

[19] 中华人民共和国国家知识产权局



[12] 发 明 专 利 说 明 书

专利号 ZL 00137112.6

[51] Int. Cl.

C12P 23/00 (2006.01)

C12N 15/09 (2006.01)

C12N 15/52 (2006.01)

C12N 15/63 (2006.01)

[45] 授权公告日 2007 年 11 月 7 日

[11] 授权公告号 CN 100347311C

[22] 申请日 2000.12.1 [21] 申请号 00137112.6

[30] 优先权

[32] 1999.12.1 [33] EP [31] 99123821.3

[73] 专利权人 DSM IP 资产公司

地址 荷兰海尔伦

[72] 发明人 星野立夫 小岛一之 濑户口丰

[56] 参考文献

WO9002804A1 1990.3.22

ARE ACTIVE OXYGEN SPECIES INVOLVDE
IN INDUCTION OF β - CAROTENE IN DU-
NALIELLA BARDAWIL? SHAISH AVIV ET
AL. PLANTA, Vol. 190 No. 3. 1993

审查员 徐 莉

[74] 专利代理机构 北京市柳沈律师事务所

代理人 巫肖南

权利要求书 2 页 说明书 36 页 附图 1 页

[54] 发明名称

类胡萝卜素尤其虾青素的重组生产及其可利用的生物物质

[57] 摘要

本发明涉及类胡萝卜素的生产方法，其包括培养重组生物并从培养物中回收类胡萝卜素，其中该重组生物中编码一或多种活性氧类物质淬灭因子的一个基因被特异于该基因的裂解盒基本上裂解，本发明还涉及可用于该方法的遗传物质，如可产生类胡萝卜素的重组生物、裂解盒、重组 DNA 序列、重组 DNA 片段；和多核苷酸。

1. 一种生产类胡萝卜素的方法,该方法包括培养重组生物体,并从培养物中回收类胡萝卜素,所述生物体中编码活性氧类物质淬灭因子的基因借助特异于该基因的裂解盒而裂解,其中所述活性氧类物质淬灭因子由选自如下的多核苷酸编码:

SEQ ID NO: 1 或 4, 或与 SEQ ID NO: 1 或 4 的互补体在高严谨杂交和洗涤条件下杂交的多核苷酸, 并且该杂交的多核苷酸编码的多肽具有线粒体超氧化物歧化酶(SOD)活性,

并且,

所述重组生物体的宿主生物体属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属。

2. 如权利要求 1 所述的方法,其中所述重组生物体的宿主生物体是 *P. rhodozyma* 的一个菌株。

3. 如权利要求 2 所述的方法,其中所述重组生物体的宿主生物体是 *P. rhodozyma* ATCC96594。

4. 一种能够产生类胡萝卜素的重组生物体,其特征在于其中编码至少一种活性氧类物质淬灭因子的基因借助特异于该基因的裂解盒而裂解,其中所述活性氧类物质淬灭因子由选自如下的多核苷酸编码:

SEQ ID NO: 1 或 4, 或与 SEQ ID NO: 1 或 4 的互补体在高严谨杂交和洗涤条件下杂交的多核苷酸, 并且该杂交的多核苷酸编码的多肽具有线粒体超氧化物歧化酶(SOD)活性,

并且,

所述重组生物体的宿主生物体属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属。

5. 如权利要求 4 所述的重组生物体,其中所述宿主生物体是 *P. rhodozyma* 的一个菌株。

6. 如权利要求 5 所述的重组生物体,其中所述宿主生物体是 *P. rhodozyma* ATCC96594。

7. 一种用于裂解活性氧类物质淬灭因子的编码基因的裂解盒,该活性氧类物质淬灭因子对在产生类胡萝卜素的生物体中的胡萝卜素形成有效,所述裂解盒包括与编码被裂解的所述活性氧类物质淬灭因子的部分DNA序列相同的部分核苷酸序列和选择性标记基因,其中被裂解的所述活性氧类物质淬灭因子由选自如下的多核苷酸编码:

SEQ ID NO: 1 或 4, 或与 SEQ ID NO: 1 或 4 的互补体在高严谨杂交和洗涤条件下杂交的多核苷酸, 并且该杂交的多核苷酸编码的多肽具有线粒体超氧化物歧化酶(SOD)活性,

并且,

所述重组生物体的宿主生物体属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Ilaematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属。

8. 如权利要求 7 所述的裂解盒,其中所述宿主生物体是 *P. rhodozyma* 的一个菌株。

9. 如权利要求 8 所述的裂解盒,其中所述宿主生物体是 *P. rhodozyma* ATCC96594。

10. 如权利要求 7-9 任一项所述的裂解盒,其中编码缺陷性氧类物质淬灭因子的部分核苷酸序列与引入了裂解盒的生物体中活性氧类物质淬灭因子的至少部分编码 DNA 序列相同。

11. 如权利要求 10 所述的裂解盒, 其中与活性氧类物质淬灭因子的部分编码DNA序列相同的部分核苷酸序列相对于相应功能基因含有至少一个缺失和/或突变。

类胡萝卜素尤其虾青素的重组 生产及其可利用的生物物质

发明领域

本发明涉及类胡萝卜素，特别是虾青素的重组生产，及可用于该生产的生物物质。

背景技术

已知虾青素分布于多种生物体如动物(例如鸟类如火烈鸟和 scarlet ibis, 和鱼如虹鳟(rainbow trout)和大麻哈鱼(salmon))、藻和微生物中。并且还认识到同大多数类胡萝卜素一样，虾青素也具有很强的针对活性氧的抗氧化特性，因此虾青素有望作为医药用途来保护活体细胞免受某些疾病如癌症的侵袭。尤其是，因为虾青素能赋予动物以鲜艳的桔红色并能在市场里吸引顾客，所以从工业应用的角度来看，特别是在鱼如大麻哈鱼的养殖业中对作为显色试剂的虾青素的需要与日俱增。

已知 *Phaffia rhodozyma* 为专门生产虾青素的产胡萝卜素酵母菌。与另一种产胡萝卜素酵母，红酵母属的种不同，*Phaffia rhodozyma* 能够发酵某些糖如 D-葡萄糖。从工业应用的角度来看，这是一个重要的性质。最近的分类学研究发现 *Phaffia rhodozyma* 具有有性繁殖并且其有性形态(telemorphic state)被命名为 *Xanthophyllomyces dendrorhous*(W. I. Golubev; 酵母 11,101-110,1995)。为了从 *Phaffia rhodozyma* 中获得虾青素的高产菌株，进行了一些菌株改造的研究，但近十年中，研究只是局限于采用传统的诱变和原生质体融合的方法。近来，Wery 等采用 *Phaffia rhodozyma* 研制了一种宿主载体系统，该载体系统中非复制质粒以多拷贝被组合到 *Phaffia rhodozyma* 基因组的核糖体 DNA 上(Wery 等，基因，184，89-97，1997)。并且 Verdoes 等报道了更多的用来获得 *Phaffia rhodozyma* 转化子及其三个产胡萝卜素基因的改造载体，所述三个基因编码能催化 3,7-二甲基-2,6-辛二烯焦磷酸生成 β -胡萝卜素的酶(国际专利 WO97/23633)。基因工程方法对于 *Phaffia rhodozyma* 的菌株改造研究的重要作用日益增加，在不久的将来将突破通过传统方法达到的产率。

如上所述，虾青素具有抗氧化剂特性，并且这一特性似乎在体内具有

防活性氧类物质如 $O_2\cdot$, H_2O_2 和 $OH\cdot$ 的重要作用, 该活性氧类物质在活细胞中不断的产生。经过传统的化学诱变后, 通过筛选抗霉素-敏感菌株, An 等从 *Phaffia rhodozyma* 中获得了虾青素高产菌株(An, G-H.等, 应用环境微生物学(Appl. Env. Microbiol.), 55(1), 116-124, 1989)。已知抗霉素是细胞色素 b 和 c_1 之间呼吸链的抑制剂(Lucchini, G.等, Mol. Gen. Genet., 177, 139-, 1979)并能增强这种抗霉素-敏感突变株中的色素形成作用。由于初级呼吸链中 bc_1 复合物阻断而产生的活性氧类物质促进了类胡萝卜素的生成(An, G-H.等, 应用环境微生物学, 55, 116-124, 1989)。实际上, 培养基中加入生成 O_2 的物质杜醌能够增加 *Phaffia rhodozyma* 中类胡萝卜素的总含量(主要的类胡萝卜素是虾青素)以及叶黄素的相对含量, 而活性氧类物质淬灭因子甘露糖醇则起相反的作用(Schroeder, W.A.等, 普通微生物学杂志(J. Gen. Microbiol.), 139, 907-912, 1993)。这些结果促使这些作者得出 *Phaffia rhodozyma* 中的虾青素具有抗氧化剂特点的结论。事实上, 如果充分诱导生物氧化作用, 就会增加指数生长后期产生虾青素的量。此外, 培养基中加入氧化底物如乙醇能增强 *Phaffia rhodozyma* 中虾青素的产生量(Gu, W-L.等, J. Ind. Microbiol. Biotechnol., 19, 114-117, 1997)。尽管 Schroeder 等试图通过比较亲本菌株与抗霉素-敏感的虾青素高产菌株之间的差异, 找到作为 *Phaffia rhodozyma* 中天然活性氧类物质淬灭因子的超氧化物歧化酶(SOD)和过氧化氢酶活性与虾青素的产率之间的关系, 但没有观察到体外活性的直接关系。

发明内容

本发明提供了活性氧类物质淬灭因子的基因和酶, 如 SOD 和过氧化氢酶, 它们是可用于改进类胡萝卜素, 特别是虾青素的生产工艺的有效生物物质。本发明涉及线粒体和细胞质的 SOD 和过氧化氢酶编码基因的克隆和测定。本发明还涉及 *Phaffia rhodozyma* 中基因裂解后产生的酶的特性。通过在适当的培养条件下于合适的培养基中培养这种转化体能够证实其裂解对胡萝卜素形成的作用。

本发明特别提供了生产类胡萝卜素的方法, 该方法包括培养重组生物体和回收培养物中的类胡萝卜素, 所述重组生物体中编码一种或多种活性氧类物质淬灭因子的基因通过借助对该基因具有特异性的裂解酶基本上裂解。所述重组生物体的宿主生物体可以属于原核生物界、原生生物界或真菌界。更优选所述重组生物体的宿主生物体可属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、

链孢霉属、红酵母属、布拉霉属或须霉属。最优选地是，所述的重组生物体是 *P. rhodozyma* 的菌株。

活性氧类物质淬灭因子是线粒体超氧化物歧化酶(SOD)，细胞质超氧化物歧化酶(SOD)和/或过氧化氢酶。

本发明还提供了一种可生产类胡萝卜素的重组生物体，其特征在于编码至少一种活性氧类物质淬灭因子的基因被借助对该基因特异性的裂解盒基本上裂解。被裂解的活性氧类物质淬灭因子是线粒体超氧化物歧化酶(SOD)，细胞质超氧化物歧化酶(SOD)和/或过氧化氢酶。

本发明进一步提供了用来制备本发明的重组生物体的裂解盒，该裂解盒用于裂解产类胡萝卜素生物体中有效的活性氧类物质淬灭因子的编码基因，该基因含有与编码活性氧类物质淬灭因子的部分 DNA 序列基本相同的部分核苷酸序列和可选择性标记基因。用于构建裂解盒的靶生物体可属于原核生物界、原生生物界或真菌界，更优选属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属。被裂解的活性氧类物质淬灭因子是线粒体超氧化物歧化酶(SOD)，细胞质超氧化物歧化酶(SOD)和/或过氧化氢酶。

如本文所用，多核苷酸或多肽序列(A)如果相对于另一序列(B)至少 75 % 相同，优选至少 85 % 相同，如至少 95 % 相同则为基本相同。

本发明进一步提供了编码活性氧类物质淬灭因子(其在产生类胡萝卜素的生物体中发挥作用)的重组 DNA 序列。DNA 序列可来源于属于原核生物界、原生生物界或真菌界的生物体，更优选属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属的生物体。特别优选 *P. rhodozyma*。由重组 DNA 序列编码的活性氧类物质淬灭因子可以是线粒体超氧化物歧化酶(SOD)，细胞质超氧化物歧化酶(SOD)和/或过氧化氢酶。

