I780847

(19)中華民國智慧財產局

(12)發明說明書公告本 **(11)證書號數：TW I780847 B**

(45)公告日： 中華民國 111 (2022) 年 10 月 11 日

(21)申請案號：110128495 (22)申請日： 中華民國 104 (2015) 年 10 月 22 日

(51)Int. Cl.： *C12N15/85 (2006.01)* *C12N15/90 (2006.01)*
C12N5/10 (2006.01)

(30)優先權：2014/10/23 美國 62/067,774

(71)申請人：美商再生元醫藥公司 (美國) REGENERON PHARMACEUTICALS, INC. (US)
美國

(72)發明人：沈盈 SHEN, YING (CN)；布拉科夫 達里亞 BURAKOV, DARYA (US)；陳 剛 CHEN, GANG (US)；凡德爾 詹姆斯 FANDL, JAMES P. (US)

(74)代理人：何愛文；王仁君

(56)參考文獻：

US 2013/0130372A1

期刊 Lattenmayer C et al. "Identification of transgene integration loci of different highly expressing recombinant CHO cell lines by FISH" Cytotechnology vol.51 Springer 2006/11/15 p.171-182

期刊 Xu X et al. "The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line" Nature biotechnology vol.29 no.8 Springer Nature Limited 2011/07/31 p.735-741

審查人員：施雅儀

申請專利範圍項數：21 項 圖式數：4 共 61 頁

(54)名稱

新穎之ＣＨＯ整合位點及其用途

(57)摘要

本發明提供用於真核表現系統的表現增高核苷酸序列，其容許重組型蛋白質在真核細胞中的增高且穩定的表現。本發明提供在真核細胞中用以表現感興趣基因之提供增高表現之基因整合位點及其使用方法。本發明提供在真核細胞中用於基因之增高與穩定表現的染色體基因座、序列及載體。

Expression-enhancing nucleotide sequences for eukaryotic expressions systems are provided that allow for enhanced and stable expression of recombinant proteins in eukaryotic cells. Genomic integration sites providing enhanced expression and methods of use thereof are provided for expression of a gene of interest in a eukaryotic cell. Chromosomal loci, sequences, and vectors are provided for enhanced and stable expression of genes in eukaryotic cells.

指定代表圖：

圖1A

圖1B

公告本

I780847

發明摘要

【發明名稱】（中文/英文）

新穎之CHO整合位點及其用途

NOVEL CHO INTEGRATION SITES AND USES THEREOF

【中文】

本發明提供用於真核表現系統的表現增高核苷酸序列，其容許重組型蛋白質在真核細胞中的增高且穩定的表現。本發明提供在真核細胞中用以表現感興趣基因之提供增高表現之基因整合位點及其使用方法。本發明提供在真核細胞中用於基因之增高與穩定表現的染色體基因座、序列及載體。

【英文】

Expression-enhancing nucleotide sequences for eukaryotic expressions systems are provided that allow for enhanced and stable expression of recombinant proteins in eukaryotic cells. Genomic integration sites providing enhanced expression and methods of use thereof are provided for expression of a gene of interest in a eukaryotic cell. Chromosomal loci, sequences, and vectors are provided for enhanced and stable expression of genes in eukaryotic cells.

【代表圖】

【本案指定代表圖】：第（ 1 ）圖。

【本代表圖之符號簡單說明】：

無

【本案若有化學式時，請揭示最能顯示發明特徵的化學式】：

無

發明專利說明書

（本說明書格式、順序，請勿任意更動）

【發明名稱】（中文/英文）

新穎之CHO整合位點及其用途

NOVEL CHO INTEGRATION SITES AND USES THEREOF

【技術領域】

【0001】 本發明在真核細胞中提供重組型蛋白質的穩定整合及/或表現。具體而言，本發明包括用於在真核細胞(尤其是中國倉鼠(灰倉鼠(*Cricetulus griseus*))細胞株)中，透過採用表現增高核苷酸序列來增進蛋白質表現的方法及組成物。本發明包括促進重組媒介之匣式交換(recombination-mediated cassette exchange, RMCE)的多核苷酸及經修飾細胞。於中國倉鼠細胞基因組中，本發明方法將外源性核酸併入特定染色體基因座，以促進由經修飾細胞提高且穩定表現重組型蛋白質。

【先前技術】

【0002】 細胞表現系統旨在為製造特定蛋白質提供可靠且有效率的來源，不論是供研究或治療使用。因為(例如)哺乳動物表現系統能夠適當地轉譯後修飾重組型蛋白質，在哺乳動物細胞中重組型蛋白質表現對於製造治療性蛋白質來說是一個偏好方法。

【0003】 數種細胞系統可用於表現蛋白質，各細胞系統含有各種順式(且在一些情況下含有反式)調節要素的組合，以在短培育時間內達到大量重組型蛋白質。儘管有許多系統可供使用，但對於重組型蛋白質的表現來說，仍存在有效率基因轉移與併入基因之穩定性的挑戰。多個局部遺傳因子將不僅決定感興趣的目標基因何時被表現，也決定了細胞能否在功能上驅動基因轉錄趨向多產，或表現能否長期持續。特定基因內或比鄰處在技

藝中已鑑定出染色體整合位點，例如中國倉鼠卵巢細胞(CHO)整合位點及基因座控制區(WO2012/138887A1；Li, Q. et al., 2002 *Blood*. 100:3077-3086)。因此，在編碼內源性蛋白質的區域中通常鑑定出靶定調節區域。但是，就長期表現目標轉基因來說，主要考量因素為細胞基因破壞最小以避免細胞株的表現型改變。

【0004】 將穩定細胞株改造成容納表現用的額外基因，諸如額外抗體鏈(呈多特異性抗體)尤其有挑戰性。整合基因的表現位準變化多。整合額外基因可能會因為局部基因環境(亦即位置效應)而造成表現變異較大與不穩定性。因此，技藝中對於改良的哺乳動物表現系統有需要。

【發明內容】

相關申請案的交叉引用

【0005】 本申請案請求於2014年10月23日提申之美國臨時申請案第62/067,774號的優先權利益，其整體內容以引用的方式併入本文中。

藉由引用併入序列表

【0006】 呈ASCII文字檔之序列表，具有28 KB且命名為32353_T0045US01_ SequenceListing.txt，是在2015年10月20日所做出並且經由EFS-Web呈交至美國專利商標局，其以引用的方式併入本文。

【0007】 在一個態樣中，本發明提供包含併入基因座內特定位點處之外源性核酸序列的細胞，其中該基因座包含與SEQ ID NO:1或SEQ ID NO: 4至少90%相同的核苷酸序列。在一些具體例中，該基因座包含與SEQ ID NO:1至少90%相同的核苷酸序列。在一些具體例中，該基因座包含與SEQ ID NO:4至少90%相同的核苷酸序列。

【0008】 在另一個態樣中，本發明提供包含併入第二核酸序列內特定位點(例如本發明基因座)中的第一核酸序列的多核苷酸。在一個具體例

中，該第二核酸序列包含SEQ ID NO:1的核苷酸序列。在另一個具體例中，該第二核酸序列包含SEQ ID NO:4的核苷酸序列。

【0009】 在一個具體例中，該第二核酸序列是表現增高序列，選自與SEQ ID NO:1或其表現增高片段具有至少90%核酸同一性(identity)的核苷酸序列。在一個具體例中，該第二核酸序列為表現增高序列，選自與SEQ ID NO:4或其表現增高片段具有至少90%核酸同一性的核苷酸序列。在另一個具體例中，該表現增高序列能夠增高由外源性核酸序列編碼之蛋白質的表現。在另一個具體例中，相較依據隨機整合至基因組一般所觀察到的表現，該表現增高序列能夠增高由外源性核酸序列編碼之蛋白質的表現至少約1.5倍至至少約3倍表現增高。

【0010】 在另一個具體例中，該外源性核酸序列併入SEQ ID NO:1或SEQ ID NO:4內任一位置處的特定位點中。

【0011】 在一些具體例中，SEQ ID NO:1內某個位置處或鄰接SEQ ID NO:1內某個位置的特定位點是選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1的10-4,000；100-3,900；200-3,800；300-3,700；400-3,600；500-3,500；600-3,400；700-3,300；800-3,200；900-3,100；1,000-3,000；1,100-2,900；1,200-2,800；1,300-2,700；1,200-2,600；1,300-2,500；1,400-2,400；1,500-2,300；1,600-2,200；1,700-2100；1,800-2050；1850-2050；1,900-2040；1950-2,025、1990-2021、2002-2021及2,010-2,015。在某些具體例中，SEQ ID NO:1內某個位置處或鄰接SEQ ID NO:1內某個位置的特定位點是選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1的1990-1991、1991-1992、1992-1993、1993-1994、1995-1996、1996-1997、1997-1998、1999-2000、2001-2002、2002-2003、2003-2004、2004-2005、2005-2006、2006-2007、2007-2008、2008-2009、2009-2010、2010-2011、

2011-2012、2012-2013、2013-2014、2014-2015、2015-2016、2016-2017、2017-2018、2018-2019、2019-2020，及2020-2021。

【0012】 在另一個具體例中，SEQ ID NO:1內某個位置處或鄰接SEQ ID NO:1內某個位置的特定位點是選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1的10-500；500-1,000；500-2,100；1,000-1,500；1,000-2,100；1,500-2,000；1,500-2,500；2,000-2,500；2,500-3,000；2,500-3,500；3,000-3,500；3,000-4,000；及3,500-4,000。在某些具體例中，該外源性核酸序列併入至上述特定位點之任一或多者處、其內或鄰近處。

【0013】 在另一個具體例中，該外源性核酸序列包含位於如上述表現增高序列內的辨識位點，前提是該表現增高序列包含與SEQ ID NO:1或SEQ ID NO:4之表現增高序列，或其表現增高片段至少約90%相同、至少約91%相同、至少約92%相同、至少約93%相同、至少約94%相同、至少約95%相同、至少約96%相同、至少約97%相同、至少約98%相同，或至少約99%相同的序列。

