



(19)中華民國智慧財產局

(12)發明說明書公開本

(11)公開編號：TW 201934752 A

(43)公開日：中華民國 108 (2019) 年 09 月 01 日

(21)申請案號：108104187

(22)申請日：中華民國 108 (2019) 年 02 月 01 日

(51)Int. Cl. : C12N15/17 (2006.01)

C12N15/81 (2006.01)

C07K14/62 (2006.01)

(30)優先權：2018/02/09 中國大陸

201810131794.3

(71)申請人：大陸商江蘇恆瑞醫藥股份有限公司(中國大陸)JIANGSU HENGRUI MEDICINE CO., LTD. (CN)

中國大陸

(72)發明人：王菲菲 WANG, FEIFEI (CN)；陳磊 CHEN, LEI (CN)；王宏偉 WANG, HONGWEI (CN)

(74)代理人：洪武雄；陳昭誠

申請實體審查：無 申請專利範圍項數：23 項 圖式數：0 共 29 頁

(54)名稱

一種密碼子優化的人胰島素類似物前體基因和信號肽基因

A CODON OPTIMIZED HUMAN INSULIN ANALOGUE PRECURSOR GENE AND A SIGNAL PEPTIDE GENE

(57)摘要

本發明涉及一種密碼子優化的人胰島素類似物前體基因和信號肽基因。具體而言，本發明涉及一種密碼子優化的用於酵母表達胰島素類似物前體的核酸分子，包含一種胰島素類似物前體核酸分子和一種酵母分泌信號肽 α -因子核酸分子的融合。該經優化的核酸分子使得胰島素類似物前體在畢赤酵母中的表達量大幅提高，降低人胰島素類似物的生產成本，進而降低治療糖尿病藥物的成本。

The present invention relates to a codon optimized human insulin analogue precursor gene and a signal peptide gene. Specifically, the present invention relates to a codon optimized nucleic acid molecule for yeast expression of insulin analogue precursors, comprising fusion of an insulin analogue precursor nucleic acid molecule and a yeast secretion signal peptide alpha-factor nucleic acid molecule. The optimized nucleic acid molecule greatly increases the expression of insulin analogue precursors in *Pichia pastoris*, reduces the production cost of human insulin analogues, and thereby reduces the cost of diabetic drugs.

【發明說明書】

【中文發明名稱】 一種密碼子優化的人胰島素類似物前體基因和信號肽基因

【英文發明名稱】 A CODON OPTIMIZED HUMAN INSULIN ANALOGUE

PRECURSOR GENE AND A SIGNAL PEPTIDE GENE

【技術領域】

【0001】 本發明涉及一種密碼子優化的人胰島素類似物前體基因和一種密碼子優化的 α -因子(α -factor)信號肽基因，並提供了該人胰島素類似物前體基因的表達方法。

【先前技術】

【0002】 人胰島素是由 51 個胺基酸組成的多肽，包含兩條鏈，分別為 A 鏈和 B 鏈。胰島素的主要藥效是調節糖代謝，在糖尿病治療中，胰島素干預作為替代或補充治療是最直接有效的方法。胰島素還有促進脂肪合成、抑制脂肪分解、減少酮體生成的作用，所以亦被用來糾正胰島素相關酮症和酸血症的各種症狀。

【0003】 胰島素最早一直由豬、牛等動物胰腺提取而得，但這些產品與人胰島素結構有所不同，故存在免疫原性。上世紀八十年代初、中期美國禮來和丹麥諾和諾德公司相繼開發出基因重組人胰島素生產技術，從此基因工程表達人胰島素及其類似物成為工業主流手段。但是人胰島素作用時間較短，使得病人必須頻繁注射，極為不便。因此，人們致力於獲得一些能夠更長時間作用於人體的胰島素類似物及其衍生物。其中，使用醯化基團修飾人胰島素或其類

似物，是一種提高其半衰期的有效方法。WO2018024186 公開了一種 B29 位被長鏈脂肪酸取代、B30 位胺基酸缺失的人胰島素類似物，公開了該人胰島素類似物的結構和生物活性。WO9507931 公開了一種 B29 位連接一個十四醯基側鏈、B30 位胺基酸缺失的胰島素類似物及其製劑。WO2005012347 公開了一種 B29 位被一個谷胺酸和長鏈脂肪酸取代、B30 位胺基酸缺失的人胰島素類似物。目前常用的表達人胰島素及其類似物的表達系統主要有是大腸桿菌、釀酒酵母和巴斯德畢赤酵母，其中大腸桿菌表達為包涵體形式，需經過包涵體裂解和複性，工藝繁瑣且得率低，釀酒酵母與畢赤酵母有著操作簡易、易於培養、對外源蛋白修飾、可以分泌表達等遺傳表達優點，但釀酒酵母分泌效率低、表達菌株不夠穩定，相比之下畢赤酵母是工業生產重組蛋白應用更廣泛的表達系統。工業上，生產過程中發酵產量是控制生產成本的關鍵因素，由於胰島素市場需求量大，作為重要生產商的諾和諾德利用酵母表達生產時罐子規模達到幾十噸，對廠房、設備要求很高，成本也高，所以提高人胰島素及其類似物的發酵產量對工業生產有著重要意義。

【0004】 遺傳密碼子是三聯體密碼，一個密碼子由信使核糖核酸（mRNA）上相鄰的三個鹼基組成。遺傳密碼共有64種，但不同的生物，甚至同種生物不同的蛋白質編碼基因，對不同密碼子使用頻率並不相同，具有一定的偏愛性。外源基因的密碼子主要在轉譯層級上影響基因的表達，很多文獻已經證明密碼子的優化對於在畢赤酵母中提高外源蛋白的表達量卓有成效。畢赤酵母表達外源基因分為胞內表達和分泌表達兩種，後者需要信號肽來引導外源基因表達產物的分泌，目前最普遍使用的信號肽為來源於釀酒酵母 α -因子信號肽，其核苷酸序列也來源於釀酒酵母，目前還沒有報道針對畢赤酵母優化的 α -

