Val Glu Asp Met Tyr Val His Phe Ser Tyr Ala Asn Glu	
[1379]	385	390 395 400
[1380]	Thr Ala Ala Thr Ile His Asn Cys Lys Gln Phe Ile Thr Asn Thr Met	
[1381]	405	410 415
[1382]	Tyr His Asn Gly Leu Arg Ser Asp Ser Ala Glu Leu Ile Ala Gln Leu	
[1383]	420	425 430
[1384]	Phe Ala Leu Arg Asp Asp Thr Ile Asp	
[1385]	435	440
[1386]	<210>22	
[1387]	<211>2418	
[1388]	<212>DNA	
[1389]	<213> 红色毛癣菌 (T. rubrum)	
[1390]	<400>22	
[1391]	atgccgccac caccggttga cacgaccag cgtctcgcaa agctcgaga gctgatggct	60
[1392]	cagaacaagg tcgatgtata tagtatgcaa ttcagataca ccattaaage tcccttgata	120
[1393]	ataacagtcg tatactcatt cttctttctt ctactcctcg ccttaaagtt gtgccttcgg	180
[1394]	aagacagcca tcagtcggag tacattgctc catgtgatgg gcgtcgaggt tagacctgtc	240
[1395]	cctccataaa agaataccta cccgtaatac cagccggcag acgctcatac gtatcactgc	300
[1396]	agctttcata tccagcttca ctggctcggc aggatgtgcc atcgtctcta tgagtaaagc	360
[1397]	tgctctgtct acagacggca gatacttcag ccaagctgca aaacagctcg atgccaactg	420
[1398]	gatcctgttg aagcgaggtg tcgagggtgt cccaacctgg gaagaatggt atatctgccc	480
[1399]	ctggtatcga cttttccggt ataatggttg acaggctgga tataggaccg ctgagcaggc	540
[1400]	cgagacacgg caaggttgtg ggtgttgacc cgtcacttat tacggcaggt gagaatctac	600
[1401]	agtatgcgct tcttacaagt gtcategtga ctaactgtat gttatagcgg atgcacgaaa	660
[1402]	gctttctcag acgttgaaga ccaccggagg ctctttggtt ggaattgatac agaacctgat	720
[1403]	tgatgccgct tggggagatg aacgtcctgc acggcctgcc aaccaaatta cggtagcacc	780
[1404]	tgttgagcgc gcgggaaagt cattcgagga gaaagtggaa gacctgcgaa aggaattgac	840
[1405]	tgcaagaag aggtctgcta tggttatttg tatgacgcta gatctatctt tgatcaaaaca	900
[1406]	tatactaaca aacgcaatat agccaccttg gatgagattg catggctctt caacctcctg	960
[1407]	ggaagcgagt aagtttctat ataaatggta tctttcactt tatacaaaaa gccatgctga	1020
[1408]	ctggtgtagt attccatata acccctctt tttctcgtac gcaattgtga cgccctcagt	1080
[1409]	tgcggaactc tatgtcgatg agagcaagct gtctccagaa gccagaaaac atctegaagg	1140
[1410]	caaggtcgtt ctcaagccat acgagtcctt cttccaagct tccaaagtcc tcgccgaatc	1200
[1411]	aaaggcatcg gctagcagcg gttcctctgg gaagttcttg ttgtctaaca aggtctcgtg	1260
[1412]	gtctttgagc ctgccctcgt gtggggaaca gaacgtcgtt gaggttcgaa gtcccatcac	1320
[1413]	tgacgcaaaa gccatcaaga acgaagtga actggaagga ttcagaaaaat gccatataccg	1380
[1414]	agacgggtgca gctctgatcg agtacttcgc ctggttgtaa aatgcattga tcaaagaagg	1440
[1415]	tgccaagcta gacgaagtag atggagccga caaactcttc gagatccgca agaaatatga	1500
[1416]	cctcttcgct gcgaactcct tcgacacat ctctctacc ggtgctaacg gtgctacat	1560

[1417]	tcattacaaa cccgagaagt caacttgcgc tatcattgac ccgaaggcta tgtacctgtg	1620
[1418]	tgactctggt ggccaatace ttgatggtac tactgatact accccaacte tccactttgg	1680
[1419]	agagcccacg gaggccaga agaaggctta tgcacttgtt ctaaaggac atatcagcat	1740
[1420]	tgacaatgcc attttcccca aaggaaccac cggatacgc attgactcgt ttgctcgaca	1800
[1421]	gcatttgtgg aaggagggtc tggattacct ccacggcacc ggtcatggtg ttggtcatt	1860
[1422]	tttggtacgg ggttcccttt ttcttttttt tttctttttt tatttttatt attacttctc	1920
[1423]	ttaggctaac acattctctc taagaacgtc catgaggac ctatgggcat aggaagccgt	1980
[1424]	gctcagtacg ctgaagtcc tctctctgcc agcaatgttc tttccaacgg taggatttct	2040
[1425]	gcatctcacc tttcttgaat cctactaatt gcaaaataga gcctggatat tatgaagacg	2100
[1426]	gcaacttcgg cattcgtctc gagagtaagt tcaatgactg cgtattctag tttttcata	2160
[1427]	ctgacggcct ctttagacct cgtaatctgc aaggaggctc agactgcaca caaattcggc	2220
[1428]	gacaagccct tctcggatt tgagtccatc acctgggtac ctttctgcca aaaactcctt	2280
[1429]	gatgcttctc tcttgaccga agctgagaga aagtgggtga atgattacca tgcgaaagtc	2340
[1430]	tgggagaaga ccagtccctt ctttgagaag gacgagttaa caaccgcctg gctaaagcgc	2400
[1431]	gagacacaac ctatttaa	2418
[1432]	<210>23	
[1433]	<211>1878	
[1434]	<212>DNA	
[1435]	<213> 红色毛癣菌 (T. rubrum)	
[1436]	<400>23	
[1437]	atgccgccac caccggttga caccgaccag cgtctcgcaa agctgcgaga gctgatggct	60
[1438]	cagaacaagg tcgatgtata tattgtgcct tcggaagaca gccatcagtc ggagtacatt	120
[1439]	gctccatgtg atgggcgtcg agctttcata tccagcttca ctggctcggc aggatgtgcc	180
[1440]	atcgtctcta tgagtaaagc tgctctgtct acagacggca gatacttcag ccaagctgca	240
[1441]	aaacagctcg atgccaactg gatcctgttg aagcgagggtg tcgagggtgt cccaacctgg	300
[1442]	gaagaatgga ccgctgagca ggccgagaca cggcaagggt gtgggtcgga tgcacgaaag	360
[1443]	ctttctcaga cgttgaagac caccggagge tccttgggtg gaattgatca gaacctgatt	420
[1444]	gatgccgtct ggggagatga acgtcctgca cggcctgcca accaaattac ggtacagcct	480
[1445]	gttgagcgcg cgggaaagtc attcagaggag aaagtggaag acctgcgaaa ggaattgact	540
[1446]	gcgaagaaga ggtctgctat ggttatttct agtaagttc tatataaatg gtatctttca	600
[1447]	ctttatacaa aaagccatgc tgactggtgt agtattccat ataaccctgt cttttctctg	660
[1448]	tacgcaattg tgacgccctc agttgcggaa ctctatgtcg atgagagcaa gctgtctcca	720
[1449]	gaagccagaa aacatctcga aggcaaggte gttctcaage catacagatc catcttccaa	780
[1450]	gcttccaaag tctcgcgca atcaaaggca tcggctagca gcggttctc tgggaagttc	840
[1451]	ttgttctcta acaaggcttc gtggtctttg agcctcgccc tcggtgggga acagaacgtc	900
[1452]	gttgaggttc gaagtcccat cactgacgcc aaagccatca agaacgaagt tgaactggaa	960
[1453]	ggattcagaa aatgcatat ccgagacggt geagctctga tcgagtactt cgcttggett	1020
[1454]	gaaaatgcat tgatcaaaga aggtgccaag ctgacgaaag tagatggagc cgacaaactc	1080
[1455]	ttcagatcc gcaagaaata tgacctctc gtcggcaact ccttcgacac catctctctt	1140

[1456] accggtgcta acggtgctac cattcattac aaacccgaga agtcaacttg cgctatcatt 1200  
 [1457] gacccgaagg ctatgtacct gtgtgactct ggtggccaat accttgatgg tactactgat 1260  
 [1458] actacccgaa ctctccactt tggagagccc acggagttec agaagaagge ttatgcactt 1320  
 [1459] gttctaaagg gacatatcag cattgacaat gccattttcc ccaagaagac caccggatac 1380  
 [1460] gccattgact cgtttgctcg acagcatttg tggaaggagg gtctggatta cctccacggc 1440  
 [1461] accggtcatg gtgttgctc atttttgaac gtccatgagg gacctatggg cataggaagc 1500  
 [1462] cgtgctcagt acgctgaagt tcctctctct gccagcaata gcctggatat tatgaagacg 1560  
 [1463] gcaacttcgg cattcgtctc gagagtaagt tcaatgactg cgtattctag ttttttcata 1620  
 [1464] ctgacggcct ctttagacct cgtaatctgc aaggaggtec agactgcaca caaattcggc 1680  
 [1465] gacaagccct tcctcggatt tgagtccatc accctggtac ctttctgcca aaaactcctt 1740  
 [1466] gatgcttctc tcttgaccga agctgagaga aagtgggtga atgattacca tgcgaaagtc 1800  
 [1467] tgggagaaga ccagtccctt ctttgagaag gacgagtaa caaccgctg gctaaagcgc 1860  
 [1468] gagacacaac ctatttaa 1878  
 [1469] <210>24  
 [1470] <211>625  
 [1471] <212>PRT  
 [1472] <213> 红色毛癣菌 (*T. rubrum*)  
 [1473] <400>24  
 [1474] Met Pro Pro Pro Pro Val Asp Thr Thr Gln Arg Leu Ala Lys Leu Arg  
 [1475] 1 5 10 15  
 [1476] Glu Leu Met Ala Gln Asn Lys Val Asp Val Tyr Ile Val Pro Ser Glu  
 [1477] 20 25 30  
 [1478] Asp Ser His Gln Ser Glu Tyr Ile Ala Pro Cys Asp Gly Arg Arg Ala  
 [1479] 35 40 45  
 [1480] Phe Ile Ser Ser Phe Thr Gly Ser Ala Gly Cys Ala Ile Val Ser Met  
 [1481] 50 55 60  
 [1482] Ser Lys Ala Ala Leu Ser Thr Asp Gly Arg Tyr Phe Ser Gln Ala Ala  
 [1483] 65 70 75 80  
 [1484] Lys Gln Leu Asp Ala Asn Trp Ile Leu Leu Lys Arg Gly Val Glu Gly  
 [1485] 85 90 95  
 [1486] Val Pro Thr Trp Glu Glu Trp Thr Ala Glu Gln Ala Glu Thr Arg Gln  
 [1487] 100 105 110  
 [1488] Gly Cys Gly Ser Asp Ala Arg Lys Leu Ser Gln Thr Leu Lys Thr Thr  
 [1489] 115 120 125  
 [1490] Gly Gly Ser Leu Val Gly Ile Asp Gln Asn Leu Ile Asp Ala Val Trp  
 [1491] 130 135 140  
 [1492] Gly Asp Glu Arg Pro Ala Arg Pro Ala Asn Gln Ile Thr Val Gln Pro  
 [1493] 145 150 155 160  
 [1494] Val Glu Arg Ala Gly Lys Ser Phe Glu Glu Lys Val Glu Asp Leu Arg