编码线粒体超氧化物歧化酶的重组 DNA 序列可以是 SEQ ID NO: 1 或 4 所示的序列，或它可以与 SEQ ID NO: 1 或 4 具有高度的同源性，足以和 SEQ ID NO: 1 或 4 中的任意一个序列发生重组。编码细胞质超氧化物歧化酶的重

组 DNA 序列可以是 SEQ ID NO: 2 或 6 所示的序列, 或它可以与 SEQ ID NO: 2 或 6 具有高度的同源性, 足以和 SEQ ID NO: 2 或 6 中的任意一个序列发生重组。编码过氧化氢酶的重组 DNA 序列可以是 SEQ ID NO: 3 或 8 所示的序列, 或它可以与 SEQ ID NO: 3 或 8 具有高度的同源性, 足以和 SEQ ID NO: 3 或 8 中的任意一个序列发生重组。

在本发明中, 只要认为适当就可利用以下杂交和洗涤条件的任意组合来鉴定同源多核苷酸序列:

高度严谨条件: 6X SSC

0.5% SDS

100 μ g/ml 变性的鲑精 DNA

50%甲酰胺

于 42 °C 轻轻旋转过夜。

高度严谨洗涤: 2X SSC 洗 1 次, 0.5% SDS 室温 15 分钟, 然后

0.1X SSC 洗 1 次, 0.5% SDS 室温 15 分钟。

低度严谨杂交: 6X SSC

0.5% SDS

100 μ g/ml 变性的鲑精 DNA

50%甲酰胺

于 37 °C 轻轻旋转过夜。

低度严谨洗涤: 0.1X SSC 洗 1 次, 0.5% SDS 室温 15 分钟。

可通过改变上述杂交反应温度和、或洗涤条件以获得中度严谨条件。

因此, 如本文所用, 如果一序列(A)与另一序列(B)在高度严谨条件(即上述高度严谨杂交和洗涤条件)下杂交, 而且如果序列 A 所编码的多肽或多肽片段具有与 B 所编码的多肽相同的活性, 则称序列(A)相对于另一序列(B)“高度同源”。

本发明进一步提供了含有位于目的蛋白编码区域上游的转运肽的编码区域的重组 DNA 片段, 该目的蛋白可以是线粒体超氧化物歧化酶。该重组 DNA 片段的表达使得目的蛋白定位于线粒体上。因此, 本发明还提供了将目的蛋白定位于线粒体上的方法, 该方法包括含有位于目的蛋白编码区域上游的转运肽的编码区域的重组 DNA 片段在合适的重组宿主生物体中的表达。

如上所述, 本发明公开了活性氧类物质淬灭因子, 如线粒体超氧化物歧

化酶、细胞质超氧化物歧化酶和过氧化氢酶的核苷酸序列。这些多聚核苷酸被用来作为克隆编码活性氧类物质淬灭因子基因的探针或引物，基于基因的同源性，该活性氧类物质淬灭因子在其它产生类胡萝卜素的生物体中发生作用。

图1表示通过使用从 ATCC96594 和它的 SOD 突变株中获得的不含细胞的提取物进行非变性聚丙烯酰胺凝胶电泳后的超氧化物歧化酶的活性染色分析。泳道 1, *P. rhodozyma* ATCC96594；泳道 2, *P. rhodozyma* ATCC96594 :: pSOD/G717(SOD1 裂解物)；泳道 3, *P. rhodozyma* ATCC96594 :: pSOD/G828(SOD2 裂解物)；泳道 4, *P. rhodozyma* ATCC96594。

主要如上述解释，本发明的目的是提供一种通过生物方法生产类胡萝卜素的新方法。所述新方法包括培养一种重组微生物，和从培养物中回收类胡萝卜素，其中该微生物中编码一种或多种活性氧类物质淬灭因子的基因被借助对该基因具有特异性的裂解盒充分裂解。

本发明的目的还在于提供一个含有编码活性氧类物质淬灭因子的开放式阅读框的重组 DNA 序列，所述活性氧类物质淬灭因子可以是一种酶，如线粒体 SOD 或细胞质 SOD。该重组 DNA 序列可以含有部分编码过氧化氢酶的基因片段。因为这些序列能与编码所述酶的天然基因重组从而特异性地裂解该基因，所以它们被用来构建所述的裂解盒。

所述重组 DNA 序列可以是一种源自属于原核生物界、原生生物界或真菌界的生物体，更优选地属于欧文氏菌属、红细菌属、粘球菌属、黄杆菌属、副球菌属、聚球蓝细菌属、集胞蓝细菌属、土壤杆菌属、链霉菌属、*Haematococcus*、*Dunaliella*、*Phaffia*、*Xanthophyllomyces*、链孢霉属、红酵母属、布拉霉属或须霉属的一种生物体的重组 DNA 序列。特别优选的生物体是 *P. rhodozyma*。由该重组 DNA 序列编码的活性氧类物质淬灭因子可以是线粒体 SOD，细胞质 SOD 和/或过氧化氢酶。该重组 DNA 序列的一个特定例子是源自于 *Phaffia rhodozyma* 的基因并选自(i)SEQ ID NO: 1 或 2 所示的 DNA 序列；(ii)SEQ ID NO: 4 或 6 所示的那些 cDNA；(iii)SEQ ID NO: 1，2，4 或 6 所示的 DNA 序列的同类编码或等位(基因)变异体，和(iv)SEQ ID NO: 1，2，4 或 6 所示的 DNA 序列中添加、插入、删除、和/或置换了一个或多个核苷酸所得到的衍生物，该衍生物编码具有所述酶活性的多肽。所述重组 DNA 序列的特征还在于(a)它编码具有选自 SEQ ID NO: 5 和 7 所示的

氨基酸序列的酶，或(b)它编码选自下组的酶的变异体(i)一个等位(基因)变异体，和(ii)具有一个或多个氨基酸添加、插入、删除、和/或置换并具有所述酶活性的酶。

如本文所用，“等位基因变异体”指在二倍体生物如 *Phaffia rhodozyma*、*Xanthophyllomyces*、*dendrorhous* 等的两个等位基因的任一个上至少有一个突变的变异体。一种指定基因的两个等位基因有相同的形状或特性，但一具体等位基因与该基因的另一等位基因在编码产物或功能上有定性或定量的差别。等位基因变异体可天然产生或经化学诱变而人工产生。野生型等位基因编码一特定生物的野生株中特定的表型特征。

如本文所用，“同类编码变异体”指下述变异体，其一个特定基因的核苷酸序列与野生型基因的序列有区别，但其翻译产物(即氨基酸序列)与野生型蛋白质相同。这是由于遗传密码的简并性造成的，并且是各种生物在密码子的利用中的差异造成的。

如本文所用，“DNA 序列的衍生物”指下述 DNA 序列，其编码的多肽具有相应 SEQ ID NO 的活性但与该 DNA 序列不同在于 1 - 20 个，优选 1 - 10 个，如 1 - 5 个核苷酸的添加、插入、删除，和/或置换。

特别涉及的上述重组 DNA 序列可以是源自 *Phaffia rhodozyma* 的基因，其选自(i)SEQ ID NO: 3 代表的 DNA 序列；(ii)SEQ ID NO: 8 所定义的 cDNA；(iii)SEQ ID NO: 3 或 8 代表的 DNA 序列的同类编码或等位(基因)变异体，和(iv)SEQ ID NO: 3 或 8 代表的 DNA 序列中添加、插入、删除、和/或置换了一个或多个核苷酸所得到的衍生物，该衍生物编码具有所述酶活性的多肽。所述重组 DNA 序列的特征还在于(a)它编码具有选自 SEQ ID NO: 9 所示的氨基酸序列中的部分氨基酸序列的酶，或(b)它编码的所述酶的一种变异体，其选自(i)一个等位(基因)变异体，和(ii)具有一个或多个氨基酸被添加、插入、删除、和/或置换并具有所述酶活性的酶。这种重组 DNA 序列优选地可以是载体的形式。

本发明还提供了用所述重组 DNA 序列转化宿主生物体以便得到这样一种生物体，该生物体中编码至少一种活性氧类物质淬灭因子的基因因引入该基因特异性裂解盒而基本上被裂解。如本文所用，一种基因如果其编码的多肽的活性相对于未被裂解的基因降低，则称该基因被“裂解”或“被基本上裂解”。优选地，活性降低 100%，优选至少 50% 如至少 75%，更优选至

少 90 % - 100 %。该重组 DNA 序列的合适形式可以是载体。通过使用该重组 DNA 序列所得到的重组生物体在其编码线粒体 SOD，细胞质 SOD，或过氧化氢酶的 DNA 序列中被裂解。被该重组 DNA 转化的宿主生物体可有效改进类胡萝卜素，特别是虾青素的生产方法。因此，本发明也提供了这样的重组生物体。

这种生产类胡萝卜素的生物方法能提高类胡萝卜素，特别是虾青素的产率。因此，以在能诱导产生类胡萝卜素的条件下培养上述重组生物体为特征的生产类胡萝卜素的方法是本发明的一个方面。这种方法可以用于虾青素的生物生产。

许多研究者指出活性氧类物质可以促进已知的产胡萝卜素生物体产生类胡萝卜素。环境的氧化压力可增强 *Haematococcus pluvialis* 的孢囊细胞对类胡萝卜素的生物合成(Kobayashi 等，应用环境微生物学，59, 867-873, 1993)。活性氧类物质可诱导类胡萝卜素的生物合成，并且所积累的类胡萝卜素可以保护 *Dunaliella bardawil* 免遭氧化压力的破坏(Shaish 等，Planta., 190, 363-368, 1993)。尽管在使活性氧类物质的产生发生改变的各种培养条件下对 *P. rhodozyma* 体内产生虾青素进行了研究，但大概由于在这种研究中仍有天然活性氧类物质淬灭因子存在并且在某种程度上弥补了活性氧类物质对类胡萝卜素生产的影响，所以还没有弄清楚活性氧的产生与类胡萝卜素产率之间的相互关系(Schroeder, W.A.等，普通微生物学杂志，139, 907-912, 1993)。

本发明中，为了排除 *P. rhodozyma* 中存在的天然活性氧类物质淬灭因子抑制活性氧对虾青素生产的正面作用的可能性，从 *P. rhodozyma* 中克隆天然活性氧类物质淬灭因子如 SOD 和过氧化氢酶，通过构建和引进基因裂解质粒来破坏天然活性氧类物质淬灭因子的表达。假设虾青素在 *P. rhodozyma* 中起着抗氧化剂的作用，天然活性氧类物质淬灭因子的失活可能影响类胡萝卜素的产生，其方式是由天然活性氧类物质淬灭因子的缺乏导致体内活性氧类物质相对增加，使作为活性氧类物质的另一种抑制因素的虾青素产生。

活性氧类物质可通过对细胞内分子如蛋白质或核酸的氧化损伤而对活细胞具有毒性。近来研究通过果蝇和线虫中超氧化物歧化酶活性的增加，存活期的增加，与氧化损伤作用的减少之间的相互关系，表明衰老是由氧化损伤作用引起的(Agarwal, S.等，美国国家科学院院报，91, 12332-12335, 1994,

Larsen, P.L. 美国国家科学院院报, 90, 8905-8909, 1993, Sohal, R.S.等, 生物化学杂志, 270, 15671-15674, 1995)。SOD 和相关的抗氧化酶以及它们的基因在原核生物和真核生物中也进行了充分的研究。

啤酒糖酵母如同大多数的真核生物一样, 在胞质溶胶中含有 Cu/Zn SOD(SOD1 基因的产物), 在线粒体中含有 MnSOD(SOD2 基因的产物)。这些酶催化 O_2 的歧化, 而生成 O_2 和 H_2O_2 。小分子抗氧化剂, 如谷胱甘肽和抗坏血酸; 其它抗氧化酶, 如过氧化氢酶和过氧化物酶; 以及金属螯合蛋白如金属硫蛋白, 可能通过减少氧化损伤使需氧微生物在有 O_2 的条件下生存。细胞质 SOD 的重要性通过缺乏 SOD 的啤酒糖酵母和大肠杆菌对双氧的高度敏感性证实。在两种生物体中, SOD 活性的丧失与有氧条件下的慢速生长, 较高的突变率和特异的生物合成缺陷有关。(sod1-酵母的需氧生长需要赖氨酸和蛋氨酸, 而 sod-大肠杆菌需要枝链氨基酸)。某些情况下, 这些效果已知是由超氧化物对铁硫簇蛋白的抑制作用而引起的(Gardner, P.R.等, 生物化学杂志, 266, 19328-19333, 1991, Kuo, C.F.等, 生物化学杂志, 262, 4724-4727, 1987)。如果, 生长在有氧条件下并以葡萄糖作为碳源, 啤酒糖酵母的 sod2 突变株不受影响。然而, 它们在要求需氧代谢的碳源中对高氧含量非常敏感而在正常氧含量(nomoxia)中生长缓慢。