【0014】 在一個具體例中，該外源性核酸序列包含一個重組酶辨識位點。在一些具體例中，該外源性核酸序列進一步包含至少一個重組酶辨識位點，該重組酶辨識位點包含獨立地選自*Lox*P位點、*Lox*511位點、*Lox*2272位點、*Lox*2372、*Lox*5171、*Lox*m2、*Lox*71、*Lox*66、*Lox*Fas及frt位點的序列。在一個具體例中，該重組酶辨識位點併入該表現增高序列內。在另一個具體例中，該整合酶辨識位點以5'方向緊鄰接基因匣5'端的末端核苷酸，或以3'方向緊鄰接基因匣3'端的末端核苷酸。在一些具體例中，該至少一個重組酶辨識位點與基因匣併入該表現增高序列內。

【0015】 在一個具體例中，至少兩個重組酶辨識位點存在於該表現增高序列內。在另一個具體例中，具有相反位向的兩個重組酶辨識位點

併入該表現增高序列內。在另一個具體例中，三個重組酶辨識位點與併入該表現增高序列內。

【0016】 在一個態樣中，提供經單離的中國倉鼠卵巢(CHO)細胞，其包含經改造之SEQ ID NO:1的表現增高序列或其表現增高片段。在一個具體例中，該表現增高序列包含SEQ ID NO:1或SEQ ID NO:4的核苷酸序列或其穩定變體，經改造以併入如上所述的外源性核酸序列。在其他具體例中，本發明提供一種包含外源性核酸序列的經單離CHO細胞，該外源性核酸序列被插入至含有SEQ ID NO:1或SEQ ID NO:4或其穩定變體之表現增高序列的基因座中。

【0017】 在一個具體例中，該CHO細胞在該表現增高序列內進一步包含至少一個重組酶辨識序列。在另一個具體例中，該至少一個重組酶辨識序列獨立地選自*Lox*P位點、*Lox*511位點、*Lox*2272位點、*Lox*2372、*Lox*5171、*Lox*m2、*Lox*71、*Lox*66、*Lox*Fas及frt位點。在另一個具體例中，該重組酶辨識位點以5'方向緊鄰接基因匣5'端的末端核苷酸，或以3'方向緊鄰接基因匣3'端的末端核苷酸。在一些具體例中，該至少一個重組酶辨識位點與基因匣併入本文所述之該CHO細胞基因組的表現增高序列內。

【0018】 在另一個具體例中，該至少一個重組辨識位點如上所述定位，前提為該基因匣包含一個表現增高序列，該序列與SEQ ID NO:1之核苷酸1001至2001 (SEQ ID NO:2)或其表現增高片段含有至少90%同一性、至少約91%同一性、至少約92%同一性、至少約93%同一性、至少約94%同一性、至少約95%同一性、至少約96%同一性、至少約97%同一性、至少約98%同一性，或至少約99%同一性。在另一個具體例中，該至少一個重組辨識位點如上所述定位，前提為該基因匣包含一個表現增高序列，該序列與SEQ ID NO:1之核苷酸2022至3022 (SEQ ID NO:3)或其表現增高片段含有至少90%

同一性、至少約91%同一性、至少約92%同一性、至少約93%同一性、至少約94%同一性、至少約95%同一性、至少約96%同一性、至少約97%同一性、至少約98%同一性，或至少約99%同一性。

【0019】 在又一個具體例中，該至少一個重組酶辨識位點於SEQ ID NO:1之下列核苷酸處或其內插入CHO細胞基因組中：1990-1991、1991-1992、1992-1993、1993-1994、1995-1996、1996-1997、1997-1998、1999-2000、2001-2002、2002-2003、2003-2004、2004-2005、2005-2006、2006-2007、2007-2008、2008-2009、2009-2010、2010-2011、2011-2012、2012-2013、2013-2014、2014-2015、2015-2016、2016-2017、2017-2018、2018-2019、2019-2020、2020-2021或2021-2022。

【0020】 在另一個具體例中，該外源性核酸於SEQ ID NO:1之下列核苷酸處或其內插入CHO基因組中：1990-1991、1991-1992、1992-1993、1993-1994、1995-1996、1996-1997、1997-1998、1999-2000、2001-2002、2002-2003、2003-2004、2004-2005、2005-2006、2006-2007、2007-2008、2008-2009、2009-2010、2010-2011、2011-2012、2012-2013、2013-2014、2014-2015、2015-2016、2016-2017、2017-2018、2018-2019、2019-2020、2020-2021或2021-2022。

【0021】 在另一個具體例中，該外源性核酸於SEQ ID NO:1之核苷酸2001-2022處或其內插入CHO基因組中。在一些具體例中，該外源性核酸於SEQ ID NO:1之核苷酸2001-2002或核苷酸2021-2022處或其內插入，並且因為這個插入而刪除SEQ ID NO:1之核苷酸2002-2021。同樣地，該外源性核酸於SEQ ID NO:4之核苷酸9302-9321處或其內插入CHO基因組中。在一些具體例中，該外源性核酸於SEQ ID NO:4之核苷酸9301-9302或核苷酸9321-9322處或其內插入，並且因為這個插入而刪除SEQ ID NO:4之核苷酸

9302-9321。

【0022】 在一些具體例中，該外源性核酸序列併入諸如SEQ ID NO:1或SEQ ID NO:4的核苷酸序列之基因座內的一個特定位點處，其包含感興趣的基因(GOI)(例如編碼感興趣蛋白質或「POI」的核苷酸序列)。在某些具體例中，該外源性核酸序列包含一或多個感興趣的基因。在一些具體例中，該一或多個感興趣基因是選自由第一GOI、第二GOI及第三GOI組成之群。

【0023】 在一些具體例中，該外源性核酸序列併入諸如SEQ ID NO:1或SEQ ID NO:4的核苷酸序列之基因座內的一個特定位點處，其包含GOI及至少一個重組酶辨識位點。在一個具體例中，第一GOI如上述插入SEQ ID NO:1或SEQ ID NO:4的表現增高序列，或與SEQ ID NO:1或SEQ ID NO:4具有至少90%核苷酸同一性的表現增高序列，或其表現增高片段內，且該第一GOI視情況可操作地連結至一啟動子，其中該啟動子連結GOI (或該GOI)的5'側為第一重組酶辨識位點而3'側為第二重組酶辨識位點。在另一個具體例中，第二GOI插入該第二重組酶辨識位點的3'，且該第二GOI的3'側為第三重組重組酶辨識位點。

【0024】 在又另一個具體例中，該GOI可操作地連結至能夠驅動GOI表現的啟動子，其中該啟動子包含受到活化因子或抑制因子調節的真核啟動子。在其他具體例中，該真核啟動子可操作地連結至原核操縱子，且該真核細胞視情況進一步包含原核抑制子蛋白。

【0025】 在另一個具體例中，在第一與第二及/或第二與第三重組酶辨識位點之間可納入一或多個可篩選標記。在一些具體例中，該第一及/或第二感興趣基因及/或一或多個可篩選標記可操作地連結至啟動子，其中該啟動子可相同或不同。在另一個具體例中，該啟動子包含真核啟動子(諸

如，例如CMV啟動子或SV40晚期啟動子)，視情況受到原核操縱子(例如tet操縱子)所控制。在其他具體例中，該細胞進一步包含編碼原核抑制子(諸如例如tet抑制子)的基因。

【0026】 在另一個具體例中，該細胞進一步包含能夠表現重組酶的基因。在一些具體例中，該重組酶為Cre重組酶。

【0027】 在一個態樣中，提供一種CHO宿主細胞，其包含選自下列的表現增高序列：SEQ ID NO:1或SEQ ID NO:4，或與SEQ ID NO:1或SEQ ID NO:4具有至少90%核苷酸同一性的表現增高序列，或其表現增高片段，該序列包含第一重組酶辨識位點，接著是第一真核啟動子、第一可篩選標記基因、第二真核啟動子、第二可篩選標記基因，以及第二重組酶辨識位點。在更多具體例中，該CHO宿主細胞進一步提供第三真核啟動子、第三標記基因，以及第三重組酶辨識位點。在一個具體例中，該表現增高序列是在如上述SEQ ID NO:1或SEQ ID NO:4內。

【0028】 在一個具體例中，該第一、第二及第三重組酶辨識位點彼此不同。在一些具體例中，該等重組酶辨識位點是選自*Lox*P位點、*Lox*511位點、*Lox*2272位點、*Lox*2372、*Lox*5171、*Lox*m2、*Lox*71、*Lox*66、*Lox*Fas及frt位點。

【0029】 在一個具體例中，該第一可篩選標記基因是藥物抗性基因。在另一個具體例中，該藥物抗性基因是新黴素抗性基因或潮黴素抗性基因。在另一個具體例中，該第二及第三可篩選標記基因編碼兩個不同的螢光蛋白質。在一個具體例中，該兩個不同的螢光蛋白質是選自由下列組成之群：珊瑚(*Discosoma* coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質

(例如mKate、mKate2、mPlum、mRaspberry或E2-crimson)。

【0030】 在一個具體例中，該第一、第二及第三啟動子相同。在另一個具體例中，該第一、第二及第三啟動子彼此不同。在另一個具體例中，該第一啟動子不同於第二及第三啟動子，而該第二及第三啟動子相同。在更多具體例中，該第一啟動子為SV40晚期啟動子，而該第二及第三啟動子各自為人類CMV啟動子。在其他具體例中，該第一及第二啟動子可操作地連結至原核操縱子。

【0031】 在一個具體例中，該宿主細胞株具有編碼重組酶之外源地添加而併入其基因組中並可操作地連結至啟動子的基因。在另一個具體例中，該重組酶為Cre重組酶。在另一個具體例中，該宿主細胞具有併入其基因組中之編碼調節蛋白質並可操作地連結至啟動子的基因。在更多具體例中，該調節蛋白質為tet壓抑子蛋白質。

【0032】 在一個具體例中，該第一GOI及第二GOI編碼抗體輕鏈或其片段，或抗體重鏈或其片段。在另一個具體例中，該第一GOI編碼抗體輕鏈而該第二GOI編碼抗體重鏈。

【0033】 在某些具體例中，該第一、第二及第三GOI編碼選自下列組成之群的多肽：第一輕鏈或其片段、第二輕鏈或其片段，以及重鏈或其片段。在又另一個具體例中，該第一、第二及第三GOI編碼選自下列組成之群的多肽：輕鏈或其片段、第一重鏈或其片段，及第二重鏈或其片段。