[1495]		165		170		175
[1496]	Lys Glu Leu Thr Ala Lys Lys Arg Ser Ala Met Val Ile Ser Ser Lys					
[1497]		180		185		190
[1498]	Phe Leu Tyr Lys Trp Tyr Leu Ser Leu Tyr Thr Lys Ser His Ala Asp					
[1499]		195		200		205
[1500]	Trp Cys Ser Ile Pro Tyr Asn Pro Val Phe Phe Ser Tyr Ala Ile Val					
[1501]		210		215		220
[1502]	Thr Pro Ser Val Ala Glu Leu Tyr Val Asp Glu Ser Lys Leu Ser Pro					
[1503]		225		230		235
[1504]	Glu Ala Arg Lys His Leu Glu Gly Lys Val Val Leu Lys Pro Tyr Glu					
[1505]		245		250		255
[1506]	Ser Ile Phe Gln Ala Ser Lys Val Leu Ala Glu Ser Lys Ala Ser Ala					
[1507]		260		265		270
[1508]	Ser Ser Gly Ser Ser Gly Lys Phe Leu Leu Ser Asn Lys Ala Ser Trp					
[1509]		275		280		285
[1510]	Ser Leu Ser Leu Ala Leu Gly Gly Glu Gln Asn Val Val Glu Val Arg					
[1511]		290		295		300
[1512]	Ser Pro Ile Thr Asp Ala Lys Ala Ile Lys Asn Glu Val Glu Leu Glu					
[1513]		305		310		315
[1514]	Gly Phe Arg Lys Cys His Ile Arg Asp Gly Ala Ala Leu Ile Glu Tyr					
[1515]		325		330		335
[1516]	Phe Ala Trp Leu Glu Asn Ala Leu Ile Lys Glu Gly Ala Lys Leu Asp					
[1517]		340		345		350
[1518]	Glu Val Asp Gly Ala Asp Lys Leu Phe Glu Ile Arg Lys Lys Tyr Asp					
[1519]		355		360		365
[1520]	Leu Phe Val Gly Asn SerPhe Asp Thr Ile Ser Ser Thr Gly Ala Asn					
[1521]		370		375		380
[1522]	Gly Ala Thr Ile His Tyr Lys Pro Glu Lys Ser Thr Cys Ala Ile Ile					
[1523]		385		390		395
[1524]	Asp Pro Lys Ala Met Tyr Leu Cys Asp Ser Gly Gly Gln Tyr Leu Asp					
[1525]		405		410		415
[1526]	Gly Thr Thr Asp Thr Thr Arg Thr Leu His Phe Gly Glu Pro Thr Glu					
[1527]		420		425		430
[1528]	Phe Gln Lys Lys Ala Tyr Ala Leu Val Leu Lys Gly His Ile Ser Ile					
[1529]		435		440		445
[1530]	Asp Asn Ala Ile Phe Pro Lys Gly Thr Thr Gly Tyr Ala Ile Asp Ser					
[1531]		450		455		460
[1532]	Phe Ala Arg Gln His Leu Trp Lys Glu Gly Leu Asp Tyr Leu His Gly					
[1533]		465		470		475
						480

[1534]	Thr Gly His Gly Val Gly Ser Phe Leu Asn Val His Glu Gly Pro Met	
[1535]		485 490 495
[1536]	Gly Ile Gly Ser Arg Ala Gln Tyr Ala Glu Val Pro Leu Ser Ala Ser	
[1537]		500 505 510
[1538]	Asn Ser Leu Asp Ile Met Lys Thr Ala Thr Ser Ala Phe Val Ser Arg	
[1539]		515 520 525
[1540]	Val Ser Ser Met Thr Ala Tyr Ser Ser Phe Phe Ile Leu Thr Ala Ser	
[1541]		530 535 540
[1542]	Leu Asp Leu Val Ile Cys Lys Glu Val Gln Thr Ala His Lys Phe Gly	
[1543]		545 550 555 560
[1544]	Asp Lys Pro Phe Leu Gly Phe Glu Ser Ile Thr Leu Val Pro Phe Cys	
[1545]		565 570 575
[1546]	Gln Lys Leu Leu Asp Ala Ser Leu Leu Thr Glu Ala Glu Arg Lys Trp	
[1547]		580 585 590
[1548]	Val Asn Asp Tyr His Ala Lys Val Trp Glu Lys Thr Ser Pro Phe Phe	
[1549]		595 600 605
[1550]	Glu Lys Asp Glu Leu Thr Thr Ala Trp Leu Lys Arg Glu Thr Gln Pro	
[1551]		610 615 620
[1552]	Ile	
[1553]	625	
[1554]	<210>25	
[1555]	<211>2344	
[1556]	<212>DNA	
[1557]	<213> 红色毛癣菌 (T. rubrum)	
[1558]	<400>25	
[1559]	atcaacctca cctcttcacc gtctcagcc cttcgtcccg tccaactctt catttcgccc	60
[1560]	tctctatgat aaccaacaaa catecgctgt tatgtaatcg aaccgcccgt tagccatccc	120
[1561]	tagccccgcg tttctccca gcatcaatac gaccgaaatg aagacagacg gggaagacga	180
[1562]	ggcaaaacaa taacacatca acaatttaac cgtttgccgt cttctaccca tcttgtctac	240
[1563]	gcatcgtcca accttttctt gccctatate agccgaacte ggccatcatg gatatccacg	300
[1564]	tcgacaaata cccggctaag agtcacgcca ggcgcgctgc cgagaagctc aaggccgcgg	360
[1565]	ggcacggctc taccggcatc atcttcgctg aaggccaaaa ggagcatatt atcgatgata	420
[1566]	gcgacgagcc gtttcacttc cggtagccg tgggaataca ctcgactggg cggaataagc	480
[1567]	taacaaaagg gtgtgatagt caacgccgaa acttctctta tctgtccggc tgtcttgagg	540
[1568]	ccgagtgctc cgttgcatac aacatcgaga aagatgagct tacattgttc attccaccag	600
[1569]	tcgaccagec ctcggttatg tggtcgggcc tccctcttga gcccgccgaa gccttgaagc	660
[1570]	agttcgatgt tgatgccgtg ctctcaciaa ctgagataaa caactatctc gcgaagtgtg	720
[1571]	ggggcgagaa ggtcttcacc attgcagaca gagtttgecc ggaggtctcc ttctcatcct	780
[1572]	tcaagcacia cgacaccgat gccctgaage ttgccatcga gtctcggcgt atagtgaaag	840

[1573]	acgagtatga aattggtctt ctccgacgtg ctaatgaggt ctccagccaa gctcatattg	900
[1574]	aagtgatgaa agccgcaacc aagtcaaaga acgagagaga gctctatgct actctcaact	960
[1575]	atgtctgcat gtctaattggc tgctccgacc agtcttacca tccaattctt gcatgtggcc	1020
[1576]	ccaatgctgc cactctccac tacaccaaga acaacggtga cctaactaac ceggctaccg	1080
[1577]	ggattaagga ccagctcgta cttatcgacg ctggatgcca gtacaaggcg tactgtgcag	1140
[1578]	atatcactcg tgcattcccc ttgtccggca aattcaccac ggagggccgc cagatctatg	1200
[1579]	atattgcctt ggagatgcag aaagtgcgct ttggcatgat caaacctaatt gttttgttcg	1260
[1580]	acgacatgca tgcctcggtc caccgggttg cgatcaaggg gctgctcaag attggcattc	1320
[1581]	tacttgctc tgaggatgag attttcgata agggaatcag cactgccttt tcccacatg	1380
[1582]	gtctaggcca ccatctcggc atggacactc acgatgttg aggaaacct aaccggctg	1440
[1583]	accgaatcg catgtttaa tacttgctc tgcgaggcac tgttcagag ggatccgtca	1500
[1584]	ttacaattga gcccggtgta agtgttgaat cgagtagttg ctccgccgaa tgtttcacat	1560
[1585]	acatttacta acccttgctc taggtctact tctgccgta catcattgag ccattcctta	1620
[1586]	ctaacccega gaccagcaag tacatcaact ccgaagtct agacaagtac tgggctgtg	1680
[1587]	gaggtgtacg tatcgaggac aacgtcgtc tccgcgcaa tggctttgag aacctgacca	1740
[1588]	cggtgccaaa ggagcccag gaggtcgaac gcattgtcca ggagggtgct aataattat	1800
[1589]	gtttttattc agtacaccga gtggtcggac acacgcagga gcatgtacat atttatgatc	1860
[1590]	taccagttg atttgctacc aaaaaagaac cgaccacagc cctatttatt gatattacat	1920
[1591]	agtaggaata aaggccactt tgcccaccgc gaataataac aataagaaaa gcaactactc	1980
[1592]	gtacaaccag cctagaaagc tctagacctc tttctcgtg ggcccttgaa tgccgggcta	2040
[1593]	ctggtgttat cacgtccct ggccctctc tcttcatgt ccaacaccg attaagcaaa	2100
[1594]	tcgaaactga actgggatg ctcaagacac aatgccttga actgctcttc agcatcatga	2160
[1595]	cgcagcacat cactcatctt agcccagaag cgagcaaccg gtctctgat agcagtgtct	2220
[1596]	tccggcgtgg tatggctgta cacgtatctc gcatactcga tctcaccgt agcactactc	2280
[1597]	tcgatgctac caatcttggt ctgagcaagc agtttgagtt tttcgtttcc gagcttttcg	2340
[1598]	gcca	2344
[1599]	<210>26	
[1600]	<211>1401	
[1601]	<212>DNA	
[1602]	<213> 红色毛癣菌 (T. rubrum)	
[1603]	<400>26	
[1604]	ccgaactcgg ccatcatgga tateccagtc gacaaatacc cggctaagag tcacgccagg	60
[1605]	cgcgtcgccg agaagctcaa ggccgcgggg cacggctcta ccggcatcat ctctcgtcga	120
[1606]	ggccaaaagg agcatattat cgatgatagc gacgagccgt ttcacttccg tcaacgccga	180
[1607]	aacttctct atctgtccgg ctgtcttgag gccgagtgt cegtgtcata caacatcgag	240
[1608]	aaagatgagc ttacattggt cattccacca gtcgaccag cctcggttat gtggctccggc	300
[1609]	ctccctcttg agcccgccga agccttgaag cagttcgatg ttgatgccgt gctcctcaca	360
[1610]	actgagataa acaactatct cgcgaagtgt gggggcgaga aggtcttcac cattgcagac	420
[1611]	agagtttgcc cggaggtctc cttctcatcc ttcaagcaca acgacaccga tgcctgaag	480