因子信號肽核苷酸序列。胰島素前體密碼子優化的相關文獻很多，如 Gurramkonda 等人優化的人胰島素前體基因序列並在畢赤酵母的表達研究 (Gurramkonda et al. Application of simple fed-batch technique to high-level secretory production of insulin precursor using *Pichia pastoris* with subsequent purification and conversion to human insulin. Microbial Cell Factories, 2010, 9:31)，專利WO1998028429公開了表達人胰島素類似物前體的基因序列，該基因編碼的胰島素前體胺基酸序列是EEGEPK-B(1-29)-AAK-A(1-21)，其中 EEGEPK為胰島素前體N端延伸，稱為間隔肽或前導肽，可以保護胰島素前體的N端免於酵母蛋白酶的水解作用，並且能提高胰島素前體的表達效率；B(1-29)是缺失B30位蘇胺酸的人胰島素B鏈；A(1-21)是人胰島素A鏈胺基酸序列；AAK為連接B鏈和A鏈的連接肽，也稱為C肽。

【0005】 為了進一步提升人胰島素及其類似物前體的產量，本發明結合畢赤酵母密碼子偏好性，優化了胰島素類似物前體基因和 α -因子信號肽基因，在畢赤酵母中進行分泌表達，並與現有技術中的人胰島素類似物前體基因作為對照，結果表明，本發明中經過密碼子優化後的基因表達人胰島素類似物前體量提高了近 2 倍，能夠大大降低後期人胰島素及其類似物工業生產的成本。

【發明內容】

【0006】 本發明的一些實施方案中提供了核酸分子，其包含如下結構：

5'—(PS)_a—(SP)_b—(LS)_c—GE—(P'S)_d—3'，其中，

PS 是編碼加工位點的核酸分子，a 是 0 或 1；

SP 是編碼信號肽的核酸分子，b 是 0 或 1；

LS 是編碼間隔肽的核酸分子，c 是 0 或 1；

GE 是編碼目標多肽的核酸分子；

P'S 是編碼加工位點的核酸分子，d 是 0 或 1。

【0007】 一些實施方案中提供了核酸分子，其包含如下結構：

5'—(PS)_a—(SP)_b—(LS)_c—GE—(P'S)_d—3'，其中，

PS 是編碼加工位點的核酸分子，a 是 0 或 1；

SP 是編碼信號肽的核酸分子，b 是 1；

LS 是編碼間隔肽的核酸分子，c 是 1；

GE 是編碼目標多肽的核酸分子；

P'S 是編碼加工位點的核酸分子，d 是 0 或 1。

【0008】 在一些實施方案中，編碼信號肽的核酸分子包含如 SEQ ID NO:1 所示序列。

【0009】 在一些實施方案中，目標多肽是人胰島素類似物前體多肽；編碼人胰島素類似物前體多肽的核酸分子包含如 SEQ ID NO:3 所示序列。

【0010】 在一些實施方案中，編碼信號肽的核酸分子 (SP) 的核酸序列如 SEQ ID NO:1 所示，胺基酸序列如 SEQ ID NO:2 所示：

ATGAGATTTCTTCTATTTTCACTGCTGTTTTGTTTGCCGCTTCCTCTGCTT
TGGCAGCTCCAGTTAATAACAACCACTGAAGATGAGACTGCTCAAATCCC
AGCCGAAGCAGTTATTGGTTACTCCGACTTGGAAGGAGATTTTGACGTC
GCTGTTTTACCATTCTCTAATTCCACTAATAACGGTCTGTTGTTTATTAATA
CTACCATTGCTTCTATCGCCGCTAAGGAGGAAGGTGTGTCCCTCGAGAA
AAGA

SEQ ID NO:1

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYS DLEGDFDVAVL
PFSNSTNNGLLFINTTIIASIAAKEEGVSLEKR

SEQ ID NO:2

【0011】 在一些實施方案中，編碼目標多肽的核酸分子（GE）可以是編碼人胰島素類似物前體的核酸分子，其可以是 B30 位蘇胺酸缺失的人胰島素。人胰島素類似物前體的核酸分子序列如 SEQ ID NO:3 所示，胺基酸序列如 SEQ ID NO:4 所示：

TTCGTCAACCAGCACTTGTGTGGTTCCCATTTGGTTGAGGCTCTGTACTT
GGTCTGTGGAGAAAGAGGTTTCTTTTACACCCCAAAGGCTGCTAAGGGT
ATCGTTGAGCAATGTTGCACCTCTATTTGTTCCCTGTATCAGTTGGAAAA
CTACTGCAACTAA

SEQ ID NO: 3

FVNQHLCGSHLVEALYLVCGERGFFYTPKAAKGIVEQCCTSICSLYQLENYC
N

SEQ ID NO:4

【0012】 在一些實施方案中，編碼人胰島素類似物前體的核酸分子的 88 至 96 位(即 GCTGCTAAG)為編碼連接肽(也稱 C 肽)的核酸分子，其可以被替代，包括但不限於如下的序列：GCCGCTAAG、GCTGCCAAG、GCTGCTAAA、GCCGCCAAG。

【0013】 在一些實施方案中，編碼間隔肽的核酸分子(LS)的序列如 SEQ ID NO:5 所示：

gaagaagggaaccaaag

SEQ ID NO:5

【0014】 在一些實施方案中，PS 和/或 P'S 是編碼酶切位點的核酸分子；

【0015】 較佳的，PS 是編碼 EcoR I 酶切位點的核酸分子，和/或 P'S 是編碼 Not I 酶切位點的核酸分子。

【0016】 在一些實施方案中，提供能夠表達人胰島素類似物前體的核酸分子，其包含編碼間隔肽的核酸分子和編碼人胰島素類似物前體的核酸分子，能夠與包含信號肽的載體重組後表達人胰島素類似物前體。

【0017】 在一些實施方案中，人胰島素類似物前體核酸分子編碼的人胰島素類似物前體胺基酸序列如下：

EEGEPK-B(1-29)-AAK-A(1-21)

