

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁶

C12N 15/54

C12N 9/10 A01H 1/00

[12] 发明专利申请公开说明书

[21] 申请号 98801214.6

[43]公开日 1999 年 12 月 1 日

[11]公开号 CN 1237206A

[22]申请日 98.7.16 [21]申请号 98801214.6

[30]优先权

[32]97.7.25 [33]JP [31]200571/97

[86]国际申请 PCT/JP98/03199 98.7.16

[87]国际公布 WO99/05287 日 99.2.4

[85]进入国家阶段日期 99.4.23

[71]申请人 三得利株式会社

地址 日本大阪

[72]发明人 水谷正子 田中良和 久住高章

齐藤和季 山崎真已 巩志忠

[74]专利代理机构 中国国际贸易促进委员会专利商标事
务所

代理人 隗永良

权利要求书 1 页 说明书 75 页 附图页数 0 页

[54]发明名称 编码具有糖转移活性的蛋白质的基因

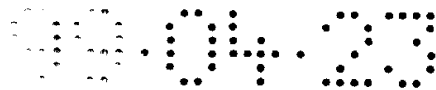
[57]摘要

提供编码具有对序列号:7~10 或 12 中任一序列记载的氨基酸序列的具有类黄酮 5 位糖转移活性的蛋白质的基因、和具有对这些氨基酸序列进行修饰的氨基酸序列、而且维持类黄酮 5 位糖转移活性的蛋白质的基因、以及使用该基因的上述蛋白质的制造方法。该基因可以用于植物的颜色的人工改良等方面。

ISSN 1 0 0 8 - 4 2 7 4

权 利 要 求 书

1. 编码具有类黄酮 5 位糖转移活性的蛋白质的基因。
2. 权利要求 1 记载的基因，是编码具有 序列 7~10 或 12 中任一个序列记载的氨基酸序列的具有类黄酮 5 位糖转移活性的蛋白质、或对这些氨基酸序列添加、缺失一个或多个氨基酸和/或具有通过其它氨基酸替代修饰的氨基酸序列、而且维持类黄酮 5 位糖转移活性的蛋白质的基因。
3. 权利要求 1 记载的基因，是编码对序列 7~10 或 12 中任一个序列记载的氨基酸序列具有 30% 以上同源性的氨基酸序列、而且具有类黄酮 5 位糖转移活性的蛋白质的基因。
4. 权利要求 1 记载的基因，是编码对序列 7~10 或 12 中任一个序列记载的氨基酸序列具有 50% 以上同源性的氨基酸序列、而且具有类黄酮 5 位糖转移活性的蛋白质的基因。
5. 权利要求 1 记载的基因，是编码对序列 7~10 或 12 中任一个序列记载的氨基酸序列的碱基序列的一部分或全部、在 5×SSC、50℃ 的条件下能杂化、而且具有类黄酮 5 位糖转移活性的蛋白质的基因。
6. 含有权利要求 1~5 任一项记载的基因的载体。
7. 利用权利要求 6 记载的载体转化的宿主。
8. 通过权利要求 1~5 任一项记载的基因编码的蛋白质。
9. 一种蛋白质的制造方法，其特征在于培养、或培育权利要求 7 记载的宿主、然后从该宿主提取具有类黄酮 5 位糖转移活性的蛋白质。
10. 权利要求 1~5 任一项记载的基因导入的植物或具有与该植物具有相同性质的其后代或它们的组织。
11. 权利要求 10 记载的植物或具有与该植物具有相同性质的其后代的切花。



说 明 书

编码具有糖转移活性的蛋白质的基因

本发明涉及编码具有类黄酮 5 位糖转移活性的蛋白质的基因以及该基因的利用方法。

花卉产业正努力开发新的且品种多样的花卉。作为培育新品种的有效方法之一往往是改变花色，利用经典的育种方法，就几乎所有商业品种花卉生成范围广泛的花色都可成功。然而，因为使用这种方法每种花的基因库受到限制，所以单一的一种品种而具有范围广泛的各種着色品种还是很稀少的。

花色主要来自两种类型的色素、即类黄酮和类胡萝卜素。类黄酮的贡献是在从黄色至红色乃至绿色范围，而类胡萝卜素的贡献是在橙色或黄色的色调。对花色起主要贡献的类黄酮分子是以花青素、花翠素、3'-甲花翠素、甲基花青素、二甲花翠素以及花葵素等糖苷的花色素，不同的花色素会带来花色的显著变化。另外，花色也受到无色的类黄酮的辅助发色、金属配合物的形成、葡萄糖基化、酰化、甲基化以及液泡的 pH 的影响 (Forkmann, 植物育种 (Plant Breeding), 106, 1, 1991)。

从苯丙氨酸开始的花色苷的生物合成途径人们已经了解的很清楚 (例如, 植物细胞 (Plant Cell), 7, 1071-1083, 1995)、与生物合成有关的基因几乎都已经被克隆了。例如、认为与作为紫苏的花色苷的丙二酰紫苏色素 (3-O-(6-O-(p-香豆酰基)- β -D-葡萄糖基)-5-O-(丙二酰- β -D-葡萄糖基)-花青素) 的生物合成有关的基因中、到现在为止还没有报道的该同源基因只是类黄酮-3'-羟基化酶基因、UDP-葡萄糖: 花色苷 (类黄酮) 5-O-葡萄糖基转移酶 (以下称 5GT) 基因、丙二酰基转移酶基因。

据推测, 其中类黄酮-3'-羟基化酶与已经知道是属于细胞色素 P450 基因家族 (植物细胞 (Plant Cell), 7, 1071-1083, 1995) 的细胞色素 P450 基因之间表现出结构的同源性。

一般来说，类黄酮的 3 位羟基是通过葡萄糖修饰的，但人们认为通过以葡萄糖基化为代表的糖的修饰会使花色苷的稳定性和溶解度增大（《黄酮》The Flavonoids, Chapman & Hall, 1994）。

编码催化该反应的 UDP-葡萄糖：花色素或类黄酮 3-葡萄糖基转移酶（以下称 3GT）的基因可以从玉米、大麦、金鱼草、龙胆等多种植物得到、相互间的氨基酸序列表现出显著的同源性。例如、单子叶植物玉米与双子叶植物龙胆的 3GT 的氨基酸序列的同源性为 32%、单子叶植物玉米与大麦的 3GT 的氨基酸序列的同源性为 73%、双子叶植物龙胆与茄子的 3GT 的氨基酸序列的同源性为 46%。

另外、编码矮牵牛花的 UDP-鼠李糖：花色素 3-葡萄糖苷鼠李糖基转移酶（3RT）的基因也已经克隆了。

可是、尽管很多植物的类黄酮的 5 位的羟基可被葡萄糖基化、但催化该反应的基因（5GT）还未得到。

另外、虽有测定将糖转移到矮牵牛花和紫罗兰的花色素 5 位上反应的例子（植物 (Planta) 160, 341-347, 1984、植物 (Planta) 168, 586-591, 1986），但这些报道只是停留在用花卉的粗提取液或部分纯化的酶研究其酶学性质方面，而将该酶纯化到纯品的例子还没有。另外、一般来说、糖转移酶从生物化学角度讲是不稳定的、酶的纯化是困难的。

通过在类黄酮分子上附加糖使其色调变化的例子几乎还没有，因为对色调有很大影响的芳香族酰基可转移到花色素内的葡萄糖分子或鼠李糖分子上、那么控制糖转移反应在控制花色素的生物合成、进而控制花色上是重要的。而作为调控糖转移酶基因的表达改变花色的例子，如在转化矮牵牛花中，就是控制矮牵牛花 3RT 引起的反应来修饰花色的。

作为可能转化的植物，例如已知有蔷薇、菊花、石竹、大丁草花、矮牵牛花、蓝竹耳、草原龙胆、カラコエ、郁金香、唐菖蒲等。

这里，本发明人等以获得编码具有类黄酮 5 位糖转移活性的蛋白质的基因作为课题，完成了本发明。

例如菊花的花色素以及蔷薇和石竹花的花色苷中的一部分结构中，5

位的羟基还没有被葡萄糖基化。通过将本发明得到的 5GT 基因导入这些植物、可以改变花色苷的结构。

通过利用国际公开公报：W096/25500 记载的酰基转移酶基因使类黄酮酰基化是有可能使花色改变、或是使类黄酮稳定、但酰基并不是直接与类黄酮结合、而是通过糖结合的、所以只是导入酰基转移酶基因的话、有时花色的变化并不充分、有时也不稳定。

然而、通过将 5GT 基因与酰基转移酶基因同时导入、可以使糖转移到类黄酮的 5 位上、进而可使其酰化、预料花色素的结构也会改变、花色也会变绿。

另外、如果利用反意义法或コサプレツシ ヲ法等抑制花色素的 5 位葡萄糖基化的植物的 5GT 基因的表达、可以抑制花色素的生物合成、其结果可以使花色改变。例如、如果在龙胆或草原龙胆中抑制 5GT 活性的话、可预料花色会变红。

本发明人利用基因重组技术、从紫苏、蓝竹耳、马鞭草以及矮牵牛花中分离出 5GT 的 cDNA、确定了结构基因的碱基序列。就是说、提供编码存在于这些植物花色苷出现的组织中的 5GT 的 DNA 序列。而且由于该酶使得花色素系色素的 5 位上转移了糖、所以可用于花色的变化、也可增加花色素的稳定性。

为了得到编码本发明的酶的 DNA、例如可以利用差别显示 (Differential display) 法。例如在紫苏 (*Perilla frutescens*) 中有蓄积花色苷的品种 (例如紫薰) 和没有蓄积花色苷的品种 (青薰)、如果克隆存在于蓄积花色素的品种而在不蓄积花色素品种中不存在的 DNA 的话、可以得到编码本发明的酶的 DNA。

具体来说、从紫薰和青薰的叶提取 RNA、按照常规方法合成 cDNA、经电泳分离、可分离出存在于来自紫薰 cDNA 文库的而不存在于来自青薰 cDNA 文库的 cDNA。然后以得到的 cDNA 作为探针、筛选来自紫薰的 cDNA 文库、得到编码本发明酶的基因。

在本发明中、作为上述的筛选的例子、利用差别显示法筛选来自紫苏的编码本发明酶的 DNA (实施例 1)、然后以得到的 DNA 作为探针、通过

筛选来自马鞭草的 cDNA、得到来自马鞭草的编码本发明酶的 DNA (实施例 2)、通过同样的操作可以得到来自蓝竹耳的编码本发明酶的基因 (实施例 3)。

然后确认了这些 DNA 能够表达具有本发明酶活性的蛋白质。

另外、得到了编码来自矮牵牛花的本发明酶的 DNA (实施例 4)。

作为本发明的 DNA、例如有编码序列 7~10 或 12 中任一个序列记载的氨基酸序列的 DNA。然而已知、具有通过添加、缺失数个氨基酸和/或通过其它氨基酸取代修饰的氨基酸序列的蛋白质也可维持与原来蛋白质同样的酶活性。因此、编码具有对序列号: 7~10 或 12 中任一个序列记载的氨基酸序列添加、缺失一个或数个氨基酸的和/或通过其它氨基酸取代修饰的氨基酸序列、而且维持类黄酮 5 位糖转移活性的蛋白质的基因也属于本发明。

本发明还涉及序列 1~4 或 6 中任一个序列记载的碱基序列或编码记载的氨基酸序列的碱基序列或对那些部分、例如对编码共有区的 6 个氨基酸以上的氨基酸部分、例如在 2 至 5×SSC、例如 5×SSC、50℃ 的条件下杂化、而且编码具有类黄酮 5 位糖转移活性的蛋白质的基因。另外、最适的杂化温度因碱基序列或其长度的不同而不同、优选的是随着碱基序列变短、杂化温度也降低, 例如编码 6 个氨基酸的碱基序列 (18 个碱基) 其杂化温度优选是在 50℃ 以下。

通过这样杂化选择的基因有来自天然的、例如来自植物的、如来自马鞭草或来自蓝竹耳的基因, 也可以是来自其它植物的、例如来自矮牵牛花、蔷薇、石竹、风信子等植物的基因。另外、通过杂化选择的基因既可以是 cDNA、也可以是基因组 DNA。

另外本发明还涉及编码对 7~10 或 12 中任一个序列记载的氨基酸序列具有 30% 以上、优选是在 50% 以上、例如 60% 或 70% 以上的同源性、有时达 90% 的同源性的氨基酸序列、而且具有类黄酮 5 位糖转移活性的蛋白质的基因。就是说、就象实施例所给出的那样、编码本发明的酶的 DNA 与其它糖转移酶基因比较表现出 20~30% 的同源性。因此、本发明也包含编码表现出与 7~10 或 12 中记载的氨基酸序列 30% 以上的同源性、而

且具有类黄酮 5 位糖转移活性的蛋白质的基因。

就象由实施例 1~4 的结果的比较所表明的那样、本发明酶的氨基酸序列因品种而异、种间的同源性在 50% 以上(参照实施例 3 和 4)、例如 60~70% (参照实施例 2)、而来自同一品种的酶的氨基酸序列的同源性在 90% 以上(参照实施例 1)。因此、编码对序列 7~10 或 12 中记载的氨基酸序列具有 50% 以上、例如 60~70% 以上、例如 60% 或 70% 以上的同源性、有时达 90% 的同源性的氨基酸序列、而且维持本发明的糖转移酶活性的蛋白质的基因也包括在本发明的范围。

具有天然的碱基序列的 DNA、就象实施例具体记载的那样, 例如可以通过 cDNA 文库筛选得到。

而编码具有修饰的氨基酸序列的酶的 DNA 是可以以具有天然的碱基序列的 DNA 为基础、利用常用的特定部位的诱导变异和 PCR 法合成。例如、将含有希望导入修饰的部位的 DNA 片段、是从上述得到的 cDNA 或基因组 DNA 经限制性内切酶消化得到的、作为模板、利用导入了所希望的变异的引物实施特定部位的诱导变异或 PCR 法、得到了导入了所希望的修饰的 DNA 片段、如果将该片段与编码目的酶的其它部分 DNA 连接就可以合成出编码具有修饰的氨基酸序列的酶的 DNA 了。

或者、为了得到编码具有缩短了氨基酸序列的酶的 DNA, 例如可以通过所希望的限制性内切酶消化编码比目的氨基酸序列长的氨基酸序列的 DNA、例如编码全长氨基酸序列的 DNA, 得到的 DNA 片段不能编码整个目的氨基酸序列时、不足部分可以通过连接合成 DNA 予以补足。

使用大肠杆菌以及酵母中的基因表达体系使该克隆表达、通过测定酶活、确认得到的基因是编码糖转移酶的基因, 通过弄清楚于类黄酮的 5 位转移糖的糖转移酶基因的翻译区, 就得到编码与本发明有关的糖转移酶的基因, 另外、通过使该基因表达可以得到作为基因产物的于类黄酮 5 位转移糖的糖转移酶目的蛋白质。

或者是利用相应于序列 7~10 或 12 中任一个序列记载的氨基酸序列的抗体也可得到上述蛋白质。

因此本发明还涉及含有上述 DNA 的重组载体、特别是表达载体、以及

通过该载体转化的宿主。作为宿主可以用原核生物或真核生物。作为原核生物可以利用细菌、例如属于 *Escherichia* 菌属的细菌、例如大肠杆菌 (*Escherichia coli*)、属于 *Bacillus* 属的微生物、例如枯草杆菌 (*Bacillus subtilis*) 等常用的宿主。

作为真核的宿主可以使用低等真核生物、例如真核微生物、例如作为真菌的酵母或霉菌。作为酵母菌有属于 *Saccharomyces* 属的微生物、例如酿酒酵母 (*Saccharomyces cerevesiae*) 等、而作为霉菌有属于 *Aspergillus* 属的微生物、例如米曲霉 (*Aspergillus oryzae*)、黑曲霉 (*Aspergillus niger*)、以及属于 *Penicillium* 属的微生物等。也可以使用动物细胞或植物细胞、作为动物细胞可以使用小鼠、大鼠、猴、人等的细胞系。而也可以使用昆虫细胞、例如蚕细胞、或蚕的成虫本身作为宿主。

本发明的表达载体依赖于应导入该载体的宿主的种类，含有表达调控区、例如启动子以及终止子、复制起点等。作为细菌用的表达载体的启动子使用常用的启动子、例如 *trc* 启动子、*tac* 启动子、*lac* 启动子等，作为酵母用的启动子、例如可使用甘油醛-3-磷酸脱氢酶启动子、*PH05* 启动子等、作为霉菌用启动子、例如可使用淀粉酶、*trp C* 等。而作为动物细胞宿主用启动子可以使用病毒性启动子、例如 *SV40* 早期启动子、*SV40* 晚期启动子等。

表达载体的制备可以使用限制性内切酶、连接酶等按照常规方法进行。另外、表达载体对宿主的转化也是按照常规方法进行。

上述蛋白质的制造方法中、通过对上述表达载体转化的宿主进行培育、栽培或饲养、按照常规方法、例如过滤、离心分离、细胞破碎、凝胶过滤层析、离子交换层析等从培养物中回收目的蛋白质、进行纯化。

另外在本说明书中虽然叙述了来自紫苏、马鞭草、蓝竹耳以及矮牵牛花的类黄酮 5 位糖转移的糖转移酶 (本发明中有时也简单称之“糖转移酶”)。可以通过直接使用或经部分改造的该酶的纯化方法纯化其它植物的糖转移酶、通过确定与该酶系有关的氨基酸序列、克隆编码该酶的基因。另外、通过使用来自与本发明有关的紫苏的糖转移酶的 cDNA 作为探

针、可以从紫苏中得到其它的糖转移酶的 cDNA、以及从其它植物得到别的糖转移酶的 cDNA。因此、如果使用糖转移酶基因的一部分或全部,可以得到其它的糖转移酶基因。

另外、就象本说明书所说明的那样,纯化来自紫苏、马鞭草、蓝竹耳以及矮牵牛花的糖转移酶、通过按照常规方法得到对应于该酶的抗体,可以克隆制造与该抗体反应的蛋白质的 cDNA 或染色体 DNA。因此、本发明并不只限制于来自紫苏、马鞭草、蓝竹耳以及矮牵牛花的糖转移酶基因,也涉及更广泛的糖转移酶。

另外本发明也涉及通过导入糖转移酶基因,颜色受到调节的植物或其后代或它们的组织,其形态即使是切花也可以。

另外、本发明中、含有花色素的类黄酮的糖转移反应中作为糖的供体有 UDP-葡萄糖。

以下根据实施例详细说明本发明。实验过程只要没有特别叙述的都是按照《分子克隆》(Molecular Cloning)(冷泉港、1989)《新生物化学实验手册》第3卷(化学同人 1996)、国际公开公报:W096/25500 记载的方法进行的。

实施例 1、在红紫苏中特异表达的基因的克隆

(1) 差别显示

紫苏(*Perilla frutescens*)中,有叶子中蓄积花色苷的品种(例如、紫薰(サカタのタネ)和叶子中不蓄积花色素的品种(例如、青薰(サカタのタネ)),据报道,主要的花色素的结构是丙二酰紫苏色苷(3-O-(p-香豆基)- β -D-葡萄糖基)-5-O-(6-O-丙二酰- β -D-葡萄糖基)-花青素(农业生物化学(Agri.Biol.Chem.) 53:197-198, 1989)。