因为编码 SOD 和过氧化氢酶的基因被从其它的种中克隆下来, 所以来源于 *P. rhodozyma* 的相应基因可采用简并 PCR 方法进行克隆。首先, 通过上述方法, 我们可以克隆含有部分 SOD 基因和 CAT 基因的一个部分基因片段。所述简并 PCR 方法是一种克隆目的基因的方法, 该目的基因与来源于其它种的功能相同或相似的已知酶具有氨基酸序列的高度同源性。通过将氨基酸序列反译成相应的(“简并”)核苷酸序列来设计简并引物, 该简并引物用来作为简并 PCR 方法中的引物。在这种简并引物中, 通常使用含有 A, C, G 或 T 中的任何碱基的混合引物, 或多义密码中含有肌苷的引物。本发明中, 采用这种混合引物作为简并引物来克隆上述基因。所采用的 PCR 反应条件随下述引物和要克隆的基因而变化。本发明中, 序列相互不同的两种 SOD 基因被从同一 PCR 带中经简并 PCR 克隆下来并被命名为 SOD1 和 SOD2 基因。

含有带内含子的编码区和调节区如启动子或终止子的全基因可通过筛选构建在合适宿主中的噬菌体载体或质粒载体上的基因组文库而从染色体上克隆下来, 方法是将上述简并 PCR 方法得到的一个部分 DNA 片段标记后

作为探针。在文库的构建及随后的遗传操作如测序、限制性消化、连接等过程中，通常采用大肠杆菌作为宿主菌，并用大肠杆菌载体、噬菌体载体如 λ 噬菌体载体、或质粒载体如 pUC 载体。本发明中，根据插入大小在 λ 载体的衍生体， λ ZAPII 和 λ DASHII 中构建 *P. rhodozyma* 的 EcoRI 基因组文库。插入大小，必须克隆的插入长度，是通过在构建文库前对每一基因进行 DNA 印记杂交来确定的。本发明中，按照厂商提供的说明书(Boehringer-Mannheim, Germany)，以地高辛配基(DIG)，即一种类固醇半抗原取代传统 ^{32}P 标记物来标记用作探针的 DNA。通过采用 DIG-标记的具有部分目的基因的 DNA 片段作为探针筛选由 *P. rhodozyma* 的染色体构建的基因组文库。挑选出杂交噬菌斑并进一步进行研究。如果使用 λ DASHII (插入大小为 9 至 23kb)，制备的 λ DNA 被 EcoRI 消化，然后将 EcoRI 插入片段克隆到质粒载体如 pUC19 或 pBluescriptIISK+ 中。当采用 λ ZAPII 构建基因组文库时，利用单链 M13 噬菌体衍生物 Ex 辅助噬菌体(Ex assist phage)(Stratagene, La Jolla, USA)，体内切除方案方便地用于克隆到质粒的后续步骤。所获质粒 DNA 经测序检查。本发明中，从 λ ZAPII 文库中分别获得 SOD1 和 SOD2 基因，而过氧化氢酶(CAT)基因则克隆自 λ DASHII 文库。

本发明中，我们使用了自动荧光 DNA 测序仪 ALFered 系统(Pharmacia, Uppsala, Sweden)，采用自动循环测序程序，其中大多数测序使用 Taq DNA 聚合酶。

本发明中，发明人确定了含有 SOD1 基因或 SOD2 基因的开放读框及其启动子和终止子序列的基因组序列。通过序列分析，根据其氨基末端存在转运肽来判断，发现 SOD1 编码线粒体 SOD。相反，SOD2 不具有这种转运肽序列并且这一事实暗示 SOD2 编码细胞质 SOD。发明人还确定了 CAT 基因的开放读框的部分基因组序列。

转运肽是一个信号序列，用来将核基因产物转移到定位于线粒体的蛋白如参与 TCA 循环的酶上，其中所述核基因在染色体上编码而其翻译蛋白在线粒体中发挥功能。在氨基末端加转运肽有利于某些蛋白质在线粒体中的表达。

本发明中，通过连接上述基因的部分片段来构建 SOD1，SOD2 和 CAT 基因的裂解质粒，该质粒不含基因的任一末端，具有自杀载体的抗药基因，由于缺乏自主复制序列而不能在 *P. rhodozyma* 中自主复制。一个编码能使宿

主在毒性抗生素存在下生存的酶的抗药基因通常被作为选择标记。定居在 pPR2T(Verdoes 等(国际专利出版物, WO97/23633))上的 G418 抗性基因是抗药基因的一个实例。这种自杀质粒不能自我复制, 只能利用载体和染色体之间的同源序列进行单交重组, 而以整合的形式存在于宿主染色体上。与目的基因重组时, 由于缺少基因的两端其基因序列不能重组到宿主菌的染色体上, 结果目的基因在如此获得的重组菌株中被裂解。裂解质粒的另一个实例是双交换型质粒。这种类型的裂解质粒含有待裂解之目的基因的两个不同片段并且筛选标记基因如抗药基因被插到这两个片段之间。受体细胞染色体和供体质粒 DNA 在该基因的两个同源区域重组后, 染色体序列被供体 DNA 置换且筛选标记基因被插入到待裂解的目的基因上。总之, 双交换型质粒比单交型载体的重组频率低。

在本发明中, 发明人通过相应基因的裂解使得目的酶失活。评价目的基因产物效果的其它方法是利用基因工程方法降低它的表达。某些方法可用于此目的。其中一种是反义方法。反义法是通过引入一段人工基因片段(其序列与目的基因的序列互补)来降低目的基因的表达。这种反义基因片段与目的基因的成熟 mRNA 片段体内形成复合体, 结果降低了 mRNA 的翻译效率。

另一方法是使启动子区突变。总之, 基因由具有不同功能的几个部分组成。在真核生物中, 编码相应蛋白质的基因被转录成前信使 RNA(pre-mRNA), 它不同于核糖体 RNA(rRNA), 小核 RNA(snRNA)和转运 RNA(tRNA)的基因。尽管 RNA 聚合酶 II(PolII)在该转录过程中起着关键作用, 但是没有能覆盖上游包含启动子之区域的顺式元件和上游活化序列(UAS), 以及反式作用蛋白因子, PolII 不能单独启动转录。首先, 由几个碱性蛋白组分组成的转录起始复合物识别表达基因的 5'-邻近区域的启动子序列。对需在某些特殊调节(如热休克反应, 或对营养缺乏的环境的适应等)下表达的基因来说, 上述过程还涉及其它因素。在这种情况下, 在 5'-非翻译上游区的启动子序列附近需要有一个 UAS 存在, 并且一些正调节或负调节蛋白识别 UAS 并与 UAS 结合。转录起始复合物与启动子序列的结合强度受启动子邻近区之反式作用因子结合的影响, 并且这能够调节转录活性。

通过磷酸化作用激活转录起始复合物后, 转录起始复合物启动从转录起始位点开始的转录。当延伸复合物从该基因的启动子区向 3'方向移动时, 转录起始复合物的某些部分脱离(这个步骤被称作启动子清除事件), 而延伸复

合物继续转录直至位于该基因 3'-侧下游区的终止序列。

为了降低目的基因的表达，通常在含有上述 UAS 序列的目的基因启动子区域进行常规化学诱变或基因定点诱变。用含有这样一个突变启动子区域的重组 DNA 转化宿主菌来获得其中目的酶表达降低的突变菌株。如上所述，这种降低基因表达的尝试也可如基因裂解一样，用于测定基因产物对活体生物体中环境的影响。

作为转化方法，LiCl 法(Wery 等，酵母，12(7)，641-651，1996)和电穿孔法(Wery 等，基因，184，89-97，1997)均被用于转化 *P. rhodozyma*，但是，这种条件下的转化效率似乎仍然很低。本发明中，Biolistic 转化方法(Johnston 等，分子生物学方法，53，147-153，1996)用于转化 *P. rhodozyma*。Biolistic 方法是一个简单可靠的方法，其中利用高压氦气将包裹在金或钨微粒表面的供体 DNA 直接射入活细胞中。用转化方法成功地转化了新型隐球菌，它和 *P. Rhodozyma* 同属担子菌纲(basidiomycetous)，并且用传统的转化方法很难转化(Toffaletti 等，细菌学，175(5)，1405-1411，1993)。本发明中，这一 Biolistic 方法被成功地用来转化 *P. rhodozyma* 细胞。

基因裂解的发生可直接通过酶活性证实，也可用所得转化体的染色体经 PCR 或 DNA 印迹杂交证实。本发明中，通过进行活性染色来直接证实 SOD 的裂解和通过类似于常用于细菌分类学的过氧化氢酶试验中采用的目测法来定性过氧化氢酶的裂解。

在合适的培养基中培养这种基因工程获得的 *P. rhodozyma* 并评估其产生虾青素的产率。

本发明用以下实施例参照附图作进一步举例说明。

实施例

实施例中使用的材料和方法描述如下：

菌株

P. rhodozyma ATCC96594(根据布达佩斯条约于 1998 年 4 月 8 日再次进行了保藏，保藏号为 ATCC74438)

大肠杆菌 DH5 α ：F⁻， ϕ 80d，lacZ Δ M15， Δ (lacZYA-argF)U169，hsd(r_K⁻，m_K⁺)，recA1，endA1，deoR，thi-1，supE44，gyrA96，relA1(Toyobo，Osaka，日本)

大肠杆菌 XL1-Blue MRF[']： Δ (mcrA)183， Δ (mcrCB-hsdSMR-mrr)173，

endA1, supE44, thi-1, recA1, gyrA96, relA1, lac[F⁻ proAB, lacI^q Z Δ M15, Tn10(ter^r)](Stratagene, La Jolla, USA)

大肠杆菌 SOLR : e14⁻ Δ (mcrA), Δ (mcrCB-hsdSMR-mrr)171, sbcC, recB, recJ, umuC :: Tn5(kan^r), uvrC, lac, gyrA96, relA1, thi-1, endA1, λ^R, [F⁻ proAB, lacI^q Z Δ M15]Su⁻(非抑制的(nonsuppressing))(Stratagene)

大肠杆菌 XL1- MRA(P2): Δ (mcrA)183, Δ (mcrCB-hsdSMR-mrr)173, endA1, supE44, thi-1, gyrA96, relA1, lac(P2 溶原菌)(Stratagene)

大肠杆菌 TOP10 : F⁻, mcrA, Δ (mrr -hsdRMS- mcrBC)φ80, Δ lacZM15, Δ lacX74, recA1, deoR, araD139, (ara-leu)7697, galU, galK, rpsL(Str^r), endA1, nupG(Invitrogen, Carlsbad, USA)

载体

λZAPII(Stratagene)

λDASHII(Stratagene)

pBluescriptII SK⁺(Stratagene)

pCR2.1TOPO(Invitrogen)

pUC19(Takara Shuzo, Otsu, 日本)

培养基

P. rhodozyma 菌株常规维持于 YPD 琼脂平板上(DIFCO, Detroit, 美国)。大肠杆菌菌株被保存在 LB 培养基中(每升培养基含有 10g 细菌用胰蛋白酶, 5g 酵母浸膏(DIFCO)和 5g NaCl)。NZY 培养基(每升培养基含有 5g NaCl, 2g MgSO₄-7H₂O, 5g 酵母浸膏(DIFCO), 10g A 型 NZ 胺(NZ amine type A)(WAKO, Osaka, 日本))被用于在软琼脂(0.7%琼脂(WAKO))中繁殖 λ 噬菌体。制备琼脂培养基需要加入 1.5%的琼脂(WAKO)。

方法

分子遗传技术的一般方法是根据“分子克隆”: 实验室手册, 第2版(冷泉港实验室出版社, 1989)中的方法。从 Takara Shuzo(日本)购买限制酶和 T4 DNA 连接酶。

用 QIAGEN 基因组试剂盒(QIAGEN, Hilden, Germany)按照厂家提供的说明书从 *P. rhodozyma* 中分离染色体 DNA。利用 DNA 自动分离系统(PI-50, Kurabo, Co. Ltd., Osaka, 日本)从大肠杆菌中小量制备质粒 DNA。