其中，“EEGEPK (SEQ ID NO: 16)”可以是胰島素前體 N 端延伸，稱為間隔肽或前導肽；“B(1-29)”可以是缺失 B30 位蘇胺酸的人胰島素 B 鏈；“A(1-21)”可以是人胰島素 A 鏈胺基酸序列，“AAK”為連接 B 鏈和 A 鏈的連接肽，也稱為 C 肽。

【0018】 在一些實施方案中，人胰島素類似物前體核酸分子的序列可以如 SEQ ID NO:6 所示，胺基酸序列如 SEQ ID NO:7 所示：

GAAGAAGGTGAACCAAAGTTCGTCAACCAGCACTTGTGTGGTTCCCATT
TGGTTGAGGCTCTGTACTTGGTCTGTGGAGAAAGAGGTTTCTTTTACACC
CCAAAGGCTGCTAAGGGTATCGTTGAGCAATGTTGCACCTCTATTTGTTC
CCTGTATCAGTTGGAAAACACTACTGCAACTAA

SEQ ID NO:6

EEGEPKFVNQHLCSHLVEALYLVCGERGFFYTPKAAKGIVEQCCTSICSLY
QLENYCN

SEQ ID NO:7

【0019】 在一些實施方案中，提供另一種能夠表達人胰島素類似物前體的核酸分子，該序列中包含信號肽序列、間隔肽序列和編碼人胰島素類似物前體的序列，其能夠與不包含信號肽的載體重組後表達人胰島素類似物前體。

【0020】 在一些實施方案中，表達人胰島素類似物前體的核酸分子的核酸序列如 SEQ ID NO:8 所示，編碼的胺基酸序列如 SEQ ID NO:9 所示：

ATGAGATTTCCTTCTATTTTCACTGCTGTTTTGTTTGCCGCTTCCTCTGCTT
TGGCAGCTCCAGTTAATAACAACCACTGAAGATGAGACTGCTCAAATCCC
AGCCGAAGCAGTTATTGGTTACTCCGACTTGGAAGGAGATTTTGACGTC
GCTGTTTTACCATTCTCTAATTCCACTAATAACGGTCTGTTGTTTATTAATA
CTACCATTGCTTCTATCGCCGCTAAGGAGGAAGGTGTGTCCCTCGAGAA
AAGAGAAGAAGGTGAACCAAAGTTCGTCAACCAGCACTTGTGTGGTTC
CCATTTGGTTGAGGCTCTGTACTTGGTCTGTGGAGAAAGAGGTTTCTTTT
ACACCCCAAAGGCTGCTAAGGGTATCGTTGAGCAATGTTGCACCTCTATT
TGTTCCCTGTATCAGTTGGAAACTACTGCAACTAA

SEQ ID NO:8

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYS DLEGDFDVAVL
PFSNSTNGLLFINTTIAASIAAKEEGVSLEKREEGEPKFVNQHLCSHLVEAL
YLVCGERGFFYTPKAAKGIVEQCCTSICSLYQLENYCN

SEQ ID NO:9

【0021】 在一些實施方案中，表達人胰島素類似物前體的核酸分子還可以連接酶切位點序列，該酶切位點較佳 EcoR I 酶切位點和 Not I 酶切位點。

【0022】 在一些實施方案中，還提供了一種能夠在真核或原核細胞中表達的載體，其能夠在原核或真核細胞中分泌表達人胰島素類似物前體。

【0023】 在一些實施方案中，還提供了一種宿主細胞，該宿主細胞較佳為酵母菌，更較佳為畢赤酵母，其能夠分泌表達人胰島素類似物前體。

【0024】 在一些實施方案中，還提供一種製備人胰島素類似物的方法，包括使用如前所述的核酸分子，載體，和/或宿主細胞。

【0025】 該方法還可以包含如下步驟：

1) 藉由編碼人胰島素類似物前體的核酸分子在真核細胞中表達人胰島素類似物前體；

2) 藉由酶切處理人胰島素類似物前體得到人胰島素類似物，該編碼人胰島素類似物前體的核酸分子可以如 SEQ ID NO:6 所示，該酶切處理胰島素前體採用本領域技術人員公知的方法。

【0026】 在一些實施方案中，步驟 1) 中包含用表達載體表達人胰島素類似物前體，所述的表達載體包含信號肽序列，該信號肽序列可以如 SEQ ID NO:1 所示。

【0027】 在一些實施方案中，該人胰島素類似物為 B30 缺失的人胰島素，所述的人胰島素類似物進一步經過醯化基團的取代。

【0028】 在一些實施方案中，該 B30 缺失的人胰島素在 B29 位的賴胺酸處被醯化基團取代。

【0029】 較佳的，上述取代的取代產物為賴胺酸 B29 (N^ε-(N^α-十六烷脂
肪二酸-L-賴胺酸-N^ε-氧代丁酰基))des(B30)人胰島素。

【0030】 縮寫和術語

“密碼子優化”是指利用宿主細胞偏愛密碼子規則，使用偏愛密碼子並避免
使用利用率低或稀有密碼子來合成基因。

【0031】 “對照 1”是“EEGEPK”的編碼核酸分子（雙下底線所示的
GAAGAAGGTGAACCAAAG）與專利 WO1998028429 中的胰島素前體基因的
編碼核酸分子相連，如下 SEQ ID NO:10 所示：

GAAGAAGGTGAACCAAAGTTCGTTAACCAACACTTGTGCGGTTCCCACT
TGGTTGAAGCTTTGTACTTGGTTTGCGGTGAAAGAGGTTTCTTCTACACT
CCTAAGGCTGCTAAGGGTATTGTGCAACAATGCTGTACCTCCATCTGCTC
CTTGTACCAATTGGAAAACACTACTGCAACTAA

SEQ ID NO:10

【0032】 “對照 2” 是“EEGEPK”的編碼核酸分子（雙下底線所示）與
Gurramkonda 等人的文獻（Microbial Cell Factories, 2010, 9:31）中經優化的胰
島素前體基因的編碼核酸分子相連，如下 SEQ ID NO:11 所示：