【0034】 在一個態樣中，提供一種製造感興趣蛋白質的方法，包含(a)將感興趣的基因(GOI)引入CHO宿主細胞中，其中該GOI併入與SEQ ID NO:1或SEQ ID NO:4包含至少90%相同之核苷酸序列的特定基因座中；(b)在容許該GOI表現的條件下培養(a)之細胞；以及(c)回收感興趣的蛋白質。在一個具體例中，感興趣的蛋白質是選自下列組成之群：免疫球蛋白質次

單位或其片段，以及受體或其配體結合片段。在某些具體例中，該感興趣的蛋白質是選自由下列組成之群：抗體輕鏈或其抗原結合片段，以及抗體重鏈或其抗原結合片段。

【0035】 在一些具體例中，採用重組酶媒介之匣式交換(RMCE)用的靶定載體將該GOI引入細胞中，且該CHO宿主細胞基因組於特定基因座內包含至少一個外源性辨識序列。在其他具體例中，該CHO宿主細胞基因組於特定基因座內包含至少一個外源性辨識序列以及一個可篩選標記，視情況連結至啟動子、IRES及/或聚腺苷酸化(聚A)序列。

【0036】 在某些具體例中，該CHO宿主細胞基因組包含一或多個如上述的重組酶辨識位點，且該GOI透過辨識重組酶辨識位點的重組酶作用而被引入至該特定基因座中。

【0037】 在另一個具體例中，採用同源性重組用的靶定載體將該GOI引入細胞中，且其中該靶定載體包含與該特定基因座內存在的序列同源的5'同源性臂、GOI以及與該特定基因座內存在的序列同源的3'同源性臂。在另一個具體例中，該靶定載體進一步包含兩個、三個、四個或五個或更多個感興趣基因。在另一個具體例中，感興趣基因中的一或多者可操作地連結至啟動子。

【0038】 在另一個態樣中，提供一種靶定載體，其中該靶定載體包含與基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的5'同源性臂、GOI以及與基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的3'同源性臂。在另一個具體例中，該靶定載體進一步包含兩個、三個、四個或五個或更多個感興趣基因。

【0039】 在另一個態樣中，提供一種修飾CHO細胞基因組以併入

外源性核酸序列的方法，包含將包括載體的媒劑引入細胞的步驟，其中該載體包含一外源性核酸序列，其中外源性核酸併入包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列之該基因組的基因座內。

【0040】 在一些具體例中，該載體包含與基因組基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的5'同源性臂、外源性核酸序列，以及與基因組基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的3'同源性臂。

【0041】 在一些具體例中，載體中的該外源性核酸序列包含一或多個辨識序列。在其他具體例中，該外源性核酸包含一或多個GOI，諸如可篩選標記或編碼POI的核酸。在又其他具體例中，該外源性核酸包含一或多個GOI以及一或多個辨識序列。

【0042】 在一個具體例中，該媒劑包含至少一個額外載體或mRNA。在另一個具體例中，該額外的載體是選自由下列組成之群：腺病毒、慢病毒、逆轉錄病毒、腺相關病毒、整合噬菌體載體、非病毒載體、轉位子及/或轉位酶、整合酶受質，以及質體。在一些具體例中，該額外載體包含編碼整合外源性核酸序列用之位點特異性核酸酶的核苷酸序列。

【0043】 在某些具體例中，該位點特異性核酸酶包含鋅手指核酸酶(ZFN)、ZFN二聚體、轉錄活化因子樣效應子核酸酶(TALEN)、TAL效應子域融合蛋白質，或RNA導引DNA內核酸酶。

【0044】 在另一個態樣中，提供一種用以修飾CHO細胞基因組以整合外源性核酸序列的媒劑，其中該媒劑包括一載體，其中該載體包含與基因組基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的5'同源性臂、外源性核酸序列，以及與基因

組基因座內存在的序列(其包含與SEQ ID NO:1或SEQ ID NO:4至少90%相同之核苷酸序列)同源的3'同源性臂。

【0045】 在一些具體例中，該外源性核酸序列包含一或多個辨識序列。在其他具體例中，該外源性核酸包含一或多個GOI，諸如可篩選標記或編碼POI的核酸。在又其他具體例中，該外源性核酸包含一或多個GOI及一或多個辨識序列。

【0046】 在又另一個態樣中，提供一種修飾CHO細胞基因組以表現治療劑的方法，包含用以將包含表現該治療劑之序列的外源性核酸引入基因組的媒劑，其中該媒劑包含與SEQ ID NO:1核苷酸序列內存在的序列同源的5'同源性臂、編碼治療劑的核酸，以及與SEQ ID NO:1或SEQ ID NO:4核苷酸序列內存在的序列同源的3'同源性臂。

【0047】 在又一個態樣中，本發明提供一種包含經修飾CHO基因組的經修飾CHO宿主細胞，其中該CHO基因組是透過在基因組的基因座內插入一與SEQ ID NO:1至少90%相同之核苷酸序列的外源性辨識序列。

【0048】 在另一個態樣中，本發明提供一種包含經修飾真核基因組的經修飾真核宿主細胞，其中該真核基因組在基因組非編碼區內的目標整合位點處被修飾以插入外源性核酸。在一些具體例中，該外源性核酸是辨識序列。在其他具體例中，該宿主細胞為哺乳動物宿主細胞，諸如CHO細胞。在其他具體例中，目標整合位點包含表現增高序列，諸如SEQ ID NO:1，前提為該序列不編碼任何內源性蛋白質。本發明亦提供製造此一經修飾真核宿主細胞的方法。

【0049】 在上述態樣及具體例的任一者中，該表現增高序列可呈指定位向位在SEQ ID NO:1中，或呈SEQ ID NO:1的反向位向。

【0050】 除非另有指明或在上下文中為明顯的，否則本發明的任

何態樣及具體例可與本發明的任何其他態樣或具體例組合使用。

【0051】 其他目的及優點將因為檢視下面詳細說明而變得清楚。

詳細說明

【0052】 在說明本方法之前，應理解本發明不囿於所述特定方法及實驗條件，因為此等方法及條件可能會改變。亦應理解，本文所用術語僅是為了描述特定具體例，而不欲具有限制性，因為本發明範疇將僅會受到隨附申請專利範圍所限制。

【0053】 除非上下文另有明確指明，否則如本說明書及隨附申請專利範圍中所用，單數形式「一(a)」及「該(the)」包括複數指涉對象。因此，例如提到「一種方法」包括一或多種方法，及/或本文所述類型的步驟及/或習於技藝者在閱讀過揭示內容之後變得清楚者。

【0054】 除非另有定義或另有指明，否則本文所用全部技術及科學術語具有與本發明所屬技藝中具有通常技術者一般所理解的相同含意。

【0055】 儘管與本文所述任一方法及材料相似或相同者可用於實施或測試本發明，現將說明具體方法及材料。本文提及的全部公開資料以全文引用的方式併入本文中。

定義

【0056】 當DNA在功能上彼此相關時，DNA可操作地連結。例如，若啟動子能夠參與序列轉錄，則啟動子可操作地連結至編碼序列；若核醣體結合位點定位成容許轉譯時，其可操作地連結至編碼序列。一般來說，可操作地連結可包括(但不必然)接近性。在序列(諸如分泌引導序列)的情況下，在閱讀框中的接近性及適當位置是典型特徵。感興趣基因座的表現增高序列可操作地連結至感興趣基因(GOI)，其在功能上與該GOI相關，例如若其存在會使得表現增高及/或GOI穩定整合。

【0057】 術語「增高」當用來說明增高表現時，包括表現增高至少約1.5倍至增高至少約3倍超過通常因為隨機整合外源性序列至基因組或整合在不同基因座所觀察到的，例如相較於帶有相同表現構築體之單一複本之隨機整合體池。採用本發明序列所觀察到的表現增高倍數是與相同基因的表現程度相比，在實質上相同的條件下於本發明序列不存在時進行測量，例如相較於在相同物種基因組內的另一基因座處進行整合。重組效率增高包括基因座的重組能力增高(例如採用重組酶辨識位點)。增高意指勝過隨機重組(例如在不採用重組酶辨識位點或類似者之情況下，通常為0.1%)的效率。較佳的增高重組效率為隨機的約10倍，或約1%。除非指明，否則所請求之本發明不限於特定重組效率。

【0058】 在提及感興趣基因座時採用片語「外源性添加的基因」或「外源性添加的核酸」時，該片語意指如該基因座在自然界中所發現之不存在於感興趣基因座內的任何DNA序列或基因。例如，CHO基因座內之「外源性添加的基因」(例如包含SEQ ID NO:1之序列的基因座)可以是在自然界中於特定CHO基因座中未被發現的倉鼠基因(意即來自倉鼠基因組中另一基因座的倉鼠基因)、來自任何其他物種的基因(例如人類基因)、嵌合基因(例如人類/小鼠)或任何其他在自然界中未被發現到存在於感興趣CHO基因座內的基因。

【0059】 當描述感興趣基因座(諸如SEQ ID NO:1或SEQ ID NO:4或其片段時)，同一性百分率(identity percentage)表示包括沿著接續同源性區域展現引用之同一性的同源性序列，但空位、刪除或插入(與比較序列無同源性)的存在不納入計算同一性百分率的考量。

【0060】 如本文所用，具有物種同源性之例如SEQ ID NO:1或其片段之間的「同一性百分率」測定不包括序列比對，其中物種同源物在比

對時不具有要比較的同源性序列(意即SEQ ID NO:1或其片段在那個點具有插入，或物種同源物具有空位或插入，隨情況而不同)。因此，「同一性百分率」不包括空位、刪除及插入的罰分。

【0061】 「同源性序列」在核酸序列上下文中意指實質上與參考核酸序列同源的序列。在一些具體例中，若兩個序列在殘基的相關延伸片段上的對應核苷酸有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高相同，則它們被視為實質上同源。在一些具體例中，該相關延伸片段為完整(意即全)序列。