[1612]	cttgccatcg agtcctgccg tatagtghaaa gacgagtatg aaattgggtct tctccgacgt	540
[1613]	gctaattgagg tctccagcca agctcatatt gaagtgatga aagccgcaac caagtcaaag	600
[1614]	aacgagagag agctctatgc tactctcaac tatgtctgca tgtctaattgg ctgctccgac	660
[1615]	cagtcttacc atccaattct tgcattgtggc cccaatgctg ccaactctcca ctacaccaag	720
[1616]	aacaacggtg acctaactaa cccggctacc gggattaagg accagctcgt acttatcgac	780
[1617]	gctggatgcc agtacaaggc gtactgtgca gatatactc gtgcattccc cttgtccggc	840
[1618]	aaattcacca cggagggccg ccagatctat gatattgcct tggagatgca gaaagtcgcg	900
[1619]	tttggcatga tcaaacctaa tgttttgttc gacgacatgc atgctgcggt ccaccgggtt	960
[1620]	gcatcaagg ggctgctcaa gattggcatt ctactggct ctgaggatga gattttcgat	1020
[1621]	aagggaatca gcaactgctt tttccacat ggtctaggcc accatctcgg catggacact	1080
[1622]	cacgatgttg gagaaacce taaccggct gaccgaatc gcatgtttaa atacttgcgt	1140
[1623]	ctgagaggca ctgtccaga gggatccgct attaccaattg agcccgggtg ctacttctgc	1200
[1624]	cgttacatca ttgagccatt ccttactaac cccgagacca gcaagtacat caactccgaa	1260
[1625]	gttctagaca agtactgggc tgttggaggt gtacgtatcg aggacaacgt cgtcgtccgc	1320
[1626]	gccaatggct ttgagaacct gaccacggtg ccaaggagc cggaggaggt cgaacgcatt	1380
[1627]	gtccaggagg gtgctaaata a	1401
[1628]	<210>27	
[1629]	<211>466	
[1630]	<212>PRT	
[1631]	<213> 红色毛癣菌 (T. rubrum)	
[1632]	<400>27	
[1633]	Pro Asn Ser Ala Ile Met Asp Ile His Val Asp Lys Tyr Pro Ala Lys	
[1634]	1                    5                    10                    15	
[1635]	Ser His Ala Arg Arg Val Ala Glu Lys Leu Lys Ala Ala Gly His Gly	
[1636]	20                    25                    30	
[1637]	Ser Thr Gly Ile Ile Phe Val Glu Gly Gln Lys Glu His Ile Ile Asp	
[1638]	35                    40                    45	
[1639]	Asp Ser Asp Glu Pro Phe His Phe Arg Gln Arg Arg Asn Phe Leu Tyr	
[1640]	50                    55                    60	
[1641]	Leu Ser Gly Cys Leu Glu Ala Glu Cys Ser Val Ala Tyr Asn Ile Glu	
[1642]	65                    70                    75                    80	
[1643]	Lys Asp Glu Leu Thr Leu Phe Ile Pro Pro Val Asp Pro Ala Ser Val	
[1644]	85                    90                    95	
[1645]	Met Trp Ser Gly Leu Pro Leu Glu Pro Ala Glu Ala Leu Lys Gln Phe	
[1646]	100                    105                    110	
[1647]	Asp Val Asp Ala Val Leu Leu Thr Thr Glu Ile Asn Asn Tyr Leu Ala	
[1648]	115                    120                    125	
[1649]	Lys Cys Gly Gly Glu Lys Val Phe Thr Ile Ala Asp Arg Val Cys Pro	
[1650]	130                    135                    140	

[1651]	Glu Val Ser Phe Ser Ser Phe Lys His Asn Asp Thr Asp Ala Leu Lys
[1652]	145 150 155 160
[1653]	Leu Ala Ile Glu Ser Cys Arg Ile Val Lys Asp Glu Tyr Glu Ile Gly
[1654]	165 170 175
[1655]	Leu Leu Arg Arg Ala Asn Glu Val Ser Ser Gln Ala His Ile Glu Val
[1656]	180 185 190
[1657]	Met Lys Ala Ala Thr Lys Ser Lys Asn Glu Arg Glu Leu Tyr Ala Thr
[1658]	195 200 205
[1659]	Leu Asn Tyr Val Cys Met Ser Asn Gly Cys Ser Asp Gln Ser Tyr His
[1660]	210 215 220
[1661]	Pro Ile Leu Ala Cys Gly Pro Asn Ala Ala Thr Leu His Tyr Thr Lys
[1662]	225 230 235 240
[1663]	Asn Asn Gly Asp Leu Thr Asn Pro Ala Thr Gly Ile Lys Asp Gln Leu
[1664]	245 250 255
[1665]	Val Leu Ile Asp Ala Gly Cys Gln Tyr Lys Ala Tyr Cys Ala Asp Ile
[1666]	260 265 270
[1667]	Thr Arg Ala Phe Pro Leu Ser Gly Lys Phe Thr Thr Glu Gly Arg Gln
[1668]	275 280 285
[1669]	Ile Tyr Asp Ile Ala Leu Glu Met Gln Lys Val Ala Phe Gly Met Ile
[1670]	290 295 300
[1671]	Lys Pro Asn Val Leu Phe Asp Asp Met His Ala Ala Val His Arg Val
[1672]	305 310 315 320
[1673]	Ala Ile Lys Gly Leu Leu Lys Ile Gly Ile Leu Thr Gly Ser Glu Asp
[1674]	325 330 335
[1675]	Glu Ile Phe Asp Lys Gly Ile Ser Thr Ala Phe Phe Pro His Gly Leu
[1676]	340 345 350
[1677]	Gly His His Leu Gly Met Asp Thr His Asp Val Gly Gly Asn Pro Asn
[1678]	355 360 365
[1679]	Pro Ala Asp Pro Asn Arg Met Phe Lys Tyr Leu Arg Leu Arg Gly Thr
[1680]	370 375 380
[1681]	Val Pro Glu Gly Ser Val Ile Thr Ile Glu Pro Gly Val Tyr Phe Cys
[1682]	385 390 395 400
[1683]	Arg Tyr Ile Ile Glu Pro Phe Leu Thr Asn Pro Glu Thr Ser Lys Tyr
[1684]	405 410 415
[1685]	Ile Asn Ser Glu Val Leu Asp Lys Tyr Trp Ala Val Gly Gly Val Arg
[1686]	420 425 430
[1687]	Ile Glu Asp Asn Val Val Val Arg Ala Asn Gly Phe Glu Asn Leu Thr
[1688]	435 440 445
[1689]	Thr Val Pro Lys Glu Pro Glu Glu Val Glu Arg Ile Val Gln Glu Gly

[1690]	450	455	460	
[1691]	Ala Lys			
[1692]	465			
[1693]	<210>28			
[1694]	<211>1730			
[1695]	<212>DNA			
[1696]	<213> 犬小孢子菌 (Microsporium canis)			
[1697]	<400>28			
[1698]	atgaagacac agttgttgag tctgggagtt gccctcacgg ccatctctca gggcgttatt			60
[1699]	gctgaggatg ccttgaactg gccattcaag ccgttggta atgctgtgag tatatacaca			120
[1700]	agatcgatcg atcgtcctct tgtccctgtc acttatcgct ctacagtaag caaaaatact			180
[1701]	ggagaatcat gtgctgatgt aatgtatag gatgacctgc aaaacaagat taagctcaag			240
[1702]	gatcttatgg ctggcgta caaactcaa gacttcgect acgctcacc tgagaagaat			300
[1703]	cgagtattcg gtgggtgctgg ccacaaggat accgctgact ggatctacaa tgagctcaag			360
[1704]	gctaccggct actacgatgt gaagatgcag ccacaagtc acctgtggc tcatgctgag			420
[1705]	gcagctgtca atgccaatgg caaggatctc actgccagtg ccatgtccta cagccctca			480
[1706]	gccgacaaga tcaactgccg gcttgtcctg gccagaaca tgggatgcaa tgctgtatgt			540
[1707]	gcgccccctt tccattctat atategactg gtcgcttga aattcagaag agctgacaat			600
[1708]	tgcaaacaga ctgattacc agagggtacc aagggaaga ttgtcctcat cgagcgtggt			660
[1709]	gtctgcagct ttggcgagaa gtccgctcag gctggcgat caaaggctat tgggtccatc			720
[1710]	gtctacaaca acgtccctgg aagcttggcc ggcaccctgg gtggccttga caaccgcat			780
[1711]	gctccaactg ctggaatctc tcaggctgat ggaaagaacc tcgctagcct tgtegcctct			840
[1712]	ggcaaggta cgcaccat gaacgttacc agcaagttg agaacaggac tacgtgagta			900
[1713]	ttgttccata ctttggta caatgatata tacacgtact aacactgctc tatagctgga			960
[1714]	acgtcattgc cgagaccaag ggaggagacc acaacaacgt catcatgctc gtttctact			1020
[1715]	ctgactctgt cgacgccggc cctggtatca acgacaacgg ctccgggtacc attggtatca			1080
[1716]	tgaccgttgc caaagccctc accaacttca aggtcaacaa cgccgtccgc ttcggctggt			1140
[1717]	ggaccgccga ggagtccggc cttctcggca gcactttcta cgtegacagc cttgacgacc			1200
[1718]	gtgaactgca caaggtaag ctgtacctca acttcgacat gattggctcc cccaacttcg			1260
[1719]	ccaaccagat ctacgacgga gacggctccg cctacaacat gactggcccc gccggatctg			1320
[1720]	ctgaaatcga gtacctgtc gagaagttct tcgatgacca gggaaatccca caccagccca			1380
[1721]	ccgccttacc cggccgtcc gactactctg ctttcaata gcgcaacgct cctgccggag			1440
[1722]	gtctgtttac tgggtctgag gtcgtcaaga ccgccgagca ggctaagcta tttggcggcg			1500
[1723]	aggctggcgt tgcttatgac aagaactacc acggcaaggg cgacactgta gacaacatca			1560
[1724]	acaagggtgc tatctacct aacactcgag gaatcgcgta tgccactgct cagtatgcta			1620
[1725]	gttcgctgag cggattccca acccgcccaa agacgggtaa gcgtgacgtg agccccctg			1680
[1726]	gccagtctat gctggtggt ggatgcggac accacagcgt cttcatgtaa			1730
[1727]	<210>29			
[1728]	<211>1488			

[1729]	<212>DNA		
[1730]	<213> 犬小孢子菌 ( <i>Microsporum canis</i> )		
[1731]	<400>29		
[1732]	atgaagacac agttgttgag tctgggagtt gccctcacgg ccattctctca gggcgttatt	60	
[1733]	gctgaggatg ccttgaactg gccattcaag ccggttggtta atgctgatga cctgcaaaac	120	
[1734]	aagattaagc tcaaggatct tatggctggc gtacagaaac tccaagactt cgcctacgct	180	
[1735]	caccctgaga agaatcgagt attcggtggt gctggccaca aggataccgt cgactggatc	240	
[1736]	tacaatgagc tcaaggctac cggctactac gatgtgaaga tgcagccaca agtccacctg	300	
[1737]	tggtctcatg ctgaggcagc tgtcaatgcc aatggcaagg atctcactgc cagtgccatg	360	
[1738]	tcctacagcc ctccagccga caagatcact gccgagcttg tcctggccaa gaacatggga	420	
[1739]	tgcaatgcta ctgattacc agagggtacc aagggaaga ttgtcctcat cgagcgtggt	480	
[1740]	gtctgcagct ttggcgagaa gtccgctcag gctggcgatg caaaggctat tgggtccatc	540	
[1741]	gtctacaaca acgtccctgg aagcttggcc ggcaccctgg gtggccttga caaccgcat	600	
[1742]	gctccaactg ctggaatctc tcaggctgat ggaaagaacc tcgctagcct tgctgcctct	660	
[1743]	ggcaaggtta ccgtcacct gaacgttata agcaagtttg agaacaggac tacctggaac	720	
[1744]	gtcattgccg agaccaaggg aggagaccac aacaacgtca tcatgctcgg ttctcactct	780	
[1745]	gactctgtcg acgccggccc tggatatcaac gacaacggct ccggtacat tggatatcatg	840	
[1746]	accgttgcca aagccctcac caacttcaag gtcaacaacg ccgtccgctt cggctggtgg	900	
[1747]	accgccgagg agttcggcct tctcggcagc actttctacg tcgacagcct tgacgaccgt	960	
[1748]	gaactgcaca aggtcaagct gtacctcaac ttgacatga ttggtcccc caacttcgcc	1020	
[1749]	aaccagatct acgacggaga cggtccgcc tacaacatga ctggccccgc cggatctgct	1080	
[1750]	gaaatcgagt acctgttcga gaagttcttc gatgaccagg gaateccaca ccagcccacc	1140	
[1751]	gccttcaccg gccgtccga ctactctgcc ttcatcaagc gcaacgtccc tgccggagg	1200	
[1752]	ctgtttactg gtgctgagg	1260	
[1753]	ctgctgagg	1320	
[1754]	ctgctgagg	1380	
[1755]	ctgctgagg	1440	
[1756]	ctgctgagg	1488	
[1757]	<210>30		
[1758]	<211>495		
[1759]	<212>PRT		
[1760]	<213> 犬小孢子菌 ( <i>Microsporum canis</i> )		
[1761]	<400>30		
[1762]	Met Lys Thr Gln Leu Leu Ser Leu Gly Val Ala Leu Thr Ala Ile Ser		
[1763]	1 5 10 15		
[1764]	Gln Gly Val Ile Ala Glu Asp Ala Leu Asn Trp Pro Phe Lys Pro Leu		
[1765]	20 25 30		
[1766]	Val Asn Ala Asp Asp Leu Gln Asn Lys Ile Lys Leu Lys Asp Leu Met		
[1767]	35 40 45		