GAAGAAGGTGAACCAAAGTTTGTTAACCAACATTTGTGTGGTTCTCATTT
GGTTGAAGCTTTGTACTTGGTTTGTGGTGAAAGAGGTTTCTTCTACACTC
CAAAGGCTGCTAAGGGTATTGTTGAACAATGTTGTACTTCTATTTGTTCTT
TGTACCAATTGGAAAACACTACTGTA ACTAA

SEQ ID NO: 11

【0033】 “IP-S”是經密碼子優化的胰島素前體基因對應的核酸分子。

【0034】 “ α -因子”是 Invitrogen 公司提供的 pPIC9K 表達載體帶有的 α -因子信號肽基因對應的核酸分子，來源於釀酒酵母。

【0035】 “ α -因子-S”是經密碼子優化的 α -因子信號肽基因對應的核酸分子。

【0036】 “載體”包括核酸分子，其能夠運輸它連接的另一核酸，包括但不局限於質粒和病毒載體。某些載體能夠在其被引入的宿主細胞中自主複製，而其它載體可在引入宿主細胞時整合到宿主細胞的基因組中，並因此與該宿主基因組一起複製。此外，某些載體能夠指導有效連接到它們的基因的表達，此類載體此處稱作“重組表達載體”（或簡單地“表達載體”），並且典型載體為本領域熟知。

【0037】 “目標多肽”是能夠在酵母中表達的多肽，包括但不限於酶、抗體、干擾素、胰島素、白介素等及其變體、前體、中間體，例如可以是胰島素前體。

【0038】 “細胞”、“宿主細胞”可互換使用。

【0039】 “多核苷酸分子”、“核酸分子”可互換使用，其序列可以是 DNA 序列。

【圖式簡單說明】

無。

【實施方式】

【0040】 以下結合實施例用於進一步描述本發明，但並非限制本發明的範圍。

【0041】 本發明實施例中所用的載體、宿主菌和培養基均購自 Invitrogen 公司，其中畢赤酵母表達載體 pPIC9K 含有醇氧化酶 AOX1 啟動子，可受甲醇誘導，該載體還含有 α -因子信號肽序列，能夠分泌表達外源蛋白；畢赤酵母表達載體 pPIC3.5K 含有醇氧化酶 AOX1 啟動子，可受甲醇誘導，該載體不含有 α -因子信號肽序列；宿主菌為畢赤酵母 GS115 菌株；培養基為畢赤酵母手冊提供的培養基配方。

【0042】 實施例 1 構建胰島素前體重組表達載體

將對照 1 (SEQ ID NO:10)、對照 2 (SEQ ID NO:11)、IP-S (SEQ ID NO:6) 的 5'端和 3'端都分別加入 EcoR I 和 Not I 酶切位點，由南京金斯瑞生物科技公司進行合成，並將合成的核酸分子序列連接到 T 載體。

GAAGAAGGTGAACCAAAGTTCGTCAACCAGCACTTGTGTGGTTCCCATT
TGGTTGAGGCTCTGTACTTGGTCTGTGGAGAAAGAGGTTTCTTTTACACC
CCAAAGGCTGCTAAGGGTATCGTTGAGCAATGTTGCACCTCTATTTGTTC
CCTGTATCAGTTGGAAAACACTACTGCAACTAA

SEQ ID NO:6

GAAGAAGGTGAACCAAAGTTCGTTAACCAACACTTGTGCGGTTCCTCACT
TGGTTGAAGCTTTGTACTTGGTTTGCGGTGAAAGAGGTTTCTTCTACACT
CCTAAGGCTGCTAAGGGTATTGTCGAACAATGCTGTACCTCCATCTGCTC
CTTGTACCAATTGGAAAACACTACTGCAACTAA

SEQ ID NO:10

GAAGAAGGTGAACCAAAGTTTGTTAACCAACATTTGTGTGGTTCTCATT
GGTTGAAGCTTTGTACTTGGTTTGTGGTGAAAGAGGTTTCTTCTACACTC
CAAAGGCTGCTAAGGGTATTGTTGAACAATGTTGTACTTCTATTTGTTCTT
TGTACCAATTGGAAAACACTACTGTAACATA

SEQ ID NO:11

【0043】 用內切酶 EcoR I 和 Not I 將上述帶有胰島素前體核酸分子的 T 載體和表達載體 pPIC9K 進行雙酶切，之後採用膠回收試劑盒分別回收目標片段和載體片段，用 T4 連接酶將酶切純化後的目標片段分別與酶切純化後的載體 pPIC9K 進行連接。

【0044】 將上述連接液分別轉化到大腸桿菌 TOP10 感受態細胞，塗布含胺苄青黴素的抗性平板，培養後，挑取純株菌提取質粒，雙酶切驗證，藉由驗證最終得到三種重組表達載體，分別含有對照 1、對照 2、IP-S 的序列。

【0045】 實施例 2 畢赤酵母重組菌表達胰島素前體

將實施例 1 中構建的三種重組表達載體分別轉化到畢赤酵母 GS115 中，表達對照 1 和對照 2 的重組菌為對照菌株，表達 IP-S 的重組菌為實驗菌株。

【0046】 將三種重組菌的菌落接種至 5mL 的 YPD 培養基中，30°C 恆溫搖床 250 rpm 振盪培養至 OD₆₀₀=10 左右（16-18 小時）。收集並重新懸浮細胞於 50mL BMGY 培養基中，30°C 恆溫搖床 250 rpm 振盪培養過夜，至 OD₆₀₀ 為 30 左右。1500rpm 離心 5 分鐘收集細胞，用 25mL BMMY 培養基重新懸浮。在培養基中加入 1/200 體積的甲醇（終濃度為 0.5%），30°C 恆溫搖床 250 rpm 振盪培養 96 小時，其中每 24 小時補加 1/200 體積的甲醇。表達結束後，10,000rpm 離心取上清，藉由 HPLC 檢測上清液中胰島素前體的產量，換算為