【0062】 「靶定插入」意指用來指引基因或核酸序列直接插入或整合至基因體上特定位置的基因靶定方法，也就是將DNA指引至接續多核苷酸鏈中兩個核苷酸間的特定位點。也可以做出靶定插入用於特定基因匣，其包括多個基因、調節要素，及/或核酸序列。「插入」及「整合」可交替使用。應理解插入基因或核酸序列(例如包含表現匣的核酸序列)可使得(或可經改造)一或多個核酸取代或刪除，取決於採用的基因編輯技術。

【0063】 「辨識位點」或「辨識序列」是核酸酶或其他酶所辨識以結合並導引位點特異性切割DNA骨架的特定DNA序列。內核酸酶在DNA分子內切割DNA。辨識位點在技藝中也可以稱為辨識目標位點。

【0064】 「重組酶辨識位點」是重組酶(諸如Cre重組酶(Cre)或翻轉酶(flippase, flp))所辨識的特定DNA序列。當位點特異性重組酶的一或多個目標辨識序列在策略上是被放入某個生物體的基因組內，其可以執行DNA重組，包括刪除、反轉及轉位。在一個實例中，Cre在其DNA目標辨識位點*lox*P處特異地媒介重組事件，*lox*P是由兩個被8-bp間隔子分開來的13-bp反轉重複序列組成。可使用超過一個重組酶辨識位點(例如)來促進DNA的重

組媒介交換。也可以採用重組酶辨識位點(例如lox位點)的變體或突變體(Araki, N. et al, 2002, *Nucleic Acids Research*, 30:19, e103)。

【0065】 「重組酶媒介之匣式交換」與將基因組目標匣準確地置換為供體匣的過程有關。一般提供來進行這個過程的分子組成物包括1) 5'與3'兩側為對特定重組酶具有特異性之辨識目標位點的基因組目標匣、2)兩側為配對辨識目標位點的供體匣，以及3)位點特異性重組酶。重組酶蛋白質為技藝中所熟知(Turan, S. and Bode J., 2011, *FASEB J.*, 25, pp. 4088–4107)且能夠準確切割特定辨識目標位點內的DNA(DNA序列)而不會獲得或損失核苷酸。常見重組酶/位點組合包括，但不限於Cre/*lox*以及Flp/*frt*。

【0066】 「媒劑」是一種由帶有外源性核酸供引入細胞用的任何多核苷酸或核苷酸組所組成的組成物。媒劑包括載體、質體及mRNA分子，它們是透過已知的轉染方法被遞送至細胞。在一個實例中，被引入至細胞中的mRNA可為短暫的且不併入基因組中，但是mRNA可帶有發生整合過程所需要的外源性核酸。

一般性說明

【0067】 本發明至少一部分是基於發現到，在基因組中的獨特序列(意即基因座)比基因組中的其他區域或序列展現更有效的重組、插入穩定性，以及更高的表現位準。本發明至少一部分也是基於發現到，當此表現增高序列被鑑定之後，適當的基因或構築體可以外源性地被添加至序列中或序列附近且該外源性地被添加之基因可有利地表現或採用供進一步基因修飾。此等序列(稱為表現增高序列)被認為是穩定且且不位在基因組的編碼區內。這些表現增高及穩定性區可經改造用於未來的選殖或基因組編輯事件。因此，在細胞的基因組骨架中建立可靠的表現系統。

【0068】 本發明亦基於外源性基因特異地靶定至整合位點。本發

明方法容許將細胞基因組有效率「轉換」至可用選殖匣中，例如透過採用重組酶媒介的匣式交換(RMCE)。為此，本發明方法採用細胞基因組重組酶辨識位點用於安置感興趣基因，以創造出高產量的重組型蛋白質生產用細胞株。

【0069】 本發明組成物也可以納入表現構築體中，例如納入表現載體供選殖及改造新細胞株用。包含本發明多核苷酸的表現載體可用於暫時表現蛋白質，或可藉由隨機或標靶重組(諸如，例如同源性重組或透過辨識特定重組位點之重組酶媒介的重組(例如Cre-lox媒介的重組))而併入基因組中。包含本發明多核苷酸的表現載體亦可用於評估其他DNA序列(例如順式作用的調節序列)的效力。

【0070】 整合位點通常是由隨機整合或分析逆轉錄病毒整合事件來加以鑑定。本文詳述之CHO整合位點是透過隨機併入編碼多鏈抗體之DNA而被鑑定且所表現的蛋白質被發現到展現之表現有所增高。

【0071】 包含一個重鏈(HC)以及兩個輕鏈(LC)複本的實例多鏈抗體以表現匣的方式被隨機地併入基因組中，該表現匣含有交替的潮黴素抗性基因(參見，例如圖1A中所示的三個相同Hyg基因)。在如SEQ ID NO:1鑑定之基因座內併入表現匣產生一個穩定且表現量高的純系。

【0072】 相較於併入CHO基因組的另一個區(對照整合位點)，實例多鏈抗體在併入SEQ ID NO:1的基因座中時展現更高的表現位準。有趣的是，就併入SEQ ID NO:1內的抗體表現多核苷酸相對於對照整合位點來說，基因複本數相當，但是併入SEQ ID NO:1內的抗體表現多核苷酸的蛋白質效價高出3倍。

【0073】 使用標靶重組方法將CHO細胞基因組轉換成含有重組酶辨識位點的選殖構築體(參見，例如圖3A-B)。

【0074】 在鑑定SEQ ID NO:1的整合位點之後，重組酶辨識位點(例如lox位點)本質上用於將包含可表現GOI (諸如可篩選標記(參見例如圖3A-B))與任何其他所需要素(諸如，例如啟動子、增強子、標記、操縱子、核醣體結合位點(例如內部核醣體進入位點)等)之表現匣引入的位點中。

【0075】 用於標靶併入SEQ ID NO:1內之lox位點的實例供體構築體的圖式說明顯示於圖1B中。供體構築體包含受到新黴素(neo)抗性基因及內部核醣體進入位點(IRES)所驅動的表現匣，其中該匣包含螢光標記(mKate)且其在5'與3'端兩側為重組酶辨識位點及5'與3'同源性臂(與SEQ ID NO:1同源)。顯示SEQ ID NO:1的基因座內的插入，其中插入造成供體neo/mKate構築體取代包含潮黴素抗性標記的表現匣，其中SEQ ID NO:1基因座內的表現匣於其5'及3'端兩側為重組辨識位點連結至5'與3'同源性臂(與SEQ ID NO:1同源)(參見圖1B)。

【0076】 提供用於將核酸序列穩定併入真核細胞中的組成物及方法，其中該核酸序列能夠透過併入SEQ ID NO:1或其表現增高片段中的特性而增高表現。提供在SEQ ID NO:1內含有重組酶辨識序列好方便插入GOI以達到從GOI表現感興趣蛋白質的細胞。亦提供用於靶定與表現構築體(例如表現載體)有關之整合位點，以及用於將外源性核酸併入感興趣CHO細胞中之組成物及方法，

CHO整合位點的物理性及功能性特徵鑑定

【0077】 SEQ ID NO:1的核酸序列(以及SEQ ID NO:4的更廣核酸序列)是透過以高位準表現蛋白質之細胞株的核酸構築體(包含表現匣)之整合位點上游與下游的序列來進行實驗鑑定。本發明的核酸序列提供具有與核酸(例如包含GOI之外源性核酸)之表現及穩定性增高有關的新穎功能性，且在不受到任何一種理論限制的情況下與先前所述順式作用要素(諸如

啟動子、增強子、位點控制區、架構附接區或基質附接區)作用相同或不同的序列。SEQ ID NO:1似乎不具有任何開放閱讀框(ORF)，使該基因座不能編碼新穎的反式作用活化因子蛋白質。在SEQ ID NO:4的基因組基基因座3'(下游)中鑑定出一個推定鋅手指蛋白質。

【0078】 關於將包含第一潮黴素(Hyg)基因、第一GOI、第二Hyg基因、第二GOI、第三Hyg基因及第三GOI編碼序列的表現匣併入CHO基因組DNA之非編碼區的特有位點內，鑑定出表現增高活性。就表現GOI的表現匣來說，包含(例如)由CHO基因組DNA之非編碼區鑑定出之5'經單離1 kb區以及3'經單離1 kb區的表現載體能夠在CHO細胞經它們轉染後賦予高位準的重組型蛋白質表現。

【0079】 本發明含括表現載體，其包含反向的SEQ ID NO:1片段或SEQ ID NO:4片段。亦可開發出本文所述片段的其他組合。亦可開發出本文所述片段之其他組合的實例包括含有本文所述多個表現增高序列多個副本的序列，或透過將所揭示之SEQ ID NO:1片段或SEQ ID NO:4片段與其他核苷酸序列組合所衍生以達到調節要素最佳組合的序列。此等組合可接續連接或排列以提供SEQ ID NO:1片段或SEQ ID NO:4片段的最佳間隔(例如透過在片段之間引入間隔子核苷酸)。也可以安置調節要素以提供SEQ ID NO:1片段與調節要素的最佳間隔。

【0080】 本文揭示的SEQ ID NO:1及SEQ ID NO:4是由CHO細胞分離。發現其他哺乳動物物種(諸如，例如人類或小鼠)與經鑑定的表現增高區具有有限的同源性，但是可在衍生自灰倉鼠的其他組織類型或其他同源性物種的細胞株中發現到同源性序列，並且可藉由技藝中熟知的技術予以分離。例如，吾人可透過物種交叉雜交或以PCR為基礎的技術來鑑定出其他同源性序列。此外，可以藉由技藝中所充分熟知的定位突變或隨機突變技

術在SEQ ID NO:1、SEQ ID NO:4或其片段所示的核苷酸序列中進行改變。接著可如所述測試所得序列變體的表現增高活性。在核酸同一性方面與SEQ ID NO:1、SEQ ID NO:4或其片段具有至少約90%相同之具有表現增高活性的DNA可藉由慣常實驗分離，且預期展現出表現增高活性。關於SEQ ID NO:1或SEQ ID NO:4的片段，同一性百分率意指在SEQ ID NO:1片段或SEQ ID NO:4片段中，所發現之參考原有序列的部分。因此，SEQ ID NO:1、SEQ ID NO:4或其片段的同源物及其變體亦受本發明具體例所含括。