[1768]	Ala Gly Val Gln Lys Leu Gln Asp Phe Ala Tyr Ala His Pro Glu Lys
[1769]	50 55 60
[1770]	Asn Arg Val Phe Gly Gly Ala Gly His Lys Asp Thr Val Asp Trp Ile
[1771]	65 70 75 80
[1772]	Tyr Asn Glu Leu Lys Ala Thr Gly Tyr Tyr Asp Val Lys Met Gln Pro
[1773]	85 90 95
[1774]	Gln Val His Leu Trp Ser His Ala Glu Ala Ala Val Asn Ala Asn Gly
[1775]	100 105 110
[1776]	Lys Asp Leu Thr Ala Ser Ala Met Ser Tyr Ser Pro Pro Ala Asp Lys
[1777]	115 120 125
[1778]	Ile Thr Ala Glu Leu Val Leu Ala Lys Asn Met Gly Cys Asn Ala Thr
[1779]	130 135 140
[1780]	Asp Tyr Pro Glu Gly Thr Lys Gly Lys Ile Val Leu Ile Glu Arg Gly
[1781]	145 150 155 160
[1782]	Val Cys Ser Phe Gly Glu Lys Ser Ala Gln Ala Gly Asp Ala Lys Ala
[1783]	165 170 175
[1784]	Ile Gly Ala Ile Val Tyr Asn Asn Val Pro Gly Ser Leu Ala Gly Thr
[1785]	180 185 190
[1786]	Leu Gly Gly Leu Asp Asn Arg His Ala Pro Thr Ala Gly Ile Ser Gln
[1787]	195 200 205
[1788]	Ala Asp Gly Lys Asn Leu Ala Ser Leu Val Ala Ser Gly Lys Val Thr
[1789]	210 215 220
[1790]	Val Thr Met Asn Val Ile Ser Lys Phe Glu Asn Arg Thr Thr Trp Asn
[1791]	225 230 235 240
[1792]	Val Ile Ala Glu Thr Lys Gly Gly Asp His Asn Asn Val Ile Met Leu
[1793]	245 250 255
[1794]	Gly Ser His Ser Asp Ser Val Asp Ala Gly Pro Gly Ile Asn Asp Asn
[1795]	260 265 270
[1796]	Gly Ser Gly Thr Ile Gly Ile Met Thr Val Ala Lys Ala Leu Thr Asn
[1797]	275 280 285
[1798]	Phe Lys Val Asn Asn Ala Val Arg Phe Gly Trp Trp Thr Ala Glu Glu
[1799]	290 295 300
[1800]	Phe Gly Leu Leu Gly Ser Thr Phe Tyr Val Asp Ser Leu Asp Asp Arg
[1801]	305 310 315 320
[1802]	Glu Leu His Lys Val Lys Leu Tyr Leu Asn Phe Asp Met Ile Gly Ser
[1803]	325 330 335
[1804]	Pro Asn Phe Ala Asn Gln Ile Tyr Asp Gly Asp Gly Ser Ala Tyr Asn
[1805]	340 345 350
[1806]	Met Thr Gly Pro Ala Gly Ser Ala Glu Ile Glu Tyr Leu Phe Glu Lys

[1807]	355	360	365
[1808]	Phe Phe Asp Asp Gln Gly Ile Pro His Gln Pro Thr Ala Phe Thr Gly		
[1809]	370	375	380
[1810]	Arg Ser Asp Tyr Ser Ala Phe Ile Lys Arg Asn Val Pro Ala Gly Gly		
[1811]	385	390	395
[1812]	Leu Phe Thr Gly Ala Glu Val Val Lys Thr Ala Glu Gln Ala Lys Leu		
[1813]	405	410	415
[1814]	Phe Gly Gly Glu Ala Gly Val Ala Tyr Asp Lys Asn Tyr His Gly Lys		
[1815]	420	425	430
[1816]	Gly Asp Thr Val Asp Asn Ile Asn Lys Gly Ala Ile Tyr Leu Asn Thr		
[1817]	435	440	445
[1818]	Arg Gly Ile Ala Tyr Ala Thr Ala Gln Tyr Ala Ser Ser Leu Arg Gly		
[1819]	450	455	460
[1820]	Phe Pro Thr Arg Pro Lys Thr Gly Lys Arg Asp Val Ser Pro Arg Gly		
[1821]	465	470	475
[1822]	Gln Ser Met Pro Gly Gly Gly Cys Gly His His Ser Val Phe Met		
[1823]	485	490	495
[1824]	<210>31		
[1825]	<211>1775		
[1826]	<212>DNA		
[1827]	<213>须癣毛癣菌 (T. mentagrophytes)		
[1828]	<400>31		
[1829]	atgaagtcgc aactgttgag cctagccgtg gccgtcacca ccatttccca gggcgttggt	60	
[1830]	ggcaagagc cctttggatg gcccttcaag cctatggatc ctcaggtgag ttgctgtcaa	120	
[1831]	cagatcgatc gatcgatcta ccttcgtccc tgccacctat aactccacag caggaccaag	180	
[1832]	aaaacacaag ttttccgggg aattcttatg tgctgatgta aatgtatagg atgacctgca	240	
[1833]	aaacaagatt aagctcaagg atatcatggc aggtgtcgag aagctgcaaa gcttttctga	300	
[1834]	tgctcatcct gaaaagaacc gagggttcgg tggtaatggc cacaaggaca ctgtcgagtg	360	
[1835]	gatctacaat gagctcaagg ccaccggcta ctacaatgtg aagaagcagg agcaggtaca	420	
[1836]	cctgtggtct cacgctgagg ccgctctcag tgccaatggc aaggacctca aggccagcgc	480	
[1837]	catgtcgtac agccctcctg ccaacaagat catggccgag cttgtcgttg ccaagaacaa	540	
[1838]	tggtgcaat gctgtaagtg ccatacactt cctatacacc acattcactt tagaatgaag	600	
[1839]	agcgcgggag aactgatttt tttttttttt tttttttttt tgtaacagac cgattacca	660	
[1840]	gagaacactc agggaaagat agtctcatt cagcgtgggtg tctgcagctt cggcgagaag	720	
[1841]	tcttctcagg ctggtgatgc gaaggctatt ggtgccgttg tctacaacaa cgccccgga	780	
[1842]	tcctttgctg gcactcttgg tgcccttgac aagcgcctat tcccaaccgc tggcttttcc	840	
[1843]	caggaggatg gaaagaatct tgctagcctc gttgcttctg gcaaggttga tgccacctg	900	
[1844]	aacgtttgca gtctgtttga gaaccgaacc acgtaagtaa ctcaacgtca tatccagcat	960	
[1845]	taatcttcag gaggatata actaattegg tatctcacag ctggaacgtc attgctgaga	1020	

[1846]	ccaagggagg agaccacaac aatgttgtca tgcttgggtgc tcaactccgac tccgtcgatg	1080
[1847]	ccggccccgg tatcaacgac aacggctccg gctccattgg tatcatgacc gttgccaaag	1140
[1848]	cccttactaa cttcaagctc aacaacgccg ttcgctttgc ctggtggacc gctgaggaat	1200
[1849]	tcggtctcct tggaagcacc ttctacgtcg acagccttga tgaccgtgag ctgcacaagg	1260
[1850]	tcaagctgta cctcaacttc gacatgatcg gctctcccaa cttcgccaac cagatctacg	1320
[1851]	acggtgacgg ttcggcctac aacatgactg gtcccgtgg ctctgctgaa atcgagtacc	1380
[1852]	tgttcgagaa gttctttgac gaccagggtc tcccacacca gccactgcc ttcaccggcc	1440
[1853]	gatccgacta ctctgcattc atcaagegca acgtccccgc tggaggtctt tteactggtg	1500
[1854]	ccgaggttgt caagaccccc gagcaagtta agctgttcgg tggtaggct ggcgttgcct	1560
[1855]	atgacaagaa ctacatggc aagggtgaca ccgttgccaa catcaacaag ggagctatct	1620
[1856]	tccttaacac tcgagcaatc gcctactctg tggccgagta tgctcgatcc ctcaagggt	1680
[1857]	tccaacccg ccaaagacc ggcaagcgtg ccgtcaacc tcagtatgct aagatgcctg	1740
[1858]	gtggtggttg cggacaccac actgtcttca tgtaa	1775
[1859]	<210>32	
[1860]	<211>1488	
[1861]	<212>DNA	
[1862]	<213>须癣毛癣菌 (T. mentagrophytes)	
[1863]	<400>32	
[1864]	atgaagtcgc aactgttgag cctagccgtg gccgtcacca ccatttccca gggcgttgtt	60
[1865]	ggtcaagagc cttttggatg gcccttcaag cctatgggtca ctcaggatga cctgcaaaac	120
[1866]	aagattaagc tcaaggatat catggcaggt gtcgagaagc tgcaaagctt ttctgatgct	180
[1867]	catcctgaaa agaaccgagt gtccggtggt aatggccaca aggacactgt cgagtggatc	240
[1868]	tacaatgagc tcaaggccac cggctactac aatgtgaaga agcaggagca ggtacacctg	300
[1869]	tggtctcacg ctgaggccgc tctcagtgcc aatggcaagg acctcaagge cagegccatg	360
[1870]	tcgtacagcc ctctgccaa caagatcatg gccgagcttg tcgttgccaa gaacaatggc	420
[1871]	tgcaatgcta ccgattacce agagaacact cagggaaga tagtctcat tcagcgtggt	480
[1872]	gtctgcagct tcggcgagaa gtcttctcag gctggtgatg cgaaggctat tgggtccgtt	540
[1873]	gtctacaaca acgtccccgg atcccttgc tggcactcttg gtggccttga caagcgcct	600
[1874]	gtcccaaccg ctggtctttc ccaggaggat ggaaagaate ttgctagcct cgttgcttct	660
[1875]	ggcaaggttg atgtcacct gaacgttgte agtctgtttg agaaccgaac cacctggaac	720
[1876]	gtcattgctg agaccaagg aggagaccac aacaatggtg tcatgcttgg tgctcactcc	780
[1877]	gactccgtcg atgccggccc cggtatcaac gacaacggct ccggtccat tggatatcatg	840
[1878]	accgttgcca aagcccttac taacttcaag ctcaacaac ccgttcgctt tgccgtggtg	900
[1879]	accgctgagg aattcggctt ctttgaagc accttctacg tcgacagcct tgatgacctg	960
[1880]	gagctgcaca aggtcaagct gtacctcaac ttcgacatga teggtctcc caacttcgcc	1020
[1881]	aaccagatct acgacggtga cggttcggcc tacaacatga ctgggtccgc tggtctgct	1080
[1882]	gaaatcgagt acctgttcga gaagtctttt gacgaccagg gtctcccaca ccageccaet	1140
[1883]	gccttaccg gccgatccga ctactctgca tteatcaagc gcaacgtccc cgtggaggt	1200
[1884]	cttttactg gtgccaggt tgtcaagacc cccgagcaag ttaagctggt cgggtggtgag	1260