差别显示是《科学》(Scienc) 257, 967-971 (1992)报道的方法,被用于获得组织特异表达的基因等研究中。

按照热酚法(《植物分子生物学手册》(Plant Molecular Biology Manual.) Kluwer Academic Publishers 1994 pp.D5/1-13)从两种紫苏的叶子中提取总 RNA。使用 mRNA 分离试剂盒(Clontech 公司)从得到的

总 RNA 中纯化聚 A+ RNA。使用附加了锚的 oligo dT 引物 (GenHunter 公司的 H-T11G、H-T11A、H-T11C) 于 $33\mu\text{l}$ 反应液中对 $0.9\mu\text{g}$ 的 oligo A+ RNA 进行逆转录, 得到单链 cDNA。以该 cDNA 作为模板、以同样附加了锚的 oligo dT 引物和合成引物 (GenHunter 公司的 H-AP1 至 8) 作为引物进行 PCR。

PCR 反应液的体积为 $20\mu\text{l}$, 含有 $2\mu\text{l}$ 的 cDNA 溶液, $0.2\mu\text{M}$ 的 H-T11G、H-T11A、H-T11C 中的任一种引物、 $0.2\mu\text{M}$ 的 H-AP1 至 8 中的任一种引物、 $0.12\mu\text{M}$ dNTP、5 或 $10\mu\text{Ci}$ $[\text{32P}]\text{dCTP}$ 、 10mM Tris-HCl (pH9.0)、 50mM KCl、 0.01% Triton X-100、 1.25mM MgCl_2 、 1u 的 Taq 聚合酶。反应条件如下。首先于 72°C 下保持 20 秒后、进行 40 个循环反应, 每一循环反应都是 94°C 30 秒、 40°C 2 分钟、 72°C 30 秒, 然后于 72°C 保持 5 分钟。

利用确定 DNA 碱基序列时用的聚丙烯酰胺凝胶分离上述扩增得到的 DNA。凝胶干燥后、对 X 胶片曝光。在得到的大约 2600 条带中、比较两个品种的紫苏、只在紫薰中看到的带有 36 条。将这些带从干燥的凝胶中切下来、溶解于 $100\mu\text{l}$ 的水中。溶出的 DNA 用乙醇沉淀, 溶解于 $20\mu\text{l}$ 的水中。其中一半量的 DNA 作为模板、分别进行上述的 PCR 反应, 可以扩增有关 33 条带的 DNA 片段。用这些 DNA 片段进行文库的筛选和 Northern 解析。

(2) Northern 解析

用以上 33 种 DNA 探针按照以下方法进行 Northern 解析。来自紫薰和青薰的 oligo A+ RNA 经含有 1.2% 琼脂糖的甲醛凝胶分离后、转移到尼龙膜上。该膜在 $5\times\text{SSPE}$ 、 $5\times\text{Denhalt}$ 液、 $0.5\%\text{SDS}$ 、 $20\mu\text{g/ml}$ 的变性鲑 DNA 存在下, 于 65°C 下与 ^{32}P 标记的上述 DNA 探针杂交过夜。将杂交的膜用 $1\times\text{SSPE}$ 、 $0.1\%\text{SDS}$ 于 65°C 下洗净, 得到放射自显影的胶片。结果、只有 5 种探针特异地出现在紫薰中。预计这些克隆是与花色素的生物合成有关的基因。

(3) cDNA 文库的筛选

利用从紫薰的叶子得到的聚 A+ RNA, 使用完全快速克隆系统 (Complete

Rapid Cloning System) λ gt10 (Amersham), 制备以 λ gt10 为载体的 cDNA 文库。该 cDNA 文库首先使用先前讲到的 5 种 DNA 片段进行筛选, 得到各个相应的 cDNA。其中取名为 3R5 的克隆由于是来自 H-T11A 和 H-AP3 的引物的 DNA 片段得到的产物, 所以表现出在已报道的玉米的类黄酮-3-O-糖转移酶的氨基酸水平上大约 26% 的同源性。

另外通过使用同样探针的文库的筛选可以得到称之 3R4 以及 3R6 的克隆, 这些克隆都表现出与 3R5 有极高的同源性。推断为 3R4 以及 3R6 的总碱基序列的氨基酸序列分别表示在序列表中的序列 1 和序列 2 中。而推断的 3R4 和 3R6 编码的蛋白质的氨基酸序列表现出 92% 的同源性。

取名为 8R6 的克隆由于是来自 H-T11G 和 H-AP8 的引物的 DNA 片段得到的产物, 所以没有表现出与到现在为止报道的 DNA 碱基序列的有意义的同源性。该序列表示在序列 5 中。虽然, 8R6 极有可能是与花色苷生物合成有关的基因, 但其结构由于与至今报道的基因没有同源性, 所以推测该基因是与花色苷生物合成有关的新的基因。

如果考虑到紫苏的花色苷(上述的丙二酰紫苏色素)的结构的话, 预料该基因是丙二酰基转移酶的基因。为了证明这一点, 只要使该基因在酵母或大肠杆菌中表达, 使表达的产物与作为底物的花色苷和丙二酰 CoA 进行反应就可证明。这样的实验可以利用诸如国际公开公报: W096/25500 记载的方法进行。丙二酰基转移酶基因在人为地改变花色素的结构上也是有用的。

(4) 3R4 的 cDNA 在酵母中的表达

利用 T4DNA 聚合酶(宝酒造公司制造)补平 p3R4 的 BstXI 的酶切部位, 再于连接体内的 BamHI 酶切部位酶切得到大约 1.5kb 的 DNA 片段, 该 DNA 片段与 pYE22m 的 EcoRI 酶切末端补平, 再经 BamHI 消化得到的大约 8kb 的 DNA 片段连接得到的质粒作为 pY3R4。

而带有 pYE22m 的大肠杆菌 JM109 菌株命名为 *Escherichia coli* SBM335, 以 FERM BP-5435 保藏于工业技术院生命工程工业技术研究所。在 pY3R4 中, 编码糖转移酶的 cDNA 连接于作为酵母的启动子组成之一的甘油醛-3 磷酸脱氢酶的启动子的下游, 通过同一个启动子控制转录。

利用伊藤等人的方法 (Ito 等, 细菌学杂志, (J. Bacteriol.), 153, 163-166, 1983) 用 pY3R4 转化酵母 *Saccharomyces cerevesiae* G1315 (Ashikari 等应用微生物学与生物技术 (Appl. Microbiol. Biotechnol.) 30, 515-520, 1989)。转化的酵母通过色氨酸合成能力恢复进行选择。得到的转化菌株于 10ml 的含有 1% 酪蛋白水解物 (Difco 公司) 的 Burkholder 培养基 (Burkholder, Amer. J. Bot. 30, 206-210) 在 30℃ 条件下振荡培育 24 小时。

同时, 作为对照实验, 也同样培养能自然恢复色氨酸合成能力的酵母。收集这些酵母菌, 将它们悬浮于悬浊缓冲液 (100mM 磷酸缓冲液 (pH8.5)、0.1%(v/v) 2- 巯基乙醇、10 μ M APMSF、100 μ M UDP - 葡萄糖), 通过加入玻璃珠 (Glass Beads 425-600microns Acid-Wash、Sigma 公司) 激烈振荡磨碎菌体。经 15000rpm 20 分钟离心, 将离心上清作为粗酶液, 用于以下酶活测定。

(5) 酶活测定

使含有 20 μ l 粗酶液的 50 μ l 反应液 (100mM 磷酸缓冲液 (pH8.5)、670 μ M 花青素 - 3 - 葡萄糖苷、1mM UDP - 葡萄糖) 于 30℃ 下反应 10 分钟后、加入含有 0.1%TFA 的 50% 乙腈溶液 50 μ l, 终止反应。将经 15000rpm 5 分钟离心得到的上清通过 Samprep LCR4(T)-LC (Millipore 公司) 除去不溶物。得到的样品利用高效液相色谱分析。分析是用反相柱 (Asahipak ODP-50, 4.6mm ϕ *250mm 昭和电工株式会社制造), 流动相为 A 溶液 0.5% TFA/H₂O、B 溶液为 0.5% TFA 50%CH₃CN、流速为 0.6 ml/min. 在进行 B20% $\rightarrow$ B100% (20min) 梯度洗脱后, 再维持 5min B100% 的洗脱。

用于分析的反应液为 20 μ l。检测是通过 A520 nm, AUFS 0.5 (岛津 SPD - 10A) 和光敏二极管阵列检测器 (岛津 SPD - M16A) 的 600-250 的吸收进行的。如果使 pY3R4 表达的酵母的粗酶液反应, 就将它加到底物花青素 - 3 - 葡萄糖苷中 (展开时间 17 分钟), 于 14.5 分钟展开时生成新出现的物质。由于该物质在对照实验的酵母粗酶液反应中没有看到, 所以认为该物质是通过来自 pY3R4 的蛋白质活性产生的物质。与花青素 - 3, 5

- 二葡萄糖苷的混合色谱分析的结果表明, 该反应生成物的展开时间与花青素-3,5-二葡萄糖苷的展开时间一致, 两者的吸收光谱也一致。由以上事实可知, 紫苏的 3R4 的 cDNA 编码 5GT。

实施例 2. 马鞭草 (*Verbena hybrida*) 的 5GT 基因的克隆

(1) cDNA 文库的制备

从马鞭草品种花球紫(三得利)收集花瓣、于液氮中用乳钵磨碎。通过利用胍硫氰酸盐/氯化铯的方法从该磨碎物中提取 RNA, 再使用 Oligotex (宝酒造制造), 用制造者推荐的方法得到聚 A+RNA。使用了胍硫氰酸盐/氯化铯的方法是依据 R. McGookin, Robert J. Slater 等人、分子生物学方法 (Methods in Molecular Biology) vol 2, (Humana Press Inc. 1984) 详细给出的方法。

以得到的聚 A+RNA 为模板, 利用 Stratagene 公司的 2AP-cDNA 合成试剂盒合成双链 cDNA, 然后利用 Uni-ZAP XR 克隆试剂盒 (Stratagene 公司) 用制造者推荐的方法制备 cDNA 文库。

(2) 5GT cDNA 的克隆

以紫苏的 p3R4 的 cDNA 作为探针, 按以下操作筛选上述得到的 λ 嗜菌体文库。将膜于 42℃ 放置于杂交缓冲液 (5×SSC, 30% 甲酰胺、50mM 磷酸钠缓冲液 (pH7.0)、1% SDS、2% Blocking reagent (Boehringer 公司)、0.1% 月桂酰肌氨酸、80g/ml 鲑鱼精子 DNA) 中 1 小时。然后将 DIG 标记的紫苏的 5GT 基因、p3R4 的 DNA 片段加入杂交溶液中, 再温育 16 小时。

用洗净液 (5×SSC 50℃、1% SDS) 将滤膜洗净、然后通过依据碱性磷酸酶标记的 DIG 特异的抗体的免疫测定法 (Boehringer 公司) 在 5-溴-4-氯-3-吡啶磷酸和硝基蓝四唑盐的发色反应中检测出探针杂交的克隆。检测方法按使用说明书进行。

检测结果得到 7 个阳性克隆。用 Stratagene 公司推荐的方法将这些 cDNA 回收到质粒 pBluescript SK 中。通过琼脂糖凝胶电泳研究 cDNA 的长度时, 确认了插入最长 2.0kb。

(3) 碱基序列的确定

从得到的克隆中提取质粒、使用测序仪 ABI373A (Perkin Elmer 公司)、

利用该公司推荐的依据荧光试剂的双脱氧测序方法确定 cDNA 的 3' 和 5' 末端附近的碱基序列。结果发现 7 个克隆中的 5 个克隆之间具有同样的碱基序列, 认为是 cDNA 长度不同的克隆。其中还确定了 pSHGT8 的整个碱基序列。碱基序列的确定是使用了 Kilo-Sequence 用 deletion 试剂盒(宝酒造制)得到一系列的缺失克隆、或利用对 pSHGT8 内部序列特异的寡核苷酸引物、按上述那样进行。

(4) 碱基序列和氨基酸序列的比较

插入 pSHGT8 的 cDNA 是 2062bp, 其中存在由 1386bp (含终止码) 构成的开放读框(ORF)。序列 3 给出了该序列。该 ORF 的氨基酸序列表现出与紫苏 p3R4 编码的 5GT 的氨基酸序列有 68%、与 p3R6 编码的有 64% 的同源性。而与单子叶植物以及双子叶植物的 3GT 有 22~25%、与矮牵牛花的 3RT 有 21% 的同源性。

(5) 在酵母中的表达和酶活的测定

将 BamHI /XhoI 消化 pSHGT8 后得到的大约 2.0kb 的 DNA 片段与 BamHI /SalI 消化 pYE22m 得到的大约 8kb 的 DNA 片段连接得到的质粒作为 pYHGT8。与实施例 1 一样、在酵母菌体内表达 pYHGT8、测定有关 pYHGT8 编码的蛋白质的酶活。结果、如果使导入 pYHGT8 的酵母粗酶液反应的话, 可生成与花青素-3,5-二葡萄糖苷的展开时间、吸收光谱都一致的产物。由这一事实可知, 马鞭草的 pSHGT8 的 cDNA 编码 5GT。

实施例 3. 蓝竹耳的 5GT 基因的克隆

(1) cDNA 文库的制备

从蓝竹耳品种 Summer Wave Blue(三得利(株))收集花瓣、于液氮中用乳钵磨碎。通过利用胍硫氰酸盐/氯化铯的方法从该磨碎物中提取 RNA, 再使用 oligotex (宝酒制造), 用制造者推荐的方法得到聚 A+RNA。使用的胍硫氰酸盐/氯化铯方法是依据 R. McGookin, Robert J. Slater 等人于分子生物学方法(Methods in Molecular Biology) vol 2, (Humana Press Inc. 1984) 中详细给出的方法。

以得到的聚 A+RNA 为模板, 利用 Stratagene 公司的 ZAP-cDNA 合成试剂盒合成双链 cDNA, 然后利用 Uni-ZAP XR 克隆试剂盒(Stratagene

公司)用制造者推荐的方法制备 cDNA 文库。

(2) 5GT cDNA 的克隆

以紫苏的 p3R4 的 cDNA 作为探针,与实施例 2 一样筛选上述得到的 λ 嗜菌体文库。结果得到 8 个阳性克隆。将 cDNA 回收到质粒 pBluescript SK 中,通过琼脂糖凝胶电泳研究该 cDNA 的长度,确认插入了最长 1.6kb。

(3) 碱基序列的确定

从得到的克隆中提取质粒,与实施例 2 一样确定两末端附近的碱基序列。结果发现其中的 6 个克隆之间具有同样的碱基序列,认为是 cDNA 长度不同的克隆。其中还确定了该 6 个克隆中 pSTGT5 的整个碱基序列。

(4) 碱基序列和氨基酸序列的比较

插入 pSTGT5 的 cDNA 是 1671bp,其中存在由 1437bp (含终止码)构成的开放读框 (ORF)。序列 4 给出了该序列。该 ORF 的氨基酸序列表现出与紫苏 p3R4 编码的 5GT 的氨基酸序列有 58%、与 p3R6 编码的有 57% 的同源性。而与马鞭草 pSHGT8 编码的有 57% 的同源性。而与单子叶植物以及双子叶植物的 3GT 有 19~23%、与矮牵牛花的 3RT 有 20% 的同源性。

(5) 5GT 基因在酵母中的表达

将 SamI /KpnI 消化 pSTGT5 后得到的大约 1.6kb 的 DNA 片段与 pYE22m 的 EcoRI 酶切部位补平后被 KpnI 消化得到的大约 8kb 的 DNA 片段连接得到的质粒作为 pYTGT5。与实施例 1 一样、在酵母菌体内表达 pYTGT5、测定有关 pSTGT5 编码的蛋白质的酶活。结果、如果使导入 pYTGT5 的酵母粗酶液反应的话,可得到与花青素-3,5-二葡萄糖苷的展开时间、吸收光谱都一致的产物。由这一事实可知,蓝竹耳的 pSTGT5 的 cDNA 编码 5GT。

实施例 4. 矮牵牛花的 5GT 基因的克隆

(1) cDNA 文库的制备

在从矮牵牛花品种 Old Glory Blue 的花瓣中提取 RNA 的基础上,按照 T.Holton 等人的报道 (植物学报 Plant Journal, 1993 4: 1003-1010) 详细记载的那样制备 cDNA 文库。

(2) 5GT cDNA 的克隆

以上述那样操作得到的紫苏、蓝竹耳、马鞭草的 cDNA 作为探针,与实

施例 2 一样进行筛选。将结果得到的阳性克隆中的 4 个克隆回收到质粒 pBluescript SK-中, 通过琼脂糖凝胶电泳研究该 cDNA 的长度, 确认插入了最长 2.0kb 的 cDNA。

(3) 碱基序列的确定

确定 5' 末端附近的碱基序列。结果表明这些克隆中的 2 个克隆的 pSPGT1 编码表现出与到目前为止得到的紫苏、蓝竹耳、马鞭草的 5GT 有很高同源性的氨基酸序列。于是确定了 pSPGT1 的整个碱基序列。

(4) 碱基序列和氨基酸序列的比较

插入 pSPGT1 的 cDNA 是 2105bp, 其中存在由 1407bp (含终止码) 构成的 ORF。序列 6 给出了该序列。该 ORF 的氨基酸序列表现出与紫苏 p3R4 编码的 5GT 的氨基酸序列有 57%、与 p3R6 编码的有 54%、与马鞭草 pSHGT8 编码的有 55%、与蓝竹耳的 pTGT5 编码的有 51% 的同源性。而与单子叶植物以及双子叶植物的 3GT 有 20~29%、与矮牵牛花的 3RT 有 20% 的同源性。由这一事实可以认为从矮牵牛花得到的 pSPGT1 的 cDNA 编码 5GT。

如上所述, 进行了编码来自紫苏、马鞭草、蓝竹耳以及矮牵牛花的向类黄酮 5 位上转移糖的酶的 cDNA 的克隆并确定了其碱基序列。通过酵母中活性的表达阐明了分离到的 cDNA 是编码 5GT 的基因。通过将该 cDNA 与适当的植物表达载体连接, 导入植物中, 使其有时具有 5GT 的活性、有时使其活性增加、有时又使其活性降低, 用来调节植物的花色成为可能。而通过利用该酶的活性, 在植物体中或在试管中改变花色素的结构可以合成更稳定的花色素。

序 列

序列: 1

序列长度: 1507

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 紫苏 (*Perilla frutescens*)