【0081】 在某些具體例中，SEQ ID NO:1的片段是選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1的10-4,000；100-3,900；200-3,800；300-3,700；400-3,600；500-3,500；600-3,400；700-3,300；800-3,200；900-3,100；1,000-3,000；1,100-2,900；1,200-2,800；1,300-2,700；1,200-2,600；1,300-2,500；1,400-2,400；1,500-2,300；1,600-2,200；1,700-2100；1,800-2050；1850-2050；1,900-2040；1950-2,025，1990-2021，2002-2021及2,010-2,015。在另一個具體例中，SEQ ID NO:1的片段是選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1的10-500；500-1,000；500-2,100；1,000-1,500；1,000-2,100；1,500-2,000；1,500-2,500；2,000-2,500；2,500-3,000；2,500-3,500；3,000-3,500；3,000-4,000；及3,500-4,000。在某些具體例中，該外源性核酸序列併入上述片段內的特定位點處或鄰接該特定位點。

【0082】 在另一個具體例中，該外源性核酸序列位在如上述SEQ ID NO:1或其片段內，或在與SEQ ID NO:1之表現增高序列或其表現增高序列至少約90%相同、至少約91%相同、至少約92%相同、至少約93%相同、至少約94%相同、至少約95%相同、至少約96%相同、至少約97%相同、至少約98%相同，或至少約99%相同的序列內。

【0083】 表現感興趣蛋白質的位準增高的細胞群可以使用本文所提供之方法來開發。絕對表現位準可能隨著特定蛋白質而改變，取決於蛋白質如何有效率地受到細胞加工處理。使用併入本發明之表現增高序列內之外源性序列來開發的細胞池隨著時間流逝是穩定的，且可以做為穩定細胞株受到處理供大多數用途。重組步驟也可以延遲，直到之後在開發本發明細胞株的程序時。

CHO表現增高基因座及其片段

【0084】 本發明含括核苷酸序列之表現增高片段，其與SEQ ID NO:1的核苷酸序列為至少約90%相同、至少約91%相同、至少約92%相同、至少約93%相同、至少約94%相同、至少約95%相同、至少約96%相同、至少約97%相同、至少約98%相同，或至少約99%相同。本發明包括含有一片段的載體，該片段包括用以暫時或穩定轉染之橫跨下列編號位置：SEQ ID NO:1的10-4,000；100-3,900；200-3,800；300-3,700；400-3,600；500-3,500；600-3,400；700-3,300；800-3,200；900-3,100；1,000-3,000；1,100-2,900；1,200-2,800；1,300-2,700；1,200-2,600；1,300-2,500；1,400-2,400；1,500-2,300；1,600-2,200；1,700-2100；1,800-2050；1850-2050、1,900-2040；1950-2,025、1990-2021、2002-2021及2,010-2,015。本發明亦包括包含此一片段的真核細胞，其中該片段對該細胞來說是外源性且併入細胞基因組中，而包含此一具有至少一個重組酶辨識位點之片段的細胞在該片段內、就在5'或就在3'。

【0085】 在一個具體例中，SEQ ID NO:1的表現增高片段是位在橫跨下列編號位置之SEQ ID NO:1內的位置：SEQ ID NO:1的10-500；500-1,000；500-2,100；1,000-1,500；1,000-2,100；1,500-2,000；1,500-2,500；2,000-2,500；2,500-3,000；2,500-3,500；3,000-3,500；3,000-4,000；或

3,500-4,000。

【0086】 當整合之多核苷酸的穩定整合及/或增高轉錄持續時，對例示之位點來說，基因座插入(亦即整合)位點的確切位置不是必須的。進一步來說，整合位點可以如本文所述在SEQ ID NO:1或SEQ ID NO:1之片段，或SEQ ID NO:4或SEQ ID NO:4之片段內或相鄰。不管是感興趣基因座內或相鄰處的特定染色體位置都能支持穩定整合且經整合之外源性基因的有效率轉錄可以依據技藝中熟知的標準程序或本文例示的方法進行測定。

【0087】 本文所考量的整合位點是位於包含SEQ ID NO:1或SEQ ID NO:4之核苷酸序列的基因座內，或是在接近感興趣的基因座內，例如染色體DNA上SEQ ID NO:1位置的上游(5')或下游(3')少於約1 kb、500鹼基對(bp)、100 bp、50 bp、25 bp、10 bp或少於約5 bp處。在又其他一些具體例中，所用整合位點是位在染色體DNA上SEQ ID NO:1或SEQ ID NO:4位置的上游(5')或下游(3')約1000、2500或更多鹼基對處。

【0088】 在技藝中應理解，大的基因組區(諸如架構/基質附接區)用於有效率地複製並轉錄染色體DNA。架構/基質附接區(S/MAR)亦已知稱為架構附接區(SAR)或基質相關或基質附接區(MAR)，是細胞核附接的真核基因組DNA區。在不受到任何一種理論囿限的情況下，S/MRAR通常繪製為非編碼區，將給定轉錄區(例如染色質域)與其鄰居分隔開來，且亦提供使轉錄能夠進行之因子的機制及/或結合的平台(諸如DNA酶或聚合酶的辨識位點)。一些S/MAR經特徵鑑定長度為約14-20 kb (Klar, et al. 2005, *Gene* 364:79-89)。因此，預期在*LOCUS 1*處(SEQ ID NO:1或SEQ ID NO:4內或鄰近處)併入基因會賦予增高的表現。

【0089】 具有技藝者將認知到，在目標基因座處可針對高轉錄活性將數個要素最佳化，以使得編碼感興趣蛋白質之插入基因有高表現。所

考量的要素包括強啟動子以驅動轉錄、適當轉錄機制以及具有開放且可進入構型的DNA。藉由靶定選自SEQ ID NO:1或SEQ ID NO:4內的整合位點，在習於技藝者的技術內可將在目標基因座內的插入最佳化。

【0090】 在一個具體例中，SEQ ID NO:1的表現增高序列用來增高GOI的表現。圖2A顯示可操作地連結至SEQ ID NO:1 (*LOCUS 1*)之GOI相較於相同GOI併入CHO細胞基因組中不同基因座(對照基因座)的結果。各細胞株測得的基因複本數相同，但實驗顯示細胞表現GOI之mRNA位準及蛋白質效價比可操作地連結至*LOCUS 1*的GOI高出3倍。

【0091】 在各種具體例中，GOI的表現可以透過將GOI安置在SEQ ID NO:1或SEQ ID NO:4內而獲得增高。在各種具體例中，表現增高至少約1.5倍至約3倍或更多。

遺傳修飾目標基因座

【0092】 可以數種方式達到在特定位置(亦即目標基因座)中遺傳改造細胞基因組的方法。使用遺傳編輯技術將核酸序列穩定地併入真核細胞中，其中核酸序列是通常不會在此等細胞中發現的外源性序列。需要純系擴增以確保細胞後代共有經改造細胞株的相同基因型及表現型特徵。在一些實例中，原細胞經同源重組技術修飾以將外源性核酸序列併入SEQ ID NO:1或SEQ ID NO:4內。在其他實例中，提供含有便於將外源性核酸序列或感興趣基因併入SEQ ID NO:1或SEQ ID NO:4內之至少一個重組酶辨識序列的細胞。

【0093】 在一些實例中，提供含有第一重組酶辨識序列及第二重組酶辨識序列的細胞，其中第一重組酶辨識序列及第二重組酶辨識序列各自選自包含下列之群：*Lox*P、*Lox*511、*Lox*5171、*Lox*2272、*Lox*2372、*Lox*m2、*Lox*-FAS、*Lox*71、*Lox*66及其變體。在這個情況下，若需要重組酶媒介的匣

式交換(RMCE)，則位點特異性重組酶為Cre重組酶或其衍生物。在其他實例中，第一重組酶辨識序列及第二重組酶辨識序列各自選自包含下列之群：FRT、F3、F5、FRT突變體-10、FRT突變體+10及其突變體，且在這個譜系中，若需要RCME，則該位點特異性重組酶為Flp重組酶或其衍生物。在又另一個實例中，第一重組酶辨識序列及第二重組酶辨識序列各自選自包含下列之群：attB、attP及其突變體，且在這個情況下若需要RMCE，則該位點特異性重組酶為phiC31整合酶或其衍生物。

【0094】 在一個態樣中，用以穩定地將核酸序列併入SEQ ID NO:1或SEQ ID NO:4或其表現增高片段內的方法是經由同源性重組。核酸分子(亦即感興趣的基因或多核苷酸)可以透過同源性重組或透過特異地靶定整合位點處之序列使用位點特異性核酸酶方法被插入至靶定基因座(亦即SEQ ID NO:1)中。關於同源性重組，同源性多核苷酸分子(亦即同源性臂)排好並交換其序列延伸段。若轉基因兩側為同源性基因組序列，則轉基因可以在這個交換期間被引入。在一個實例中，重組酶辨識位點可以在整合位點處被引入至宿主細胞基因組中。

【0095】 在真核細胞中，同源性重組可以透過於整合位點處在染色體DNA中引入斷裂而被促進。模型系統已證明，在基因靶定期間，若雙股斷裂被引入染色體目標序列內，則同源性重組的頻率增加。這可以透過將某些核酸酶靶定至特定整合位點而達成。在目標基因座處辨識DNA序列的DNA結合蛋白質為技藝中已知的。也採用基因靶定載體來促進同源性重組。在用於同源性指引之修復的基因靶定載體不存在時，細胞常常透過非同源性末端接合(NHEJ)來封閉雙股斷裂，其可能會導致在裂解位點處的多個核苷酸缺失或插入。若發生插入或刪除(InDel)，則少數核苷酸於斷裂位點處隨機被插入或被刪除且這些InDel可能改變或中斷目標基因座處內基因

的任何開放閱讀框(ORF)。應理解，鑑別為SEQ ID NO:1 (或SEQ ID NO:4)的基因座並非基因編碼區。因此，預見在這個基因座處不會有因為插入及/或刪除而產生內源性基因轉錄中斷。