通过 QIAGEN 柱(QIAGEN)从大肠杆菌转化体中量制备质粒 DNA。利用 Wizard lambda 制备 DNA 纯化系统(Promega, Madison, USA)按照厂家提供的说明书分离λDNA。采用 QIAquick 或 QIAEX II(QIAGEN)在琼脂糖上分离和纯化 DNA 片段。λ 噬菌体衍生体的操作按照厂家制定的说明书进行(Stratagene)。

通过使用 Isogen(Nippon Gene, Toyama, 日本), 利用酚法从 *P. rhodozyma* 中分离总 RNA。使用 mRNA 分离试剂盒(Clontech, Palo Alto, USA)从如此获得的总 RNA 中纯化出 mRNA。使用 CapFinder cDNA 构建试剂盒(Clontech)合成 cDNA。

使用 GigapackIII 金包装提取物(Stratagene)进行体外包装。

聚合酶链反应(PCR)在 Perkin Elmer 2400 型的热循环仪中进行。实施例中描述了每一 PCR 反应条件。PCR 引物从商家购买或采用 DNA 合成仪合成(392 型, Perkin Elmer, 日本, Urayasu, 日本)。用于 DNA 测序的荧光 DNA 引物购自 Pharmacia 公司。利用荧光 DNA 自动测序仪(ALFred, Pharmacia)进行 DNA 测序。

大肠杆菌 DH5α 的感受态细胞购自 Toyobo(日本)。

P. rhodozyma 进行 biolistic 转化所使用的设备和试剂购自 Nippon Bio-Red Laboratories(日本, 东京)。

实施例 1：从 *P. rhodozyma* 中分离 mRNA 并构建 cDNA 文库

为了构建 *P. rhodozyma* 的 cDNA 文库, 细胞破碎后, 紧接着采用酚提取法分离总 RNA, 利用 mRNA 分离试剂盒(Clontech)从 *P. rhodozyma* 的 ATCC96594 菌株中纯化出 mRNA。

首先, 通过离心(1500 x g 离心 10 分钟)从用 YPD 培养基培养了两天的 10ml 培养物中收集 ATCC96594 菌株的细胞并用提取缓冲液(10mM 柠檬酸钠/HCl(PH6.2)含有 0.7M KCl)洗涤一次。将细胞悬浮在 2.5ml 的提取缓冲液中, 用弗氏压碎匀浆器(Ohtake Works Corp., 日本, 东京)按照生产厂家提供的方法以 1500kgf/cm² 的条件将细胞破碎并且马上与两倍体积的 isogen(Nippon gene)混合。通过这一步骤, 回收 400μg 的总 RNA。

然后, 用 mRNA 分离试剂盒(Clontech)按照厂家提供的方法纯化该总 RNA。最后, 得到 16μg *P. rhodozyma* ATCC96594 菌株的 mRNA。

用 CapFinder PCR cDNA 构建试剂盒(Clontech)按照厂家提供的方法获

得 *P. rhodozyma* 的 cDNA。用 1 μ g 的纯 mRNA 经 PCR 扩增而合成第一条链。该 PCR 扩增反应后，得到了 1mg 的 cDNA 库。

实施例 2：两种 *P. rhodozyma* 的部分 SOD 基因的克隆

采用简并 PCR 方法克隆 *P. rhodozyma* 的一个部分 SOD 基因。两种混合引物根据其它菌种的已知超氧化物歧化酶基因的共同序列进行设计和合成，该两种混合引物的核苷酸序列见表 1。

表 1

用于克隆 SOD1 和 SOD2 基因的引物序列

Sod1 ; AARCAYCAYCARACNTAYGTNAA(有义引物)(SEQ ID NO: 10)

Sod4 ; GCCCANCCNGANCCYTGNACNCC(反义引物)(SEQ ID NO: 11)

(R=A 或 G ; Y=C 或 T ; N=A, C, G 或 T)

以 ExTaq(Takara Shuzo)为 DNA 聚合酶和以实施例 1 得到的 cDNA 库为模板进行 25 轮 PCR，每一轮为 94 $^{\circ}$ C 15 秒，46 $^{\circ}$ C 30 秒，72 $^{\circ}$ C 为 15 秒，然后使反应混合物在琼脂糖凝胶上电泳。回收具有所需长度的 PCR 带并利用 QIAquick(QIAGEN)按照生产厂家提供的方法进行纯化，然后连接到 pCR2.1-TOPO(Invitrogen)上。对感受态大肠杆菌 TOP10 进行转化后，挑选出 6 个白色菌落，然后利用 DNA 自动分离系统(Kurabo PI-50)分离质粒。测序后，发现两种克隆具有互不相同的序列，其推导出的氨基酸序列分别与已知 SOD 基因相类似。这些分离出的 cDNA 克隆被命名为 pSOD614#2 和 pSOD614#3 并被用来作进一步的研究。

实施例 3：从 *P. rhodozyma* 中分离基因组 DNA

利用 QIAGEN 基因组试剂盒按照生产厂家提供的方法从 *P. rhodozyma* 中分离基因组 DNA。

首先，通过离心(1500 x g 离心 10 分钟)从用 YPD 培养基过夜培养的 100ml 培养物中收集 *P. rhodozyma* ATCC96594 菌株的细胞并用 TE 缓冲液(含有 1mM EDTA 的 10mM Tris/HCl(pH 8.0))洗涤一次。将细胞悬浮在 QIAGEN 基因组试剂盒的 Y1 缓冲液 8ml 中，加入浓度为 2mg/ml 的溶细胞酶(SIGMA)经酶降解来破碎细胞并将反应混合物于 30 $^{\circ}$ C 温育 90 分钟，然后进行下一步的提取步骤。最后得到 20 μ g 的基因组 DNA。

实施例 4：以 pSOD614#2 和 pSOD614#3 作为探针的 DNA 印迹杂交

进行 DNA 印迹杂交以便克隆含 *P. rhodozyma* 的 SOD 基因的基因组片段。用 EcoRI 消化 2 μ g 的基因组 DNA 并进行琼脂糖凝胶电泳，然后进行酸性和碱性处理。经 transblot(Joto Rika, 日本, 东京)运行 1 小时使变性的 DNA 向尼龙膜(HybondN+, Amersham)转移。对转移到尼龙膜上的 DNA 进行热处理固定(80 °C, 90 分钟)。按照 DIG 多引发方法(DIG multipriming)(Boehringer Mannheim)标记模板 DNA(EcoRI 消化的 pSOD614#2 和 pSOD614#3)来制备探针。按照厂家提供的方法进行杂交。结果，与由 pSOD614#2 制备的探针杂交的带显示 7.5 千碱基(kb)的长度，而与由 pSOD614#3 制备的探针杂交的带显示 8.0 千碱基(kb)的长度。

实施例 5：含有序列互不相同的 SOD 基因的基因组片段的克隆

用 EcoRI 消化 4 μ g 的基因组 DNA 并进行琼脂糖凝胶电泳。然后，利用透析膜、通过传统的洗脱方法回收长度为 7 至 9kb 的 DNA。纯化的 DNA 与 1 μ g 的经 EcoRI-消化和 CIAP(牛小肠碱性磷酸酶)-处理的 λ ZAPII(Strangene)于 16 °C 下过夜以连接，并用 Gigapack III 金包装提取物(Strangene)进行包装。用包装的提取物感染大肠杆菌 XL1Blue MRF⁻菌株，然后涂布于 NZY 培养基(其在 LB 琼脂培养基上)。以 EcoRI 消化的 pSOD614#2 和 pSOD614#3 为探针筛选出大约 6000 个噬菌斑。6 个噬菌斑与标记的 pSOD614#2 探针进行了杂交，2 个噬菌斑与标记的 pSOD614#3 探针进行了杂交。然后，按照生产厂家提供的方法(Strangene)对杂交噬菌斑进行体内切除。以 sod1 和 sod4 为引物进行 PCR 分析，结果发现从 6 个 pSOD614#2-杂交噬菌斑分离的一个质粒与 pSOD614 #2 具有相同的片段。这一质粒被命名为 pSOD703。在以 sod1 和 sod4 为引物的 PCR 分析结果中，还发现从 2 个 pSOD614#3-杂交噬菌斑分离的两个质粒与 pSOD614#3 具有相同的片段。其中一个质粒被命名为 pSOD626 并用于进一步的研究。

实施例 6：从 *P. rhodozyma* 中得到的两种 MnSOD 基因的序列分析

采用引物-步行(primer-walking)法对 pSOD703 和 pSOD626 进行测序来确定完整核苷酸序列。位于 pSOD703 上的 SOD1 基因和位于 pSOD626 上的 SOD2 基因的核苷酸及其推导的氨基酸序列为序列表中的 SEQ ID NO: 1, NO: 2, NO: 5 和 NO: 7。

BLAST 分析(Altschul.S.F.等, 分子生物学杂志 215, 403-410, 1990)的结

果表明, SOD1 和 SOD2 基因的推导氨基酸序列与从其它菌种中得到的已知 MnSOD 同源, 但不与 Cu/Zn SOD 或 Fe SOD 同源。

SOD1 基因具有 7 个内含子和 8 个外显子, 其推导的开放阅读框由 223 个氨基酸组成。另一方面, SOD2 基因具有 10 个内含子和 11 个外显子, 其推导的开放阅读框由 199 个氨基酸组成。两种分离的 SOD 基因之间的最大差别在于 SOD1 基因氨基酸末端的延伸区, 该序列可能作为到线粒体的转运肽。

事实上, Schroeder 等已报导了两种 SOD, 通过非变性聚丙烯酰胺凝胶电泳(PAGE)的活性染色, 分析出该两种 SOD 是源自 *P. rhodozyma* 的抗 KCN 和 H₂O₂ 的 SOD。他们指出两种 MnSOD 是凝集体或同工酶, 但没有提到它们的精确特征和亚细胞定位。如下所述, 这两种抗 KCN 和 H₂O₂ 的 SOD(即 MnSOD)是 SOD1 和 SOD2 基因的产物。如“发明详述”部分所述, 大多数真核生物在其细胞内不同部位具有不同种类的 SOD(MnSOD 在线粒体中, Cu/ZnSOD 在细胞质中)。这是两种 Mn SOD 在不同亚细胞位置上发挥作用的第一个实例。

实施例 7: SOD1 和 SOD2 基因的裂解质粒的构建

如“发明详述”部分所述, 通过将 G418 抗性结构基因插入到甘油醛-3-磷酸脱氢酶(GAP)基因的启动子和终止子之间并连接到 KpnI 和 HindIII 消化的 pUC19 上来构建具有抗药标记盒的质粒。这一质粒被命名为 pUC19-G418 并用来作进一步的研究。通过以表 2 所示序列为引物的 PCR 方法体外合成 SOD1 和 SOD2 基因的部分片段用于同源重组。

表 2

Sod14:GGTACCTCCGATGATAGGAATGTGAG(有义引物)(SEQ ID NO: 12)

Sod15:GAATTCAGTTCAACGGAGGAGGACAC(反义引物)(SEQ ID NO: 13)

Sod47:GAATTCGGAGGAGGACACATCAACCG(有义引物)(SEQ ID NO: 14)

Sod48:GGTACCTGTACTGGAGGTAGAAAGCG(反义引物)(SEQ ID NO: 15)

PCR 的条件是: 25 个循环反应, 每个循环反应为 94 °C 维持 15 秒钟,

50 °C 维持 30 秒钟和 72 °C 维持 15 分钟。以 0.1µg 由实施例 3 得到的基因组 DNA 为模板,以 ExTaq 为 DNA 聚合酶。分别以 sod14 和 sod15 为引物经 PCR 扩增 SOD1 的一个部分片段,以 sod47 和 sod48 为引物经 PCR 扩增 SOD2 基因的一个部分片段。回收扩增的每一个 0.65kb 片段,并按照厂家提供的方案将其克隆到 pCR2.1-TOPO(Invitrogen)上。从大肠杆菌 TOP10 转化体的白色菌落中挑选 6 个单菌落并从那些转化体中制备质粒。通过限制性分析和测序,一个具有部分 SOD1 基因的克隆被筛选出来作进一步研究并被命名为 pSOD715。以类似的方式,一个具有部分 SOD2 基因的克隆被筛选出来并被命名为 pSOD826。