相對於對照菌株胰島素前體表達量的百分比，胰島素前體的表達量百分比如表 1 所示。

表 1

菌株	載體	所用信號肽	表達基因	產量百分比(%)
對照菌	pPIC9K	α -因子	對照 1	100
對照菌	pPIC9K	α -因子	對照 2	125
實驗菌	pPIC9K	α -因子	IP-S	225

【0047】 表 1 的數據顯示，經過優化的胰島素前體基因表達的胰島素前體量與兩個對照組相比提高到 1.8 到 2.25 倍，可見經過優化的胰島素前體基因有更好的表達效果，其表達的胰島素前體的產量顯著提高。

【0048】 實施例 3 構建融合了不同 α -因子的胰島素前體基因的重組表達載體

在 IP-S 前分別加入 α -因子 (SEQ ID NO:12) 和 α -因子-S (SEQ ID NO:1)，融合後的核酸分子序列分別如 SEQ ID NO:13 和 SEQ ID NO:8 所示，進行合成，合成後的核酸分子的 5'端和 3'端分別帶有 EcoRI 和 NotI 酶切位點，並將合成的核酸分子連接到 T 載體。

ATGAGATTTCCTTCAATTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCA
TTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTC
CGGCTGAAGCTGTCATCGGTACTCAGATTAGAAAGGGGATTTTCGATGTT
GCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTATTGTTTATAAAT

ACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGA
AAAGA

SEQ ID NO:12

ATGAGATTTCCTTCTATTTTCACTGCTGTTTTGTTTGCCGCTTCCTCTGCTT
TGGCAGCTCCAGTTAATACAACCACTGAAGATGAGACTGCTCAAATCCC
AGCCGAAGCAGTTATTGGTTACTCCGACTTGGAAGGAGATTTTGACGTC
GCTGTTTTACCATTCTCTAATTCCACTAATAACGGTCTGTTGTTTATTAATA
CTACCATTGCTTCTATCGCCGCTAAGGAGGAAGGTGTGTCCCTCGAGAA
AAGA

SEQ ID NO:1

ATGAGATTTCCTTCAATTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCA
TTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAATTC
CGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTTCGATGTT
GCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTATTGTTTATAAAT
ACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGA
AAAGAGAAGAAGGTGAACCAAAGTTCGTCAACCAGCACTTGTGTGGTT
CCCATTGTTGAGGCTCTGTACTTGGTCTGTGGAGAAAGAGGTTTCTTT
TACACCCCAAAGGCTGCTAAGGGTATCGTTGAGCAATGTTGCACCTCTAT
TTGTTCCCTGTATCAGTTGGAAAACACTGCAACTAA

SEQ ID NO:13

ATGAGATTTCCTTCTATTTTCACTGCTGTTTTGTTTGCCGCTTCCTCTGCTT
TGGCAGCTCCAGTTAATACAACCACTGAAGATGAGACTGCTCAAATCCC
AGCCGAAGCAGTTATTGGTTACTCCGACTTGGAAGGAGATTTTGACGTC
GCTGTTTTACCATTCTCTAATTCCACTAATAACGGTCTGTTGTTTATTAATA

CTACCATTGCTTCTATCGCCGCTAAGGAGGAAGGTGTGTCCCTCGAGAA
AAGAGAAGAAGGTGAACCAAAGTTCGTCAACCAGCACTTGTGTGGTTC
CCATTTGGTTGAGGCTCTGTACTTGGTCTGTGGAGAAAGAGGTTTCTTTT
ACACCCCAAAGGCTGCTAAGGGTATCGTTGAGCAATGTTGCACCTCTATT
TGTTCCCTGTATCAGTTGGAAACTACTGCAACTAA

SEQ ID NO:8

【0049】 將對照 1 前分別加入 α -因子 (SEQ ID NO: 12) 和 α -因子-S (SEQ ID NO: 1)，融合後的核酸分子序列分別如 SEQ ID NO:14 和 SEQ ID NO:15 所示，進行合成，合成後的核酸分子 5'端和 3'端分別帶有 EcoRI 和 NotI 酶切位點，並將合成的核酸分子連接到 T 載體。

ATGAGATTTCCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCA
TTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAATTC
CGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTCGATGTT
GCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTATTGTTTATAAAT
ACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGA
AAAGAGAAGAAGGTGAACCAAAGTTCGTTAACCAACACTTGTGCGGTT
CCCACTTGGTTGAAGCTTTGTACTTGGTTTGCGGTGAAAGAGGTTTCTTC
TACACTCCTAAGGCTGCTAAGGGTATTGTCGAACAATGCTGTACCTCCAT
CTGCTCCTTGTACCAATTGGAAACTACTGCAACTAA

SEQ ID NO:14

ATGAGATTTCCTTCTATTTTCACTGCTGTTTTGTTTGCCGCTTCCTCTGCTT
TGGCAGCTCCAGTTAATAACAACCACTGAAGATGAGACTGCTCAAATCCC
AGCCGAAGCAGTTATTGGTTACTCCGACTTGGAAGGAGATTTTGACGTC

GCTGTTTTACCATTCTCTAATTCCACTAATAACGGTCTGTTGTTTATTAATA
CTACCATTGCTTCTATCGCCGCTAAGGAGGAAGGTGTGTCCCTCGAGAA
AAGAGAAGAAGGTGAACCAAAGTTCGTTAACCAACACTTGTGCGGTTC
CCACTTGGTTGAAGCTTTGTACTTGGTTTGCGGTGAAAGAGGTTTCTTCT
ACACTCCTAAGGCTGCTAAGGGTATTGTCTGAACAATGCTGTACCTCCATC
TGCTCCTTGTACCAATTGGAAACTACTGCAACTAA

SEQ ID NO:15

【0050】 用內切酶EcoR I和Not I將上述 T 載體和表達載體 pPIC3.5K 進行雙酶切，之後採用膠回收試劑盒分別回收目標片段和載體片段，用 T4 連接酶將酶切純化後的目標片段分別與酶切純化後的載體 pPIC3.5K 進行連接。