组织种类: 叶

直接起源

文库名: cDNA library

克隆名: p3R4

序列

GAAAATTTCC ACAAAA ATG GTC CGC CGC CGC GTG CTG CTA GCA ACG TTT	49
Met Val Arg Arg Arg Val Leu Leu Ala Thr Phe	
1 5 10	
CCT GCG CAA GGC CAC ATA AAT CCC GCC CTC CAA TTC GCC AAG AGA CTC	97
Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala Lys Arg Leu	
15 20 25	
CTA AAA GCC GGC ACT GAC GTC ACA TTT TTC ACG AGC GTT TAT GCA TGG	145
Leu Lys Ala Gly Thr Asp Val Thr Phe Phe Thr Ser Val Tyr Ala Trp	
30 35 40	
CGC CGC ATG GCC AAC ACA GCC TCC GCC GCT GCC GGA AAC CCA CCG GGC	193
Arg Arg Met Ala Asn Thr Ala Ser Ala Ala Ala Gly Asn Pro Pro Gly	
45 50 55	

CTC GAC TTC GTG GCG TTC TCC GAC GGC TAC GAC GAC GGG CTG AAG CCC	241
Leu Asp Phe Val Ala Phe Ser Asp Gly Tyr Asp Asp Gly Leu Lys Pro	
60 65 70 75	
TGC GGC GAC GGG AAG CGC TAC ATG TCC GAG ATG AAA GCC CGC GGC TCC	289
Cys Gly Asp Gly Lys Arg Tyr Met Ser Glu Met Lys Ala Arg Gly Ser	
80 85 90	
GAG GCC TTA AGA AAC CTC CTT CTC AAC AAC CAC GAC GTC ACG TTC GTC	337
Glu Ala Leu Arg Asn Leu Leu Leu Asn Asn His Asp Val Thr Phe Val	
95 100 105	
GTC TAC TCC CAC CTC TTT GCA TGG GCG GCG GAG GTG GCG CGT GAG TCC	385
Val Tyr Ser His Leu Phe Ala Trp Ala Ala Glu Val Ala Arg Glu Ser	
110 115 120	
CAG GTC CCG AGC GCC CTT CTC TGG GTC GAG CCC GCC ACC GTG CTG TGC	433
Gln Val Pro Ser Ala Leu Leu Trp Val Glu Pro Ala Thr Val Leu Cys	
125 130 135	
ATA TAT TAC TTC TAC TTC AAC GGC TAC GCA GAC GAG ATC GAC GCC GGT	481
Ile Tyr Tyr Phe Tyr Phe Asn Gly Tyr Ala Asp Glu Ile Asp Ala Gly	
140 145 150 155	
TCC GAC GAA ATT CAG CTC CCT CGG CTT CCA CCC CTG GAG CAG CGC AGT	529
Ser Asp Glu Ile Gln Leu Pro Arg Leu Pro Pro Leu Glu Gln Arg Ser	
160 165 170	
CTT CCG ACC TTT CTG CTG CCG GAG ACA CCG GAG AGA TTC CGG TTG ATG	577
Leu Pro Thr Phe Leu Leu Pro Glu Thr Pro Glu Arg Phe Arg Leu Met	
175 180 185	
ATG AAG GAG AAG CTG GAA ACT TTA GAC GGT GAA GAG AAG GCG AAA GTG	625
Met Lys Glu Lys Leu Glu Thr Leu Asp Gly Glu Glu Lys Ala Lys Val	
190 195 200	

TTG GTG AAC ACG TTT GAT GCG TTG GAG CCC GAT GCA CTC ACG GCT ATT	673
Leu Val Asn Thr Phe Asp Ala Leu Glu Pro Asp Ala Leu Thr Ala Ile	
205 210 215	
GAT AGG TAT GAG TTG ATC GGG ATC GGG CCG TTG ATT CCC TCC GCC TTC	721
Asp Arg Tyr Glu Leu Ile Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe	
220 225 230 235	
TTG GAC GGC GGA GAT CCC TCC GAA ACG TCT TAC GGC GGC GAT CTT TTC	769
Leu Asp Gly Gly Asp Pro Ser Glu Thr Ser Tyr Gly Gly Asp Leu Phe	
240 245 250	
GAA AAA TCG GAG GAG AAT AAC TGC GTG GAG TGG TTG GAC ACG AAG CCG	817
Glu Lys Ser Glu Glu Asn Asn Cys Val Glu Trp Leu Asp Thr Lys Pro	
255 260 265	
AAA TCT TCG GTG GTG TAT GTG TCG TTT GGG AGC GTT TTG AGG TTT CCA	865
Lys Ser Ser Val Val Tyr Val Ser Phe Gly Ser Val Leu Arg Phe Pro	
270 275 280	
AAG GCA CAA ATG GAA GAG ATT GGG AAA GGG CTA TTA GCC TGC GGA AGG	913
Lys Ala Gln Met Glu Glu Ile Gly Lys Gly Leu Leu Ala Cys Gly Arg	
285 290 295	
CCG TTT TTA TGG ATG ATA CGA GAA CAG AAG AAT GAC GAC GGC GAA GAA	961
Pro Phe Leu Trp Met Ile Arg Glu Gln Lys Asn Asp Asp Gly Glu Glu	
300 305 310 315	
GAA GAA GAA GAG TTG AGT TGC ATT GGG GAA TTG AAA AAA ATG GGG AAA	1009
Glu Glu Glu Glu Leu Ser Cys Ile Gly Glu Leu Lys Lys Met Gly Lys	
320 325 330	
ATA GTT TCG TGG TGC TCG CAG TTG GAG GTT CTG GCG CAC CCT GCG TTG	1057
Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu Ala His Pro Ala Leu	
335 340 345	

GGA TGT TTC GTG ACG CAT TGT GGG TGG AAC TCG GCT GTG GAG AGC TTG 1105
 Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser Ala Val Glu Ser Leu
 350 355 360
 AGT TGC GGG GTT CCG GTG GTG GCG GTG CCG CAG TGG TTT GAT CAG ACG 1153
 Ser Cys Gly Val Pro Val Val Ala Val Pro Gln Trp Phe Asp Gln Thr
 365 370 375
 ACG AAT GCG AAG CTG ATT GAG GAT GCG TGG GGG ACA GGG GTG AGA GTG 1201
 Thr Asn Ala Lys Leu Ile Glu Asp Ala Trp Gly Thr Gly Val Arg Val
 380 385 390 395
 AGA ATG AAT GAA GGG GGT GGG GTT GAT GGA TCT GAG ATA GAG AGG TGT 1249
 Arg Met Asn Glu Gly Gly Gly Val Asp Gly Ser Glu Ile Glu Arg Cys
 400 405 410
 GTG GAG ATG GTG ATG GAT GGG GGT GAG AAG AGC AAA CTA GTG AGA GAA 1297
 Val Glu Met Val Met Asp Gly Gly Glu Lys Ser Lys Leu Val Arg Glu
 415 420 425
 AAT GCC ATA AAA TGG AAG ACT TTG GCC AGA GAA GCC ATG GGA GAG GAT 1345
 Asn Ala Ile Lys Trp Lys Thr Leu Ala Arg Glu Ala Met Gly Glu Asp
 430 435 440
 GGA TCT TCA CTC AAG AAT CTC AAC GCC TTT CTT CAT CAA GTT GCA CGT 1393
 Gly Ser Ser Leu Lys Asn Leu Asn Ala Phe Leu His Gln Val Ala Arg
 445 450 455
 GCT TAATACACAA AATGGCTTTC CACTTTTAAT CTACTCAAAC ACCGGTTCAA 1446
 Ala
 460
 ATAAATATCC CCTTCCACTT CTTTCTATTT CACTATCACA TTTATAATTT TAGTAACAAA 1506
 A 1507

序列: 2

序列长度: 1470

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 紫苏 (*Perilla frutescens*)

组织种类: 叶

直接起源

文库名: cDNA library

克隆名: p3R6

序列

ACCAAACCAA AACAAAATTT CCACAAAA ATG GTC CGC CGC CGC GTG CTG CTA 48

Met Val Arg Arg Arg Val Leu Leu

1

5

GCA ACG TTT CCG GCG CAA GGC CAC ATA AAT CCC GCC CTC CAA TTC GCC 96

Ala Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala

10

15

20

AAG AGA CTC CTA AAA GCC GGC ACT GAC GTC ACG TTT TTC ACG AGC GTT 144

Lys Arg Leu Leu Lys Ala Gly Thr Asp Val Thr Phe Phe Thr Ser Val

25

30

35

40

TAT GCA TGG CGC CGC ATG GCC AAC ACA GCC TCC GCC GCT GCC GGA AAC 192

Tyr Ala Trp Arg Arg Met Ala Asn Thr Ala Ser Ala Ala Ala Gly Asn

45

50

55

CCA CCG GGC CTC GAC TTC GTG GCG TTC TCC GAC GGC TAC GAC GAC GGG 240

Pro Pro Gly Leu Asp Phe Val Ala Phe Ser Asp Gly Tyr Asp Asp Gly

60

65

70

CTG AAG CCC GGC GGC GAC GGG AAG CGC TAC ATG TCC GAG ATG AAA GCC	288
Leu Lys Pro Gly Gly Asp Gly Lys Arg Tyr Met Ser Glu Met Lys Ala	
75 80 85	
CGC GGC TCC GAG GCC TTA AGA AAC CTC CTT CTC AAC AAC GAC GAC GTC	336
Arg Gly Ser Glu Ala Leu Arg Asn Leu Leu Leu Asn Asn Asp Asp Val	
90 95 100	
ACT TTC GTC GTC TAC TCC CAC CTC TTT GCA TGG GCG GCG GAG GTG GCG	384
Thr Phe Val Val Tyr Ser His Leu Phe Ala Trp Ala Ala Glu Val Ala	
105 110 115 120	
CGT TTG TCC CAC GTC CCG ACC GCC CTT CTC TGG GTC GAG CCC GCC ACC	432
Arg Leu Ser His Val Pro Thr Ala Leu Leu Trp Val Glu Pro Ala Thr	
125 130 135	
GTG CTG TGC ATA TAC CAC TTC TAC TTC AAC GGC TAC GCA GAC GAG ATC	480
Val Leu Cys Ile Tyr His Phe Tyr Phe Asn Gly Tyr Ala Asp Glu Ile	
140 145 150	
GAC GCC GGT TCC AAT GAA ATT CAG CTC CCT CGG CTT CCA TCC CTG GAG	528
Asp Ala Gly Ser Asn Glu Ile Gln Leu Pro Arg Leu Pro Ser Leu Glu	
155 160 165	
CAG CGC AGT CTT CCG ACG TTT CTG CTG CCT GCG ACG CCG GAG AGA TTC	576
Gln Arg Ser Leu Pro Thr Phe Leu Leu Pro Ala Thr Pro Glu Arg Phe	
170 175 180	
CGG TTG ATG ATG AAG GAG AAG CTG GAA ACT TTA GAC GGT GAA GAG AAG	624
Arg Leu Met Met Lys Glu Lys Leu Glu Thr Leu Asp Gly Glu Glu Lys	
185 190 195 200	
GCG AAA GTA TTG GTG AAC ACG TTT GAT GCG TTG GAG CCC GAT GCA CTC	672
Ala Lys Val Leu Val Asn Thr Phe Asp Ala Leu Glu Pro Asp Ala Leu	
205 210 215	

ACG GCT ATT GAT AGG TAT GAG TTG ATC GGG ATC GGG CCG TTG ATT CCC	720
Thr Ala Ile Asp Arg Tyr Glu Leu Ile Gly Ile Gly Pro Leu Ile Pro	
220 225 230	
TCC GCC TTC TTG GAC GGC GAA GAT CCC TCC GAA ACG TCT TAC GGC GGC	768
Ser Ala Phe Leu Asp Gly Glu Asp Pro Ser Glu Thr Ser Tyr Gly Gly	
235 240 245	
GAT CTT TTC GAA AAA TCG GAG GAG AAT AAC TGC GTG GAG TGG TTG AAC	816
Asp Leu Phe Glu Lys Ser Glu Glu Asn Asn Cys Val Glu Trp Leu Asn	
250 255 260	
TCG AAG CCG AAA TCT TCG GTG GTG TAT GTG TCG TTT GGG AGC GTT TTG	864
Ser Lys Pro Lys Ser Ser Val Val Tyr Val Ser Phe Gly Ser Val Leu	
265 270 275 280	
AGG TTT CCA AAG GCA CAA ATG GAA GAG ATT GGG AAA GGG CTA TTA GCC	912
Arg Phe Pro Lys Ala Gln Met Glu Glu Ile Gly Lys Gly Leu Leu Ala	
285 290 295	
TGC GGA AGG CCC TTT TTA TGG ATG ATA CGA GAA CAG AAG AAT GAC GAC	960
Cys Gly Arg Pro Phe Leu Trp Met Ile Arg Glu Gln Lys Asn Asp Asp	
300 305 310	
GGC GAA GAA GAA GAA GAA GAA GAA GAG TTG AGT TGC ATT GGG GAA TTG	1008
Gly Glu Glu Glu Glu Glu Glu Glu Glu Leu Ser Cys Ile Gly Glu Leu	
315 320 325	
AAA AAA ATG GGG AAA ATA GTG TCG TGG TGC TCG CAG TTG GAG GTT CTG	1056
Lys Lys Met Gly Lys Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu	
330 335 340	
GCG CAC CCT GCG TTG GGA TGT TTC GTG ACG CAT TGT GGG TGG AAC TCG	1104
Ala His Pro Ala Leu Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser	
345 350 355 360	

```

GCT GTG GAG AGC TTG AGT TGC GGG ATT CCG GTG GTG GCG GTG CCG CAG   1152
Ala Val Glu Ser Leu Ser Cys Gly Ile Pro Val Val Ala Val Pro Gln
          365                370                375

TGG TTT GAT CAG ACG ACG AAT GCG AAG CTG ATT GAG GAT GCG TGG GGG   1200
Trp Phe Asp Gln Thr Thr Asn Ala Lys Leu Ile Glu Asp Ala Trp Gly
          380                385                390

ACA GGG GTG AGA GTG AGA ATG AAT GAA GGG GGT GGG GTT GAT GGA TGT   1248
Thr Gly Val Arg Val Arg Met Asn Glu Gly Gly Gly Val Asp Gly Cys
          395                400                405

GAG ATA GAA AGG TGT GTG GAG ATG GTG ATG GAT GGG GGT GAC AAG ACC   1296
Glu Ile Glu Arg Cys Val Glu Met Val Met Asp Gly Gly Asp Lys Thr
          410                415                420

AAA CTA GTG AGA GAA AAT GCC ATC AAA TGG AAG ACT TTG GCC AGA CAA   1344
Lys Leu Val Arg Glu Asn Ala Ile Lys Trp Lys Thr Leu Ala Arg Gln
          425                430                435                440

GCC ATG GGA TAGGATGGAT CTTCACTCAA CAATCTCAAC GCCTTTCTTC           1393
Ala Met Gly
          443

GTCAAGTTGC ACACTTTTAA TCTGCTCAAA CAGCGGTTCA AATAAATATC CCCTTCCACT   1453
TAAAAAAAAA AAAAAAA           1470

```

序列: 3

序列长度: 2062

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 马鞭草 (Verbena hybrida)

组织种类: 花瓣

直接起源

文库名: cDNA library

克隆名: pSHGT8

序列

```

ATTTTACCAA AAAAATAAAA AAAAA ATG AGC AGA GCT CAC GTC CTC TTG GCC      52
                               Met Ser Arg Ala His Val Leu Leu Ala
                               1             5

ACA TTC CCA GCA CAG GGA CAC ATA AAT CCC GCC CTT CAA TTC GCC AAG      100
Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala Lys
10             15             20             25

CGT CTC GCA AAT GCC GAC ATT CAA GTC ACA TTC TTC ACC AGC GTC TAC      148
Arg Leu Ala Asn Ala Asp Ile Gln Val Thr Phe Phe Thr Ser Val Tyr
30             35             40

GCA TGG CGC CGC ATG TCC AGA ACC GCC GCT GGC TCA AAC GGG CTC ATC      196
Ala Trp Arg Arg Met Ser Arg Thr Ala Ala Gly Ser Asn Gly Leu Ile
45             50             55

AAT TTT GTG TCG TTT TCC GAC GGG TAT GAC GAC GGG TTA CAG CCC GGA      244
Asn Phe Val Ser Phe Ser Asp Gly Tyr Asp Asp Gly Leu Gln Pro Gly
60             65             70

GAC GAT GGG AAG AAC TAC ATG TCG GAG ATG AAA AGC AGA GGT ATA AAA      292
Asp Asp Gly Lys Asn Tyr Met Ser Glu Met Lys Ser Arg Gly Ile Lys
75             80             85

GCC TTG AGC GAT ACT CTT GCA GCC AAT AAT GTC GAT CAA AAA AGC AGC      340
Ala Leu Ser Asp Thr Leu Ala Ala Asn Asn Val Asp Gln Lys Ser Ser
90             95             100             105

```

AAA ATC ACG TTC GTG GTG TAC TCC CAC CTC TTT GCA TGG GCG GCC AAG	388
Lys Ile Thr Phe Val Val Tyr Ser His Leu Phe Ala Trp Ala Ala Lys	
110 115 120	
GTG GCG CGT GAG TTC CAT CTC CGG AGC GCG CTA CTC TGG ATT GAG CCA	436
Val Ala Arg Glu Phe His Leu Arg Ser Ala Leu Leu Trp Ile Glu Pro	
125 130 135	
GCT ACG GTG TTG GAT ATA TTT TAC TTT TAT TTC AAC GGC TAT AGC GAC	484
Ala Thr Val Leu Asp Ile Phe Tyr Phe Tyr Phe Asn Gly Tyr Ser Asp	
140 145 150	
GAA ATC GAT GCG GGT TCG GAT GCT ATT CAC TTG CCC GGA GGA CTC CCA	532
Glu Ile Asp Ala Gly Ser Asp Ala Ile His Leu Pro Gly Gly Leu Pro	
155 160 165	
GTG CTG GCC CAG CGT GAT TTA CCG TCT TTC CTT CTT CCT TCC ACG CAT	580
Val Leu Ala Gln Arg Asp Leu Pro Ser Phe Leu Leu Pro Ser Thr His	
170 175 180 185	
GAG AGA TTC CGT TCA CTG ATG AAG GAG AAA TTG GAA ACT TTA GAA GGT	628
Glu Arg Phe Arg Ser Leu Met Lys Glu Lys Leu Glu Thr Leu Glu Gly	
190 195 200	
GAA GAA AAA CCT AAG GTC TTG GTG AAC AGC TTT GAT GCG TTG GAG CCT	676
Glu Glu Lys Pro Lys Val Leu Val Asn Ser Phe Asp Ala Leu Glu Pro	
205 210 215	
GAT GCG CTC AAG GCC ATT GAT AAG TAC GAG ATG ATT GCA ATC GGG CCG	724
Asp Ala Leu Lys Ala Ile Asp Lys Tyr Glu Met Ile Ala Ile Gly Pro	
220 225 230	
TTG ATT CCT TCC GCA TTC TTG GAC GGT AAA GAT CCT TCG GAC AGG TCT	772
Leu Ile Pro Ser Ala Phe Leu Asp Gly Lys Asp Pro Ser Asp Arg Ser	
235 240 245	