然后,用 EcoRI 和 KpnI 消化 pSOD715 和 pSOD826,所产生的 0.65kb 片段采用 QIAquick 方案纯化,并连接到 EcoRI-和 KpnI-消化的 pUC19-G418 上。从大肠杆菌 DH5α 转化体的抗氨苄青霉素菌落中挑选 6 个单菌落并从那些转化体中制备质粒。通过限制性分析和测序,获得了一个克隆,其中一个部分 SOD1 被融合到 G418 抗性盒的克隆并被命名为 pSOD/G717。以类似的方式,获得了一个部分 SOD2 被融合到 G418 抗性盒的克隆并被命名为 pSOD/G828。

实施例 8 : 用 biolistic 方法转化 *P. rhodozyma* ATCC96594

按照“分子生物学方法”(Johnston 等, 53 ; 147-153, 1996)中描述的方法制定转化方案。作为宿主菌的 *P. rhodozyma* ATCC96594 在 YPD 培养基中培养到稳定期。离心培养液后,用无菌水将细胞浓缩 10 倍,然后取 200µl 的细胞悬浮液铺在含有 100µg 的遗传霉素, 0.75M 的甘露醇和山梨糖醇的 YPD 培养基上。将 5µg 来自 pSOD/G717 和 pSOD/G828 的环形 DNA 包裹上 1.5mg 的 0.9µm 金微粒来作为 biolistic 转化的供体 DNA。20 °C 下培养一星期后产生了几百个遗传霉素-抗性克隆。挑选 4 个这种转化体并制备其染色体。通过 PCR 和 DNA 印迹杂交分析,其中一个转化体被确定具有 SOD1 或 SOD2 基因的染色体裂解结构,并用来作进一步的研究。

实施例 9 : 对从 SOD1 和 SOD2 裂解物的候选者中获得的粗提物进行非变性聚丙烯酰胺凝胶电泳的活性染色分析

ATCC96594 菌株和从 ATCC96594 的 biolistic 转化中获得的候选者用 YPD 培养基培养两天后,在 4 °C 下以 3000 x g 离心 10 分钟进行收集。用 Tris-HCl 缓冲液(10mM/PH8.0)洗涤后,用同样的缓冲液将细胞浓缩 10 倍。

用弗氏压榨匀浆器(Ohtake Works), 在 $1500\text{kg}/\text{cm}^2$ 下破碎细胞, 并在 1500rpm 下微量离心(TOMY, MRX150)匀浆级分来制备粗提物。

采用 PIERCE(Rockford, 美国)的 BCA 蛋白分析试剂测定如此获得的粗提物中的蛋白质浓度。按照 Schroeder W.A.等描述的方法(普通微生物学杂志, 139, 907-912, 1993), 对相当于 $300\mu\text{g}$ 蛋白质的粗提物进行非变性聚丙烯酰胺凝胶电泳。活性染色分析参考 Flohe 等描述的方法(酶学方法, 105, 93-104, 1984)。

活性染色的结果见图 1。亲代菌株, ATCC96594 的提取物中, 观察到两条带在暗背景下为透明带, 该透明带具有 SOD 活性。相反, 其中 SOD1 基因被裂解的 ATCC96594 :: pSOD/G717 菌株在不变性聚丙烯酰胺凝胶电泳中不具有高迁移率的活性带, 而其中 SOD2 基因被裂解的 ATCC96594 :: pSOD/G828 菌株在不变性聚丙烯酰胺凝胶电泳中不具有低迁移率的活性带。从这一结果发现, *P. rhodozyma* 的粗提物中的两种 MnSOD 是 SOD1 和 SOD2 基因的产物, 并且, 在不变性聚丙烯酰胺凝胶电泳中具有高迁移率和低迁移率的 SOD 分别对应于 SOD1 和 SOD2 基因的产物。

实施例 10 : 克隆 *P. rhodozyma* 的部分过氧化氢酶(CAT)基因

采用简并 PCR 方法克隆 *P. rhodozyma* 的部分 CAT 基因。如表 3 所示的两种混合引物是根据源自其它菌种的已知过氧化氢酶基因的共有序列进行设计和合成的。

表 3

用于克隆 CAT 基因的引物序列

Cat2 ; MGNTTYTCNACNGTNGGNGGNGA(有义引物)(SEQ ID NO: 16)

Cat5 ; CKRTGNCKYTGNGTRTCNGGRTA(反义引物)(SEQ ID NO: 17)
(M=A 或 C ; N=A, C, G 或 T ; Y=C 或 T ; K=G 或 T ; R=A 或 G)

以 ExTaq(Takara Shuzo)为 DNA 聚合酶和以实施例 3 得到的基因组 DNA 为模板进行 25 轮 PCR, 每轮为 94°C 维持 15 秒钟, 45°C 维持 30 秒钟, 72°C 维持 15 秒钟, 反应混合物进行琼脂糖凝胶电泳。回收具有 1.0kb 长度的一条 PCR 带并利用 QIAquick(QIAGEN)按照生产厂家提供的方法进行纯化, 然后连接到 pCR2.1-TOPO(Invitrogen)上。对感受态大肠杆菌 TOP10 进行转化后, 挑选出 6 个白色菌落, 然后利用 DNA 自动分离系统分离质粒。测序后,

发现两种克隆的序列推导出的氨基酸序列与已知的 CAT 基因相类似。这些分离的 DNA 克隆中的一个被命名为 pCAT702 并被用来作进一步的研究。

实施例 11：含有 CAT 基因的基因组片段的克隆

以与实施例 4 相似的方式，采用 pCAT702 作为探针进行 DNA 印迹杂交的研究。结果，观察到了大小为 9kb 至 23kb 的杂交带。接着，用 EcoRI 消化 4 μ g 的基因组 DNA 并进行琼脂糖凝胶电泳。然后，利用透析膜、通过传统的洗脱方法回收长度为 9kb 至 23kb 的 DNA。纯化的 DNA 与 1 μ g 的经 EcoRI-消化和 CIAP(牛小肠碱性磷酸酶)-处理的 λ DASHII(Strangene)于 16 $^{\circ}$ C 过夜以进行连接，并用 Gigapack III 金包装提取物(Strangene)进行包装。用包装的提取物感染大肠杆菌 XL1Blue MRA(P2)菌株，然后涂布于 NZY 培养基(其在 LB 琼脂培养基上)。以 EcoRI 消化的 pCAT702 为探针筛选出大约 8000 个噬菌斑。6 个噬菌斑与标记的 pCAT702 探针进行了杂交。从每一 λ 克隆中制备 λ DNA，并且从以 Cat2 和 Cat5 为引物的 PCR 和序列分析结果中发现 6 个克隆中的 4 个具有与插入的 pCAT702 相同的片段。CAT 基因的一个部分核苷酸及其推导出的氨基酸序列为序列表部分的 SEQ ID NO: 3 和 SEQ ID NO: 9。

实施例 12：CAT 基因的裂解质粒的构建

以与实施例 7 相似的方式，构建 CAT 基因的裂解质粒。首先，SacI 接头被插入到 pUC19-G418 的 HindIII 位点上，其中定位着 G418-抗性盒的终止子区，并且限制性分析结果表明产生了 pUC19-G418Sa，其在 G418-抗性盒的末端具有 SacI 位点。然后，pUC19-G419Sa 的 KpnI-和 SacI-片段被连接到 KpnI-和 SacI-消化的 pCAT702 上，并产生 pCAT/706，其中一个部分基因组 CAT 基因与 G418-抗性盒融合。

实施例 13 以 pCAT/706 作为供体质粒转化 *P. rhodozyma* ATCC 96594

以类似于实施例 8 的方式，采用 CAT-裂解质粒，pCAT/706 转化 *P. rhodozyma* ATCC 96594。20 $^{\circ}$ C 下培养一星期后产生了几百个遗传霉素-抗性克隆。挑选 4 个这种转化体并制备其染色体。通过 PCR 和 DNA 印迹杂交分析，其中一个转化体被确定具有染色体 CAT 基因的裂解结构，并用来作进一步的研究。

然后采用过氧化氢酶试验定性 CAT 裂解物的两个候选者，该方法通常用来进行细菌分类的研究。将一接种环的 *P. rhodozyma* 细胞浸入 3%的 H₂O₂

溶液中，然后观察到氧气的生成。尽管当 ATCC 96594 细胞一进行过氧化氢酶试验，就立刻产生了 O_2 泡沫，而两株 ATCC 96594 :: pCAT/706 突变体被浸入到 H_2O_2 溶液中时，却很久才产生 O_2 泡沫。这一结果意味着 CAT 基因发生了裂解，但残余的微弱活性表明存在着另一种催化消除 H_2O_2 的物质如源自 *P. rhodozyma* 的过氧化物酶。

实施例 14 评价 *P. rhodozyma* 的 SOD1，SOD2 和 CAT 裂解体的虾青素产生

SOD1，SOD2 和 CAT 的基因裂解对虾青素之产生的影响用 YPD 培养基中的振荡来评价。将生长在 YPD 琼脂培养基上的细胞悬浮到 YPD 培养基中，然后将一部分细胞悬浮液接种在 500ml 带挡板摇瓶中的 50ml YPD 培养基中。细胞在 200r.p.m. 和 20 °C 的条件下培养 84 小时。间隔适当的时间，取 3ml 培养液并分析细胞的产量，葡萄糖的消耗及虾青素的含量。

细胞产量以 66nm 处的光密度计量，也可以细胞干重表示，干重是使细胞经 0.45 μ m 醋酸纤维膜加硝酸纤维膜(HAWP04700, Millipore, Bedford, 美国)过滤并于 80 °C 下加热过夜后测得。从以玻璃珠破碎的 *P. rhodozyma* 细胞中提取类胡萝卜素后，采用 HPLC 测定 *P. rhodozyma* 的虾青素含量。提取后，加入 5ml 含有适当浓度的胭脂树橙作为内标的丙酮/BHT/水。采用下述 HPLC 系统分析上清液中虾青素的含量。

HPLC 柱：YMC-Pak ODS-A(6mm, 150mm)

温度：室温

洗脱液：乙腈/甲醇/异丙醇(85/10/5)

进样体积：10 μ l

流速：2.0ml/分钟

检测：471nm 处的 UV

表 4 概述了从 84 小时-培养物中获得的结果。

表 4 SOD 和 CAT 突变对 *P. rhodozyma* 的总类胡萝卜素和虾青素产率的影响

菌株	总类胡萝卜素 (mg/g 细胞干重)	虾青素 (mg/g 细胞干重)
ATCC96594	0.169	0.111
Δ SOD1	0.259	0.146

△ SOD2	0.202	0.129
△ CAT	0.229	0.144

SOD1 和 SOD2 裂解体显示出高于其宿主菌， ATCC96594 的总类胡萝卜素和虾青素的产率。尤其是 SOD1 裂解体的类胡萝卜素和虾青素产量分别明显地增加了 53.3%和 31.5%。

根据推导的其氨基酸末端的转运肽序列来判断， SOD1 似乎是线粒体酶并且可能起着清除超氧化物自由基的作用，该超氧化物自由基是线粒体的呼吸链上出现的一种活性氧类物质。这些数据表明，由于胞内活性氧的产生而补偿了天然活性氧类物质淬灭因子 SOD1 的缺乏，从而促进了虾青素的产生。

序列表

<110> F. HOFFMANN-LA ROCHE AG

<120> 类胡萝卜素尤其虾青素的重组生产及其可利用的生物物质

<130> SOD

<140>

<141>

<160> 17

<170> PatentIn Ver. 2.0

<210> 1

<211> 3632

<212> DNA

<213> *Phaffia rhodozyma*

<220>

<221> 5'UTR

<222> (922)..(923)

<223> 实验的

<220>

<221> 外显子

<222> (986)..(1096)

<220>

<221> 内含子

<222> (1097)..(1306)

<220>

<221> 外显子

<222> (1307)..(1456)

<220>

<221> 内含子

<222> (1457)..(1555)

<220>

<221> 外显子

<222> (1556)..(1589)

<220>

<221> 内含子

<222> (1590)..(1694)

<220>

<221> 外显子

<222> (1695)..(1799)

<220>

<221> 内含子

<222> (1800)..(1920)

<220>

<221> 外显子

<222> (1921)..(1982)

<220>

<221> 内含子

<222> (1983)..(2076)

<220>
 <221> 外显子
 <222> (2077) .. (2140)

<220>
 <221> 内含子
 <222> (2141) .. (2246)

<220>
 <221> 外显子
 <222> (2247) .. (2272)

<220>
 <221> 内含子
 <222> (2273) .. (2390)

<220>
 <221> 外显子
 <222> (2391) .. (2507)