【0051】 將上述連接液分別轉化到大腸桿菌 TOP10 感受態細胞，塗布含胺苄青黴素的抗性平板，培養後，挑取純株菌提取質粒，雙酶切驗證，藉由驗證最終得到四種重組表達載體，分別融合了 α -因子信號肽和 α -因子-S 信號肽，表達不同核苷酸序列的胰島素前體基因。

【0052】 實施例 4 畢赤酵母重組菌表達優化前後的胰島素前體

將實施例 3 中構建的重組表達載體分別轉化到畢赤酵母 GS115 中。

【0053】 將重組菌菌落接種至 5mL 的 YPD 培養基中，30℃恆溫搖床 250 rpm 振盪培養至 OD₆₀₀=10 左右（16-18 小時）。收集並重新懸浮細胞於 50mL BMGY 培養基中，30℃恆溫搖床 250 rpm 振盪培養過夜，至 OD₆₀₀ 為 30 左右。1500rpm 離心 5 分鐘收集細胞，用 25mL BMMY 培養基重新懸浮。在培養基中加入 1/200 體積的甲醇（終濃度為 0.5%），30℃恆溫搖床 250 rpm 振盪培養 96

小時，其中每 24 小時補加 1/200 體積的甲醇。表達結束後，10,000rpm 離心取上清，藉由 HPLC 檢測上清液中胰島素前體的產量。

【0054】 將表達融合 α -因子的對照 1 基因的重組菌作為對照菌，將其他菌株的胰島素前體產量換算為相對於對照菌株產量的百分比，如表 2 所示。

表 2

菌株	載體	所用信號肽	表達基因	產量百分比(%)
對照菌	pPIC3.5K	α -因子	對照 1	100
實驗菌	pPIC3.5K	α -因子-S	對照 1	150
實驗菌	pPIC3.5K	α -因子	IP-S	225
實驗菌	pPIC3.5K	α -因子-S	IP-S	275

【0055】 表 2 的數據顯示，僅信號肽優化的核酸分子序列表達胰島素前體的量提高到 1.5 倍，僅胰島素前體基因優化的核酸分子序列表達胰島素前體的量提高到 2.25 倍，而經信號肽和胰島素前體基因同時優化的核酸分子序列表達胰島素前體量提高到 2.75 倍。綜合以上，密碼子優化使得胰島素前體的表達量增加到 1.5 至 2.75 倍。

【符號說明】

無。

【 序列表 】

<110> 江蘇恆瑞醫藥股份有限公司
<120> 一種密碼子優化的人胰島素類似物前體基因和信號肽基因
<130> 719003CPCT
<150> CN201810131794.3
<151> 2018-02-09
<160> 16
<170> SIPOSequenceListing 1.0
<210> 1
<211> 255
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>
<221> 基因
<222> (1)..(255)
<223> 編碼信號肽的核酸分子

<400> 1
atgagatttc cttctat ttt cactgctgtt ttgtttgccg cttcctctgc tttggcagct 60
ccagttaata caaccactga agatgagact gctcaaatcc cagccgaagc agttattggt 120
tactccgact tggaaggaga ttttgacgtc gctgttttac cattctctaa ttccactaat 180
aacgggtctgt tgtttattaa tactaccatt gcttctatcg ccgctaagga ggaagggtgtg 240
tcctctcgaga aaaga 255

<210> 2
<211> 85
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<221> 肽
<222> (1)..(85)
<223> 信號肽

<400> 2
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
65 70 75 80
Ser Leu Glu Lys Arg
85

<210> 3

<211> 162
<212> DNA
<213> 人(Homo sapiens)

<220>
<221> 基因
<222> (1)..(162)
<223> 編碼人胰島素類似物前體的核酸分子

<400> 3
ttcgtcaacc agcacttggtg tggttcccat ttggttgagg ctctgtactt ggtctgtgga 60
gaaagagggtt tcttttacac cccaaaggct gctaagggtg tcgttgagca atgttgaccc 120
tctatttgtt cctgtatca gttggaaaac tactgcaact aa 162

<210> 4
<211> 53
<212> PRT
<213> 人(Homo sapiens)

<220>
<221> 肽
<222> (1)..(53)
<223> 人胰島素類似物前體

<400> 4
Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15
Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys
20 25 30
Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
35 40 45
Glu Asn Tyr Cys Asn
50

<210> 5
<211> 18
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>
<221> 基因
<222> (1)..(18)
<223> 編碼間隔肽的核酸分子

<400> 5
gaagaagggtg aaccaaag 18

<210> 6
<211> 180
<212> DNA
<213> 人(Homo sapiens)

<400> 6
gaagaagggtg aaccaaagtt cgtcaaccag cacttggtgtg gttcccattt ggttgaggct 60
ctgtacttgg tctgtggaga aagaggtttc tttaacccc caaaggctgc taagggtatc 120
gttgagcaat gttgcacctc tatttgttcc ctgtatcagt tggaaaacta ctgcaactaa 180

<210> 7
<211> 59
<212> PRT
<213> 人(Homo sapiens)

<400> 7
Glu Glu Gly Glu Pro Lys Phe Val Asn Gln His Leu Cys Gly Ser His
1 5 10 15
Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Tyr
20 25 30
Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile
35 40 45
Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
50 55

<210> 8
<211> 435
<212> DNA
<213> 人(Homo sapiens)

<400> 8
atgagatttc cttctatttt cactgetggt ttgtttgceg ctctctctgc ttggcagct 60
ccagttaata caaccactga agatgagact gctcaaatcc cagccgaagc agttattggt 120
tactccgact tggaaggaga tttgacgtc gctgttttac cattctctaa ttccactaat 180
aacgggtctgt tgtttattaa tactaccatt gcttctatcg ccgctaagga ggaagggtgtg 240
tccctcgaga aaagagaaga aggtgaacca aagtctgtca accagcactt gtgtggttcc 300
catttggttg aggtctctgta cttggtctgt ggagaaagag gttctcttta caccctaaag 360
gctgctaagg gtatcgttga gcaatgttgc acctctattt gttccctgta tcagttggaa 420
aactactgca actaa 435

<210> 9
<211> 144
<212> PRT
<213> 人(Homo sapiens)