TTC GGC GGA GAT TTG TTC GAG AAA GGG TCG AAT GAC GAC GAT TGC CTC	820
Phe Gly Gly Asp Leu Phe Glu Lys Gly Ser Asn Asp Asp Asp Cys Leu	
250 255 260 265	
GAA TGG TTG AGC ACG AAT CCT CGA TCT TCG GTG GTT TAC GTT TCG TTC	868
Glu Trp Leu Ser Thr Asn Pro Arg Ser Ser Val Val Tyr Val Ser Phe	
270 275 280	
GGA AGC TTC GTT AAT ACG ACG AAG TCG CAA ATG GAA GAG ATA GCA AGA	916
Gly Ser Phe Val Asn Thr Thr Lys Ser Gln Met Glu Glu Ile Ala Arg	
285 290 295	
GGG CTG TTA GAT TGT GGG AGG CCG TTT TTG TGG GTG GTA AGA GTA AAC	964
Gly Leu Leu Asp Cys Gly Arg Pro Phe Leu Trp Val Val Arg Val Asn	
300 305 310	
GAA GGA GAA GAG GTA TTG ATA AGT TGC ATG GAG GAG TTG AAA CGA GTG	1012
Glu Gly Glu Glu Val Leu Ile Ser Cys Met Glu Glu Leu Lys Arg Val	
315 320 325	
GGG AAA ATT GTA TCT TGG TGT TCT CAA TTG GAA GTC CTG ACG CAT CCC	1060
Gly Lys Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu Thr His Pro	
330 335 340 345	
TCG TTG GGA TGT TTC GTG ACA CAC TGC GGG TGG AAT TCG ACT CTA GAG	1108
Ser Leu Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser Thr Leu Glu	
350 355 360	
AGT ATA TCT TTC GGG GTT CCG ATG GTG GCT TTT CCG CAG TGG TTC GAT	1156
Ser Ile Ser Phe Gly Val Pro Met Val Ala Phe Pro Gln Trp Phe Asp	
365 370 375	
CAA GGG ACG AAT GCG AAG CTG ATG GAG GAT GTG TGG AGG ACG GGT GTG	1204
Gln Gly Thr Asn Ala Lys Leu Met Glu Asp Val Trp Arg Thr Gly Val	
380 385 390	

AGA GTG AGA GCT AAT GAG GAG GGT AGC GTC GTT GAT GGT GAT GAA ATT 1252
 Arg Val Arg Ala Asn Glu Glu Gly Ser Val Val Asp Gly Asp Glu Ile
 395 400 405

AGG AGA TGT ATT GAG GAG GTT ATG GAT GGG GGA GAA AAG AGT AGG AAA 1300
 Arg Arg Cys Ile Glu Glu Val Met Asp Gly Gly Glu Lys Ser Arg Lys
 410 415 420 425

CTT AGA GAG AGT GCT GGC AAG TGG AAG GAT TTG GCA AGA AAA GCT ATG 1348
 Leu Arg Glu Ser Ala Gly Lys Trp Lys Asp Leu Ala Arg Lys Ala Met
 430 435 440

GAG GAA GAT GGA TCT TCA GTT AAC AAC CTC AAG GTC TTT CTT GAT GAG 1396
 Glu Glu Asp Gly Ser Ser Val Asn Asn Leu Lys Val Phe Leu Asp Glu
 445 450 455

GTT GTA GGT ATC TAAAGACGTA AATGAGGTCC CCATAGGCAA AATTGCAAAT 1448
 Val Val Gly Ile
 460 461

TTCATCTCGT AAGTTGAATA CTTTTTGGCT TTAATTTTGT TCGAGTTTGT TTTTCAAAT 1508
 TTATCTTGTA ATTTTACATT GAGTGTAAT TTAGTCTGAT TTAACTGGA AAAATATAAA 1568
 ATTCATTGTT GAGACTCTTC ATCAAAATCA TCTGATTTC TTTATTGTCT TGGTCAAAT 1628
 TCTCATATCA ATTGGAAAAA ATAAATTTCA AAATCGTCCA ATTTTGAACC AAGAAAGAAG 1688
 TATAATTTGA CCAAAATAAT AAAAGGATTC AAGTGATCTT GATGAAGTGT CTGAGCGACG 1748
 AGTTCTATAT TTTTCCACCG AATTTCTAAC GAGTTTTTGA ATTTTTTTTA GCCAAAATCG 1808
 GACTAACTTT GTACAAAATG AAAAGTTATA TGATGAAATT TAAAAAACA AACTCAGACA 1868
 ATAATAAAGC CCGAAAGTAG TAAAATTACC TGACGAAATT TGCAATTTTC CCTCCTATTT 1928
 TAATTTTTTT GGTGTGTTTA ATAAATCGGT TATTTTACTT TTAATTAAAA TAAAAGTGAG 1988
 ATGCATGATA GCTTGGTGAG TATATATGAG TTGATGGTAA TGTACGATAT TTTCTAAAAA 2048
 AAAAAAAAAA AAAA 2062

序列: 4

序列长度: 1671

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 蓝竹耳

组织种类: 花瓣

直接起源

文库名: cDNA library

克隆名: pSTGT5

序列

```

AACACATAAA AAAAAATAA AAGAAGAAAT AATTAAAAAA AAAA ATG GTT AAC      53
                                     Met Val Asn
                                     1
AAA CGC CAT ATT CTA CTA GCA ACA TTC CCA GCA CAA GGC CAC ATA AAC      101
Lys Arg His Ile Leu Leu Ala Thr Phe Pro Ala Gln Gly His Ile Asn
      5              10              15
CCT TCT CTC GAG TTC GCC AAA AGG CTC CTC AAC ACC GGA TAC GTC GAC      149
Pro Ser Leu Glu Phe Ala Lys Arg Leu Leu Asn Thr Gly Tyr Val Asp
      20              25              30              35
CAA GTC ACA TTC TTC ACG AGT GTA TAC GCA TTG AGA CGC ATG CGC TTC      197
Gln Val Thr Phe Phe Thr Ser Val Tyr Ala Leu Arg Arg Met Arg Phe
              40              45              50
GAA ACC GAT CCG AGC AGC AGA ATC GAT TTC GTG GCA TKT YCA GAT TCT      245
Glu Thr Asp Pro Ser Ser Arg Ile Asp Phe Val Ala X X Asp Ser
              55              60              65

```

TAC GAT GAT GGC TTA AAG AAA GGC GAC GAT GGC AAA AAC TAC ATG TCG	293
Tyr Asp Asp Gly Leu Lys Lys Gly Asp Asp Gly Lys Asn Tyr Met Ser	
70 75 80	
GAG ATG AGA AAG CGC GGA ACG AAG GCC TTA AAG GAC ACT CTT ATT AAG	341
Glu Met Arg Lys Arg Gly Thr Lys Ala Leu Lys Asp Thr Leu Ile Lys	
85 90 95	
CTC AAC GAT GCT GCG ATG GGA AGT GAA TGT TAC AAT CGC GTG AGC TTT	389
Leu Asn Asp Ala Ala Met Gly Ser Glu Cys Tyr Asn Arg Val Ser Phe	
100 105 110 115	
GTG GTG TAC TCT CAT CTA TTT TCG TGG GCA GCT GAA GTG GCG CGT GAA	437
Val Val Tyr Ser His Leu Phe Ser Trp Ala Ala Glu Val Ala Arg Glu	
120 125 130	
GTC GAC GTG CCG AGT GCC CTT CTT TGG ATT GAA CCG GCT ACG GTT TTC	485
Val Asp Val Pro Ser Ala Leu Leu Trp Ile Glu Pro Ala Thr Val Phe	
135 140 145	
GAT GTG TAC TAT TTT TAC TTC AAT GGG TAT GCC GAT GAT ATC GAT GCG	533
Asp Val Tyr Tyr Phe Tyr Phe Asn Gly Tyr Ala Asp Asp Ile Asp Ala	
150 155 160	
GGC TCA GAT CAA ATC CAA CTG CCC AAT CTT CCG CAG CTC TCC AAG CAA	581
Gly Ser Asp Gln Ile Gln Leu Pro Asn Leu Pro Gln Leu Ser Lys Gln	
165 170 175	
GAT CTC CCC TCT TTC CTA CTC CCT TCG AGC CCC GCG AGA TTC CGA ACC	629
Asp Leu Pro Ser Phe Leu Leu Pro Ser Ser Pro Ala Arg Phe Arg Thr	
180 185 190 195	
CTA ATG AAA GAA AAG TTC GAC ACG CTC GAC AAA GAA CCG AAA GCG AAG	677
Leu Met Lys Glu Lys Phe Asp Thr Leu Asp Lys Glu Pro Lys Ala Lys	
200 205 210	

GTC TTG ATA AAC ACG TTC GAC GCA TTA GAA ACC GAA CAA CTC AAA GCC	725
Val Leu Ile Asn Thr Phe Asp Ala Leu Glu Thr Glu Gln Leu Lys Ala	
215 220 225	
ATC GAC AGG TAT GAA CTA ATA TCC ATC GGC CCA TTA ATC CCA TCA TCG	773
Ile Asp Arg Tyr Glu Leu Ile Ser Ile Gly Pro Leu Ile Pro Ser Ser	
230 235 240	
ATA TTC TCA GAT GGC AAC GAC CCC TCA TCA AGC AAC AAA TCC TAC GGT	821
Ile Phe Ser Asp Gly Asn Asp Pro Ser Ser Ser Asn Lys Ser Tyr Gly	
245 250 255	
GGA GAC CTC TTC AGA AAA GCC GAT GAA ACT TAC ATG GAC TGG CTA AAC	869
Gly Asp Leu Phe Arg Lys Ala Asp Glu Thr Tyr Met Asp Trp Leu Asn	
260 265 270 275	
TCA AAA CCC GAA TCA TCG GTC GTT TAC GTT TCG TTC GGG AGC CTC CTG	917
Ser Lys Pro Glu Ser Ser Val Val Tyr Val Ser Phe Gly Ser Leu Leu	
280 285 290	
AGG CTC CCG AAA CCC CAA ATG GAA GAA ATA GCA ATA GGG CTT TCA GAC	965
Arg Leu Pro Lys Pro Gln Met Glu Glu Ile Ala Ile Gly Leu Ser Asp	
295 300 305	
ACC AAA TCG CCA GTT CTC TGG GTG ATA AGA AGA AAC GAA GAG GGC GAC	1013
Thr Lys Ser Pro Val Leu Trp Val Ile Arg Arg Asn Glu Glu Gly Asp	
310 315 320	
GAA CAA GAG CAA GCA GAA GAA GAA GAG AAG CTG CTG AGC TTC TTT GAT	1061
Glu Gln Glu Gln Ala Glu Glu Glu Glu Lys Leu Leu Ser Phe Phe Asp	
325 330 335	
CGT CAC GGA ACT GAA CGA CTC GGG AAA ATC GTG ACA TGG TGC TCA CAA	1109
Arg His Gly Thr Glu Arg Leu Gly Lys Ile Val Thr Trp Cys Ser Gln	
340 345 350 355	

TTG GAT GTT CTG ACG CAT AAG TCG GTG GGA TGC TTC GTG ACG CAT TGC	1157
Leu Asp Val Leu Thr His Lys Ser Val Gly Cys Phe Val Thr His Cys	
360 365 370	
GGT TGG AAT TCT GCT ATC GAG AGC CTG GCT TGT GGT GTG CCC GTG GTG	1205
Gly Trp Asn Ser Ala Ile Glu Ser Leu Ala Cys Gly Val Pro Val Val	
375 380 385	
TGC TTT CCT CAA TGG TTC GAT CAA GGG ACT AAT GCG AAG ATG ATC GAA	1253
Cys Phe Pro Gln Trp Phe Asp Gln Gly Thr Asn Ala Lys Met Ile Glu	
390 395 400	
GAT GTG TGG AGG AGT GGT GTG AGA GTC AGA GTG AAT GAG GAA GGC GGC	1301
Asp Val Trp Arg Ser Gly Val Arg Val Arg Val Asn Glu Glu Gly Gly	
405 410 415	
GTT GTT GAT AGG CGT GAG ATT AAG AGG TGC GTC TCG GAG GTT ATA AAG	1349
Val Val Asp Arg Arg Glu Ile Lys Arg Cys Val Ser Glu Val Ile Lys	
420 425 430 435	
AGT CGA GAG TTG AGA GAA AGC GCA ATG ATG TGG AAG GGT TTG GCT AAA	1397
Ser Arg Glu Leu Arg Glu Ser Ala Met Met Trp Lys Gly Leu Ala Lys	
440 445 450	
GAA GCT ATG GAT GAA GAA CGT GGA TCA TCA ATG AAC AAT CTG AAG AAT	1445
Glu Ala Met Asp Glu Glu Arg Gly Ser Ser Met Asn Asn Leu Lys Asn	
455 460 465	
TTT ATT ACT AGG ATT ATT AAT GAA AAT GCC TCA TAAGTTGTAC	1488
Phe Ile Thr Arg Ile Ile Asn Glu Asn Ala Ser	
470 475 478	
TATATATGTT ATTATTGTTG TTATGGACGT CGAATTAAGT ATTAGTTAAA TGATATGTAT	1548
TTAGAGGAAG GCCAAAACGG GCTACACCCG GCAGGCCACG GGTGGAAGAA GCCCGCCATG	1608
ATTTAAATA TATATTTTAA AATAAATATT TTCTACTATT AACTAAAAA AAAAAAAAAA	1668
AAA	1671

序列: 5

序列长度: 1437

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 紫苏 (*Perilla frutescens*)

组织种类: 叶

直接起源

文库名: cDNA library

克隆名: p8R6

序列

```

TTCAAACTC ATAACGTGAT TGAGCTAATG TGCACATCTT CCTCTTCAAA GTCTACAGTG      60
TCATCCTACC AGCATCATCA TGATCAATCT CTTTATAATG AGGAGAATGG AGTAACAAGG      120
AGTGGGTTTT GTTACTCAGC TTCAACCTAC GTACGTACTA CTACTGACTC AACTCTCAAG      180
AGAATGAATA TAATATATAA TGGGCGATAG ATCTTTGTAG ATATGTAGGT GTAGCCTGCA      240
GGTGGTTAAT TAATTTCCGG TGTGGGAAAA TAAATAAATA AATAAATATA GCG ATG AGC      299

```

Met Ser

1

```

AGC AGC AGC AGC AGA AGG TGG AGA GAG AAT GAG GGG ATG CGA AGG ACA      347
Ser Ser Ser Ser Arg Arg Trp Arg Glu Asn Glu Gly Met Arg Arg Thr

```

5

10

15

```

TTG CTG GGG TTG GGT TTG GGG CAG TTG GTT TCT TTC GAT TTG GCT ATC      395
Leu Leu Gly Leu Gly Leu Gly Gln Leu Val Ser Phe Asp Leu Ala Ile

```

20

25

30

```

ATG ACC TTT TCT GCT TCT TTG GTT TCA ACC ACA GTG GAT GCA CCA CTT      443

```

Met Thr Phe Ser Ala Ser Leu Val Ser Thr Thr Val Asp Ala Pro Leu
 35 40 45 50
 ACT ATG TCG TTC ACT ACA TAC ACT GTT GTG GCC CTG CTC TAT GGA ACC 491
 Thr Met Ser Phe Thr Thr Tyr Thr Val Val Ala Leu Leu Tyr Gly Thr
 55 60 65
 ATC TTG CTT TAC CGC CGC CAC AAA TTC TTG GTT CCA TGG TAC TGG TAT 539
 Ile Leu Leu Tyr Arg Arg His Lys Phe Leu Val Pro Trp Tyr Trp Tyr
 70 75 80
 GCT CTC CTG GGG TTC GTG GAC GTC CAC GGC AAT TAT CTT GTT AAT AAA 587
 Ala Leu Leu Gly Phe Val Asp Val His Gly Asn Tyr Leu Val Asn Lys
 85 90 95
 GCA TTC GAG TTG ACA TCG ATT ACG AGT GTG AGC ATA CTG GAT TGT TGG 635
 Ala Phe Glu Leu Thr Ser Ile Thr Ser Val Ser Ile Leu Asp Cys Trp
 100 105 110
 ACA ATC GTG TGG TCC ATC ATC TTT ACA TGG ATG TTC CTA GGC ACA AAA 683
 Thr Ile Val Trp Ser Ile Ile Phe Thr Trp Met Phe Leu Gly Thr Lys
 115 120 125 130
 TAC TCT GTA TAC CAG TTT GTC GGT GCT GCT ATT TGT GTA GGA GGC CTC 731
 Tyr Ser Val Tyr Gln Phe Val Gly Ala Ala Ile Cys Val Gly Gly Leu
 135 140 145
 CTC CTC GTG CTT CTT TCC GAC TCA GGC GTC ACT GCT GCT GGT TCG AAT 779
 Leu Leu Val Leu Leu Ser Asp Ser Gly Val Thr Ala Ala Gly Ser Asn
 150 155 160
 CCT CTT TTG GGT GAT TTT CTT GTC ATA ACA GGC TCT ATT TTG TTC ACA 827
 Pro Leu Leu Gly Asp Phe Leu Val Ile Thr Gly Ser Ile Leu Phe Thr
 165 170 175

CTC AGC ACT GTT GGT CAG GAA TAC TGC GTG AAG AGG AAA GAT CGT ATT 875
 Leu Ser Thr Val Gly Gln Glu Tyr Cys Val Lys Arg Lys Asp Arg Ile
 180 185 190
 GAA GTA GTA GCA ATG ATC GGT GTA TTT GGT ATG CTC ATC AGT GCA ACC 923
 Glu Val Val Ala Met Ile Gly Val Phe Gly Met Leu Ile Ser Ala Thr
 195 200 205 210
 GAG ATT ACT GTG CTG GAG AGG AAT GCC CTC TCA TCA ATG CAG TGG TCT 971
 Glu Ile Thr Val Leu Glu Arg Asn Ala Leu Ser Ser Met Gln Trp Ser
 215 220 225
 ACT GGA CTT TTG GCA GCC TAT GTT GTT TAT GCA CTG TCC AGC TTC CTC 1019
 Thr Gly Leu Leu Ala Ala Tyr Val Val Tyr Ala Leu Ser Ser Phe Leu
 230 235 240
 TTC TGC ACA CTC ACC CCT TTT CTT CTC AAG ATG AGT GGC GCT GCA TTT 1067
 Phe Cys Thr Leu Thr Pro Phe Leu Leu Lys Met Ser Gly Ala Ala Phe
 245 250 255
 TTC AAT CTT TCC ATG CTT ACA TCT GAT ATG TGG GCT GTT GCA ATT AGG 1115
 Phe Asn Leu Ser Met Leu Thr Ser Asp Met Trp Ala Val Ala Ile Arg
 260 265 270
 ACA TTC ATA TAC AAC CAG GAG GTT GAT TGG TTA TAC TAT TTG GCC TTT 1163
 Thr Phe Ile Tyr Asn Gln Glu Val Asp Trp Leu Tyr Tyr Leu Ala Phe
 275 280 285 290
 TGT CTC GTT GTT GTT GGA ATA TTC ATA TAT ACA AAA ACA GAG AAG GAT 1211
 Cys Leu Val Val Val Gly Ile Phe Ile Tyr Thr Lys Thr Glu Lys Asp
 295 300 305
 CCT AAC AAT ACG AGA GCC CTT GAG AAT GGA AAC TTG GAT CAT GAA TAT 1259
 Pro Asn Asn Thr Arg Ala Leu Glu Asn Gly Asn Leu Asp His Glu Tyr
 310 315 320