【0096】 同源性定向修復(或同源性定向重組)(HDR)尤其可用於將基因插入並併入目標基因座處。供體構築體包含如本文所述之衍生自SEQ ID NO:1或SEQ ID NO:4的同源性臂。

【0097】 基因靶定載體構築及核酸酶選定為本發明所屬技藝中具有通常技術者的技能。

【0098】 在一些實例中，鋅手指核酸酶(ZFN)(其具有模組結構且含有單獨鋅手指域)辨識目標序列中的特定3-核苷酸序列(例如靶定整合位點)。一些具體例可採用ZFN與靶定多目標序列之個別鋅手指域的組合。

【0099】 轉錄活化因子樣(TAL)效應子核酸酶(TALEN)亦可用於位點特異性基因組編輯。TAL效應子蛋白質DNA結合域通常與限制核酸酶(諸如FokI)的非特異性切割域組合使用。在一些具體例中，包含TAL效應子蛋白質DNA結合域及限制核酸酶切割域的融合蛋白質被用來在本發明基因座內的目標序列處辨識並切割DNA (Boch J et al., 2009 *Science* 326:1509-1512)。

【0100】 RNA導引的內核酸酶(RGEN)是程式化基因組改造工具，它是由細菌適應性免疫機制發展而來。在這個系統(簇集的規律間隔短回文重複序列(CRISPR)/CRISPR相關(Cas)免疫反應)中，當蛋白質Cas9與兩個RNA複合時形成序列特異性內核酸酶，RNA中之一者導引目標選擇。RGEN由組份(Cas9與tracrRNA)及目標特異性CRISPR RNA (crRNA)組成。DNA目標切割的效率及切割位點位置會依據前間區序列鄰近模體(PAM)的位置、其他關於目標辨識的要求而改變(Chen, H. et al, *J. Biol. Chem.*

published online March 14, 2014, as Manuscript M113.539726)。

【0101】 用於鑑定對SEQ ID NO:1之特異性靶定基因座具有獨特性之序列的策略為技藝中已知，但是，許多這些序列與CHO基因組的比對揭示，潛在脫靶位點有16-17個鹼基對配對。由SEQ ID NO:5中所示序列(對應於SEQ ID NO:1的核苷酸1990-2001)編碼的一個實例20 bp導引RNA可用於SEQ ID NO:1或SEQ ID NO:4的RNA導引CRISPR/Cas基因編輯。包含驅動小導引RNA及tracrRNA (例如SEQ ID NO:6)表現之啟動子，且帶有受啟動子控制之適當Cas酶的質體可與透過此方法與採用靶定整合之供體載體(帶有感興趣基因，兩側為5'及3'同源性臂)一起共轉染。除了上文所述以外之RNA分子的各種修飾及變體對於習於技藝者來說是明顯的且意欲落入本發明範疇內。

【0102】 在一些具體例中，用於將含有編碼感興趣基因或辨識序列或基因匣之序列的外源性核酸引入基因組中的媒劑視情況包含帶有外源性核酸及一或多個額外載體或mRNA的載體。在一個具體例中，該一或多個額外載體或mRNA包含編碼位點特異性核酸酶的核苷酸序列，包括(但不限於)鋅手指核酸酶(ZFN)、ZFN二聚體、轉錄活化因子樣效應子核酸酶(TALEN)、TAL效應子域融合蛋白質，以及RNA導引DNA內核酸酶。在某些具體例中，該一或多個載體或mRNA包含含有導引RNA、tracrRNA及編碼Cas酶之核苷酸序列的第一載體，以及含有供體(外源性)核苷酸序列的第二載體。此供體序列包含編碼欲用於靶定插入之感興趣基因，或辨識序列，或含有此等外源性要素中任一者之基因匣的核苷酸序列。若使用mRNA，則mRNA可以藉由習於技藝者已知的常用轉染方法被轉染至細胞中，且可編碼酶(例如轉位酶或內核酸酶)。儘管被引入細胞中的mRNA可能是暫時的且不會併入至基因組中，但mRNA可能帶有對於發生整合來說必須或有利的外源

性核酸。在一些情況下，選擇mRNA以降低附加性多核苷酸的長期持續副作用，其中僅需要短期表現來達成GOI的所需整合。

【0103】 對於習於技藝者來說，仍有其他同源性重組方法可供使用，諸如具有準確DNA結合特異性的BuD-衍生的核酸酶(BuDN)(Stella, S. et al. *Acta Cryst.* 2014, D70, 2042-2052)。依據與SEQ ID NO:1內獨特目標序列可相容的工具來選擇精確的基因組修飾方法，以避免細胞表現型的破壞。

基因靶定構築體

【0104】 要被併入宿主基因組中的多核苷酸序列可以是任何產業上可用的DNA序列，諸如辨識序列，以供產生細胞表現系統。要被併入宿主基因組的多核苷酸序列可編碼如本文所述任一種治療上或產業上可用蛋白質。鑑定要併入外源性核酸序列之目標基因座內的目標序列仰賴於一些因素。取決於採用的同源性重組方法，習於技藝者能適當選擇與SEQ ID NO:1或SEQ ID NO:4同源的序列。當使用位點特異性核酸酶載體時，需要辨識欲用於DNA切割之特定位點的額外組份(序列組成物)。

【0105】 因此，靶定基因的構築體通常併入促使外源性核酸序列靶定整合至感興趣的基因座中的此等核苷酸序列。在一些具體例中，該構築體包含第一同源性臂及第二同源性臂。在其他具體例中，該構築體(例如基因匣)包含衍生自SEQ ID NO:1或SEQ ID NO:4的同源性臂。在一些具體例中，該同源性臂包含與SEQ ID NO:1或SEQ ID NO:4中存在的核苷酸序列同源的核苷酸序列。在特定具體例中，該構築體包含具有SEQ ID NO:2的核苷酸序列(對應於SEQ ID NO:1的核苷酸1001-2001)之5'同源性臂，以及具有SEQ ID NO:3的核苷酸序列(對應於SEQ ID NO:1的核苷酸2022-2001)之3'同源性臂。同源性臂(例如第一同源性臂(又稱為5'同源性臂)以及第二同源性臂(又稱為3'同源性臂))與基因座內的靶定序列同源。5'至3'的同源性臂可能是

基因座內延伸的區域或靶定序列，其含有至少1 kb，或至少約2 kb，或至少約 3 kb，或至少約4 kb，或至少5 kb，或至少約10 kb。在其他具體例中，針對第一級第二同源性臂所選定的靶定序列之核苷酸總數包含至少1 kb，或至少約2 kb，或至少約3 kb，或至少約4 kb，或至少5 kb，或至少約10 kb。在一些情況下，5'同源性臂與3'同源性臂(與靶定序列同源)之間的距離包含至少5 bp、10 bp、20 bp、30 bp、40 bp、50 bp、60 bp、70 bp、80 bp、90 bp、100 bp、200 bp、300 bp、400 bp、500 bp、600 bp、700 bp、800 bp、900 bp、或至少1 kb、或至少約2 kb、或至少約3 kb、或至少約4 kb、或至少5 kb，或至少約10 kb。在SEQ ID NO:2與SEQ ID NO:3被選作為5'同源性臂及3'同源性臂的情況下，兩個同源性臂之間的距離可為20個核苷酸(對應於SEQ ID NO:1的核苷酸2002-2021)；且此等同源性臂可媒介外源性核酸序列併入包含SEQ ID NO:1的基因座中，例如SEQ ID NO:1的核苷酸1990-2021或2002-2021中，並同時刪除SEQ ID NO:1的核苷酸2002-2021。

【0106】 在其他具體例中，該構築體包含第一同源性臂及第二同源性臂，其中合併的該第一同源性臂及第二同源性臂包含靶定序列，其取代基因座內的內源性序列。在又其他具體例中，該第一同源性臂及第二同源性臂包含靶定序列，其併入或插入基因座內的內源性序列中。

【0107】 經修飾細胞株可以透過將一或多個重組酶辨識位點併入SEQ ID NO:1內的一個位置處而做出。這些經修飾細胞株也可以包括額外外源性基因用於正向或負向篩選感興趣的表現基因。

【0108】 本發明提供用以修飾CHO細胞基因組的方法，包含將一或多個媒劑引入細胞中，其中該一或多個媒劑包含含有併入用序列的外源性核酸、與SEQ ID NO:1之核苷酸序列中存在的序列同源的5'同源性臂，以及與SEQ ID NO:1之核苷酸序列中存在的序列同源的3'同源性臂。在一些具

體例中，該等方法進一步提供一或多種媒劑，其包含用於在整合位點處之位點特異性DNA切割的核酸酶及組成物。

【0109】 經修飾的細胞株可用作為重組酶媒介之匣式交換(RMCE)用的便利且穩定表現系統。編碼感興趣蛋白質的核酸序列可透過例如RMCE過程簡便地併入至含有SEQ ID NO:1或其表現增高片段(其具有至少一個重組酶辨識位點)之經修飾細胞。

【0110】 重組型表現載體可包含編碼蛋白質之可操作地連結至適當轉錄及/或轉譯調節要素(衍生自哺乳動物、病毒或昆蟲基因)的合成或cDNA衍生的DNA片段。此等調節要素包括轉錄啟動子、增強子、編碼適當mRNA核醣體結合位點的序列，以及控制轉錄與轉譯中止的序列，如下文詳述。哺乳動物表現載體亦可包含非轉錄要素，諸如複製源點、其他5'或3'側翼非轉錄序列，以及5'或3'非轉譯序列(諸如剪接供體與受體位點)。亦可併入幫助辨識轉染體的可篩選標記基因。

【0111】 螢光標記是適當的可篩選標記基因，視情況用於辨識已成功或尚未成功插入及/或取代的基因匣。螢光標記的實例為技藝中熟知，包括(但不限於) 珊瑚(*Discosoma* coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質(例如mKate、mKate2、mPlum、mRaspberry或E2-crimson)。亦參見例如Nagai, T., et al. 2002 *Nature Biotechnology* 20:87-90；Heim, R. et al. 23 February 1995 *Nature* 373:663-664；及Strack, R.L. et al. 2009 *Biochemistry* 48:8279-81。