<400> 9
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15
Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45
Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
65 70 75 80
Ser Leu Glu Lys Arg Glu Glu Gly Glu Pro Lys Phe Val Asn Gln His
85 90 95
Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu
100 105 110
Arg Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln
115 120 125
Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
130 135 140

<210> 10
<211> 180
<212> DNA

<213> 人工序列(Artificial Sequence)		
<220>		
<221> misc_feature		
<222> (1)..(18)		
<223> 編碼間隔肽的核酸分子		
<400> 10		
gaagaagg	tg aaccaaag	tt cgttaac
caacaa	cacttg	tgcg gttcc
cactt	gactt	ggttga
agct	ttgtact	tg tttgc
ggtga	aagaggt	ttc ttctac
ctc	ctaagg	ctgc taagg
gtatt	gtcgaac	aat gctgt
acctc	catctg	ctcc ttgtac
caat	tggaaa	acta ctgca
actaa		180
<210> 11		
<211> 180		
<212> DNA		
<213> 人工序列(Artificial Sequence)		
<220>		
<221> misc_feature		
<222> (1)..(18)		
<223> 編碼間隔肽的核酸分子		
<400> 11		
gaagaagg	tg aaccaaag	tt tgtaac
caacaa	catttg	tggt gttct
cattt	gactt	ggttga
agct	ttgtact	tg tttgc
ggtga	aagaggt	ttc ttctac
ctc	caaagg	ctgc taagg
gtatt	ggtgaac	aat gttgt
acttc	tatttgt	ctc ttgtac
caat	tggaaa	acta ctgta
actaa		180
<210> 12		
<211> 255		
<212> DNA		
<213> 人工序列(Artificial Sequence)		
<220>		
<221> 基因		
<222> (1)..(255)		
<223> 編碼 α -因子的核酸		
<400> 12		
atgagatt	ctc ottca	atttt tactgc
cagtt	ttattgc	cag catcct
cgc	attagct	gct 60
ccagtc	aaac	ctacaac
aga	agatgaa	acg gcacaa
attc	cggtga	agc tgtcat
cggt	tactcag	att taga
agggga	tttcgat	gtt gctgt
tttgc	cattttc	caa cagc
caaat	aacggg	tat tgttt
ataaaa	tactact	att gccag
cattg	ctgctaa	aga aga
aggggt	tctctcg	aga aaaga
aaaga		255
<210> 13		
<211> 435		
<212> DNA		
<213> 人工序列(Artificial Sequence)		
<220>		
<221> 基因		
<222> (1)..(435)		
<223> 編碼 α -因子的核酸+編碼胰島素類似物前體的核酸		
<400> 13		

atgagatttc cttcaat ttt tactgcagtt ttattgcag catctccgc attagctgct 60
ccagtcaca ctacaacaga agatgaaacg gcacaaatc cggctgaagc tgcacgcgt 120
tactcagatt tagaagggga ttccgatgtt gctgtttgc cttttccaa cagcacaat 180
aacgggttat tgtttataaa tactactatt gccagcattg ctgctaaaga agaaggggta 240
tctctcgaga aaagagaaga aggtgaacca aagttcgta accagcactt gtgtgggtcc 300
catttggttg aggcctgtgta cttgggtgtt ggagaaagag gtttcttta caccctaaag 360
gctgctaagg gtatcggtga gcaatgttg accctctatt gtccctgta tcagttggaa 420
aactactgca actaa 435

<210> 14
<211> 435
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>
<221> 基因
<222> (1)..(435)
<223> 編碼 α -因子的核酸+編碼胰島素類似物前體的核酸

<400> 14
atgagatttc cttcaat ttt tactgcagtt ttattgcag catctccgc attagctgct 60
ccagtcaca ctacaacaga agatgaaacg gcacaaatc cggctgaagc tgcacgcgt 120
tactcagatt tagaagggga ttccgatgtt gctgtttgc cttttccaa cagcacaat 180
aacgggttat tgtttataaa tactactatt gccagcattg ctgctaaaga agaaggggta 240
tctctcgaga aaagagaaga aggtgaacca aagttcgta accaactt gtgcgggtcc 300
catttggttg aagcttgta cttgggttg ggtgaaagag gtttcttta cactcctaag 360
gctgctaagg gtattgtcga acaatgtgt accctcatct gtccttgta ccaattggaa 420
aactactgca actaa 435

<210> 15
<211> 435
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>
<221> 基因
<222> (1)..(435)
<223> 編碼 α -因子的核酸+編碼胰島素類似物前體的核酸

<400> 15
atgagatttc cttctat ttt cactgctgtt ttgtttgcc cttctctgc tttggcagct 60
ccagtttaata caaccactga agatgagact gctcaaatcc cagccgaagc agttattggt 120
tactccgact tggaaggaga tttgacgtc gctgttttac cttctctaa ttccactaat 180
aacgggtctgt tgtttataaa tactaccatt gcttctatcg ccgctaagga ggaaggtgtg 240
tccctcgaga aaagagaaga aggtgaacca aagttcgta accaactt gtgcgggtcc 300
catttggttg aagcttgta cttgggttg ggtgaaagag gtttcttta cactcctaag 360
gctgctaagg gtattgtcga acaatgtgt accctcatct gtccttgta ccaattggaa 420
aactactgca actaa 435

<210> 16
<211> 6
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>
<221> 肽
<222> (1)..(6)

<223> 間隔肽

<400> 16
Glu Glu Gly Glu Pro Lys
1 5

【發明摘要】

【中文發明名稱】 一種密碼子優化的人胰島素類似物前體基因和信號肽基因

【英文發明名稱】 A CODON OPTIMIZED HUMAN INSULIN ANALOGUE

PRECURSOR GENE AND A SIGNAL PEPTIDE GENE

【中文】

本發明涉及一種密碼子優化的人胰島素類似物前體基因和信號肽基因。具體而言，本發明涉及一種密碼子優化的用於酵母表達胰島素類似物前體的核酸分子，包含一種胰島素類似物前體核酸分子和一種酵母分泌信號肽 α -因子核酸分子的融合。該經優化的核酸分子使得胰島素類似物前體在畢赤酵母中的表達量大幅提高，降低人胰島素類似物的生產成本，進而降低治療糖尿病藥物的成本。

【英文】

The present invention relates to a codon optimized human insulin analogue precursor gene and a signal peptide gene. Specifically, the present invention relates to a codon optimized nucleic acid molecule for yeast expression of insulin analogue precursors, comprising fusion of an insulin analogue precursor nucleic acid molecule and a yeast secretion signal peptide alpha-factor nucleic acid molecule. The optimized nucleic acid molecule greatly increases the expression of insulin analogue precursors in *Pichia pastoris*, reduces the production cost of human insulin analogues, and thereby reduces the cost of diabetic drugs.