AGT CTC CTT GAG GAT CAA GAT GAC ACA CCA AGA AAA CCA TAGCTAGCTT 1308
 Ser Leu Leu Glu Asp Gln Asp Asp Thr Pro Arg Lys Pro
 325 330 335
 TGCCCACAAT CTTTTCATCA ACAGTTTTAA ATAATTCGTG AGGGGGAGAG AGATCGAGAT 1368
 ACTAATTAAT GGACGTCTAT TATATAGTTG GAGGTTTTTG TTTTATTTAT TTATTTGAGT 1428
 AAAAAAAAAA 1437

序列: 6

序列长度: 2105

序列类型: 核酸

链数: 双链

拓扑类型: 直链形

起源

生物名称: 矮牵牛花

组织种类: 花瓣

直接起源

文库名: cDNA library

克隆名: pSPGT1

序列

AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC 60
 TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA 120
 AACAGCTATG ACCATGATTA CGCCAAGCTC GAAATTAACC CTCACTAAAG GGAACAAAAG 180
 CTGGAGCTCC ACGCGGTGGC GGCCGCTCTA GAACTAGTGG ATCCCCCGGG CTGCAGGAAT 240
 TCCGTTGCTG TCGCCACAAT TTACAAACCA AGAAATTAAG CATCCCTTTC CCCCCCTTAA 300
 AAAACATACA AGTTTTTAAT TTTTCACTAA GCAAGAAAAT ATG GTG CAG CCT CAT GTC 358

Met Val Gln Pro His Val

ATC TTA ACA ACA TTT CCA GCA CAA GGC CAT ATT AAT CCA GCA CTT CAA	406
Ile Leu Thr Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln	
10 15 20	
TTT GCC AAG AAT CTT GTC AAG ATG GGC ATA GAA GTG ACA TTT TCT ACA	454
Phe Ala Lys Asn Leu Val Lys Met Gly Ile Glu Val Thr Phe Ser Thr	
25 30 35	
AGC ATT TAT GCC CAA AGC CGT ATG GAT GAA AAA TCC ATT CTT AAT GCA	502
Ser Ile Tyr Ala Gln Ser Arg Met Asp Glu Lys Ser Ile Leu Asn Ala	
40 45 50	
CCA AAA GGA TTG AAT TTC ATT CCA TTT TCC GAT GGC TTT GAT GAA GGT	550
Pro Lys Gly Leu Asn Phe Ile Pro Phe Ser Asp Gly Phe Asp Glu Gly	
55 60 65 70	
TTT GAT CAT TCA AAA GAC CCT GTA TTT TAC ATG TCA CAA CTT CGT AAA	598
Phe Asp His Ser Lys Asp Pro Val Phe Tyr Met Ser Gln Leu Arg Lys	
75 80 85	
TGT GGA AGT GAA ACT GTC AAA AAA ATA ATT CTC ACT TGC TCT GAA AAT	646
Cys Gly Ser Glu Thr Val Lys Lys Ile Ile Leu Thr Cys Ser Glu Asn	
90 95 100	
GGA CAG CCT ATA ACT TGC CTA CTT TAC TCC ATT TTC CTT CCT TGG GCA	694
Gly Gln Pro Ile Thr Cys Leu Leu Tyr Ser Ile Phe Leu Pro Trp Ala	
105 110 115	
GCA GAG GTA GCA CGT GAA GTT CAC ATC CCT TCT GCT CTT CTT TGG AGT	742
Ala Glu Val Ala Arg Glu Val His Ile Pro Ser Ala Leu Leu Trp Ser	
120 125 130	
CAA CCA GCA ACA ATA TTG GAC ATA TAT TAC TTC AAC TTT CAT GGA TAT	790
Gln Pro Ala Thr Ile Leu Asp Ile Tyr Tyr Phe Asn Phe His Gly Tyr	
135 140 145 150	

GAA AAA GCT ATG GCT AAT GAA TCC AAT GAT CCA AAT TGG TCC ATT CAA	838
Glu Lys Ala Met Ala Asn Glu Ser Asn Asp Pro Asn Trp Ser Ile Gln	
155 160 165	
CTT CCC GGG CTT CCA CTA CTG GAA ACT CGA GAT CTT CCT TCA TTT TTA	886
Leu Pro Gly Leu Pro Leu Leu Glu Thr Arg Asp Leu Pro Ser Phe Leu	
170 175 180	
CTT CCT TAT GGT GCA AAA GGG AGT CTT CGA GTT GCA CTT CCA CCA TTC	934
Leu Pro Tyr Gly Ala Lys Gly Ser Leu Arg Val Ala Leu Pro Pro Phe	
185 190 195	
AAA GAA TTG ATA GAC ACA TTA GAT GCT GAA ACC ACT CCT AAG ATT CTT	982
Lys Glu Leu Ile Asp Thr Leu Asp Ala Glu Thr Thr Pro Lys Ile Leu	
200 205 210	
GTG AAT ACA TTT GAT GAA TTA GAG CCT GAG GCA CTC AAT GCA ATT GAA	1030
Val Asn Thr Phe Asp Glu Leu Glu Pro Glu Ala Leu Asn Ala Ile Glu	
215 220 225 230	
GGT TAT AAG TTT TAT GGA ATT GGA CCG TTG ATT CCT TCT GCT TTC TTG	1078
Gly Tyr Lys Phe Tyr Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe Leu	
235 240 245	
GGT GGA AAT GAC CCT TTA GAT GCT TCA TTT GGT GGT GAT CTT TTT CAA	1126
Gly Gly Asn Asp Pro Leu Asp Ala Ser Phe Gly Gly Asp Leu Phe Gln	
250 255 260	
AAT TCA AAT GAC TAT ATG GAA TGG TTA AAC TCA AAG CCA AAT TCA TCA	1174
Asn Ser Asn Asp Tyr Met Glu Trp Leu Asn Ser Lys Pro Asn Ser Ser	
265 270 275	
GTT GTT TAT ATA TCT TTT GGG AGT CTA ATG AAT CCA TCT ATT AGC CAA	1222
Val Val Tyr Ile Ser Phe Gly Ser Leu Met Asn Pro Ser Ile Ser Gln	
280 285 290	

ATG GAG GAG ATA TCA AAA GGG TTG ATA GAC ATA GGA AGG CCG TTT TTA 1270
 Met Glu Glu Ile Ser Lys Gly Leu Ile Asp Ile Gly Arg Pro Phe Leu
 295 300 305 310
 TGG GTG ATA AAA GAA AAT GAA AAA GGC AAA GAA GAA GAG AAT AAA AAG 1318
 Trp Val Ile Lys Glu Asn Glu Lys Gly Lys Glu Glu Glu Asn Lys Lys
 315 320 325
 CTT GGT TGT ATT GAA GAA TTG GAA AAA ATA GGA AAA ATA GTT CCA TGG 1366
 Leu Gly Cys Ile Glu Glu Leu Glu Lys Ile Gly Lys Ile Val Pro Trp
 330 335 340
 TGT TCA CAA CTT GAA GTT CTA AAA CAT CCA TCT TTA GGA TGT TTT GTT 1414
 Cys Ser Gln Leu Glu Val Leu Lys His Pro Ser Leu Gly Cys Phe Val
 345 350 355
 TCT CAT TGT GGA TGG AAT TCA GCC TTA GAG AGT TTA GCT TGT GGA GTG 1462
 Ser His Cys Gly Trp Asn Ser Ala Leu Glu Ser Leu Ala Cys Gly Val
 360 365 370
 CCA GTT GTG GCA TTT CCT CAA TGG ACA GAT CAA ATG ACA AAT GCC AAA 1510
 Pro Val Val Ala Phe Pro Gln Trp Thr Asp Gln Met Thr Asn Ala Lys
 375 380 385 390
 CAA GTT GAA GAT GTG TGG AAA AGT GGA GTA AGA GTG AGA ATA AAT GAA 1558
 Gln Val Glu Asp Val Trp Lys Ser Gly Val Arg Val Arg Ile Asn Glu
 395 400 405
 GAT GGT GTT GTT GAA AGT GAG GAA ATC AAA AGG TGT ATT GAA TTG GTA 1606
 Asp Gly Val Val Glu Ser Glu Glu Ile Lys Arg Cys Ile Glu Leu Val
 410 415 420
 ATG GAT GGA GGA GAG AAA GGG GAA GAA TTG AGA AAG AAT GCT AAG AAA 1654
 Met Asp Gly Gly Glu Lys Gly Glu Glu Leu Arg Lys Asn Ala Lys Lys
 425 430 435

TGG AAA GAA TTG GCT AGA GAA GCT GTG AAG GAA GGT GGA TCT TCA CAC	1702
Trp Lys Glu Leu Ala Arg Glu Ala Val Lys Glu Gly Gly Ser Ser His	
440 445 450	
AAG AAT TTA AAG GCT TTT ATT GAT GAT GTT GCC AAA GGG TTT TAATATTTAC	1754
Lys Asn Leu Lys Ala Phe Ile Asp Asp Val Ala Lys Gly Phe	
455 460 465 468	
AGGCTTTTGC CGTGATATTA CTTCCCCTAG TTGGCGATTC ACTCTTTGTG GACTTGCTTG	1814
ACAAAAAACT GAGGGAATGT GCTAAGACAC GCTAATGCTT TAAGAAGTCA TTTCCAAGGC	1874
TTGAAGCCTG CTTTTAAAAC TTATTAGCCA GTAATCTATA GGGTTCTCTT CTATTTTCT	1934
CTGTCTCTCT TTTAGCCTT TTTCTTTCCA AGGTTTAAGA ATAGCGTGAA CATAGCTTAG	1994
TACGTAGTCT TGGTATCTCT ATCTTACCAA GTGCAAGATT ATGCTTATGC TGTCTCCTA	2054
AATTTCTTAA TAAAATGCAA GATGAAAAG TACAAAAAA AAAAAAAAAA A	2105

序 列 表

< 1 1 0 > Suntory Limited

< 1 2 0 > 编码具有糖转移活性的蛋白质的基因

< 1 3 0 > STY-F846-PCT

< 1 5 0 > JP PH9-200571

< 1 5 1 > 1997-07-25

< 1 6 0 > 1 1

< 2 1 0 > 1

< 2 1 1 > 1 5 0 7

< 2 1 2 > DNA

< 2 1 3 > *Perilla frutescens*

< 4 0 0 > 1

gaaaatttcc acaaaa atg gtc cgc cgc cgc gtg ctg cta gca acg ttt 49

Met Val Arg Arg Arg Val Leu Leu Ala Thr Phe

1

5

10

cct gcg caa ggc cac ata aat ccc gcc ctc caa ttc gcc aag aga ctc 97

Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala Lys Arg Leu

15

20

25

cta aaa gcc ggc act gac gtc aca ttt ttc acg agc gtt tat gca tgg 145

Leu Lys Ala Gly Thr Asp Val Thr Phe Phe Thr Ser Val Tyr Ala Trp

30

35

40

cgc cgc atg gcc aac aca gcc tcc gcc gct gcc gga aac cca ccg ggc 193

Arg Arg Met Ala Asn Thr Ala Ser Ala Ala Ala Gly Asn Pro Pro Gly

45

50

55

ctc gac ttc gtg gcg ttc tcc gac ggc tac gac gac ggg ctg aag ccc 241
 Leu Asp Phe Val Ala Phe Ser Asp Gly Tyr Asp Asp Gly Leu Lys Pro
 60 65 70 75
 tgc ggc gac ggg aag cgc tac atg tcc gag atg aaa gcc cgc ggc tcc 289
 Cys Gly Asp Gly Lys Arg Tyr Met Ser Glu Met Lys Ala Arg Gly Ser
 80 85 90
 gag gcc tta aga aac ctc ctt ctc aac aac cac gac gtc acg ttc gtc 337
 Glu Ala Leu Arg Asn Leu Leu Leu Asn Asn His Asp Val Thr Phe Val
 95 100 105
 gtc tac tcc cac ctc ttt gca tgg gcg gcg gag gtg gcg cgt gag tcc 385
 Val Tyr Ser His Leu Phe Ala Trp Ala Ala Glu Val Ala Arg Glu Ser
 110 115 120
 cag gtc ccg agc gcc ctt ctc tgg gtc gag ccc gcc acc gtg ctg tgc 433
 Gln Val Pro Ser Ala Leu Leu Trp Val Glu Pro Ala Thr Val Leu Cys
 125 130 135
 ata tat tac ttc tac ttc aac ggc tac gca gac gag atc gac gcc ggt 481
 Ile Tyr Tyr Phe Tyr Phe Asn Gly Tyr Ala Asp Glu Ile Asp Ala Gly
 140 145 150 155
 tcc gac gaa att cag ctc cct cgg ctt cca ccc ctg gag cag cgc agt 529
 Ser Asp Glu Ile Gln Leu Pro Arg Leu Pro Pro Leu Glu Gln Arg Ser
 160 165 170
 ctt ccg acc ttt ctg ctg ccg gag aca ccg gag aga ttc cgg ttg atg 577
 Leu Pro Thr Phe Leu Leu Pro Glu Thr Pro Glu Arg Phe Arg Leu Met
 175 180 185
 atg aag gag aag ctg gaa act tta gac ggt gaa gag aag gcg aaa gtg 625
 Met Lys Glu Lys Leu Glu Thr Leu Asp Gly Glu Glu Lys Ala Lys Val
 190 195 200

ttg gtg aac acg ttt gat gcg ttg gag ccc gat gca ctc acg gct att	673
Leu Val Asn Thr Phe Asp Ala Leu Glu Pro Asp Ala Leu Thr Ala Ile	
205 210 215	
gat agg tat gag ttg atc ggg atc ggg ccg ttg att ccc tcc gcc ttc	721
Asp Arg Tyr Glu Leu Ile Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe	
220 225 230 235	
ttg gac ggc gga gat ccc tcc gaa acg tct tac ggc ggc gat ctt ttc	769
Leu Asp Gly Gly Asp Pro Ser Glu Thr Ser Tyr Gly Gly Asp Leu Phe	
240 245 250	
gaa aaa tcg gag gag aat aac tgc gtg gag tgg ttg gac acg aag ccg	817
Glu Lys Ser Glu Glu Asn Asn Cys Val Glu Trp Leu Asp Thr Lys Pro	
255 260 265	
aaa tct tcg gtg gtg tat gtg tcg ttt ggg agc gtt ttg agg ttt cca	865
Lys Ser Ser Val Val Tyr Val Ser Phe Gly Ser Val Leu Arg Phe Pro	
270 275 280	
aag gca caa atg gaa gag att ggg aaa ggg cta tta gcc tgc gga agg	913
Lys Ala Gln Met Glu Glu Ile Gly Lys Gly Leu Leu Ala Cys Gly Arg	
285 290 295	
ccg ttt tta tgg atg ata cga gaa cag aag aat gac gac ggc gaa gaa	961
Pro Phe Leu Trp Met Ile Arg Glu Gln Lys Asn Asp Asp Gly Glu Glu	
300 305 310 315	
gaa gaa gaa gag ttg agt tgc att ggg gaa ttg aaa aaa atg ggg aaa	1009
Glu Glu Glu Glu Leu Ser Cys Ile Gly Glu Leu Lys Lys Met Gly Lys	
320 325 330	
ata gtt tcg tgg tgc tcg cag ttg gag gtt ctg gcg cac cct gcg ttg	1057
Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu Ala His Pro Ala Leu	
335 340 345	

gga tgt ttc gtg acg cat tgt ggg tgg aac tcg gct gtg gag agc ttg 1105
 Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser Ala Val Glu Ser Leu
 350 355 360
 agt tgc ggg gtt ccg gtg gtg gcg gtg ccg cag tgg ttt gat cag acg 1153
 Ser Cys Gly Val Pro Val Val Ala Val Pro Gln Trp Phe Asp Gln Thr
 365 370 375
 acg aat gcg aag ctg att gag gat gcg tgg ggg aca ggg gtg aga gtg 1201
 Thr Asn Ala Lys Leu Ile Glu Asp Ala Trp Gly Thr Gly Val Arg Val
 380 385 390 395
 aga atg aat gaa ggg ggt ggg gtt gat gga tct gag ata gag agg tgt 1249
 Arg Met Asn Glu Gly Gly Gly Val Asp Gly Ser Glu Ile Glu Arg Cys
 400 405 410
 gtg gag atg gtg atg gat ggg ggt gag aag agc aaa cta gtg aga gaa 1297
 Val Glu Met Val Met Asp Gly Gly Glu Lys Ser Lys Leu Val Arg Glu
 415 420 425
 aat gcc ata aaa tgg aag act ttg gcc aga gaa gcc atg gga gag gat 1345
 Asn Ala Ile Lys Trp Lys Thr Leu Ala Arg Glu Ala Met Gly Glu Asp
 430 435 440
 gga tct tca ctc aag aat ctc aac gcc ttt ctt cat caa gtt gca cgt 1393
 Gly Ser Ser Leu Lys Asn Leu Asn Ala Phe Leu His Gln Val Ala Arg
 445 450 455
 gct taatacaciaa aatggctttc cacttttaat ctactcaaac accggttcaa 1446
 Ala
 460
 ataaatatcc ccttcacatt ctttctatit cactatcaca ttitataatit tagtaacaaa 1506
 a 1507