【0112】 在表現載體中，可用於轉染脊椎動物細胞之轉錄及轉譯控制序列可透過病毒來源提供。例如，常用啟動子與增強子是衍生自病毒，

諸如多瘤病毒、腺病毒2、猴病毒40 (SV40)與人類巨大細胞病毒(CMV)。病毒性基因組啟動子、控制及/或訊號序列可用來驅動表現，前提是此等控制序列與選定的宿主細胞相容。亦可使用非病毒性細胞啟動子(例如β-球蛋白或EF-1α啟動子)，取決於重組型蛋白質要在其中表現的細胞類型。

【0113】 衍生自SV40病毒性基因組的DNA序列(例如SV40源點、早期與晚期啟動子、增強子、剪接、以及聚腺苷酸化位點)可用於提供其他用以表現異源性DNA序列的遺傳要素。早期與晚期啟動子尤其有用，因為它們均是以片段形式簡易得自於SV40病毒，且亦含有SV40病毒性複製源點(Fiers *et al.*, Nature 273:113, 1978)。也可以使用更小或更大的SV40片段。通常納入在SV40複製源點中自Hind III位點往BglI位點延伸的約250 bp序列。

【0114】 先前已說明過用於表現多個轉錄本的雙順向作用表現載體(Kim S. K. and Wold B. J., Cell 42:129, 1985)且可與本發明之表現增高序列(例如SEQ ID NO:1或其片段)組合使用。其它類型的表現載體亦為可用的，例如彼等描述於美國專利第4,634,665號(Axel *et al.*)以及美國專利第4,656,134號(Ringold *et al*)中者。

感興趣的蛋白質

【0115】 可使用適於在真核細胞中表現的任何感興趣的蛋白質。例如，感興趣的蛋白質包括(但不限於)抗體或其抗原結合片段、嵌合抗體或其抗原結合片段、ScFv或其片段、Fc融合蛋白質或其片段、生長因子或其片段、細胞激素或其片段，或細胞表面受體之胞外域或其片段。感興趣的蛋白質可以是由單個次單位組成的簡單多鏈，或由兩個或更多個次單位組成的複合多次單位蛋白質。

宿主細胞及轉染

【0116】 用於本發明方法中的宿主細胞為哺乳動物宿主細胞，包括(例如)中國倉鼠卵巢(CHO)細胞及小鼠細胞。在一個較佳具體例中，本發明提供一種具有SEQ ID NO:1的核酸序列片段，其在CHO細胞中編碼表現增高序列。在SEQ ID NO:1或SEQ ID NO:1的任一片段內可發現一個整合位點。整合位點可以是(例如)位於SEQ ID NO:1或SEQ ID NO:1的任一片段內的整合酶辨識位點。適宜整合位點的一個實例為*Lox*P位點。適當整合位點的另一個實例是兩個重組酶辨識位點，例如選自由*Lox*P位點、*Lox*511位點、*Lox*2272位點、*Lox*2372位點、*Lox*m2位點、*Lox*71位點、*Lox*66位點及*Lox*5171位點組成之群。在其他具體例中，整合位點位於選自由橫跨下列編號位置之核苷酸組成之群的序列內或鄰近選自由橫跨下列編號位置之核苷酸組成之群的序列內的位置：SEQ ID NO:1的10-4,000；100-3,900；200-3,800；300-3,700；400-3,600；500-3,500；600-3,400；700-3,300；800-3,200；900-3,100；1,000-3,000；1,100-2,900；1,200-2,800；1,300-2,700；1,200-2,600；1,300-2,500；1,400-2,400；1,500-2,300；1,600-2,200；1,700-2100；1,800-2050；1850-2050，1,900-2040；1950-2,025，1990-2021，2002-2021及2,010-2,015。在某些具體例中，SEQ ID NO:1內之一個位置處或鄰近SEQ ID NO:1內之一個位置的整合位點選自由橫跨下列編號位置之核苷酸組成之群：SEQ ID NO:1 的 1990-1991、1991-1992、1992-1993、1993-1994、1995-1996、1996-1997、1997-1998、1999-2000、2001-2002、2002-2003、2003-2004、2004-2005、2005-2006、2006-2007、2007-2008、2008-2009、2009-2010、2010-2011、2011-2012、2012-2013、2013-2014、2014-2015、2015-2016、2016-2017、2017-2018、2018-2019、2019-2020，及2020-2021。

【0117】 本發明包括一種經本發明表現載體或mRNA轉染的哺乳動物宿主細胞。儘管可使用任何一種哺乳動物細胞，但在一個特定具體

例中宿主細胞為CHO細胞。

【0118】 經轉染細胞包括經包含編碼蛋白質或多肽之序列的表現載體或mRNA分子轉染的細胞。所表現蛋白質可被分泌至培養基內，取決於選定而核酸序列而定，但也可留在細胞中或存放在細胞膜內。可採用各種哺乳動物細胞培養系統來表現重組型蛋白質。針對特定篩選或擴增方案而發展的其他細胞株亦可與本文提供之方法與組成物一起使用，前提為已鑑定出與SEQ ID NO:1具有至少80%同源性的目標基因座。一個具體細胞株是命名為K1的CHO細胞株。為達到大量生產重組型蛋白質，宿主細胞株在一些適當情況下可以預先適應生物反應器培養基。

【0119】 技藝中已知數種轉染程序，且如在Kaufman (1988) Meth. Enzymology 185:537中所回顧。所選轉染程序將取決於宿主細胞種類以及GOI的本質，而且可根據慣常實驗來選擇。任何此等程序的基本要求首先是將編碼感興趣蛋白質的DNA引入適當宿主細胞內，且接著鑑定並分離已以相對穩定、可表現的方式併入異源性DNA的宿主細胞。編碼蛋白質之可用以併入宿主細胞基因組或其他功能的mRNA分子可以是暫時的並且因而是限時的。

【0120】 轉染程序以及將多肽或多核苷酸序列引入至細胞的程序可能會有所變化。非限制性轉染方法包括以化學品為主的轉染方法，包括使用脂質體；奈米顆粒；磷酸鈣(Graham et al. (1973). *Virology* 52 (2): 456–67，Bacchetti et al. (1977) *Proc Natl Acad Sci USA* 74 (4): 1590–4與Kriegler, M (1991). Transfer and Expression: A Laboratory Manual. New York: W. H. Freeman and Company. pp. 96–97)；樹枝狀聚合物；或陽離子聚合物，諸如DEAE-聚葡萄糖或聚乙烯亞胺。非化學方法包括電穿孔；Sono-穿孔；及光學轉染。以顆粒為基礎的轉染包括使用基因槍、磁鐵輔助的轉染(Bertram, J.

(2006) *Current Pharmaceutical Biotechnology* 7, 277–28)。病毒性方法亦可用於轉染。mRNA遞送包括使用TransMessenger™與TransIT®的方法(Bire et al. *BMC Biotechnology* 2013, 13:75)。

【0121】 將異源性DNA引入至細胞中的一種常用方法是磷酸鈣沉澱，例如如由Wigler *et al*所述(Proc. Natl. Acad. Sci. USA 77:3567, 1980)。透過這個方法被引入至宿主細胞中的DNA通常經歷重組，使得這個程序可用於共轉染獨立基因。

【0122】 聚乙烯誘導之細菌原生質體與哺乳動物細胞融合(Schaffner *et al.*, (1980) Proc. Natl. Acad. Sci. USA 77:2163)是引入異源性DNA的另一種可用方法。原生質體融合程序常會產生多個質體DNA複本併入哺乳動物宿主細胞基因組中，且這個技術需要與GOI在相同質體上的篩選和擴增標記。

【0123】 電穿孔亦可用於將DNA直接引入宿主細胞的細胞質中，例如如Potter *et al.*所述(Proc. Natl. Acad. Sci. USA 81:7161, 1988)或由Shigekawa *et al.*所述(BioTechniques 6:742, 1988)。不同於原生質體融合，電穿孔不需要篩選標記與GOI在相同質體上。

【0124】 其他可用於將異源性DNA引入哺乳動物細胞中的試劑已被描述，諸如Lipofectin™ Reagent與Lipofectamine™Reagent (Gibco BRL, Gaithersburg, Md.)。這些商用試劑均用於形成脂質-核酸複合物(或脂質體)，其在施加至培養細胞時會促進核酸吸收至細胞內。

【0125】 在一個具體例中，將一或多個多核苷酸引入至細胞中是由電穿孔、由細胞質內注射、由病毒感染、由腺病毒、由慢病毒、由逆轉錄病毒、由轉染、由脂質-媒介之轉染媒介或經Nucleofection™媒介。

【0126】 用於擴增GOI的一個方法對於表現重組型蛋白質來說也

是需要的，而且通常涉及使用篩選標記(在上文Kaufman中回顧)。對細胞毒性藥物的抗性是作為篩選標記最常使用的特徵，而且可能是顯性性狀(例如可獨立於宿主細胞類型來使用)或隱性性狀(例如可在特定宿主細胞類型中使用，在不論篩選哪種活性都是缺乏的)。數種可擴增標記適宜用於表現本發明的載體(例如，如Sambrook, Molecular Biology: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, 1989; pgs 16.9-16.14中所述)。

【0127】 在藥物抗性哺乳動物細胞中可用的基因擴增用可篩選標記顯示於上文Kaufman, R. J.的表1中，且包括DHFR-MTX抗性、P-醣蛋白及多重藥物抗性(MDR)-各種親脂性細胞毒性劑(例如阿黴素、秋水仙素、長春新鹼)，以及腺苷酸去胺酶(ADA)-Xyl-A或腺苷酸與2'-去氧輔間型黴素。

【0128】 其他顯性可篩選標記包括微生物衍生的抗生素抗性基因，例如新黴素、康黴素或潮黴素抗性。但是，這些篩選標記未曾顯示為可擴增(Kaufman, R. J.，上文)。針對哺乳動物宿主，存在有數種適宜篩選系統(Sambrook上文, pgs 16.9-16.15)。採用兩種顯性可篩選標記的共轉染程序亦已被描述(Okayama and Berg, Mol. Cell Biol 5:1136, 1985)。