【指定代表圖】 無

【代表圖之符號簡單說明】 無

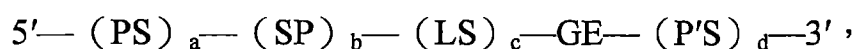
本案無圖式。

【特徵化學式】

無。

【發明申請專利範圍】

【第1項】一種核酸分子，其包含具有以下通式的分子或結構：



其中，

PS 是編碼加工位點的核酸分子，a 是 0 或 1；

SP 是編碼信號肽的核酸分子，b 是 1；

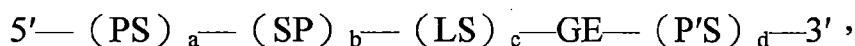
LS 是編碼間隔肽的核酸分子，c 是 0 或 1；

GE 是編碼目標多肽的核酸分子；

P'S 是編碼加工位點的核酸分子，d 是 0 或 1；

該編碼信號肽的核酸分子包含如 SEQ ID NO:1 所示序列。

【第2項】一種核酸分子，其包含具有以下通式的分子或結構：



其中，

PS 是編碼加工位點的核酸分子，a 是 0 或 1；

SP 是編碼信號肽的核酸分子，b 是 0 或 1；

LS 是編碼間隔肽的核酸分子，c 是 0 或 1；

GE 是編碼人胰島素類似物前體多肽的核酸分子；

P'S 是編碼加工位點的核酸分子，d 是 0 或 1；

該編碼人胰島素類似物前體多肽的核酸分子包含如 SEQ ID NO:3 所示序列。

【第3項】如申請專利範圍第1項所述的核酸分子，其中，該目標多肽是人胰島素類似物前體，編碼該人胰島素類似物前體的核酸分子包含編碼SEQ ID NO:4 所示胺基酸序列的核酸分子。

【第4項】 如申請專利範圍第3項所述的核酸分子，其中，編碼該人胰島素類似物前體的核酸分子包含如SEQ ID NO:3所示核酸序列。

【第5項】 如申請專利範圍第2項所述的核酸分子，其中編碼該信號肽的核酸分子包含編碼SEQ ID NO:2所示胺基酸序列的核酸分子。

【第6項】 如申請專利範圍第5項所述的核酸分子，其中，編碼該信號肽的核酸分子包含如SEQ ID NO：1或SEQ ID NO:12所示核酸序列。

【第7項】 如申請專利範圍第2至6項中任一項所述的核酸分子，其中，該編碼人胰島素類似物前體的核酸分子在相當於SEQ ID NO:3的第88-96位被替換。

【第8項】 如申請專利範圍第7項所述的核酸分子，其中，該編碼人胰島素類似物前體的核酸分子在相當於SEQ ID NO:3的第88至96位被被GCCGCTAAG、GCTGCCAAG、GCTGCTAAA或GCCGCCAAG替換。

【第9項】 如申請專利範圍第1至8項中任一項所述的核酸分子，其中，該間隔肽的胺基酸序列包含EEGEPK（Glu-Glu-Gly-Glu-Pro-Lys）。

【第10項】 如申請專利範圍第9項所述的核酸分子，其中，該編碼間隔肽的核酸分子包含如SEQ ID NO:5所示序列。

【第11項】 如申請專利範圍第1至10項中任一項所述的核酸分子，其包含如SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:13、SEQ ID NO:15所示序列中的任意一個；或為SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:13、SEQ ID NO:15所示序列中的任意一個。

【第12項】 如申請專利範圍第1至11項中任一項所述的核酸分子，其中，該加工位點是連接酶切位點。

【第13項】 如申請專利範圍第12項所述的核酸分子，其中，PS是編碼EcoRI酶切位點的核酸分子和/或P'S是編碼Not I酶切位點的核酸分子。

【第14項】 一種載體，其包含如申請專利範圍第1至13項中任一項所述的核酸分子。

【第15項】 如申請專利範圍第14項所述的載體，其是真核細胞表達載體或原核細胞表達載體。

【第16項】 一種宿主細胞，其含有申請專利範圍第1至13項中任一項任一項所述的核酸分子和/申請專利範圍第14或15項所述的載體。

【第17項】 如申請專利範圍第16項所述的宿主細胞，其為酵母菌。

【第18項】 如申請專利範圍第17項所述的宿主細胞，其為畢赤酵母。

【第19項】 一種製備人胰島素類似物的方法，包括使用申請專利範圍第1至13項中任一項所述的核酸分子，申請專利範圍第14或15項所述的載體，和/或申請專利範圍第16至18項中任一項所述的宿主細胞。

【第20項】 如申請專利範圍第19項所述的方法，其還包括如下步驟：

- 1) 表達人胰島素類似物前體；
- 2) 酶切步驟 1) 獲得的人胰島素類似物前體，得到人胰島素類似物。

【第21項】 如申請專利範圍第19或20項所述的方法，該人胰島素類似物為B30缺失的人胰島素，和/或進一步經過醯化基團的取代。

【第22項】 如申請專利範圍第21項所述的方法，其中，該取代是B29位的賴胺酸被取代。

【第23項】 如申請專利範圍第22項所述的方法，其中，該取代是取代的產物為賴胺酸B29 (N^{ϵ} -(N^{α} -十六烷脂肪二酸-L-賴胺酸- N^{ϵ} -氧代丁醯基))des(B30)人胰島素。