< 2 1 0 > 2

< 2 1 1 > 1 4 7 0

< 2 1 2 > D N A

< 2 1 3 > *Perilla frutescens*

< 4 0 0 > 2

accaaacc aaacaaaattt ccacaaaa atg gtc cgc cgc cgc gtg ctg cta 48

Met Val Arg Arg Arg Val Leu Leu

1 5

gca acg ttt ccg gcg caa ggc cac ata aat ccc gcc ctc caa ttc gcc 96

Ala Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala

10 15 20

aag aga ctc cta aaa gcc ggc act gac gtc acg ttt ttc acg agc gtt 144

Lys Arg Leu Leu Lys Ala Gly Thr Asp Val Thr Phe Phe Thr Ser Val

25 30 35 40

tat gca tgg cgc cgc atg gcc aac aca gcc tcc gcc gct gcc gga aac 192

Tyr Ala Trp Arg Arg Met Ala Asn Thr Ala Ser Ala Ala Ala Gly Asn

45 50 55

cca ccg gcc ctc gac ttc gtg gcg ttc tcc gac gcc tac gac gac ggg 240

Pro Pro Gly Leu Asp Phe Val Ala Phe Ser Asp Gly Tyr Asp Asp Gly

60 65 70

ctg aag ccc gcc ggc gac ggg aag cgc tac atg tcc gag atg aaa gcc 288

Leu Lys Pro Gly Gly Asp Gly Lys Arg Tyr Met Ser Glu Met Lys Ala

75 80 85

cgc gcc tcc gag gcc tta aga aac ctc ctt ctc aac aac gac gac gtc 336

Arg Gly Ser Glu Ala Leu Arg Asn Leu Leu Leu Asn Asn Asp Asp Val

90 95 100

act ttc gtc gtc tac tcc cac ctc ttt gca tgg gcg gcg gag gtg gcg 384
 Thr Phe Val Val Tyr Ser His Leu Phe Ala Trp Ala Ala Glu Val Ala
 105 110 115 120
 cgt ttg tcc cac gtc ccg acc gcc ctt ctc tgg gtc gag ccc gcc acc 432
 Arg Leu Ser His Val Pro Thr Ala Leu Leu Trp Val Glu Pro Ala Thr
 125 130 135
 gtg ctg tgc ata tac cac ttc tac ttc aac ggc tac gca gac gag atc 480
 Val Leu Cys Ile Tyr His Phe Tyr Phe Asn Gly Tyr Ala Asp Glu Ile
 140 145 150
 gac gcc ggt tcc aat gaa att cag ctc cct cgg ctt cca tcc ctg gag 528
 Asp Ala Gly Ser Asn Glu Ile Gln Leu Pro Arg Leu Pro Ser Leu Glu
 155 160 165
 cag cgc agt ctt ccg acg ttt ctg ctg cct gcg acg ccg gag aga ttc 576
 Gln Arg Ser Leu Pro Thr Phe Leu Leu Pro Ala Thr Pro Glu Arg Phe
 170 175 180
 cgg ttg atg atg aag gag aag ctg gaa act tta gac ggt gaa gag aag 624
 Arg Leu Met Met Lys Glu Lys Leu Glu Thr Leu Asp Gly Glu Glu Lys
 185 190 195 200
 gcg aaa gta ttg gtg aac acg ttt gat gcg ttg gag ccc gat gca ctc 672
 Ala Lys Val Leu Val Asn Thr Phe Asp Ala Leu Glu Pro Asp Ala Leu
 205 210 215
 acg gct att gat agg tat gag ttg atc ggg atc ggg ccg ttg att ccc 720
 Thr Ala Ile Asp Arg Tyr Glu Leu Ile Gly Ile Gly Pro Leu Ile Pro
 220 225 230
 tcc gcc ttc ttg gac ggc gaa gat ccc tcc gaa acg tct tac ggc ggc 768
 Ser Ala Phe Leu Asp Gly Glu Asp Pro Ser Glu Thr Ser Tyr Gly Gly
 235 240 245

gat ctt ttc gaa aaa tcg gag gag aat aac tgc gtg gag tgg ttg aac	816
Asp Leu Phe Glu Lys Ser Glu Glu Asn Asn Cys Val Glu Trp Leu Asn	
250 255 260	
tcg aag ccg aaa tct tcg gtg gtg tat gtg tcg ttt ggg agc gtt ttg	864
Ser Lys Pro Lys Ser Ser Val Val Tyr Val Ser Phe Gly Ser Val Leu	
265 270 275 280	
agg ttt cca aag gca caa atg gaa gag att ggg aaa ggg cta tta gcc	912
Arg Phe Pro Lys Ala Gln Met Glu Glu Ile Gly Lys Gly Leu Leu Ala	
285 290 295	
tgc gga agg ccc ttt tta tgg atg ata cga gaa cag aag aat gac gac	960
Cys Gly Arg Pro Phe Leu Trp Met Ile Arg Glu Gln Lys Asn Asp Asp	
300 305 310	
ggc gaa gaa gaa gaa gaa gaa gaa gag ttg agt tgc att ggg gaa ttg	1008
Gly Glu Glu Glu Glu Glu Glu Glu Glu Glu Leu Ser Cys Ile Gly Glu Leu	
315 320 325	
aaa aaa atg ggg aaa ata gtg tcg tgg tgc tcg cag ttg gag gtt ctg	1056
Lys Lys Met Gly Lys Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu	
330 335 340	
gcg cac cct gcg ttg gga tgt ttc gtg acg cat tgt ggg tgg aac tcg	1104
Ala His Pro Ala Leu Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser	
345 350 355 360	
gct gtg gag agc ttg agt tgc ggg att ccg gtg gtg gcg gtg ccg cag	1152
Ala Val Glu Ser Leu Ser Cys Gly Ile Pro Val Val Ala Val Pro Gln	
365 370 375	
tgg ttt gat cag acg acg aat gcg aag ctg att gag gat gcg tgg ggg	1200
Trp Phe Asp Gln Thr Thr Asn Ala Lys Leu Ile Glu Asp Ala Trp Gly	
380 385 390	

aca ggg gtg aga gtg aga atg aat gaa ggg ggt ggg gtt gat gga tgt 1248

Thr Gly Val Arg Val Arg Met Asn Glu Gly Gly Gly Val Asp Gly Cys

395

400

405

gag ata gaa agg tgt gtg gag atg gtg atg gat ggg ggt gac aag acc 1296

Glu Ile Glu Arg Cys Val Glu Met Val Met Asp Gly Gly Asp Lys Thr

410

415

420

aaa cta gtg aga gaa aat gcc atc aaa tgg aag act ttg gcc aga caa 1344

Lys Leu Val Arg Glu Asn Ala Ile Lys Trp Lys Thr Leu Ala Arg Gln

425

430

435

440

gcc atg gga taggatggat cttcactcaa caatctcaac gcctttcttc 1393

Ala Met Gly

443

gtcaagttgc acacttttaa tctgctcaaa cagcgggttca aataaatatc cccttcact 1453

taaaaaaaaa aaaaaaa 1470

< 2 1 0 > 3

< 2 1 1 > 2 0 6 2

< 2 1 2 > D N A

< 2 1 3 > Verbena hybrida

< 4 0 0 > 3

attttaccaa aaaaataaaa aaaaa atg agc aga gct cac gtc ctc ttg gcc 52

Met Ser Arg Ala His Val Leu Leu Ala

1

5

aca ttc cca gca cag gga cac ata aat ccc gcc ctt caa ttc gcc aag 100

Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln Phe Ala Lys

10

15

20

25

cgt ctc gca aat gcc gac att caa gtc aca ttc ttc acc agc gtc tac	148
Arg Leu Ala Asn Ala Asp Ile Gln Val Thr Phe Phe Thr Ser Val Tyr	
30 35 40	
gca tgg cgc cgc atg tcc aga acc gcc gct ggc tca aac ggg ctc atc	196
Ala Trp Arg Arg Met Ser Arg Thr Ala Ala Gly Ser Asn Gly Leu Ile	
45 50 55	
aat ttt gtg tcg ttt tcc gac ggg tat gac gac ggg tta cag ccc gga	244
Asn Phe Val Ser Phe Ser Asp Gly Tyr Asp Asp Gly Leu Gln Pro Gly	
60 65 70	
gac gat ggg aag aac tac atg tcg gag atg aaa agc aga ggt ata aaa	292
Asp Asp Gly Lys Asn Tyr Met Ser Glu Met Lys Ser Arg Gly Ile Lys	
75 80 85	
gcc ttg agc gat act ctt gca gcc aat aat gtc gat caa aaa agc agc	340
Ala Leu Ser Asp Thr Leu Ala Ala Asn Asn Val Asp Gln Lys Ser Ser	
90 95 100 105	
aaa atc acg ttc gtg gtg tac tcc cac ctc ttt gca tgg gcg gcc aag	388
Lys Ile Thr Phe Val Val Tyr Ser His Leu Phe Ala Trp Ala Ala Lys	
110 115 120	
gtg gcg cgt gag ttc cat ctc cgg agc gcg cta ctc tgg att gag cca	436
Val Ala Arg Glu Phe His Leu Arg Ser Ala Leu Leu Trp Ile Glu Pro	
125 130 135	
gct acg gtg ttg gat ata ttt tac ttt tat ttc aac ggc tat agc gac	484
Ala Thr Val Leu Asp Ile Phe Tyr Phe Tyr Phe Asn Gly Tyr Ser Asp	
140 145 150	
gaa atc gat gcg ggt tcg gat gct att cac ttg ccc gga gga ctc cca	532
Glu Ile Asp Ala Gly Ser Asp Ala Ile His Leu Pro Gly Gly Leu Pro	
155 160 165	

gtg	ctg	gcc	cag	cgt	gat	tta	ccg	tct	ttc	ctt	ctt	cct	tcc	acg	cat	580
Val	Leu	Ala	Gln	Arg	Asp	Leu	Pro	Ser	Phe	Leu	Leu	Pro	Ser	Thr	His	
170						175				180					185	
gag	aga	ttc	cgt	tca	ctg	atg	aag	gag	aaa	ttg	gaa	act	tta	gaa	ggt	628
Glu	Arg	Phe	Arg	Ser	Leu	Met	Lys	Glu	Lys	Leu	Glu	Thr	Leu	Glu	Gly	
				190					195					200		
gaa	gaa	aaa	cct	aag	gtc	ttg	gtg	aac	agc	ttt	gat	gcg	ttg	gag	cct	676
Glu	Glu	Lys	Pro	Lys	Val	Leu	Val	Asn	Ser	Phe	Asp	Ala	Leu	Glu	Pro	
			205					210					215			
gat	gcg	ctc	aag	gcc	att	gat	aag	tac	gag	atg	att	gca	atc	ggg	ccg	724
Asp	Ala	Leu	Lys	Ala	Ile	Asp	Lys	Tyr	Glu	Met	Ile	Ala	Ile	Gly	Pro	
			220					225					230			
ttg	att	cct	tcc	gca	ttc	ttg	gac	ggg	aaa	gat	cct	tcg	gac	agg	tct	772
Leu	Ile	Pro	Ser	Ala	Phe	Leu	Asp	Gly	Lys	Asp	Pro	Ser	Asp	Arg	Ser	
			235					240					245			
ttc	ggc	gga	gat	ttg	ttc	gag	aaa	ggg	tcg	aat	gac	gac	gat	tgc	ctc	820
Phe	Gly	Gly	Asp	Leu	Phe	Glu	Lys	Gly	Ser	Asn	Asp	Asp	Asp	Cys	Leu	
250						255				260				265		
gaa	tgg	ttg	agc	acg	aat	cct	cga	tct	tcg	gtg	ggt	tac	ggt	tcg	ttc	868
Glu	Trp	Leu	Ser	Thr	Asn	Pro	Arg	Ser	Ser	Val	Val	Tyr	Val	Ser	Phe	
				270					275					280		
gga	agc	ttc	ggt	aat	acg	acg	aag	tcg	caa	atg	gaa	gag	ata	gca	aga	916
Gly	Ser	Phe	Val	Asn	Thr	Thr	Lys	Ser	Gln	Met	Glu	Glu	Ile	Ala	Arg	
				285				290					295			
ggg	ctg	tta	gat	tgt	ggg	agg	ccg	ttt	ttg	tgg	gtg	gta	aga	gta	aac	964
Gly	Leu	Leu	Asp	Cys	Gly	Arg	Pro	Phe	Leu	Trp	Val	Val	Arg	Val	Asn	
				300				305					310			

gaa gga gaa gag gta ttg ata agt tgc atg gag gag ttg aaa cga gtg 1012
 Glu Gly Glu Glu Val Leu Ile Ser Cys Met Glu Glu Leu Lys Arg Val
 315 320 325
 ggg aaa att gta tct tgg tgt tct caa ttg gaa gtc ctg acg cat ccc 1060
 Gly Lys Ile Val Ser Trp Cys Ser Gln Leu Glu Val Leu Thr His Pro
 330 335 340 345
 tcg ttg gga tgt ttc gtg aca cac tgc ggg tgg aat tcg act cta gag 1108
 Ser Leu Gly Cys Phe Val Thr His Cys Gly Trp Asn Ser Thr Leu Glu
 350 355 360
 agt ata tct ttc ggg gtt ccg atg gtg gct ttt ccg cag tgg ttc gat 1156
 Ser Ile Ser Phe Gly Val Pro Met Val Ala Phe Pro Gln Trp Phe Asp
 365 370 375
 caa ggg acg aat gcg aag ctg atg gag gat gtg tgg agg acg ggt gtg 1204
 Gln Gly Thr Asn Ala Lys Leu Met Glu Asp Val Trp Arg Thr Gly Val
 380 385 390
 aga gtg aga gct aat gag gag ggt agc gtc gtt gat ggt gat gaa att 1252
 Arg Val Arg Ala Asn Glu Glu Gly Ser Val Val Asp Gly Asp Glu Ile
 395 400 405
 agg aga tgt att gag gag gtt atg gat ggg gga gaa aag agt agg aaa 1300
 Arg Arg Cys Ile Glu Glu Val Met Asp Gly Gly Glu Lys Ser Arg Lys
 410 415 420 425
 ctt aga gag agt gct ggc aag tgg aag gat ttg gca aga aaa gct atg 1348
 Leu Arg Glu Ser Ala Gly Lys Trp Lys Asp Leu Ala Arg Lys Ala Met
 430 435 440
 gag gaa gat gga tct tca gtt aac aac ctc aag gtc ttt ctt gat gag 1396
 Glu Glu Asp Gly Ser Ser Val Asn Asn Leu Lys Val Phe Leu Asp Glu
 445 450 455

ggt gta ggt atc taaagacgta aatgaggtcc ccataggcaa aattgcaaat 1448

Val Val Gly Ile

460 461

tatcatctcgt aagttgaata ctttttggct ttaattttgt tcgagtttgt ttttcaaaat 1508

ttatcttgta attttacatt gagtgtaaat ttagtctgat ttttaactgga aaaatataaa 1568

attcattggt gagactcttc atcaaaatca tctgatttcc tttattgtct tgggtcaaaat 1628

tctcatatca attggaaaaa ataaatttca aaatcgtcca attttgaacc aagaagaag 1688

tataatttga ccaaaataat aaaaggattc aagtgatctt gatgaagtgt ctgagcgacg 1748

agttctatat ttttccaccg aatttctaac gagtttttga atttttttta gccaaaatcg 1808

gactaacttt gtacaaaatg aaaagttata tgatgaaatt ttaaaaaaca aactcagaca 1868

ataataaagc ccgaaagtag taaaattacc tgacgaaatt tgcaatttcg cctcctat 1928

taattttttt ggtgtgttta ataaatcgtt tattttactt ttaattaaaa taaaagtga 1988

atgcatgata gcttgggtgag tatatatgag ttgatggtaa tgtacgatat tttctaaaaa 2048

aaaaaaaaaa aaaa 2062

< 2 1 0 > 4

< 2 1 1 > 1 6 7 1

< 2 1 2 > D N A

< 2 1 3 > Torenia hybrira

< 2 2 0 >

< 2 2 1 >

< 2 2 2 >

< 2 2 3 > Xaa (64) is Cys or Phe, Xaa (65) is Ser or Pro.

< 4 0 0 > 4

aacacataaa aaaaaaataa aagaagaat aattaaaaaa aaaa atg gtt aac 53

Met Val Asn

aaa cgc cat att cta cta gca aca ttc cca gca caa ggc cac ata aac	101
Lys Arg His Ile Leu Leu Ala Thr Phe Pro Ala Gln Gly His Ile Asn	
5 10 15	
cct tct ctc gag ttc gcc aaa agg ctc ctc aac acc gga tac gtc gac	149
Pro Ser Leu Glu Phe Ala Lys Arg Leu Leu Asn Thr Gly Tyr Val Asp	
20 25 30 35	
caa gtc aca ttc ttc acg agt gta tac gca ttg aga cgc atg cgc ttc	197
Gln Val Thr Phe Phe Thr Ser Val Tyr Ala Leu Arg Arg Met Arg Phe	
40 45 50	
gaa acc gat ccg agc agc aga atc gat ttc gtg gca tkt yca gat tct	245
Glu Thr Asp Pro Ser Ser Arg Ile Asp Phe Val Ala Xaa Xaa Asp Ser	
55 60 65	
tac gat gat ggc tta aag aaa ggc gac gat ggc aaa aac tac atg tcg	293
Tyr Asp Asp Gly Leu Lys Lys Gly Asp Asp Gly Lys Asn Tyr Met Ser	
70 75 80	
gag atg aga aag cgc gga acg aag gcc tta aag gac act ctt att aag	341
Glu Met Arg Lys Arg Gly Thr Lys Ala Leu Lys Asp Thr Leu Ile Lys	
85 90 95	
ctc aac gat gct gcg atg gga agt gaa tgt tac aat cgc gtg agc ttt	389
Leu Asn Asp Ala Ala Met Gly Ser Glu Cys Tyr Asn Arg Val Ser Phe	
100 105 110 115	
gtg gtg tac tct cat cta ttt tcg tgg gca gct gaa gtg gcg cgt gaa	437
Val Val Tyr Ser His Leu Phe Ser Trp Ala Ala Glu Val Ala Arg Glu	
120 125 130	
gtc gac gtg ccg agt gcc ctt ctt tgg att gaa ccg gct acg gtt ttc	485
Val Asp Val Pro Ser Ala Leu Leu Trp Ile Glu Pro Ala Thr Val Phe	
135 140 145	

gat gtg tac tat ttt tac ttc aat ggg tat gcc gat gat atc gat gcg	533
Asp Val Tyr Tyr Phe Tyr Phe Asn Gly Tyr Ala Asp Asp Ile Asp Ala	
150 155 160	
ggc tca gat caa atc caa ctg ccc aat ctt ccg cag ctc tcc aag caa	581
Gly Ser Asp Gln Ile Gln Leu Pro Asn Leu Pro Gln Leu Ser Lys Gln	
165 170 175	
gat ctc ccc tct ttc cta ctc cct tcg agc ccc gcg aga ttc cga acc	629
Asp Leu Pro Ser Phe Leu Leu Pro Ser Ser Pro Ala Arg Phe Arg Thr	
180 185 190 195	
cta atg aaa gaa aag ttc gac acg ctc gac aaa gaa ccg aaa gcg aag	677
Leu Met Lys Glu Lys Phe Asp Thr Leu Asp Lys Glu Pro Lys Ala Lys	
200 205 210	
gtc ttg ata aac acg ttc gac gca tta gaa acc gaa caa ctc aaa gcc	725
Val Leu Ile Asn Thr Phe Asp Ala Leu Glu Thr Glu Gln Leu Lys Ala	
215 220 225	
atc gac agg tat gaa cta ata tcc atc ggc cca tta atc cca tca tcg	773
Ile Asp Arg Tyr Glu Leu Ile Ser Ile Gly Pro Leu Ile Pro Ser Ser	
230 235 240	
ata ttc tca gat ggc aac gac ccc tca tca agc aac aaa tcc tac ggt	821
Ile Phe Ser Asp Gly Asn Asp Pro Ser Ser Ser Asn Lys Ser Tyr Gly	
245 250 255	
gga gac ctc ttc aga aaa gcc gat gaa act tac atg gac tgg cta aac	869
Gly Asp Leu Phe Arg Lys Ala Asp Glu Thr Tyr Met Asp Trp Leu Asn	
260 265 270 275	
tca aaa ccc gaa tca tcg gtc gtt tac gtt tcg ttc ggg agc ctc ctg	917
Ser Lys Pro Glu Ser Ser Val Val Tyr Val Ser Phe Gly Ser Leu Leu	
280 285 290	