<220>
 <221> 聚腺苷酸位点
 <222> (2663) .. (2664)
 <223> 实验的

<400> 1
 tcctgttgat aatcttttcta acgccttgta ctttgaccaa ggcgtttgtc cgaaattttg 60
 caaacttagt gttggtcgca tggacggctc tcggatccag aactgacggc tcgccaataa 120
 agtatgacga tggtagaggt gaaggaggga accacaggtt gaccagtctc aaagagtgtc 180
 gatgtgcgcg aggatttgct attaaatggt gttgtatatg ctagagccaa gagaagacat 240
 ttggttttgg ttttggtttt gcatttgatg agatgtgtca cgattgaaga cgggaggagg 300
 ctactaacc caagaagcca ggatcaggag gaatgcctcc cctttttcat caagatcttt 360
 ctccatcgca acatttgaca ttctctttag tacccttcta tccttttctt ccaacttctc 420
 ccattgtatc gactttgctc gacttgctct tcttatctct gagcagagat gggcattcca 480
 atatcgaagg agcgacacaa gaccttgag tttgggtaac agatgaagag gggccgaggt 540
 ggatggggct gtaggaagta gctgatcgat gagttcctgg atgatgatag gcgaaggaac 600
 agacatagga tctctgtctc gtcctggaat tactgagtct tgtatccagc gtgttcttgt 660
 ctcaagaag ccttcaagat cgatgtaaga taagacaggc aatgaggacg gacgaacgaa 720
 cgaacgaaaa gaacagaaga gctggtaagt cagtcagtca gtcagtcagt caatcaaaca 780
 ctggtgtcta gggttatagc tcgacgcgac ggcacgcgtt tgagacgca tatgcttacg 840
 taatacctgg cgctcatccc ccagccgagg caagagccga gccgctcgtg aacgacaaaa 900
 ttcaaaaggc tttctccatc ttaagctcat tctcatctaa ccgactcatc tcgttcccat 960
 cattcccatc attctaccgc catcc atg tct gtt cga gca tcc ctc tct tcc 1012
 gtg tct aga cag act ttc gtc gct cct gct gct ttc cag atc agg gca 1060
 aag cat acc ctg cct gag ctt cct tac gct tac gat gtaagacttt 1106
 tccgtgttct cctattcgtc gctttcttgg ttttttctgt cttegcctc tagctcttct 1166

tcgtcctttc tgtcctgctc tttgttggtg atattcagct cgatagacta acccatctca 1226
 tctcctggac attcttttac tggaaacgta tcttgctcctt gggttttctt ggctttgggtt 1286
 gaaaattcct ctccactcag gcc ctg gag ccc tcc atc tcc aag gag atc atg 1339
 acc ctt cac cac acc aag cac cat cag act tat gtt aac ggc ctc aac 1387
 gct gcc gag gag agc tac tgg gcc gct gtg ggc aag gag gat gtg ctt 1435
 acc cag gtt aag ctt cag tct gtacgtctga ccgttttttt atcgaccgga 1486
 acgcctggtg aggagggaga tgaagtttga tgagcgctca tcgtctagca cgttgacccg 1546
 atcatacag gct ctc aag ttc aac gga gga gga cac atc aat c 1589
 gtcagtata ttcttcaaac tcttgctgag caagtcaggt caagctgact gtttcgcttt 1649
 gtttctgcgg atctatctca tccttgattt ggcatgatga aacag ac tct ctg ttc 1705
 tgg aag aac ttg gct ccc tat gga tcc gag gag gct acc ctc tct gaa 1753
 gga cct ctc aag aag gct atc gag gaa tct ttt ggt tct ttc gag g 1799
 gtccgtccat ctatcttctt attcagttgt gtttggttcc ggtatactca tctgttttgt 1859
 ttccccacaa aataaaaata aaaatcttgt cctctccggg ggttcgactg cacgttcata 1919
 g cc ttc aag aag aag ttc aac gct gac acc gct gct gtc caa gga tcc 1967
 gga tgg ggc tgg ctt gtatgtatca tatcctttcc atctcaaact cttctcagag 2022
 tctttttctt tgagacttca aactgactat acatgtttct acaacaaaca acag ggc 2079
 ttg aac ccg ctt act aag aag ctg gaa gtc acc acg acc gcc aac cag 2127
 gac cct ctg ctt a gtaagttgtt tctacatgat tttctatctc aacgcgatct 2180
 gcatgattcg tcaactgattc actggattct cttgtttcgt ttttctcggg atgatttcat 2240
 aaacag ct cac att cct atc atc gga gtt gac gtgcgtatct ttcttgaata 2292
 gtgtagcgt ctgatctcgt tttattgact gacgtgttgc ttctgtccaa atcattaaaa 2352
 aaaatgaaaa caaataatcg attgaccgac gaaaacag atc tgg gag cac gct ttc 2408
 tac ctt cag tac aag aac gtc aag cct gac tat ctc gct gct gtt tgg 2456
 tcc gtt atc aac tac aag gag gca gag gcc cga ttg cag gct gct ctc 2504
 taa gcggggacgaa aagtaacgac atatgaaggg aggatcaa atcgtttctt 2557
 cataaacaac tttcgaggca gatgggagag tacgtacaag agaggtttgt atggagaatt 2617
 gagtttggtg acggttagca ggttatgata tatgtagcta tagtctagtc taaatctgaa 2677
 agaagagaac aagatgggtt gtccgaagag attgagagat caagcccgtt catctgatgt 2737
 cgaacaaaca tgccctggtc tgccaacagt ttctagcaca ttatgaccat gttcatgtgt 2797
 aaattgggaa atgagccaga aaggtttatt atctaattca ttgattcatg cgactatgga 2857
 tacatatggg atttccagaa caaacagatg caacaaagca cggcattttc caaagatcga 2917
 gtctctccac aagtatgcgg caaggtttgt tgtaagaga tataaaagca gacgacaaaa 2977

caaatacgttt atcgaccctg tgcaccaaca ccgtgaccgt ttgacgagtt ggtagagttg 3037
tagttgttgc tgttcaaagg agctccagac tggacgcttc caagcttcaa caacttctcg 3097
gcagcgtcgc tgttcgggaa aagaaaaggc aaaaaggaac agagcgataa gcatatgtga 3157
ttctctactt cttataggct cttagctcaa gtcaactcac atgtcttttg cggtagcgaa 3217
gacgttctca agctgctgct tggaaagcttt tccgagcttg ccagtaggtc cctggttggg 3277
gaagaagatg tcgaaggcta agggcgatga aaagcatgaa gatattagct atcggcgcgga 3337
taaaagtgtg acgagatgaa aatggagaaa agatgattcg caccatcgac gacctcgacc 3397
aaaggaatgg aggtgtcacc ggccttcac ttcttgtact cctcaacgtt gacgaagatg 3457
acgaagcagt cgggtggcctt agcctcgggc tcgtagatgc tgatgaaaca caataggtag 3517
taggagagga gaaagagaag atgatgagat gtcaggatgc ttgcttcact gtagatggag 3577
gaagaagata tgcaagcaa gacatacact ttggaaagag cttgaacat tgtag 3632

<210> 2
<211> 3375
<212> DNA
<213> *Phaffia rhodozyma*

<220>
<221> 5'UTR
<222> (974)..(975)
<223> 实验的

<220>
<221> 外显子
<222> (1040)..(1063)

<220>
<221> 内含子
<222> (1064)..(1237)

<220>
<221> 外显子
<222> (1238)..(1280)

<220>
<221> 内含子
<222> (1281)..(1358)

<220>
<221> 外显子
<222> (1359)..(1402)

<220>
<221> 内含子
<222> (1403)..(1483)

<220>
<221> 外显子
<222> (1484)..(1589)

<220>
<221> 内含子
<222> (1590)..(1674)

<220>
 <221> 外显子
 <222> (1675) .. (1826)

<220>
 <221> 内含子
 <222> (1827) .. (1907)

<220>
 <221> 外显子
 <222> (1908) .. (1922)

<220>
 <221> 内含子
 <222> (1923) .. (1992)

<220>
 <221> 外显子
 <222> (1993) .. (2056)

<220>
 <221> 内含子
 <222> (2057) .. (2131)

<220>
 <221> 外显子
 <222> (2132) .. (2157)

<220>
 <221> 内含子
 <222> (2158) .. (2238)

<220>
 <221> 外显子
 <222> (2239) .. (2293)

<220>
 <221> 内含子
 <222> (2294) .. (2376)

<220>
 <221> 外显子
 <222> (2377) .. (2426)

<220>
 <221> 内含子
 <222> (2427) .. (2524)

<220>
 <221> 外显子
 <222> (2525) .. (2542)

<220>
 <221> 聚腺苷酸位点
 <222> (2667) .. (2668)
 <223> 实验的

<400> 2
 cttatccttc tgccgctggt ctctgtctgt cgagtgtgtg ggatggtttt ggatatgttc 60
 ctatacgaaa ggtagcgcag agcaaagctg acagtattaa gcaagacaag agcttccttc 120
 tgttgacaga tgaaaggacg aactatgaag ctgtccatgc tccccaaacc gattgacaca 180
 ccgccgtcag gcaacgcaga atttctcact gcttcgacgt cacaccaaca tcgatacctcc 240

```

atacctaataa gcagatcgag acacattggt ggtcgccatg ttggatggat gtacatcaaa 300
ccacagcat atatcactca catgtgagaa ctccgtagcc tctaccttct tgtctctcaa 360
tctgaatgtc tcgttgagag gtggaatgaa tgtttacagt ttgagaagac gaaagaaaga 420
aagagaagag aagagaggaa tacgtacgac gaagttatca tcgtatggga acttttctaa 480
aaaactgcct atagtagaga cgatctctgg aggaaagctc tgtagtatga tagtgaagag 540
cgagcaagtc tgggcaagtg catccttcgt ctacaagaaa gagaccagga aatgaaggag 600
agacgagtaa gcaggtacct accgatattg gatcgttctc tctaccagc gatgccttca 660
ccacgcgttc tatctcttct tgggatggca gatacatact taacgagagc aatctgatgt 720
ataccgaact tcgaacggaa tgatcccaga atcctcttga acccttgaac ccttgaaccc 780
tggaaaccaag taccaaccga gcaacacgcc gatacgttcc acaccacaga accacacgcc 840
ctcgtcatta aaggtgggac gcgccgatgc tggttacgtt cggcccaatc cggaagttac 900
cggcttggac gtgcctgtaa ccatgccttg acggtatttc gccttcagct aactccatct 960
catctttttc ctttactacc acaaccacc cttgaacctt cttccccggc ttttttacta 1020
tatccatcta tcaatcacc atg gct cct tac act ctt ccc gac gtaagcttaa 1073
agtttgagct gtgtgtgctt atctcaatct tggagttgaa ctcaccgttt tttgtttttg 1133
cttcttggtt tttttatcgg catccctcct ttttttccc tcgtggtcgc atatgatttg 1193
ctcatcaatc ggcgtttccc atgcattctt gtcattcgtt tcag ctt cct tac gct 1249
tac gat gcc ttg gag cct tac atc tct aag g gtgagattct tagtcagact 1300
gttgttccgg ttgcacacga tagctaactg tctctcgttc ctcaatatga acatgcag 1358
aa atc atg atc ctt cac cac tcc aag cac cat cag act tac gtc 1402
gtacgtaatc taaagggtcat ctccgtctac atggccggat caacttgctc atagatcttc 1462
cttctgttcg gcgctacgta g acc aac ctc aac gcc gct atc cag gct ttc 1513
tcc cag acc aat gac atc aag gcc cag atc gct ctt cag agc gct ctc 1561
aag ttc aac gga gga gga cac atc aac c gtacgatcat tctccctctt 1609
ctggcttata atatgtgttg cttgtcacta acacgcacgc aaccccgga tatctcacc 1669
tgtag ac tcc ctc ttc tgg aag aac atg gct cct gcc gac tct gct gat 1718
gcc aag ctc acc gag gga tcg ctc aag act gcc atc gac aag gac ttt 1766
gga tcc ttc gag gag ttc aag aag aag ttc aac act gct act ctc ggt 1814
gtc cag gga tct gtcagtatct cgtttgcttc gacatactct cagctttcct 1866
tccgtaaaact gacgaatagt ttttcggaca tgtacttgta g gga tgg gga tgg ctc 1922
gtgcgtttga cttttttcca ctttgaacat tagcgatagt gatacctaac aactgtgaat 1982
tggaatatag gga tac aac acc gct acc aag cac ctc gag atc gcc acc 2031
acc gcc aac cag gat ccc ctt atc a gtagtgact tctctcgtgt 2076