[1885]	gctggcgttg cctatgacaa gaactacat ggcaagggtg acaccgttgc caacatcaac	1320
[1886]	aaggagcta tcttccttaa cactcgagca atcgctact ctgtggccga gtatgctcga	1380
[1887]	tccctcaagg gcttccaac ccgcccagg accggcaagc gtgccgtcaa ccctcagtat	1440
[1888]	gctaagatgc ctggtggtg ttgcggacac cacactgtct tcatgtaa	1488
[1889]	<210>33	
[1890]	<211>495	
[1891]	<212>PRT	
[1892]	<213>须癣毛癣菌 (T. mentagrophytes)	
[1893]	<400>33	
[1894]	Met Lys Ser Gln Leu Leu Ser Leu Ala Val Ala Val Thr Thr Ile Ser	
[1895]	1 5 10 15	
[1896]	Gln Gly Val Val Gly Gln Glu Pro Phe Gly Trp Pro Phe Lys Pro Met	
[1897]	20 25 30	
[1898]	Val Thr Gln Asp Asp Leu Gln Asn Lys Ile Lys Leu Lys Asp Ile Met	
[1899]	35 40 45	
[1900]	Ala Gly Val Glu Lys Leu Gln Ser Phe Ser Asp Ala His Pro Glu Lys	
[1901]	50 55 60	
[1902]	Asn Arg Val Phe Gly Gly Asn Gly His Lys Asp Thr Val Glu Trp Ile	
[1903]	65 70 75 80	
[1904]	Tyr Asn Glu Leu Lys Ala Thr Gly Tyr Tyr Asn Val Lys Lys Gln Glu	
[1905]	85 90 95	
[1906]	Gln Val His Leu Trp Ser His Ala Glu Ala Ala Leu Ser Ala Asn Gly	
[1907]	100 105 110	
[1908]	Lys Asp Leu Lys Ala Ser Ala Met Ser Tyr Ser Pro Pro Ala Asn Lys	
[1909]	115 120 125	
[1910]	Ile Met Ala Glu Leu Val Val Ala Lys Asn Asn Gly Cys Asn Ala Thr	
[1911]	130 135 140	
[1912]	Asp Tyr Pro Glu Asn Thr Gln Gly Lys Ile Val Leu Ile Gln Arg Gly	
[1913]	145 150 155 160	
[1914]	ValCys Ser Phe Gly Glu Lys Ser Ser Gln Ala Gly Asp Ala Lys Ala	
[1915]	165 170 175	
[1916]	Ile Gly Ala Val Val Tyr Asn Asn Val Pro Gly Ser Leu Ala Gly Thr	
[1917]	180 185 190	
[1918]	Leu Gly Gly Leu Asp Lys Arg His Val Pro Thr Ala Gly Leu Ser Gln	
[1919]	195 200 205	
[1920]	Glu Asp Gly Lys Asn Leu Ala Ser Leu Val Ala Ser Gly Lys Val Asp	
[1921]	210 215 220	
[1922]	Val Thr Met Asn Val Val Ser Leu Phe Glu Asn Arg Thr Thr Trp Asn	
[1923]	225 230 235 240	



[1924]	Val Ile Ala Glu Thr Lys Gly Gly Asp His Asn Asn Val Val Met Leu		
[1925]		245	250 255
[1926]	Gly Ala His Ser Asp Ser Val Asp Ala Gly Pro Gly Ile Asn Asp Asn		
[1927]		260	265 270
[1928]	Gly Ser Gly Ser Ile Gly Ile Met Thr Val Ala Lys Ala Leu Thr Asn		
[1929]		275	280 285
[1930]	Phe Lys Leu Asn Asn Ala Val Arg Phe Ala Trp Trp Thr Ala Glu Glu		
[1931]		290	295 300
[1932]	Phe Gly Leu Leu Gly Ser Thr Phe Tyr Val Asp Ser Leu Asp Asp Arg		
[1933]		305	310 315 320
[1934]	Glu Leu His Lys Val Lys Leu Tyr Leu Asn Phe Asp Met Ile Gly Ser		
[1935]		325	330 335
[1936]	Pro Asn Phe Ala Asn Gln Ile Tyr Asp Gly Asp Gly Ser Ala Tyr Asn		
[1937]		340	345 350
[1938]	Met Thr Gly Pro Ala Gly Ser Ala Glu Ile Glu Tyr Leu Phe Glu Lys		
[1939]		355	360 365
[1940]	Phe Phe Asp Asp Gln Gly Leu Pro His Gln Pro Thr Ala Phe Thr Gly		
[1941]		370	375 380
[1942]	Arg Ser Asp Tyr Ser Ala Phe Ile Lys Arg Asn Val Pro Ala Gly Gly		
[1943]		385	390 395 400
[1944]	Leu Phe Thr Gly Ala Glu Val Val Lys Thr Pro Glu Gln Val Lys Leu		
[1945]		405	410 415
[1946]	Phe Gly Gly Glu Ala Gly Val Ala Tyr Asp Lys Asn Tyr His Gly Lys		
[1947]		420	425 430
[1948]	Gly Asp Thr Val Ala Asn Ile Asn Lys Gly Ala Ile Phe Leu Asn Thr		
[1949]		435	440 445
[1950]	Arg Ala Ile Ala Tyr Ser Val Ala Glu Tyr Ala Arg Ser Leu Lys Gly		
[1951]		450	455 460
[1952]	Phe Pro Thr Arg Pro Lys Thr Gly Lys Arg Ala Val Asn Pro Gln Tyr		
[1953]		465	470 475 480
[1954]	Ala Lys Met Pro Gly Gly Gly Cys Gly His His Thr Val Phe Met		
[1955]		485	490 495
[1956]	<210>34		
[1957]	<211>2328		
[1958]	<212>DNA		
[1959]	<213> 红色毛癣菌 (T. rubrum)		
[1960]	<400>34		
[1961]	atgaagetcc tctcgctact tatgctggcg ggcacgccc aagccatcgt tctcctcgt		60
[1962]	gagcccggtt caccaactgg tggeggcaac aagctgttga cctacaagga gtgtgtcct		120

[1963]	agagctacta tctctccaag gtcgacgtcc cttgcctgga ttaacagtga agaagatggc	180
[1964]	cggtacatct cccagtcgca cgatggagca ttgatcctcc agaacatcgt cacgaacacc	240
[1965]	aacaagactc tcgtggccgc agacaaggta cccaagggtt actatgacta ctggttcaag	300
[1966]	ccagaccttt ctgctgtctt atgggcaacc aattacacca agcagtaccg tcaactttac	360
[1967]	tttgccaact acttcattct agacatcaaa aagggatcgt tgaccctct agcccaggac	420
[1968]	caggctggtg acatccagta tgctcaatgg agcccatga acaactctat cgcctatgtc	480
[1969]	cgtggaaacg acctgtatat ctggaacaat ggcaagacca agcgtattac cgaaaatggc	540
[1970]	ggccccgata tcttcaatgg tgctccctgac tgggtatacg aggaagaaat cttcggggac	600
[1971]	cggttcgctc tttggtctc acctgacggt gaataccttg cgtacctccg ctttaacgag	660
[1972]	actggagtcc cgacctacac tattccgtac tacaagaaca agcaaaagat tgcccctgcc	720
[1973]	taccaaggg agctggagat ccgttacct aaagtctctg cgaagaacce aaccgtgcag	780
[1974]	ttccacctgt taaacattgc ttcateccag gagacaacta tcccagttac tgcgttcccg	840
[1975]	gaaaacgatc ttgtgatcgg tgaggttgct tggctcagca gtggccatga tagtgtagca	900
[1976]	tatcgtgctt tcaacctgt ccaggataga gaaaagattg tcagcgtcaa ggttgagtcc	960
[1977]	aaggaatcca aggtattcg cgaaagagat ggcaccgacg gctggatcga caacctctc	1020
[1978]	tccatgtcat atatcggaac cgtaaacggc aaggagtact acgtcgatat atctgatgct	1080
[1979]	tctggctggg cacatatcta cctctaccg gttgatggag gaaaggagat tgactaaca	1140
[1980]	aagggagaat ggaagtcgt tgccattctc aaggttgaca cgaagaagaa gctgatctac	1200
[1981]	ttcacctcta ccaaatatca cagcaccact cgacacgtct actctgtctc gtatgacaca	1260
[1982]	aaggtcatga ccctctcgt caacgataag gaggtgcgt actacactgc atccttctcg	1320
[1983]	gccaagggtg gttactatat cttgtcctac caaggcca atgttccata ccaagaactt	1380
[1984]	tactccacca aggacagtaa gaagcctctc aagacaatca ctagcaatga tgcattgctc	1440
[1985]	gagaagctga aggagtacaa gctccccaa gttagcttct ttgagatcaa gttccatct	1500
[1986]	ggtgaaacce ttaatgttaa gcaacgccta ccacctaact tcaaccaca caagaagtac	1560
[1987]	cccgtctct tcaactccgta tgggtggcct ggtgccaag aggtaagcca ggcatggaat	1620
[1988]	tcattggact tcaagtccta cattacatct gacctgac ttgaatacgt tacctggact	1680
[1989]	gttgacaacc gtggaaccgg ctacaagggc cgcaagtcc gcagcgccgt agctaagcgt	1740
[1990]	ctcggtttc tcgaagccca ggaccaggtc tttgctgcta aggagggtgct gaaaaaccgt	1800
[1991]	tgggctgata aggaccatat tggaatctgg ggtggagct atggcggctt cctgaccgt	1860
[1992]	aagaccctcg agaccgacag tgggtgtattc acttttgta tcagtactgc tctgtctct	1920
[1993]	gatttcagac tctacgacag catgtacact gagcgttaca tgaagaccgt tgaactaaac	1980
[1994]	gctgacggt acagtgagac cgccgtgcac aaggttgatg gctttaagaa cctcaaaggt	2040
[1995]	cattacttea tccagcatgg aaccggtgac gacaacgtcc acttccaaaa cgccgtgtc	2100
[1996]	ctttccaaca cctgatgaa cggcgggtgta actgcagaca agttgactac tcagtggttt	2160
[1997]	actgactcgg accacggcat cagatacgt atggactcca cttaccagta caagcagctt	2220
[1998]	tctaagatgg tctacgacca gaagcaacga aggccagaaa gcccaccaat gcaccaatgg	2280
[1999]	agcaagagag ttttggctgc cctgtttggt gagagggcag aggaatga	2328
[2000]	<210>35	
[2001]	<211>775	