agg ctc ccg aaa ccc caa atg gaa gaa ata gca ata ggg ctt tca gac 965
 Arg Leu Pro Lys Pro Gln Met Glu Glu Ile Ala Ile Gly Leu Ser Asp
 295 300 305
 acc aaa tcg cca gtt ctc tgg gtg ata aga aga aac gaa gag ggc gac 1013
 Thr Lys Ser Pro Val Leu Trp Val Ile Arg Arg Asn Glu Glu Gly Asp
 310 315 320
 gaa caa gag caa gca gaa gaa gaa gag aag ctg ctg agc ttc ttt gat 1061
 Glu Gln Glu Gln Ala Glu Glu Glu Glu Lys Leu Leu Ser Phe Phe Asp
 325 330 335
 cgt cac gga act gaa cga ctc ggg aaa atc gtg aca tgg tgc tca caa 1109
 Arg His Gly Thr Glu Arg Leu Gly Lys Ile Val Thr Trp Cys Ser Gln
 340 345 350 355
 ttg gat gtt ctg acg cat aag tcg gtg gga tgc ttc gtg acg cat tgc 1157
 Leu Asp Val Leu Thr His Lys Ser Val Gly Cys Phe Val Thr His Cys
 360 365 370
 ggt tgg aat tct gct atc gag agc ctg gct tgt ggt gtg ccc gtg gtg 1205
 Gly Trp Asn Ser Ala Ile Glu Ser Leu Ala Cys Gly Val Pro Val Val
 375 380 385
 tgc ttt cct caa tgg ttc gat caa ggg act aat gcg aag atg atc gaa 1253
 Cys Phe Pro Gln Trp Phe Asp Gln Gly Thr Asn Ala Lys Met Ile Glu
 390 395 400
 gat gtg tgg agg agt ggt gtg aga gtc aga gtg aat gag gaa ggc ggc 1301
 Asp Val Trp Arg Ser Gly Val Arg Val Arg Val Asn Glu Glu Gly Gly
 405 410 415
 gtt gtt gat agg cgt gag att aag agg tgc gtc tcg gag gtt ata aag 1349
 Val Val Asp Arg Arg Glu Ile Lys Arg Cys Val Ser Glu Val Ile Lys
 420 425 430 435

agt cga gag ttg aga gaa agc gca atg atg tgg aag ggt ttg gct aaa 1397
 Ser Arg Glu Leu Arg Glu Ser Ala Met Met Trp Lys Gly Leu Ala Lys
 440 445 450
 gaa gct atg gat gaa gaa cgt gga tca tca atg aac aat ctg aag aat 1445
 Glu Ala Met Asp Glu Glu Arg Gly Ser Ser Met Asn Asn Leu Lys Asn
 455 460 465
 ttt att act agg att att aat gaa aat gcc tca taagttgtac 1488
 Phe Ile Thr Arg Ile Ile Asn Glu Asn Ala Ser
 470 475 478
 tatatatgtt attattgttg ttatggacgt cgaattaagt attagttaaa tgatatgtat 1548
 ttagaggaag gccaaaacgg gctacacccg gcaggccacg ggttggaaaa gcccgccatg 1608
 atttaaaata tatattttta aataaatatt ttctactatt aaactaaaaa aaaaaaaaaa 1668
 aaa 1671

< 2 1 0 > 5

< 2 1 1 > 1 4 3 7

< 2 1 2 > D N A

< 2 1 3 > *Perilla frutescens*

< 4 0 0 > 5

ttcaaaactc ataacgtgat tgagctaattg tgcacatctt cctcttcaaa gtctacagtg 60
 tcctcctacc agcatcatca tgatcaatct ctttataatg aggagaatgg agtaacaagg 120
 agtggggtttt gttactcagc ttcaacctac gtacgtacta ctactgactc aactctcaag 180
 agaatgaata taatatataa tgggcgatag atctttgtag atatgtaggt gtagcctgca 240
 ggtgggttaat taatttccgg tgtgggaaaa taaataaata aataaatata gcg atg agc 299

Met Ser

agc agc agc agc aga agg tgg aga gag aat gag ggg atg cga agg aca 347
 Ser Ser Ser Ser Arg Arg Trp Arg Glu Asn Glu Gly Met Arg Arg Thr
 5 10 15
 ttg ctg ggg ttg ggt ttg ggg cag ttg gtt tct ttc gat ttg gct atc 395
 Leu Leu Gly Leu Gly Leu Gly Gln Leu Val Ser Phe Asp Leu Ala Ile
 20 25 30
 atg acc ttt tct gct tct ttg gtt tca acc aca gtg gat gca cca ctt 443
 Met Thr Phe Ser Ala Ser Leu Val Ser Thr Thr Val Asp Ala Pro Leu
 35 40 45 50
 act atg tcg ttc act aca tac act gtt gtg gcc ctg ctc tat gga acc 491
 Thr Met Ser Phe Thr Thr Tyr Thr Val Val Ala Leu Leu Tyr Gly Thr
 55 60 65
 atc ttg ctt tac cgc cgc cac aaa ttc ttg gtt cca tgg tac tgg tat 539
 Ile Leu Leu Tyr Arg Arg His Lys Phe Leu Val Pro Trp Tyr Trp Tyr
 70 75 80
 gct ctc ctg ggg ttc gtg gac gtc cac ggc aat tat ctt gtt aat aaa 587
 Ala Leu Leu Gly Phe Val Asp Val His Gly Asn Tyr Leu Val Asn Lys
 85 90 95
 gca ttc gag ttg aca tcg att acg agt gtg agc ata ctg gat tgt tgg 635
 Ala Phe Glu Leu Thr Ser Ile Thr Ser Val Ser Ile Leu Asp Cys Trp
 100 105 110
 aca atc gtg tgg tcc atc atc ttt aca tgg atg ttc cta ggc aca aaa 683
 Thr Ile Val Trp Ser Ile Ile Phe Thr Trp Met Phe Leu Gly Thr Lys
 115 120 125 130
 tac tct gta tac cag ttt gtc ggt gct gct att tgt gta gga ggc ctc 731
 Tyr Ser Val Tyr Gln Phe Val Gly Ala Ala Ile Cys Val Gly Gly Leu
 135 140 145

ctc ctc gtg ctt ctt tcc gac tca ggg gtc act gct gct ggt tgc aat	779
Leu Leu Val Leu Leu Ser Asp Ser Gly Val Thr Ala Ala Gly Ser Asn	
150 155 160	
cct ctt ttg ggt gat ttt ctt gtc ata aca ggc tct att ttg ttc aca	827
Pro Leu Leu Gly Asp Phe Leu Val Ile Thr Gly Ser Ile Leu Phe Thr	
165 170 175	
ctc agc act gtt ggt cag gaa tac tgc gtg aag agg aaa gat cgt att	875
Leu Ser Thr Val Gly Gln Glu Tyr Cys Val Lys Arg Lys Asp Arg Ile	
180 185 190	
gaa gta gta gca atg atc ggt gta ttt ggt atg ctc atc agt gca acc	923
Glu Val Val Ala Met Ile Gly Val Phe Gly Met Leu Ile Ser Ala Thr	
195 200 205 210	
gag att act gtg ctg gag agg aat gcc ctc tca tca atg cag tgg tct	971
Glu Ile Thr Val Leu Glu Arg Asn Ala Leu Ser Ser Met Gln Trp Ser	
215 220 225	
act gga ctt ttg gca gcc tat gtt gtt tat gca ctg tcc agc ttc ctc	1019
Thr Gly Leu Leu Ala Ala Tyr Val Val Tyr Ala Leu Ser Ser Phe Leu	
230 235 240	
ttc tgc aca ctc acc cct ttt ctt ctc aag atg agt ggc gct gca ttt	1067
Phe Cys Thr Leu Thr Pro Phe Leu Leu Lys Met Ser Gly Ala Ala Phe	
245 250 255	
ttc aat ctt tcc atg ctt aca tct gat atg tgg gct gtt gca att agg	1115
Phe Asn Leu Ser Met Leu Thr Ser Asp Met Trp Ala Val Ala Ile Arg	
260 265 270	
aca ttc ata tac aac cag gag gtt gat tgg tta tac tat ttg gcc ttt	1163
Thr Phe Ile Tyr Asn Gln Glu Val Asp Trp Leu Tyr Tyr Leu Ala Phe	
275 280 285 290	

tgt ctc gtt gtt gtt gga ata ttc ata tat aca aaa aca gag aag gat 1211

Cys Leu Val Val Val Gly Ile Phe Ile Tyr Thr Lys Thr Glu Lys Asp

295

300

305

cct aac aat acg aga gcc ctt gag aat gga aac ttg gat cat gaa tat 1259

Pro Asn Asn Thr Arg Ala Leu Glu Asn Gly Asn Leu Asp His Glu Tyr

310

315

320

agt ctc ctt gag gat caa gat gac aca cca aga aaa cca tagctagctt 1308

Ser Leu Leu Glu Asp Gln Asp Asp Thr Pro Arg Lys Pro

325

330

335

tgccacaat cttttcatca acagtttttaa ataattcgtg aggggggagag agatcgagat 1368

actaattaat ggacgtctat tatatagttg gaggtttttg ttttatttat ttatttgagt 1428

aaaaaaaaa 1437

< 2 1 0 > 6

< 2 1 1 > 2 1 0 5

< 2 1 2 > D N A

< 2 1 3 > Petunia hybrida

< 4 0 0 > 6

agtgagcgca acgcaattaa tgtgagttag ctactcatt aggcacccca ggctttacac 60

tttatgcttc cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga 120

aacagctatg accatgatta cgccaagctc gaaattaacc ctactaaag ggaacaaaag 180

ctggagctcc acgcggtggc ggccgctcta gaactagtgg atccccggg ctgcaggaat 240

tccgttgctg tcgccacaat ttacaaacca agaaattaag catcccttc ccccccttaa 300

aaaacataca agtttttaat ttttactaa gcaagaaaat atg gtg cag cct cat gtc 358

Met Val Gln Pro His Val

1

5

atc tta aca aca ttt cca gca caa ggc cat att aat cca gca ctt caa	406
Ile Leu Thr Thr Phe Pro Ala Gln Gly His Ile Asn Pro Ala Leu Gln	
10 15 20	
ttt gcc aag aat ctt gtc aag atg ggc ata gaa gtg aca ttt tct aca	454
Phe Ala Lys Asn Leu Val Lys Met Gly Ile Glu Val Thr Phe Ser Thr	
25 30 35	
agc att tat gcc caa agc cgt atg gat gaa aaa tcc att ctt aat gca	502
Ser Ile Tyr Ala Gln Ser Arg Met Asp Glu Lys Ser Ile Leu Asn Ala	
40 45 50	
cca aaa gga ttg aat ttc att cca ttt tcc gat ggc ttt gat gaa ggt	550
Pro Lys Gly Leu Asn Phe Ile Pro Phe Ser Asp Gly Phe Asp Glu Gly	
55 60 65 70	
ttt gat cat tca aaa gac cct gta ttt tac atg tca caa ctt cgt aaa	598
Phe Asp His Ser Lys Asp Pro Val Phe Tyr Met Ser Gln Leu Arg Lys	
75 80 85	
tgt gga agt gaa act gtc aaa aaa ata att ctc act tgc tct gaa aat	646
Cys Gly Ser Glu Thr Val Lys Lys Ile Ile Leu Thr Cys Ser Glu Asn	
90 95 100	
gga cag cct ata act tgc cta ctt tac tcc att ttc ctt cct tgg gca	694
Gly Gln Pro Ile Thr Cys Leu Leu Tyr Ser Ile Phe Leu Pro Trp Ala	
105 110 115	
gca gag gta gca cgt gaa gtt cac atc cct tct gct ctt ctt tgg agt	742
Ala Glu Val Ala Arg Glu Val His Ile Pro Ser Ala Leu Leu Trp Ser	
120 125 130	
caa cca gca aca ata ttg gac ata tat tac ttc aac ttt cat gga tat	790
Gln Pro Ala Thr Ile Leu Asp Ile Tyr Tyr Phe Asn Phe His Gly Tyr	
135 140 145 150	

gaa aaa gct atg gct aat gaa tcc aat gat cca aat tgg tcc att caa	838
Glu Lys Ala Met Ala Asn Glu Ser Asn Asp Pro Asn Trp Ser Ile Gln	
155 160 165	
ctt ccc ggg ctt cca cta ctg gaa act cga gat ctt cct tca ttt tta	886
Leu Pro Gly Leu Pro Leu Leu Glu Thr Arg Asp Leu Pro Ser Phe Leu	
170 175 180	
ctt cct tat ggt gca aaa ggg agt ctt cga gtt gca ctt cca cca ttc	934
Leu Pro Tyr Gly Ala Lys Gly Ser Leu Arg Val Ala Leu Pro Pro Phe	
185 190 195	
aaa gaa ttg ata gac aca tta gat gct gaa acc act cct aag att ctt	982
Lys Glu Leu Ile Asp Thr Leu Asp Ala Glu Thr Thr Pro Lys Ile Leu	
200 205 210	
gtg aat aca ttt gat gaa tta gag cct gag gca ctc aat gca att gaa	1030
Val Asn Thr Phe Asp Glu Leu Glu Pro Glu Ala Leu Asn Ala Ile Glu	
215 220 225 230	
ggg tat aag ttt tat gga att gga ccg ttg att cct tct gct ttc ttg	1078
Gly Tyr Lys Phe Tyr Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe Leu	
235 240 245	
ggg gga aat gac cct tta gat gct tca ttt ggt ggt gat ctt ttt caa	1126
Gly Gly Asn Asp Pro Leu Asp Ala Ser Phe Gly Gly Asp Leu Phe Gln	
250 255 260	
aat tca aat gac tat atg gaa tgg tta aac tca aag cca aat tca tca	1174
Asn Ser Asn Asp Tyr Met Glu Trp Leu Asn Ser Lys Pro Asn Ser Ser	
265 270 275	
gtt gtt tat ata tct ttt ggg agt cta atg aat cca tct att agc caa	1222
Val Val Tyr Ile Ser Phe Gly Ser Leu Met Asn Pro Ser Ile Ser Gln	
280 285 290	

atg gag gag ata tca aaa ggg ttg ata gac ata gga agg ccg ttt tta	1270
Met Glu Glu Ile Ser Lys Gly Leu Ile Asp Ile Gly Arg Pro Phe Leu	
295 300 305 310	
tgg gtg ata aaa gaa aat gaa aaa ggc aaa gaa gaa gag aat aaa aag	1318
Trp Val Ile Lys Glu Asn Glu Lys Gly Lys Glu Glu Glu Asn Lys Lys	
315 320 325	
ctt ggt tgt att gaa gaa ttg gaa aaa ata gga aaa ata gtt cca tgg	1366
Leu Gly Cys Ile Glu Glu Leu Glu Lys Ile Gly Lys Ile Val Pro Trp	
330 335 340	
tgt tca caa ctt gaa gtt cta aaa cat cca tct tta gga tgt ttt gtt	1414
Cys Ser Gln Leu Glu Val Leu Lys His Pro Ser Leu Gly Cys Phe Val	
345 350 355	
tct cat tgt gga tgg aat tca gcc tta gag agt tta gct tgt gga gtg	1462
Ser His Cys Gly Trp Asn Ser Ala Leu Glu Ser Leu Ala Cys Gly Val	
360 365 370	
cca gtt gtg gca ttt cct caa tgg aca gat caa atg aca aat gcc aaa	1510
Pro Val Val Ala Phe Pro Gln Trp Thr Asp Gln Met Thr Asn Ala Lys	
375 380 385 390	
caa gtt gaa gat gtg tgg aaa agt gga gta aga gtg aga ata aat gaa	1558
Gln Val Glu Asp Val Trp Lys Ser Gly Val Arg Val Arg Ile Asn Glu	
395 400 405	
gat ggt gtt gtt gaa agt gag gaa atc aaa agg tgt att gaa ttg gta	1606
Asp Gly Val Val Glu Ser Glu Glu Ile Lys Arg Cys Ile Glu Leu Val	
410 415 420	
atg gat gga gga gag aaa ggg gaa gaa ttg aga aag aat gct aag aaa	1654
Met Asp Gly Gly Glu Lys Gly Glu Glu Leu Arg Lys Asn Ala Lys Lys	
425 430 435	

tgg aaa gaa ttg gct aga gaa gct gtg aag gaa ggt gga tct tca cac 1702
 Trp Lys Glu Leu Ala Arg Glu Ala Val Lys Glu Gly Gly Ser Ser His
 440 445 450
 aag aat tta aag gct ttt att gat gat gtt gcc aaa ggg ttt taatatttac 1754
 Lys Asn Leu Lys Ala Phe Ile Asp Asp Val Ala Lys Gly Phe
 455 460 465 468
 aggcttttgc cgtgatatta ctccccctag ttggcgattc actctttgtg gacttgcttg 1814
 acaaaaaact gagggaatgt gctaagacac gctaattgctt taagaagtca tttccaaggc 1874
 ttgaagcctg cttttaaaac ttattagcca gtaattctata gggttctctt ctatttttct 1934
 ctgtctctct ttttagcctt tttctttcca aggtttaaga atagcgtgaa catagcttag 1994
 tacgtagtct tggatatctt atcttaccac gtcgaagatt atgcttatgc tgcctccta 2054
 aatttcttaa taaaatgcaa gatgaaaaag tacaaaaaaaa aaaaaaaaaa a 2105

< 2 1 0 > 7

< 2 1 1 > 4 6 0

< 2 1 2 > P R T

< 2 1 3 > *Perilla frutescens*

< 4 0 0 > 7

Met Val Arg Arg Arg Val Leu Leu Ala Thr Phe Pro Ala Gln Gly His

1 5 10 15

Ile Asn Pro Ala Leu Gln Phe Ala Lys Arg Leu Leu Lys Ala Gly Thr

20 25 30

Asp Val Thr Phe Phe Thr Ser Val Tyr Ala Trp Arg Arg Met Ala Asn

35 40 45

Thr Ala Ser Ala Ala Ala Gly Asn Pro Pro Gly Leu Asp Phe Val Ala

50 55 60

Phe Ser Asp Gly Tyr Asp Asp Gly Leu Lys Pro Cys Gly Asp Gly Lys

65 70 75 80

09.04.23

Arg Tyr Met Ser Glu Met Lys Ala Arg Gly Ser Glu Ala Leu Arg Asn			
	85	90	95
Leu Leu Leu Asn Asn His Asp Val Thr Phe Val Val Tyr Ser His Leu			
	100	105	110
Phe Ala Trp Ala Ala Glu Val Ala Arg Glu Ser Gln Val Pro Ser Ala			
	115	120	125
Leu Leu Trp Val Glu Pro Ala Thr Val Leu Cys Ile Tyr Tyr Phe Tyr			
	130	135	140
Phe Asn Gly Tyr Ala Asp Glu Ile Asp Ala Gly Ser Asp Glu Ile Gln			
145	150	155	160
Leu Pro Arg Leu Pro Pro Leu Glu Gln Arg Ser Leu Pro Thr Phe Leu			
	165	170	175
Leu Pro Glu Thr Pro Glu Arg Phe Arg Leu Met Met Lys Glu Lys Leu			
	180	185	190
Glu Thr Leu Asp Gly Glu Glu Lys Ala Lys Val Leu Val Asn Thr Phe			
	195	200	205
Asp Ala Leu Glu Pro Asp Ala Leu Thr Ala Ile Asp Arg Tyr Glu Leu			
	210	215	220
Ile Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe Leu Asp Gly Gly Asp			
225	230	235	240
Pro Ser Glu Thr Ser Tyr Gly Gly Asp Leu Phe Glu Lys Ser Glu Glu			
	245	250	255
Asn Asn Cys Val Glu Trp Leu Asp Thr Lys Pro Lys Ser Ser Val Val			
	260	265	270
Tyr Val Ser Phe Gly Ser Val Leu Arg Phe Pro Lys Ala Gln Met Glu			
	275	280	285
Glu Ile Gly Lys Gly Leu Leu Ala Cys Gly Arg Pro Phe Leu Trp Met			