```

```

ggtcaccata agccagttgc tgacacattt cgttcgctgt ctctcgactt cgtag ct 2133
ttg act ccc atc att ggt ctt gac gttagtaatt ctatctagtg attggagtcg 2187
agttctgaac ttgccttgat ctcaaacgaa tgaatcaatt tcttttggtg g atc tgg 2244
gag cac gct ttc tac ctc cag tac aag aat gtc aag cct gat tac ctt g 2293
gtacgtaatt ctctattcgt ttgccccggt ttgatctttg actcactctt caaaatgttt 2353
tcgtttgta actttgaaaa acag cc gct ttc tgg aac gtc tgc aac ttt gct 2405
gag gct cag cga agg ttt gat gtgagtacag gcgctacccc tacggaggaa 2456
gcgaagggtga gctgaccact ttttatcttt ctgatttgga atgaacgata cgatgatcaa 2516
acaaacag gct gct gtc aag gct taa tgggccatt tatctctttg attcgacggc 2572
gatgacggct ttctcgcatc cgaagaaggc aaggctatga ttactgttat tctgccatgt 2632
tgcttgcttt gctatgctct atgttctttt cttttgcctc tcttcaaagc caaggcggtta 2692
aggaaggccc ttcagtctgt ttacatatg cacatataca tgagaacata tcacggactc 2752
ggcggctggt ggtcctcttg agcgtcggct tcaagattag tgtccacacg tgaagcggtc 2812
ggtgccatcc aacctggtag gaatcccat cgggcgggaa tccaattatc aattggcgggt 2872
cggccagatt cgagctcggg tatctcagaa gcgtcaagcg ggcgcatctc caggccttta 2932
agaggagcaa atttaatccg cctgggtggt cagcgagaca cgaacagttt gaaacagagt 2992
ctgcttggtga gttactcggc gagatcactg aggactaaac tttctcagct cgtggacgaa 3052
aagaacgaac caaacggtct tccctgtatc tcgaccatct ccttctccat ctcttacaac 3112
acctcggatg aactccaagg cttgctttcc aaagttcaaa caaactccgg gttgccatcc 3172
acctgggttg tctctaacga gccgagggat atccatcggt cggaacgttt gaacagactg 3232
gatggtaggt ggccggtcgc ttcggaagcc aatcataatg gtgggaatcg agagaaggaa 3292
tgattgggcc cagtgtttaa gacttggtgt tgttggcaga gtacggacgg aaagtaggac 3352
agacttaatc aaggcgagcc aag 3375

```

<210> 3
 <211> 951
 <212> DNA
 <213> *Phaffia rhodozyma*

<220>
 <221> 外显子
 <222> (1) .. (34)

<220>
 <221> 内含子
 <222> (35) .. (115)

<220>
 <221> 外显子
 <222> (117) .. (159)

<220>
 <221> 内含子

<222> (160)..(242)

<220>

<221> 外显子

<222> (243)..(363)

<220>

<221> 内含子

<222> (364)..(436)

<220>

<221> 外显子

<222> (437)..(951)

<400> 3

```
tcc gga agc tca gat acc gct cga gat cct cga g gtttgtgtgc 44
tttcgctttg ttgcgatgga tgaagctgtt aacttaaaaa aatcctcgtg tttctctttg 104
tttcaacata g gt ttc tct ctt aag gtc aag acc tct gag gga aac tgg 153
gac ttt gtacgtattc ttatcgactg agtcatcaag ctcgttatcg ctctcttacc 209
ctcatccttt tgtgtctctg tctacacctc tag gtc gga aac aac act ccc atc 263
ttt ttc ttg aga gac cca gcc aag ttt ccg atc ttc att cac acc cag 311
aag agg aac ccg cag aca aac tct aaa gac aag gac gct ttc tgg gac 359
tac c gttegtata accttgtcac tccctgcgtg ccgctctgat tcatgttgac 412
cttgtctttg atataatttt atag ta tcc caa aac ccc gag tcc gtg cat cag 465
gtg ctg cac ctg ttc agt gat cga gga acc cct gct tct tac cga cac 513
atg cat ggt tac tct gga cac acc ttc aag atg gtc aac agg aac ggt 561
gac tgg aat tat gtc cag att cac atg cgc acc gat cag ggt gtc aag 609
act cac acc aat gaa gag gct tcg aaa ctc gac gcc tcc aat ccc gat 657
tca aac gga gac gac ttg ttc gac gca atc aag aat gga gac ttc cct 705
agc tgg acg gtt cag gta cag gta atg tct cct gag cag gcc cag aag 753
ttc aga tac aac att ctg gat ctc acc aag gtc tgg tcc cac aag gag 801
ttc cca ctt aag acg att gga aag ttc act ttg aac cga aac gtg gat 849
aac tat ttc gca gag gtt gaa cag ctc gcc ttt gct cct tcc cat ctg 897
cct cct gga atc gag ccc tcg aac gat ccc gtc ctt cag gct cga cta 945
ttc tcc 951
```

<210> 4

<211> 669

<212> DNA

<213> *Phaffia rhodozyma*

<220>

<221> CDS

<222> (1)..(669)

<400> 4

```

atg tct gtt cga gca tcc ctc tct tcc gtg tct aga cag act ttc gtc 48
Met Ser Val Arg Ala Ser Leu Ser Ser Val Ser Arg Gln Thr Phe Val
1 5 10 15

gct cct gct gct ttc cag atc agg gca aag cat acc ctg cct gag ctt 96
Ala Pro Ala Ala Phe Gln Ile Arg Ala Lys His Thr Leu Pro Glu Leu
20 25 30

cct tac gct tac gat gcc ctg gag ccc tcc atc tcc aag gag atc atg 144
Pro Tyr Ala Tyr Asp Ala Leu Glu Pro Ser Ile Ser Lys Glu Ile Met
35 40 45

acc ctt cac cac acc aag cac cat cag act tat gtt aac ggc ctc aac 192
Thr Leu His His Thr Lys His His Gln Thr Tyr Val Asn Gly Leu Asn
50 55 60

gct gcc gag gag agc tac tcg gcc gct gtg ggc aag gag gat gtg ctt 240
Ala Ala Glu Glu Ser Tyr Ser Ala Ala Val Gly Lys Glu Asp Val Leu
65 70 75 80

acc cag gtt aag ctt cag tct gct ctc aag ttc aac gga gga gga cac 288
Thr Gln Val Lys Leu Gln Ser Ala Leu Lys Phe Asn Gly Gly Gly His
85 90 95

atc aat cac tct ctg ttc tgg aag aac ttg gct ccc tat gga tcc gag 336
Ile Asn His Ser Leu Phe Trp Lys Asn Leu Ala Pro Tyr Gly Ser Glu
100 105 110

gag gct acc ctc tct gaa gga cct ctc aag aag gct atc gag gaa tct 384
Glu Ala Thr Leu Ser Glu Gly Pro Leu Lys Lys Ala Ile Glu Glu Ser
115 120 125

ttt ggt tct ttc gag gcc ttc aag aag aag ttc aac gct gac acc gct 432
Phe Gly Ser Phe Glu Ala Phe Lys Lys Lys Phe Asn Ala Asp Thr Ala
130 135 140

gct gtc caa gga tcc gga tgg ggc tgg ctt ggc ttg aac ccg ctt act 480
Ala Val Gln Gly Ser Gly Trp Gly Trp Leu Gly Leu Asn Pro Leu Thr
145 150 155 160

aag aag ctg gaa gtc acc acg acc gcc aac cag gac cct ctg ctt act 528
Lys Lys Leu Glu Val Thr Thr Thr Ala Asn Gln Asp Pro Leu Leu Thr
165 170 175

cac att cct atc atc gga gtt gac atc tgg gag cac gct ttc tac ctt 576
His Ile Pro Ile Ile Gly Val Asp Ile Trp Glu His Ala Phe Tyr Leu
180 185 190

cag tac aag aac gtc aag cct gac tat ctc gct gct gtt tgg tcc gtt 624
Gln Tyr Lys Asn Val Lys Pro Asp Tyr Leu Ala Ala Val Trp Ser Val
195 200 205

atc aac tac aag gag gca gag gcc cga ttg cag gct gct ctc taa 669
Ile Asn Tyr Lys Glu Ala Glu Ala Arg Leu Gln Ala Ala Leu
210 215 220

```

<210> 5

<211> 222

<212> PRT

<213> Phaffia rhodozyma

<400> 5

```

Met Ser Val Arg Ala Ser Leu Ser Ser Val Ser Arg Gln Thr Phe Val
1 5 10 15

```

Ala Pro Ala Ala Phe Gln Ile Arg Ala Lys His Thr Leu Pro Glu Leu
 20 25 30
 Pro Tyr Ala Tyr Asp Ala Leu Glu Pro Ser Ile Ser Lys Glu Ile Met
 35 40 45
 Thr Leu His His Thr Lys His His Gln Thr Tyr Val Asn Gly Leu Asn
 50 55 60
 Ala Ala Glu Glu Ser Tyr Ser Ala Ala Val Gly Lys Glu Asp Val Leu
 65 70 75 80
 Thr Gln Val Lys Leu Gln Ser Ala Leu Lys Phe Asn Gly Gly Gly His
 85 90 95
 Ile Asn His Ser Leu Phe Trp Lys Asn Leu Ala Pro Tyr Gly Ser Glu
 100 105 110
 Glu Ala Thr Leu Ser Glu Gly Pro Leu Lys Lys Ala Ile Glu Glu Ser
 115 120 125
 Phe Gly Ser Phe Glu Ala Phe Lys Lys Lys Phe Asn Ala Asp Thr Ala
 130 135 140
 Ala Val Gln Gly Ser Gly Trp Gly Trp Leu Gly Leu Asn Pro Leu Thr
 145 150 155 160
 Lys Lys Leu Glu Val Thr Thr Thr Ala Asn Gln Asp Pro Leu Leu Thr
 165 170 175
 His Ile Pro Ile Ile Gly Val Asp Ile Trp Glu His Ala Phe Tyr Leu
 180 185 190
 Gln Tyr Lys Asn Val Lys Pro Asp Tyr Leu Ala Ala Val Trp Ser Val
 195 200 205
 Ile Asn Tyr Lys Glu Ala Glu Ala Arg Leu Gln Ala Ala Leu
 210 215 220

<210> 6
 <211> 597
 <212> DNA
 <213> Phaffia rhodozyma

<220>
 <221> CDS
 <222> (1)..(597)

<400> 6
 atg gct cct tac act ctt ccc gac ctt cct tac gct tac gat gcc ttg 48
 Met Ala Pro Tyr Thr Leu Pro Asp Leu Pro Tyr Ala Tyr Asp Ala Leu
 1 5 10 15
 gag cct tac atc tct aag gaa atc atg atc ctt cac cac tcc aag cac 96
 Glu Pro Tyr Ile Ser Lys Glu Ile Met Ile Leu His His Ser Lys His
 20 25 30
 cat cag act tac gtc acc aac ctc aac gcc gct atc cag gct ttc tcc 144
 His Gln Thr Tyr Val Thr Asn Leu Asn Ala Ala Ile Gln Ala Phe Ser
 35 40 45
 cag acc aat gac atc aag gcc cag atc gct ctt cag agc gct ctc aag 192
 Gln Thr Asn Asp Ile Lys Ala Gln Ile Ala Leu Gln Ser Ala Leu Lys
 50 55 60

```

ttc aac gga gga gga cac atc aac cac tcc ctc ttc tgg aag aac atg 240
Phe Asn Gly Gly Gly His Ile Asn His Ser Leu Phe Trp Lys Asn Met
65 70 75 80

gct cct gcc gac tct gct gat gcc aag ctc acc gag gga tcg ctc aag 288
Ala Pro Ala Asp Ser Ala Asp Ala Lys Leu Thr Glu Gly Ser Leu Lys
85 90 95

act gcc atc gac aag gac ttt gga tcc ttc gag gag ttc aag aag aag 336
Thr Ala Ile Asp Lys Asp Phe Gly Ser Phe Glu Glu Phe Lys Lys Lys
100 105 110

ttc aac act gct act ctc ggt gtc cag gga tct gga tgg gga tgg ctc 384
Phe Asn Thr Ala Thr Leu Gly Val Gln Gly Ser Gly Trp Gly Trp Leu
115 120 125

gga tac aac acc gct acc aag cac ctc gag atc gcc acc acc gcc aac 432
Gly Tyr Asn Thr Ala Thr Lys His Leu Glu Ile Ala Thr Thr Ala Asn
130 135 140

cag gat ccc ctt atc act ttg act ccc atc att ggt ctt gac atc tgg 480
Gln Asp Pro Leu Ile Thr Leu Thr Pro Ile Ile Gly Leu Asp Ile Trp
145 150 155 160

gag cac gct ttc tac ctc cag tac aag aat gtc aag cct gat tac ctt 528
Glu His Ala Phe Tyr Leu Gln Tyr Lys Asn Val Lys Pro Asp Tyr Leu
165 170 175

gcc gct ttc tgg aac gtc tgc aac ttt gct gag gct cag cga agg ttt 576
Ala Ala Phe Trp Asn Val Cys Asn Phe Ala Glu Ala Gln Arg Arg Phe
180 185 190

gat gct gct gtc aag gct taa 597
Asp Ala Ala Val Lys Ala
195