[2002] <212>PRT  
 [2003] <213> 红色毛癣菌 (*T. rubrum*)  
 [2004] <220>  
 [2005] <221>MISC FEATURE  
 [2006] <222>(162).. (612)  
 [2007] <223>X  
 [2008] <400>35  
 [2009] Met Lys Leu Leu Ser Leu Leu Met Leu Ala Gly Ile Ala Gln Ala Ile  
 [2010] 1 5 10 15  
 [2011] Val Pro Pro Arg Glu Pro Arg Ser Pro Thr Gly Gly Gly Asn Lys Leu  
 [2012] 20 25 30  
 [2013] Leu Thr Tyr Lys Glu Cys Val ProArg Ala Thr Ile Ser Pro Arg Ser  
 [2014] 35 40 45  
 [2015] Thr Ser Leu Ala Trp Ile Asn Ser Glu Glu Asp Gly Arg Tyr Ile Ser  
 [2016] 50 55 60  
 [2017] Gln Ser Asp Asp Gly Ala Leu Ile Leu Gln Asn Ile Val Thr Asn Thr  
 [2018] 65 70 75 80  
 [2019] Asn Lys Thr Leu Val Ala Ala Asp Lys Val Pro Lys Gly Tyr Tyr Asp  
 [2020] 85 90 95  
 [2021] Tyr Trp Phe Lys Pro Asp Leu Ser Ala Val Leu Trp Ala Thr Asn Tyr  
 [2022] 100 105 110  
 [2023] Thr Lys Gln Tyr Arg His Ser Tyr Phe Ala Asn Tyr Phe Ile Leu Asp  
 [2024] 115 120 125  
 [2025] Ile Lys Lys Gly Ser Leu Thr Pro Leu Ala Gln Asp Gln Ala Gly Asp  
 [2026] 130 135 140  
 [2027] Ile Gln Tyr Ala Gln Trp Ser Pro Met Asn Asn Ser Ile Ala Tyr Val  
 [2028] 145 150 155 160  
 [2029] Arg Xaa Asn Asp Leu Tyr Ile Trp Asn Asn Gly Lys Thr Lys Arg Ile  
 [2030] 165 170 175  
 [2031] Thr Glu Asn Gly Gly Pro Asp Ile Phe Asn Gly Val Pro Asp Trp Val  
 [2032] 180 185 190  
 [2033] Tyr Glu Glu Glu Ile Phe Gly Asp Arg Phe Ala Leu Trp Phe Ser Pro  
 [2034] 195 200 205  
 [2035] Asp Gly Glu Tyr Leu Ala Tyr Leu Arg Phe Asn Glu Thr Gly Val Pro  
 [2036] 210 215 220  
 [2037] Thr Tyr Thr Ile Pro Tyr Tyr Lys Asn Lys Gln Lys Ile Ala Pro Ala  
 [2038] 225 230 235 240  
 [2039] Tyr Pro Arg Glu Leu Glu Ile Arg Tyr Pro Lys Val Ser Ala Lys Asn  
 [2040] 245 250 255

[2041]	Pro Thr Val Gln Phe His Leu Leu Asn Ile Ala Ser Ser Gln Glu Thr
[2042]	260 265 270
[2043]	Thr Ile Pro Val Thr Ala Phe Pro Glu Asn Asp Leu Val Ile Gly Glu
[2044]	275 280 285
[2045]	Val Ala Trp Leu Ser Ser Gly His Asp Ser Val Ala Tyr Arg Ala Phe
[2046]	290 295 300
[2047]	Asn Arg Val Gln Asp Arg Glu Lys Ile Val Ser Val Lys Val Glu Ser
[2048]	305 310 315 320
[2049]	Lys Glu Ser Lys Val Ile Arg Glu Arg Asp Gly Thr Asp Gly Trp Ile
[2050]	325 330 335
[2051]	Asp Asn Leu Leu Ser Met Ser Tyr Ile Gly Asn Val Asn Gly Lys Glu
[2052]	340 345 350
[2053]	Tyr Tyr Val Asp Ile Ser Asp Ala Ser Gly Trp Ala His Ile Tyr Leu
[2054]	355 360 365
[2055]	Tyr Pro Val Asp Gly Gly Lys Glu Ile Ala Leu Thr Lys Gly Glu Trp
[2056]	370 375 380
[2057]	Glu Val Val Ala Ile Leu Lys Val Asp Thr Lys Lys Lys Leu Ile Tyr
[2058]	385 390 395 400
[2059]	Phe Thr Ser Thr Lys Tyr His Ser Thr Thr Arg His Val Tyr Ser Val
[2060]	405 410 415
[2061]	Ser Tyr Asp Thr Lys Val Met Thr Pro Leu Val Asn Asp Lys Glu Ala
[2062]	420 425 430
[2063]	Ala Tyr Tyr Thr Ala Ser Phe Ser Ala Lys Gly Gly Tyr Tyr Ile Leu
[2064]	435 440 445
[2065]	Ser Tyr Gln Gly Pro Asn Val Pro Tyr Gln Glu Leu Tyr Ser Thr Lys
[2066]	450 455 460
[2067]	Asp Ser Lys Lys Pro Leu Lys Thr Ile Thr Ser Asn Asp Ala Leu Leu
[2068]	465 470 475 480
[2069]	Glu Lys Leu Lys Glu Tyr Lys Leu Pro Lys Val Ser Phe Phe Glu Ile
[2070]	485 490 495
[2071]	Lys Leu Pro Ser Gly Glu Thr Leu Asn Val Lys Gln Arg Leu Pro Pro
[2072]	500 505 510
[2073]	Asn Phe Asn Pro His Lys Lys Tyr Pro Val Leu Phe Thr Pro Tyr Gly
[2074]	515 520 525
[2075]	Gly Pro Gly Ala Gln Glu Val Ser Gln Ala Trp Asn Ser Leu Asp Phe
[2076]	530 535 540
[2077]	Lys Ser Tyr Ile Thr Ser Asp Pro Glu Leu Glu Tyr Val Thr Trp Thr
[2078]	545 550 555 560
[2079]	Val Asp Asn Arg Gly Thr Gly Tyr Lys Gly Arg Lys Phe ArgSer Ala

[2080]		565		570		575										
[2081]	Val	Ala	Lys	Arg	Leu	Gly	Phe	Leu	Glu	Ala	Gln	Asp	Gln	Val	Phe	Ala
[2082]			580						585							590
[2083]	Ala	Lys	Glu	Val	Leu	Lys	Asn	Arg	Trp	Ala	Asp	Lys	Asp	His	Ile	Gly
[2084]			595						600							605
[2085]	Ile	Trp	Gly	Xaa	Ser	Tyr	Gly	Gly	Phe	Leu	Thr	Ala	Lys	Thr	Leu	Glu
[2086]			610						615							620
[2087]	Thr	Asp	Ser	Gly	Val	Phe	Thr	Phe	Gly	Ile	Ser	Thr	Ala	Pro	Val	Ser
[2088]			625						630							635
[2089]	Asp	Phe	Arg	Leu	Tyr	Asp	Ser	Met	Tyr	Thr	Glu	Arg	Tyr	Met	Lys	Thr
[2090]									645							650
[2091]	Val	Glu	Leu	Asn	Ala	Asp	Gly	Tyr	Ser	Glu	Thr	Ala	Val	His	Lys	Val
[2092]									660							665
[2093]	Asp	Gly	Phe	Lys	Asn	Leu	Lys	Gly	His	Tyr	Leu	Ile	Gln	His	Gly	Thr
[2094]									675							680
[2095]	Gly	Asp	Asp	Asn	Val	His	Phe	Gln	Asn	Ala	Ala	Val	Leu	Ser	Asn	Thr
[2096]									690							695
[2097]	Leu	Met	Asn	Gly	Gly	Val	Thr	Ala	Asp	Lys	Leu	Thr	Thr	Gln	Trp	Phe
[2098]									705							710
[2099]	Thr	Asp	Ser	Asp	His	Gly	Ile	Arg	Tyr	Asp	Met	Asp	Ser	Thr	Tyr	Gln
[2100]																725
[2101]	Tyr	Lys	Gln	Leu	Ser	Lys	Met	Val	Tyr	Asp	Gln	Lys	Gln	Arg	Arg	Pro
[2102]																740
[2103]	Glu	Ser	Pro	Pro	Met	His	Gln	Trp	Ser	Lys	Arg	Val	Leu	Ala	Ala	Leu
[2104]									755							760
[2105]	Phe	Gly	Glu	Arg	Ala	Glu	Glu									
[2106]																770