290	295	300	
Ile Arg Glu Gln Lys Asn Asp Asp Gly Glu Glu Glu Glu Glu Leu			
305	310	315	320
Ser Cys Ile Gly Glu Leu Lys Lys Met Gly Lys Ile Val Ser Trp Cys			
	325	330	335
Ser Gln Leu Glu Val Leu Ala His Pro Ala Leu Gly Cys Phe Val Thr			
	340	345	350
His Cys Gly Trp Asn Ser Ala Val Glu Ser Leu Ser Cys Gly Val Pro			
	355	360	365
Val Val Ala Val Pro Gln Trp Phe Asp Gln Thr Thr Asn Ala Lys Leu			
	370	375	380
Ile Glu Asp Ala Trp Gly Thr Gly Val Arg Val Arg Met Asn Glu Gly			
385	390	395	400
Gly Gly Val Asp Gly Ser Glu Ile Glu Arg Cys Val Glu Met Val Met			
	405	410	415
Asp Gly Gly Glu Lys Ser Lys Leu Val Arg Glu Asn Ala Ile Lys Trp			
	420	425	430
Lys Thr Leu Ala Arg Glu Ala Met Gly Glu Asp Gly Ser Ser Leu Lys			
	435	440	445
Asn Leu Asn Ala Phe Leu His Gln Val Ala Arg Ala			
	450	455	460

< 2 1 0 > 8

< 2 1 1 > 4 4 3

< 2 1 2 > P R T

< 2 1 3 > *Perilla frutescens*

< 4 0 0 > 8

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

Asp Ala Leu Glu Pro Asp Ala Leu Thr Ala Ile Asp Arg Tyr Glu Leu
 210 215 220
 Ile Gly Ile Gly Pro Leu Ile Pro Ser Ala Phe Leu Asp Gly Glu Asp
 225 230 235 240
 Pro Ser Glu Thr Ser Tyr Gly Gly Asp Leu Phe Glu Lys Ser Glu Glu
 245 250 255
 Asn Asn Cys Val Glu Trp Leu Asn Ser Lys Pro Lys Ser Ser Val Val
 260 265 270
 Tyr Val Ser Phe Gly Ser Val Leu Arg Phe Pro Lys Ala Gln Met Glu
 275 280 285
 Glu Ile Gly Lys Gly Leu Leu Ala Cys Gly Arg Pro Phe Leu Trp Met
 290 295 300
 Ile Arg Glu Gln Lys Asn Asp Asp Gly Glu Glu Glu Glu Glu Glu
 305 310 315 320
 Glu Leu Ser Cys Ile Gly Glu Leu Lys Lys Met Gly Lys Ile Val Ser
 325 330 335
 Trp Cys Ser Gln Leu Glu Val Leu Ala His Pro Ala Leu Gly Cys Phe
 340 345 350
 Val Thr His Cys Gly Trp Asn Ser Ala Val Glu Ser Leu Ser Cys Gly
 355 360 365
 Ile Pro Val Val Ala Val Pro Gln Trp Phe Asp Gln Thr Thr Asn Ala
 370 375 380
 Lys Leu Ile Glu Asp Ala Trp Gly Thr Gly Val Arg Val Arg Met Asn
 385 390 395 400
 Glu Gly Gly Gly Val Asp Gly Cys Glu Ile Glu Arg Cys Val Glu Met
 405 410 415

Val Met Asp Gly Gly Asp Lys Thr Lys Leu Val Arg Glu Asn Ala Ile

420

425

430

Lys Trp Lys Thr Leu Ala Arg Gln Ala Met Gly

435

440

443

< 2 1 0 > 9

< 2 1 1 > 4 6 1

< 2 1 2 > P R T

< 2 1 3 > Verbena hybrida

< 4 0 0 > 9

Met Ser Arg Ala His Val Leu Leu Ala Thr Phe Pro Ala Gln Gly His

1

5

10

15

Ile Asn Pro Ala Leu Gln Phe Ala Lys Arg Leu Ala Asn Ala Asp Ile

20

25

30

Gln Val Thr Phe Phe Thr Ser Val Tyr Ala Trp Arg Arg Met Ser Arg

35

40

45

Thr Ala Ala Gly Ser Asn Gly Leu Ile Asn Phe Val Ser Phe Ser Asp

50

55

60

Gly Tyr Asp Asp Gly Leu Gln Pro Gly Asp Asp Gly Lys Asn Tyr Met

65

70

75

80

Ser Glu Met Lys Ser Arg Gly Ile Lys Ala Leu Ser Asp Thr Leu Ala

85

90

95

Ala Asn Asn Val Asp Gln Lys Ser Ser Lys Ile Thr Phe Val Val Tyr

100

105

110

Ser His Leu Phe Ala Trp Ala Ala Lys Val Ala Arg Glu Phe His Leu

115

120

125

Arg Ser Ala Leu Leu Trp Ile Glu Pro Ala Thr Val Leu Asp Ile Phe

130

135

140

Tyr Phe Tyr Phe Asn Gly Tyr Ser Asp Glu Ile Asp Ala Gly Ser Asp			
145	150	155	160
Ala Ile His Leu Pro Gly Gly Leu Pro Val Leu Ala Gln Arg Asp Leu			
	165	170	175
Pro Ser Phe Leu Leu Pro Ser Thr His Glu Arg Phe Arg Ser Leu Met			
	180	185	190
Lys Glu Lys Leu Glu Thr Leu Glu Gly Glu Glu Lys Pro Lys Val Leu			
	195	200	205
Val Asn Ser Phe Asp Ala Leu Glu Pro Asp Ala Leu Lys Ala Ile Asp			
	210	215	220
Lys Tyr Glu Met Ile Ala Ile Gly Pro Leu Ile Pro Ser Ala Phe Leu			
	225	230	235
Asp Gly Lys Asp Pro Ser Asp Arg Ser Phe Gly Gly Asp Leu Phe Glu			
	245	250	255
Lys Gly Ser Asn Asp Asp Asp Cys Leu Glu Trp Leu Ser Thr Asn Pro			
	260	265	270
Arg Ser Ser Val Val Tyr Val Ser Phe Gly Ser Phe Val Asn Thr Thr			
	275	280	285
Lys Ser Gln Met Glu Glu Ile Ala Arg Gly Leu Leu Asp Cys Gly Arg			
	290	295	300
Pro Phe Leu Trp Val Val Arg Val Asn Glu Gly Glu Glu Val Leu Ile			
	305	310	315
Ser Cys Met Glu Glu Leu Lys Arg Val Gly Lys Ile Val Ser Trp Cys			
	325	330	335
Ser Gln Leu Glu Val Leu Thr His Pro Ser Leu Gly Cys Phe Val Thr			
	340	345	350

His Cys Gly Trp Asn Ser Thr Leu Glu Ser Ile Ser Phe Gly Val Pro
 355 360 365

Met Val Ala Phe Pro Gln Trp Phe Asp Gln Gly Thr Asn Ala Lys Leu
 370 375 380

Met Glu Asp Val Trp Arg Thr Gly Val Arg Val Arg Ala Asn Glu Glu
 385 390 395 400

Gly Ser Val Val Asp Gly Asp Glu Ile Arg Arg Cys Ile Glu Glu Val
 405 410 415

Met Asp Gly Gly Glu Lys Ser Arg Lys Leu Arg Glu Ser Ala Gly Lys
 420 425 430

Trp Lys Asp Leu Ala Arg Lys Ala Met Glu Glu Asp Gly Ser Ser Val
 435 440 445

Asn Asn Leu Lys Val Phe Leu Asp Glu Val Val Gly Ile
 450 455 460 461

< 2 1 0 > 1 0

< 2 1 1 > 4 7 8

< 2 1 2 > P R T

< 2 1 3 > *Torenia hybrida*

< 2 2 0 >

< 2 2 1 >

< 2 2 2 >

< 2 2 3 > Xaa (64) is Cys or Phe, Xaa (65) is Ser or Pro.

< 4 0 0 > 1 0

Met Val Asn Lys Arg His Ile Leu Leu Ala Thr Phe Pro Ala Gln Gly
 1 5 10 15

His Ile Asn Pro Ser Leu Glu Phe Ala Lys Arg Leu Leu Asn Thr Gly
 20 25 30

Tyr	Val	Asp	Gln	Val	Thr	Phe	Phe	Thr	Ser	Val	Tyr	Ala	Leu	Arg	Arg
	35							40					45		
Met	Arg	Phe	Glu	Thr	Asp	Pro	Ser	Ser	Arg	Ile	Asp	Phe	Val	Ala	Xaa
	50							55					60		
Xaa	Asp	Ser	Tyr	Asp	Asp	Gly	Leu	Lys	Lys	Gly	Asp	Asp	Gly	Lys	Asn
	65					70				75				80	
Tyr	Met	Ser	Glu	Met	Arg	Lys	Arg	Gly	Thr	Lys	Ala	Leu	Lys	Asp	Thr
					85					90				95	
Leu	Ile	Lys	Leu	Asn	Asp	Ala	Ala	Met	Gly	Ser	Glu	Cys	Tyr	Asn	Arg
			100					105						110	
Val	Ser	Phe	Val	Val	Tyr	Ser	His	Leu	Phe	Ser	Trp	Ala	Ala	Glu	Val
			115					120						125	
Ala	Arg	Glu	Val	Asp	Val	Pro	Ser	Ala	Leu	Leu	Trp	Ile	Glu	Pro	Ala
			130					135					140		
Thr	Val	Phe	Asp	Val	Tyr	Tyr	Phe	Tyr	Phe	Asn	Gly	Tyr	Ala	Asp	Asp
	145					150				155				160	
Ile	Asp	Ala	Gly	Ser	Asp	Gln	Ile	Gln	Leu	Pro	Asn	Leu	Pro	Gln	Leu
					165				170					175	
Ser	Lys	Gln	Asp	Leu	Pro	Ser	Phe	Leu	Leu	Pro	Ser	Ser	Pro	Ala	Arg
			180					185						190	
Phe	Arg	Thr	Leu	Met	Lys	Glu	Lys	Phe	Asp	Thr	Leu	Asp	Lys	Glu	Pro
			195					200						205	
Lys	Ala	Lys	Val	Leu	Ile	Asn	Thr	Phe	Asp	Ala	Leu	Glu	Thr	Glu	Gln
			210					215						220	
Leu	Lys	Ala	Ile	Asp	Arg	Tyr	Glu	Leu	Ile	Ser	Ile	Gly	Pro	Leu	Ile
	225					230				235				240	

Pro Ser Ser Ile Phe Ser Asp Gly Asn Asp Pro Ser Ser Ser Asn Lys			
245	250	255	
Ser Tyr Gly Gly Asp Leu Phe Arg Lys Ala Asp Glu Thr Tyr Met Asp			
260	265	270	
Trp Leu Asn Ser Lys Pro Glu Ser Ser Val Val Tyr Val Ser Phe Gly			
275	280	285	
Ser Leu Leu Arg Leu Pro Lys Pro Gln Met Glu Glu Ile Ala Ile Gly			
290	295	300	
Leu Ser Asp Thr Lys Ser Pro Val Leu Trp Val Ile Arg Arg Asn Glu			
305	310	315	320
Glu Gly Asp Glu Gln Glu Gln Ala Glu Glu Glu Glu Lys Leu Leu Ser			
325	330	335	
Phe Phe Asp Arg His Gly Thr Glu Arg Leu Gly Lys Ile Val Thr Trp			
340	345	350	
Cys Ser Gln Leu Asp Val Leu Thr His Lys Ser Val Gly Cys Phe Val			
355	360	365	
Thr His Cys Gly Trp Asn Ser Ala Ile Glu Ser Leu Ala Cys Gly Val			
370	375	380	
Pro Val Val Cys Phe Pro Gln Trp Phe Asp Gln Gly Thr Asn Ala Lys			
385	390	395	400
Met Ile Glu Asp Val Trp Arg Ser Gly Val Arg Val Arg Val Asn Glu			
405	410	415	
Glu Gly Gly Val Val Asp Arg Arg Glu Ile Lys Arg Cys Val Ser Glu			
420	425	430	
Val Ile Lys Ser Arg Glu Leu Arg Glu Ser Ala Met Met Trp Lys Gly			
435	440	445	

Leu Ala Lys Glu Ala Met Asp Glu Glu Arg Gly Ser Ser Met Asn Asn
 450 455 460
 Leu Lys Asn Phe Ile Thr Arg Ile Ile Asn Glu Asn Ala Ser
 465 470 475 478
 < 2 1 0 > 1 1
 < 2 1 1 > 3 3 5
 < 2 1 2 > P R T
 < 2 1 3 > *Perilla frutescens*
 < 4 0 0 > 1 1
 Met Ser Ser Ser Ser Ser Arg Arg Trp Arg Glu Asn Glu Gly Met Arg
 1 5 10 15
 Arg Thr Leu Leu Gly Leu Gly Leu Gly Gln Leu Val Ser Phe Asp Leu
 20 25 30
 Ala Ile Met Thr Phe Ser Ala Ser Leu Val Ser Thr Thr Val Asp Ala
 35 40 45
 Pro Leu Thr Met Ser Phe Thr Thr Tyr Thr Val Val Ala Leu Leu Tyr
 50 55 60
 Gly Thr Ile Leu Leu Tyr Arg Arg His Lys Phe Leu Val Pro Trp Tyr
 65 70 75 80
 Trp Tyr Ala Leu Leu Gly Phe Val Asp Val His Gly Asn Tyr Leu Val
 85 90 95
 Asn Lys Ala Phe Glu Leu Thr Ser Ile Thr Ser Val Ser Ile Leu Asp
 100 105 110
 Cys Trp Thr Ile Val Trp Ser Ile Ile Phe Thr Trp Met Phe Leu Gly
 115 120 125
 Thr Lys Tyr Ser Val Tyr Gln Phe Val Gly Ala Ala Ile Cys Val Gly
 130 135 140

Gly Leu Leu Leu Val Leu Leu Ser Asp Ser Gly Val Thr Ala Ala Gly
145 150 155 160

Ser Asn Pro Leu Leu Gly Asp Phe Leu Val Ile Thr Gly Ser Ile Leu
165 170 175

Phe Thr Leu Ser Thr Val Gly Gln Glu Tyr Cys Val Lys Arg Lys Asp
180 185 190

Arg Ile Glu Val Val Ala Met Ile Gly Val Phe Gly Met Leu Ile Ser
195 200 205

Ala Thr Glu Ile Thr Val Leu Glu Arg Asn Ala Leu Ser Ser Met Gln
210 215 220

Trp Ser Thr Gly Leu Leu Ala Ala Tyr Val Val Tyr Ala Leu Ser Ser
225 230 235 240

Phe Leu Phe Cys Thr Leu Thr Pro Phe Leu Leu Lys Met Ser Gly Ala
245 250 255

Ala Phe Phe Asn Leu Ser Met Leu Thr Ser Asp Met Trp Ala Val Ala
260 265 270

Ile Arg Thr Phe Ile Tyr Asn Gln Glu Val Asp Trp Leu Tyr Tyr Leu
275 280 285

Ala Phe Cys Leu Val Val Val Gly Ile Phe Ile Tyr Thr Lys Thr Glu
290 295 300

Lys Asp Pro Asn Asn Thr Arg Ala Leu Glu Asn Gly Asn Leu Asp His
305 310 315 320

Glu Tyr Ser Leu Leu Glu Asp Gln Asp Asp Thr Pro Arg Lys Pro
325 330 335

< 2 1 0 > 1 2

< 2 1 1 > 4 6 8

< 2 1 2 > P R T

< 2 1 3 > *Petunia hybrida*

< 4 0 0 > 1 2

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

Val Ala Leu Pro Pro Phe Lys Glu Leu Ile Asp Thr Leu Asp Ala Glu

195

200

205

Thr Thr Pro Lys Ile Leu Val Asn Thr Phe Asp Glu Leu Glu Pro Glu

210

215

220

Ala Leu Asn Ala Ile Glu Gly Tyr Lys Phe Tyr Gly Ile Gly Pro Leu

225

230

235

240

Ile Pro Ser Ala Phe Leu Gly Gly Asn Asp Pro Leu Asp Ala Ser Phe

245

250

255

Gly Gly Asp Leu Phe Gln Asn Ser Asn Asp Tyr Met Glu Trp Leu Asn

260

265

270

Ser Lys Pro Asn Ser Ser Val Val Tyr Ile Ser Phe Gly Ser Leu Met

275

280

285

Asn Pro Ser Ile Ser Gln Met Glu Glu Ile Ser Lys Gly Leu Ile Asp

290

295

300

Ile Gly Arg Pro Phe Leu Trp Val Ile Lys Glu Asn Glu Lys Gly Lys

305

310

315

320

Glu Glu Glu Asn Lys Lys Leu Gly Cys Ile Glu Glu Leu Glu Lys Ile

325

330

335

Gly Lys Ile Val Pro Trp Cys Ser Gln Leu Glu Val Leu Lys His Pro

340

345

350

Ser Leu Gly Cys Phe Val Ser His Cys Gly Trp Asn Ser Ala Leu Glu

355

360

365

Ser Leu Ala Cys Gly Val Pro Val Val Ala Phe Pro Gln Trp Thr Asp

370

375

380

Gln Met Thr Asn Ala Lys Gln Val Glu Asp Val Trp Lys Ser Gly Val

385

390

395

400

1004-23

Arg Val Arg Ile Asn Glu Asp Gly Val Val Glu Ser Glu Glu Ile Lys

405

410

415

Arg Cys Ile Glu Leu Val Met Asp Gly Gly Glu Lys Gly Glu Glu Leu

420

425

430

Arg Lys Asn Ala Lys Lys Trp Lys Glu Leu Ala Arg Glu Ala Val Lys

435

440

445

Glu Gly Gly Ser Ser His Lys Asn Leu Lys Ala Phe Ile Asp Asp Val

450

455

460

Ala Lys Gly Phe

465

468