```

<210> 7
 <211> 198
 <212> PRT
 <213> Phaffia rhodozyma

```

<400> 7
Met Ala Pro Tyr Thr Leu Pro Asp Leu Pro Tyr Ala Tyr Asp Ala Leu
1 5 10 15

Glu Pro Tyr Ile Ser Lys Glu Ile Met Ile Leu His His Ser Lys His
20 25 30

His Gln Thr Tyr Val Thr Asn Leu Asn Ala Ala Ile Gln Ala Phe Ser
35 40 45

Gln Thr Asn Asp Ile Lys Ala Gln Ile Ala Leu Gln Ser Ala Leu Lys
50 55 60

Phe Asn Gly Gly Gly His Ile Asn His Ser Leu Phe Trp Lys Asn Met
65 70 75 80

Ala Pro Ala Asp Ser Ala Asp Ala Lys Leu Thr Glu Gly Ser Leu Lys
85 90 95

Thr Ala Ile Asp Lys Asp Phe Gly Ser Phe Glu Glu Phe Lys Lys Lys
100 105 110

Phe Asn Thr Ala Thr Leu Gly Val Gln Gly Ser Gly Trp Gly Trp Leu

```

```

115              120              125
Gly Tyr Asn Thr Ala Thr Lys His Leu Glu Ile Ala Thr Thr Ala Asn
130              135              140
Gln Asp Pro Leu Ile Thr Leu Thr Pro Ile Ile Gly Leu Asp Ile Trp
145              150              155              160
Glu His Ala Phe Tyr Leu Gln Tyr Lys Asn Val Lys Pro Asp Tyr Leu
165              170              175
Ala Ala Phe Trp Asn Val Cys Asn Phe Ala Glu Ala Gln Arg Arg Phe
180              185              190
Asp Ala Ala Val Lys Ala
195

<210> 8
<211> 714
<212> DNA
<213> Phaffia rhodozyma

<220>
<221> CDS
<222> (1) .. (714)

<400> 8
tcc gga agc tca gat acc gct cga gat cct cga ggt ttc tct ctt aag 48
Ser Gly Ser Ser Asp Thr Ala Arg Asp Pro Arg Gly Phe Ser Leu Lys
1 5 10 15
gtc aag acc tct gag gga aac tgg gac ttt gtc gga aac aac act ccc 96
Val Lys Thr Ser Glu Gly Asn Trp Asp Phe Val Gly Asn Asn Thr Pro
20 25 30
atc ttt ttc ttg aga gac cca gcc aag ttt ccg atc ttc att cac acc 144
Ile Phe Phe Leu Arg Asp Pro Ala Lys Phe Pro Ile Phe Ile His Thr
35 40 45
cag aag agg aac ccg cag aca aac tct aaa gac aag gac gct ttc tgg 192
Gln Lys Arg Asn Pro Gln Thr Asn Ser Lys Asp Lys Asp Ala Phe Trp
50 55 60
gac tac cta tcc caa aac ccc gag tcc gtg cat cag gtg ctg cac ctg 240
Asp Tyr Leu Ser Gln Asn Pro Glu Ser Val His Gln Val Leu His Leu
65 70 75 80
ttc agt gat cga gga acc cct gct tct tac cga cac atg cat ggt tac 288
Phe Ser Asp Arg Gly Thr Pro Ala Ser Tyr Arg His Met His Gly Tyr
85 90 95
tct gga cac acc ttc aag atg gtc aac agg aac ggt gac tgg aat tat 336
Ser Gly His Thr Phe Lys Met Val Asn Arg Asn Gly Asp Trp Asn Tyr
100 105 110
gtc cag att cac atg cgc acc gat cag ggt gtc aag act cac acc aat 384
Val Gln Ile His Met Arg Thr Asp Gln Gly Val Lys Thr His Thr Asn
115 120 125
gaa gag gct tcg aaa ctc gac gcc tcc aat ccc gat tca aac gga gac 432
Glu Glu Ala Ser Lys Leu Asp Ala Ser Asn Pro Asp Ser Asn Gly Asp
130 135 140
gac ttg ttc gac gca atc aag aat gga gac ttc cct agc tgg acg gtt 480

```

```

Asp Leu Phe Asp Ala Ile Lys Asn Gly Asp Phe Pro Ser Trp Thr Val
145                      150                      155                      160

cag gta cag gta atg tct cct gag cag gcc cag aag ttc aga tac aac      528
Gln Val Gln Val Met Ser Pro Glu Gln Ala Gln Lys Phe Arg Tyr Asn
                      165                      170                      175

att ctg gat ctc acc aag gtc tgg tcc cac aag gag ttc cca ctt agg      576
Ile Leu Asp Leu Thr Lys Val Trp Ser His Lys Glu Phe Pro Leu Arg
                      180                      185                      190

acg att gga aag ttc act ttg aac cga aac gtg gat aac tat ttc gca      624
Thr Ile Gly Lys Phe Thr Leu Phe Asn Arg Asn Val Asp Asn Tyr Phe Ala
                      195                      200                      205

gag gtt gaa cag ctc gcc ttt gct cct tcc cat ctg cct cct gga atc      672
Glu Val Glu Gln Leu Ala Phe Ala Pro Ser His Leu Pro Pro Gly Ile
                      210                      215                      220

gag ccc tcg aac gat ccc gtc ctt cag gct cga cta ttc tcc              714
Glu Pro Ser Asn Asp Pro Val Leu Gln Ala Arg Leu Phe Ser
225                      230                      235

<210> 9
<211> 238
<212> PRT
<213> Phaffia rhodozyma

<400> 9
Ser Gly Ser Ser Asp Thr Ala Arg Asp Pro Arg Gly Phe Ser Leu Lys
1                      5                      10                      15

Val Lys Thr Ser Glu Gly Asn Trp Asp Phe Val Gly Asn Asn Thr Pro
20                      25                      30

Ile Phe Phe Leu Arg Asp Pro Ala Lys Phe Pro Ile Phe Ile His Thr
35                      40                      45

Gln Lys Arg Asn Pro Gln Thr Asn Ser Lys Asp Lys Asp Ala Phe Trp
50                      55                      60

Asp Tyr Leu Ser Gln Asn Pro Glu Ser Val His Gln Val Leu His Leu
65                      70                      75                      80

Phe Ser Asp Arg Gly Thr Pro Ala Ser Tyr Arg His Met His Gly Tyr
85                      90                      95

Ser Gly His Thr Phe Lys Met Val Asn Arg Asn Gly Asp Trp Asn Tyr
100                      105                      110

Val Gln Ile His Met Arg Thr Asp Gln Gly Val Lys Thr His Thr Asn
115                      120                      125

Glu Glu Ala Ser Lys Leu Asp Ala Ser Asn Pro Asp Ser Asn Gly Asp
130                      135                      140

Asp Leu Phe Asp Ala Ile Lys Asn Gly Asp Phe Pro Ser Trp Thr Val
145                      150                      155                      160

Gln Val Gln Val Met Ser Pro Glu Gln Ala Gln Lys Phe Arg Tyr Asn
165                      170                      175

Ile Leu Asp Leu Thr Lys Val Trp Ser His Lys Glu Phe Pro Leu Arg
180                      185                      190

```

Thr Ile Gly Lys Phe Thr Leu Asn Arg Asn Val Asp Asn Tyr Phe Ala
195 200 205

Glu Val Glu Gln Leu Ala Phe Ala Pro Ser His Leu Pro Pro Gly Ile
210 215 220

Glu Pro Ser Asn Asp Pro Val Leu Gln Ala Arg Leu Phe Ser
225 230 235

<210> 10

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列: sod1(克隆SOD基因的有义引物)的描述

<400> 10

aarcaycayc aracntaygt naa

23

<210> 11

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列: sod4(克隆SOD基因的反义引物)的描述

<400> 11

gcccancng anccytgnac ncc

23

<210> 12

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列: sod14(构建SOD1-裂解质粒的有义引物)的描述

<400> 12

ggtacctccg atgtaggaa tgtgag

26

<210> 13

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列: sod15(构建SOD1-裂解质粒的反义引物)的描述

<400> 13

gaattcagtt caacggagga ggacac

26

<210> 14

<211> 26

<212> DNA
<213> 人工序列

<220>
<223> 人工序列: sod47(构建SOD2-裂解质粒的有义引物)的描述

<400> 14
gaattcggag gaggacacat caaccg 26

<210> 15
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列: sod48(构建SOD2-裂解质粒的反义引物)的描述

<400> 15
ggtagcctgta ctggaggtag aaagcg 26

<210> 16
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列: sod2(克隆CAT基因的有义引物)的描述

<400> 16
mgnttytcna cngtnggngg nga 23

<210> 17
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列: cat5(克隆CAT基因的反义引物)的描述

<400> 17
ckrtgnckyt gngtrtcngg rta 23

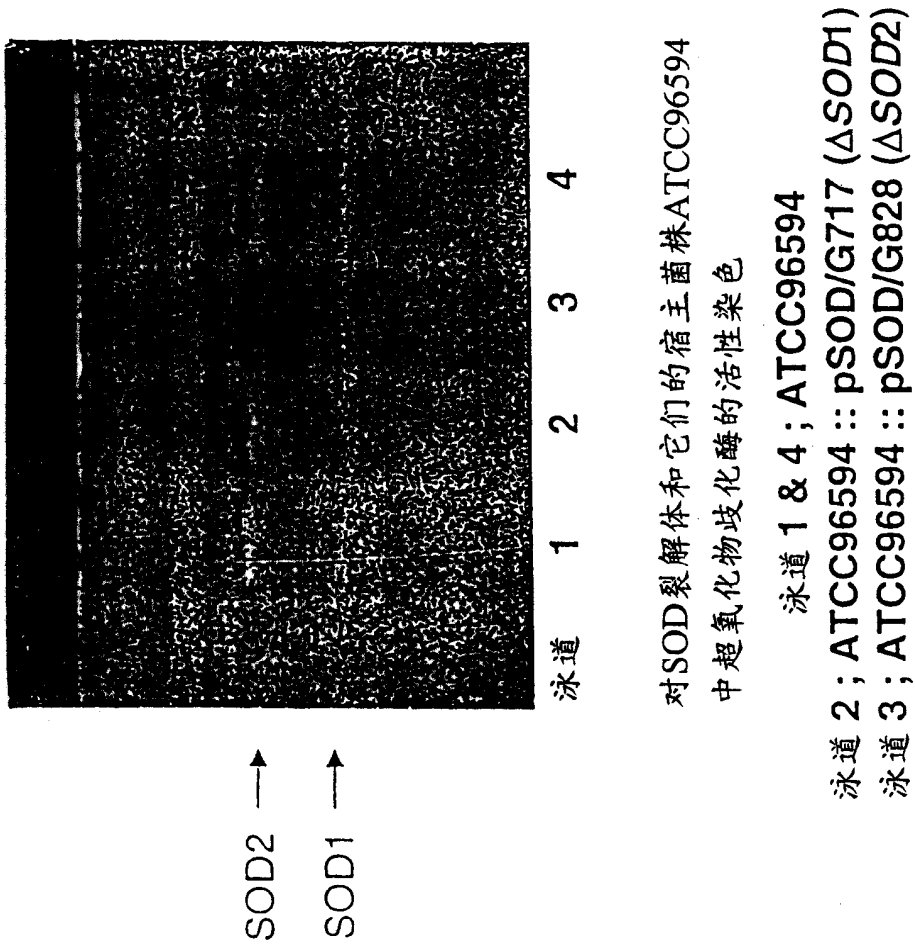


图 1