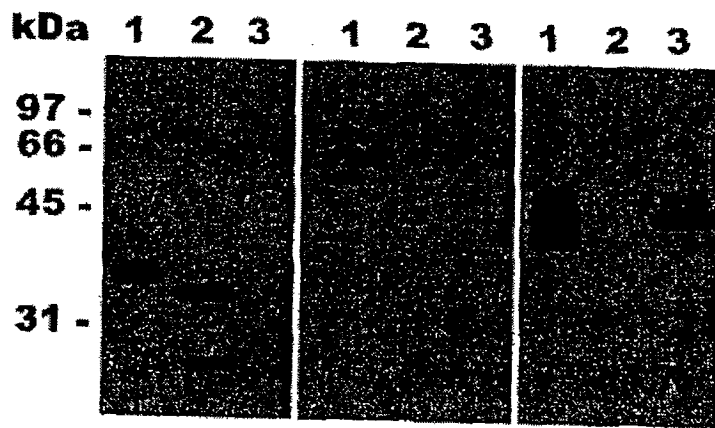


图 1

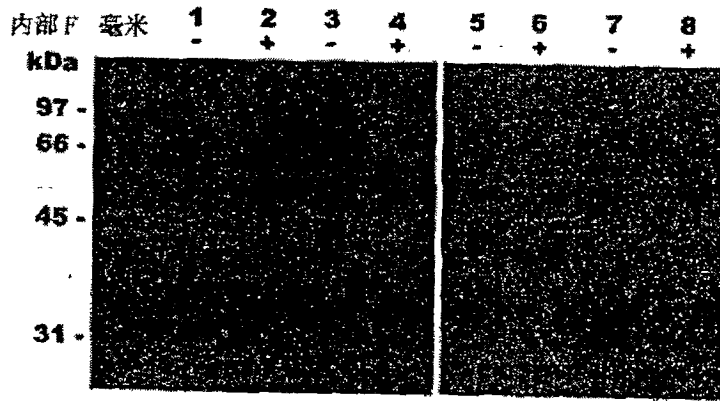


图 2

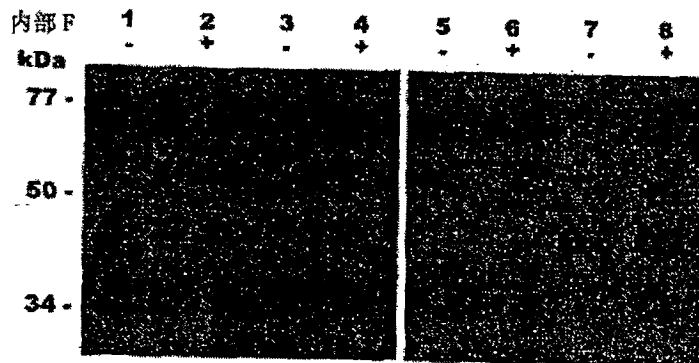


图 3

图 4 :以 Lys (Abz)-Pro-Pro-pNA 为底物,红色毛癣菌 AMPP 在不同 pHs 下的活性

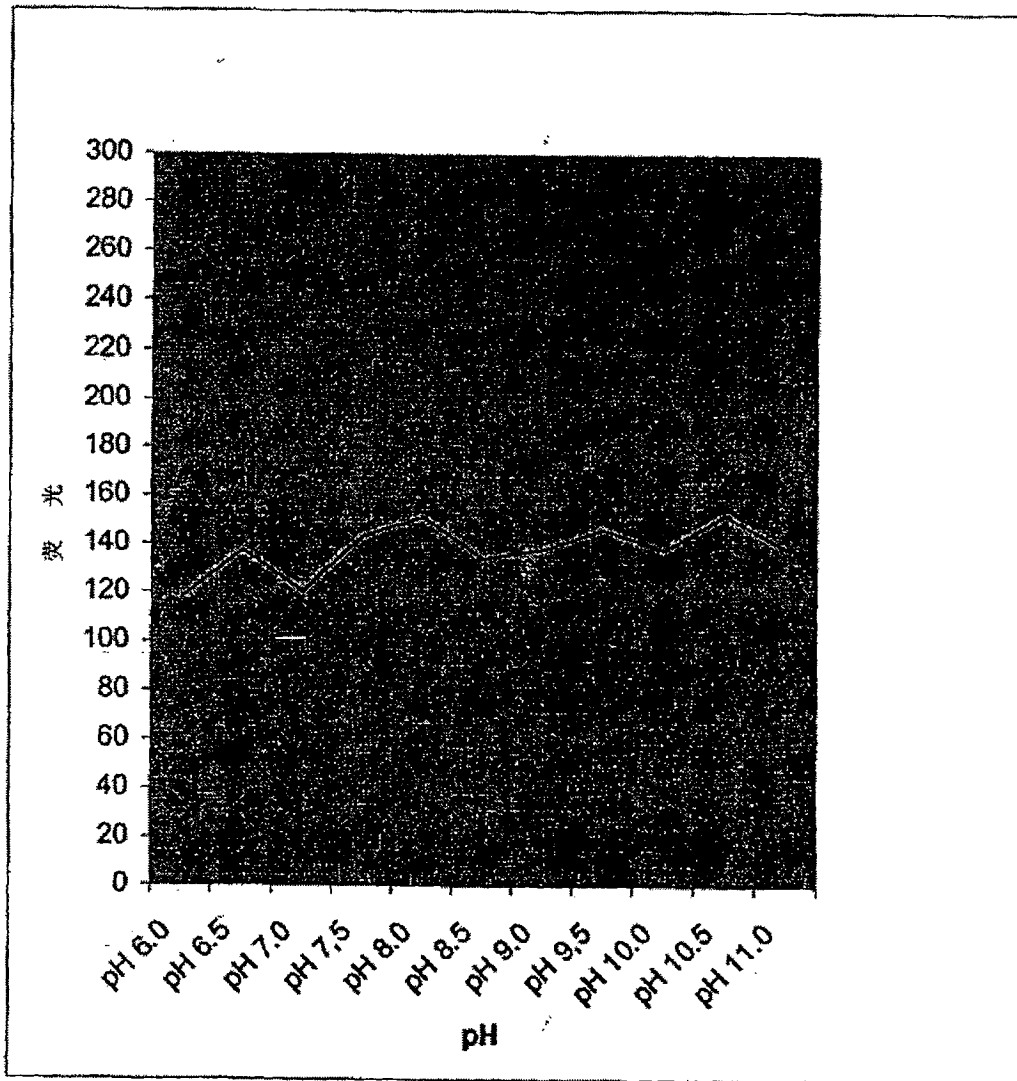


图 4

图 5: 红色毛癣菌 AMPP 在不同温度下的酶活性

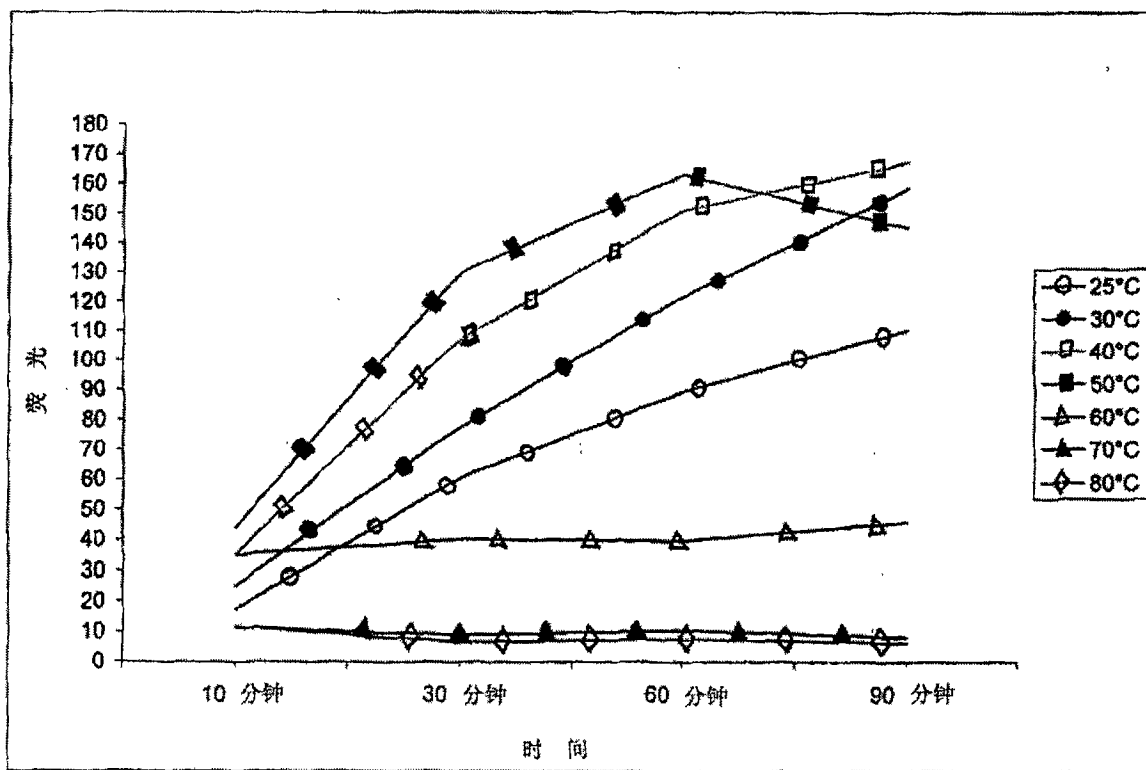


图 5

图 6: 在 37°C, 4h, 不含 ruLAP2(A) 和含有 ruLAP2(B) 以 E/S 比率 1/50(w : w) 分解 gliadin14mer 的结果图



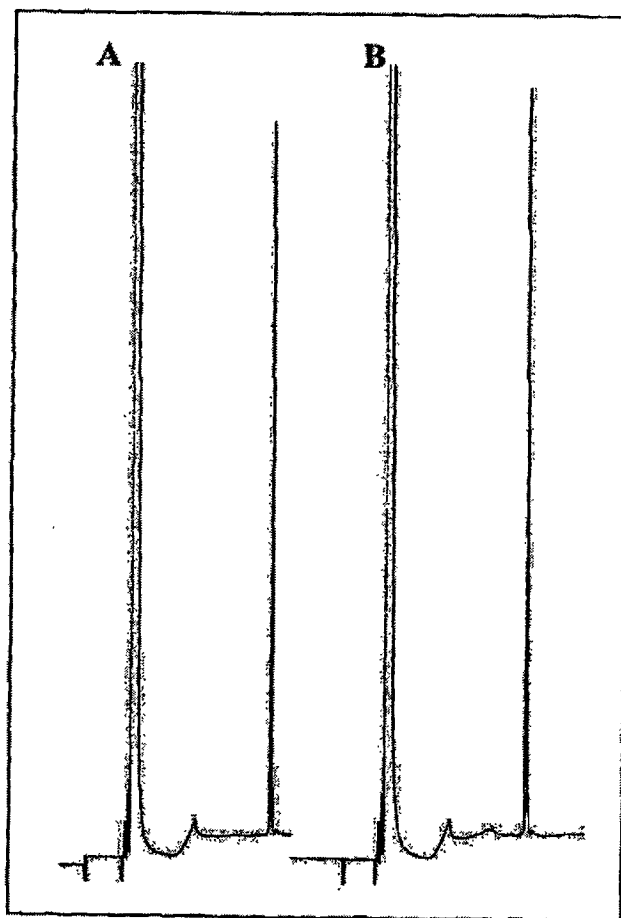


图 6

图 7 :DPPIV 以 E/S 比率 1/25(w : w) 单独和 DPPIV/ruLAP2 的酶混和物以 E/S 比率 1/50(w : w) 分解 gliadin 14mer

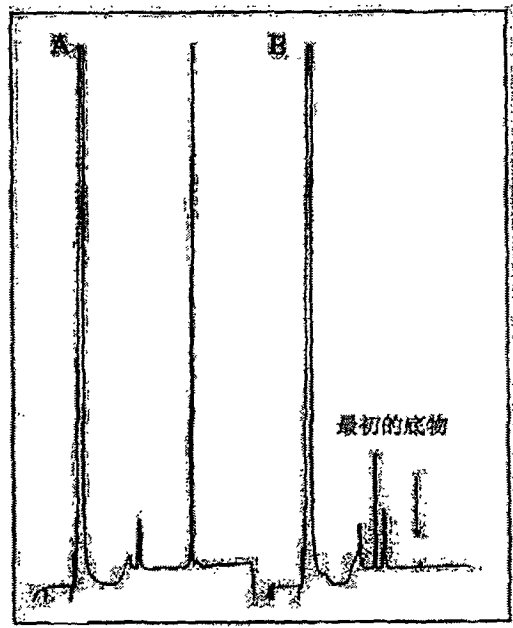


图 7

图 8 :在 37°C, 4h, 不含 ruDPPIV 和含有 ruDPPIV 以 E/S 比率 1/50 (w : w) 分解 gliadin 33mer

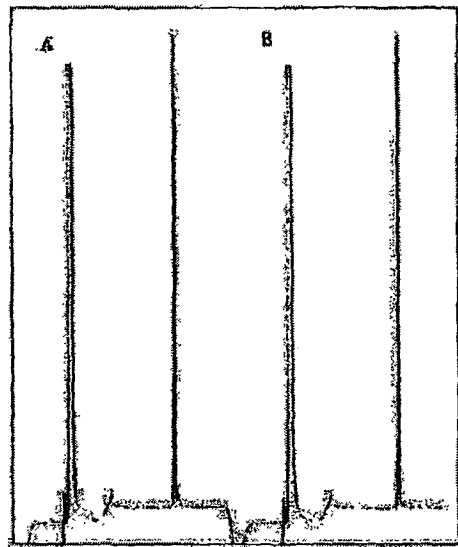


图 8

图 9 :在 37°C, 4h, ruLAP2/DPPIV 的酶混和物以 E/S 比率 1/50 (w : w) 分解 gliadin 33mer

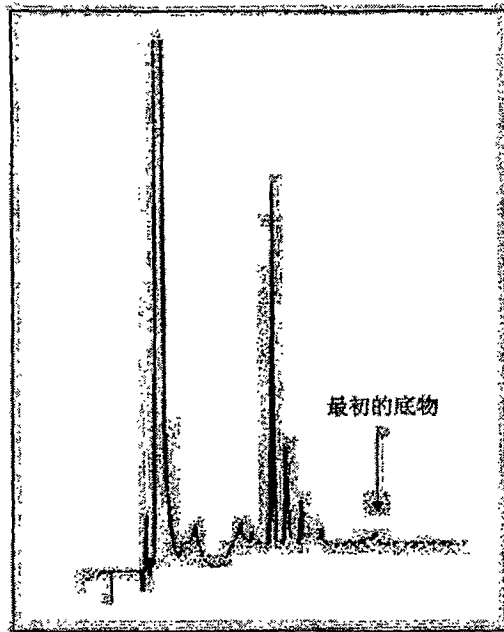


图 9

图 10 :ruLAP2 分解 Gly-Ser-proNPY (分子量经计算为 8203. 12Da) (A) 前和 (B) 后的质谱图, 其中是去除了 Gly-Ser (144. 1Da) 的结果

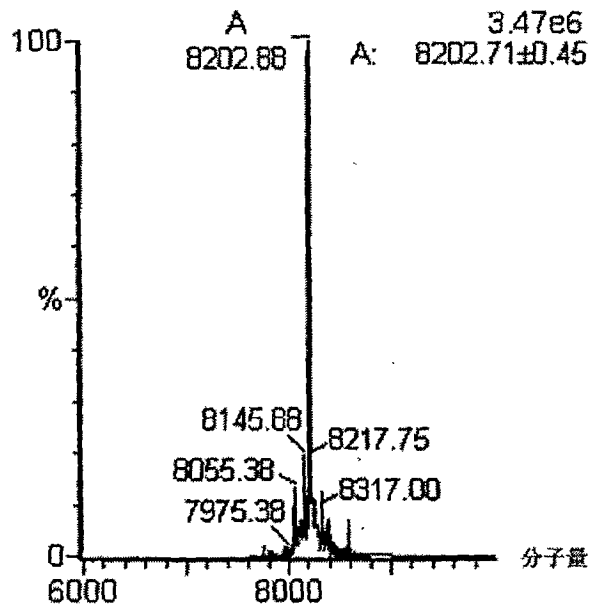


图 10-A

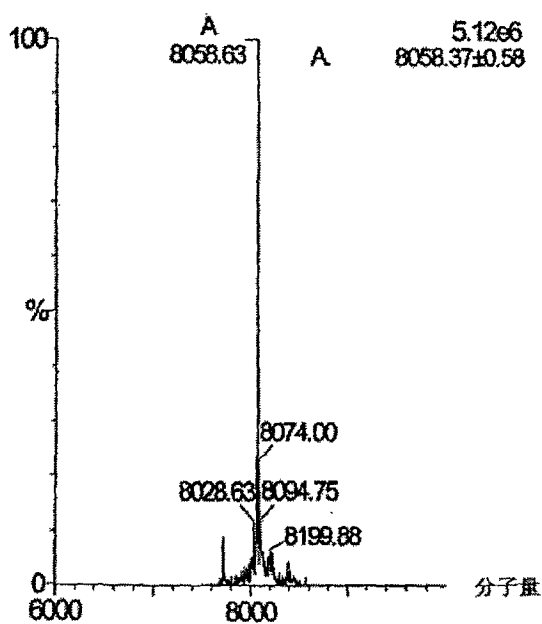


图 10-B

图 11 : 丙氨酸 - 前神经肽 Y (A, 分子量经计算为 8130.0 $\mu$ ma) 和 ruLAP2 分解 (B) 后的质谱图

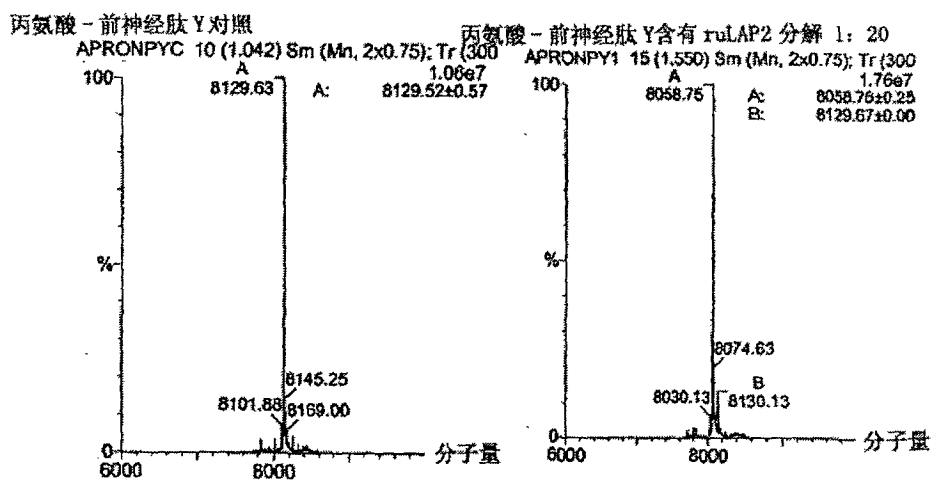


图 11A

图 11B

图 11A 图 11B

图 12 : (A) TG47 (分子量经计算为 18894.9 $\mu$ ma) 和 ruLAP2 分解后 (B) 的质谱图, 其中是去除了 N- 末端甲硫氨酸的结果

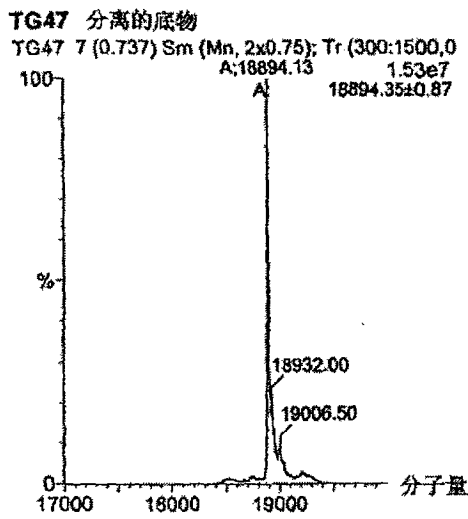


图 12A

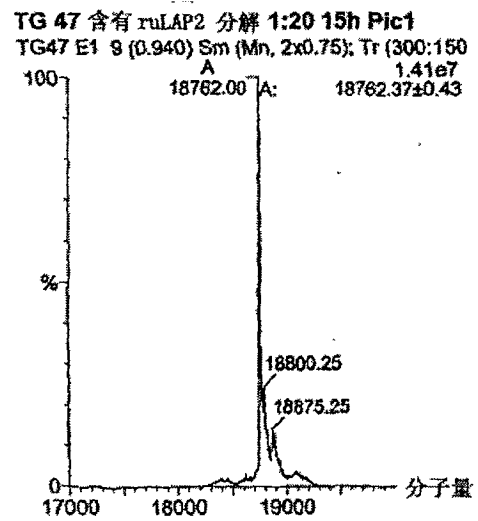


图 12B

图 13A 和 13B 表示不含 (A) 和含有 (B) DPPIV 分解 desMet-G-CSF 的 ESI-MS 特征图

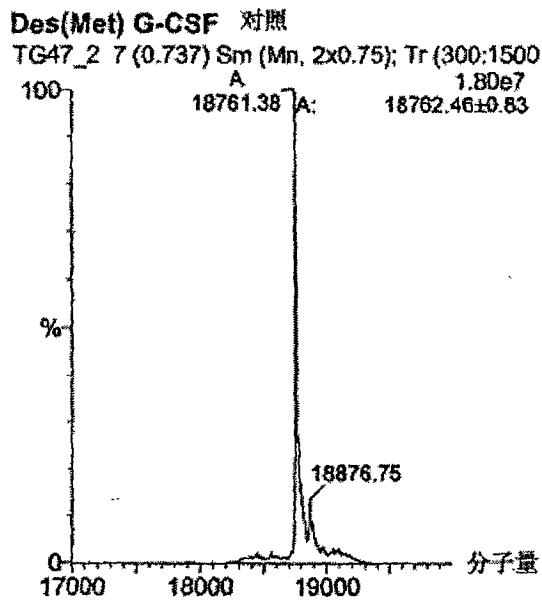


图 13A

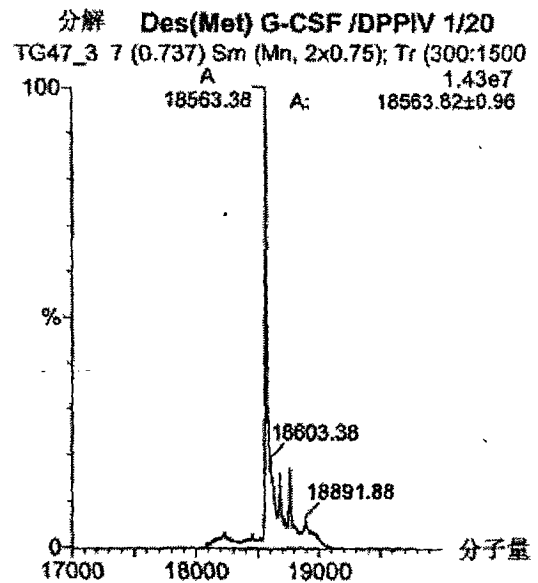


图 13B

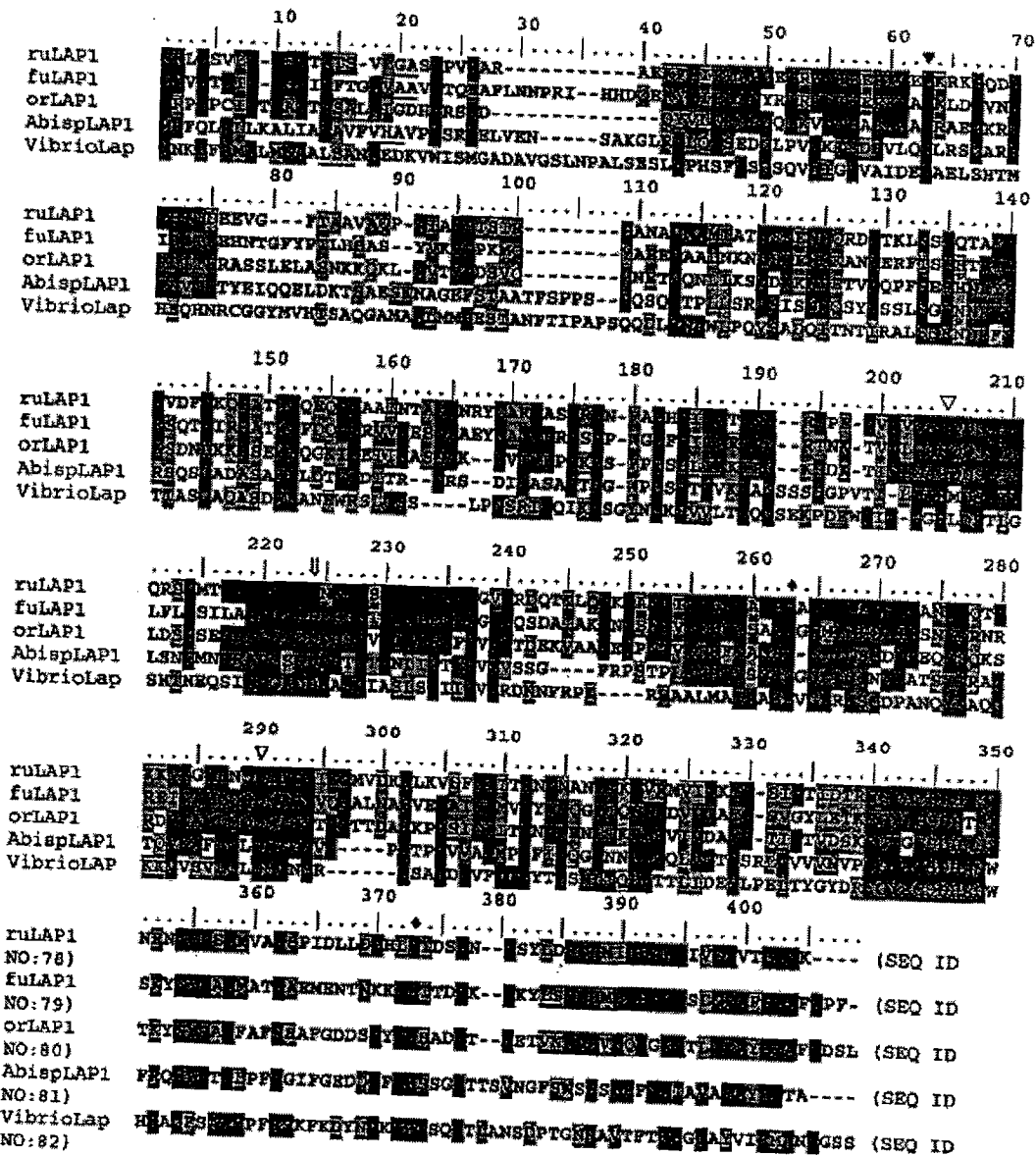


图 14