【0129】 先前所述或技藝中熟知的可用調節要素亦可被納入要用來轉染哺乳動物細胞的核酸構築體中。針對其中用途所選轉染程序及所篩選要素將取決於所使用的宿主細胞類型。基於所用細胞培養系統的需求，彼等習於技藝者將了解到有許多不同程序及宿主細胞，且可選出適當系統用以表現所要蛋白質。

【0130】 本發明的其他特徵將在以下用以說明本發明而非意欲限制本發明所提供的例示性具體例說明中變得清楚。

【圖式簡單說明】

【0131】 圖1A及1B. 圖1A：可操作構築體的簡圖，採用將表現

GOI (例如多鏈抗體)以及多個篩選標記複本之核酸分子隨機引入細胞基因組(例如用於鑑定目標基因座的CHO基因組)中。例示構築體包括：重鏈(HC)；第一複本篩選標記，諸如潮黴素抗性基因(Hyg)；第一複本輕鏈(LC)；第二複本篩選標記(例如Hyg)，第二複本輕鏈(LC)；第三複本篩選標記(例如Hyg)。圖1B：實例螢光標記供體載體，其經由同源性重組用以整合至經鑑定為衍生自SEQ ID NO:1之SEQ ID NO:1 5'同源性臂與3'同源性臂之原基因座中。

【0132】 圖2A至2C說明可操作地連結至感興趣基因的SEQ ID NO:1的基因座(*LOCUS 1*)相較於未可操作地連結至*LOCUS 1*而是連結至對照基因座的相同GOI的mRNA表現增高。圖2A：依據qPCR對基因複本數之分析顯示編碼感興趣之抗體基因(亦即一個重鏈(HC)與兩個輕鏈(LC)可操作地連結至對照基因座vs. *LOCUS 1*之細胞所表現的相同基因複本數。圖2B：依據qPCR對相對mRNA位準之分析顯示在*LOCUS 1*中表現之GOI的mRNA位準相較於對照基因座mRNA更高。圖2C：表現在*LOCUS 1*中之GOI的細胞的生物反應器蛋白質效價相較於表現在對照基因座中之相同GOI的細胞所生產的生物反應器蛋白質效價高出3倍。

【0133】 圖3A及3B說明包含併入*LOCUS 1* (例如要與eYFP交換之兩側為lox位點的mKate以及GOI)處之螢光標記以及GOI的實例匣與併入對照基因座(與如dsRed2之不同螢光標記交換，兩側為lox位點)處的相同匣相比較，其中此整合採用Cre重組酶以及重組酶媒介之匣式交換(RMCE)。在實驗中，此等匣被用來測量重組效率及GOI轉錄。

【0134】 圖4顯示感興趣基因(GOI)如在表現*LOCUS 1* (SEQ ID NO:1)中之GOI的CHO細胞池測得的mRNA位準相較於在相同調節條件下，但表現併入對照基因座(亦即EESYR)內之相同GOI的CHO細胞池的mRNA更高。此數據顯示經改造LOCUS 1支持RMCE以及高位準轉錄。

【實施方式】

實例

【0135】 提出下列實例以提供彼等習於技藝者如何製作並使用本文所述方法及組成物，且不意欲限制本發明範疇。已盡力確保所用數字(例如數量、溫度等)的準確性，但應解釋一些為實驗誤差與偏差。除非另有指明，否則份為重量份、分子量為平均分子量、溫度為攝氏度，而壓力為大氣壓或近大氣壓。

實例1. 鑑定感興趣之基因座並特徵鑑定整合位點

【0136】 以兩個含有抗體序列及可篩選抗生素抗性基因作為可篩選標記的質體轉染CHO K1細胞。在抗生素存在下，透過擴增細胞來實施穩定轉染株的篩選。表現大量抗體的個別細胞純系是使用FASTR®選別技術來分離(參見美國專利第8673589B2號)。鑑定展現出最高抗體表現量的數個純系。

【0137】 這些純系的基因組DNA是使用Covaris Adaptive Focused Acoustics (AFA)™技術(Fisher, S. et al. 2011, *Genome Biology* 12:R1)予以片段化。產生DNA庫(Agilent SureSelectXT #G9612A)並且與客製生物素化RNA餌(Agilent SureSelectXT #5190-4811)(對抗被引入至CHO細胞中之整個質體序列所設計)一起培育。含有質體序列的基因組DNA片段是使用磁性卵白素珠粒予以增濃並且經過Illumina MiSeq定序以鑑定質體整合位點。分析含有質體序列與CHO基因組序列兩者的融合序列並與CHO基因組比對。透過南方墨點分析及PCR，之後定序來確認單一整合位點。具有SEQ ID NO:1核苷酸序列的整合位點被鑑定為表現熱點(亦參見GenBank Locus ID No. AFTD01150902.1，nt35529:39558)。分析整合位點以確定其用於進一步生成細胞株的適宜性。需要整合位點是位在不會打斷細胞正常基因組機制(例如轉譯蛋白質或改變細胞表現型)的非編碼區內。

【0138】 由Blat研究(Kent WJ., BLAT - the BLAST-like alignment tool. *Genome Res*. 2002 Apr;12(4):656-64)比對，SEQ ID NO:1與小鼠及人類基因組序列共有非常低的同源性。SEQ ID NO:1對CHO-1[ATCC]refseq_transcript (www.chogenome.org)的序列blast揭示，經鑑定的基因座序列不含有任何已知基因的任何編碼區。SEQ ID NO:4的較廣序列(其含括SEQ ID NO:1)亦經鑑定為適於靶定整合的基因座。

【0139】 確定位在CHO與小鼠基因組非編碼區中的整合位點序列，並進一步用於下述實驗中。

實例2. 外源性DNA有效率地併入宿主細胞整合位點中

【0140】 靶定插入外源性基因至CHO基因組內的特定基因座(經鑑定為SEQ ID NO:1)是透過採用TALE核酸酶(TALEN)來完成。如實例1中，含有隨機併入細胞基因組之抗體重鏈與輕鏈序列的構築體是受到TALEN所靶定。TALEN被靶定至抗體表現構築體的三個相同Hyg基因內之位置(參見圖1A)。Hyg序列的TALEN目標切割位點是以ZiFit.partners.org (ZiFit Targeter Version 4.2)為基礎。TALEN是根據已知方法(Boch J et al., 2009 Science 326:1509-1512)來設計。

【0141】 使用標準Lipofectin程序(LIPOFECTAMINE, Life Technologies, Gaithersburg, Md.)將供體mKate載體(參見圖1B)及TALEN編碼載體轉染至CHO宿主細胞中。培養細胞並且藉由FACS分離並分選具有所要特徵的穩定純系。所要基因座中的單個整合是透過南方墨點與PCR來確認。

實例3. 在感興趣的基因座處透過RMCE的經改造細胞的靶定重組

【0142】 篩選表現大量螢光基因(例如mKate)的CHO細胞株用於分離，其中該基因兩側為感興趣基因座內的lox位點。第二個CHO細胞株表現第二種螢光基因(*dsRed*)，其中該基因兩側為位在對照基因座內的lox位點

(亦即EESYR)(美國專利第8389239B2號，2013年3月5日發證)。

【0143】 經轉染CHO細胞適應生長在無血清生產培養基的懸浮液中。接著在10公分盤中以供體表現載體與編碼Cre重組酶的質體轉染細胞。供體表現載體含有編碼Fc融合蛋白質的感興趣基因，其兩側為Lox位點(參見圖3A或3B)。轉染之後，在具有400 μg/ml潮黴素的培養基中培養細胞兩週，並且使用流式細胞儀分離表現eYFP而非mKate(或在EESYR基因座整合的情況下為dsRed)的細胞。在無血清生產培養基的懸浮培養物中擴增表現eYFP的細胞，並且藉由qRT-PCR使用標準程序針對各個編碼Fc融合蛋白質的細胞池測定mRNA位準(參見圖4)。

【0144】 比較細胞池之間的重組交換效率(表現來自供體匣標記(亦即eYFP)之存活細胞群百分率，因為與紅色標記(亦即mKate或dsRed)交換)(表1)。在各基因座處觀察到高重組交換效率。

表1：重組效率

	紅色標記	交換效率(%) (紅色標記+/eYFP-)	隨機整合 (%) (紅色標記+/eYFP+)
LOCUS 1 (SEQ ID NO:1)	mKate	72	27
對照基因座 (EESYR)	dsRed	92	7

【0145】 在具有經改造LOCUS1的細胞池中，觀察到比對照基因座有更高比率(1.5倍高)的轉錄(圖4)。

【0146】 本發明不囿於本文所述特定具體例的範疇。甚至除本文所述的彼等修飾以外，本發明的各種修飾對於習於技藝者來說將因為先前說明與隨附圖式而變得清楚。此等修飾意欲落在隨附申請專利範圍的範疇內。

【符號說明】

無

【生物材料寄存】

國內寄存資訊【請依寄存機構、日期、號碼順序註記】

無

國外寄存資訊【請依寄存國家、機構、日期、號碼順序註記】

無

【序列表】(請換頁單獨記載)

序列表

<110> 美商再生元醫藥公司(Regeneron Pharmaceuticals, Inc.)

<120> 新穎之CHO整合位點及其用途

<150> 62/067,774
<151> 2014-10-23

<160> 6

<170> PatentIn第3.5版

<210> 1
<211> 4001
<212> DNA
<213> 灰倉鼠

<400> 1
ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtgtaatat 60

tattcagttt ttaagaaaaa tgaaattatg taataagcat gtaaatggat atatcttgaa 120

acaaccattc cccattatat tacctaaaca ttgaaagtcc aaaatcatat gatcttttta 180

gtggatctac taatcttttg ctatatgtat tttattgaac tacccatgga tgtgagataa 240

ttggtaacaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca 300

ttttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg 360

atggttagag aaggcaatat gaggagggat aactagcact tagggccttt tgaaaaagac 420

atagagaaaa tactattgta gaaacttcct ataattggtg tatagttata tacaccaaag 480

agctcagatg gagttaccct ataatggaaa tattaactac tttttatcac tgtgataaaa 540

catcctgaac agagcaacat agattgggaa gcatttactt tggcttacag ttctaacggg 600

ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga 660

agcaggtgag agctcacatc ttgaagtgta agaatgtagc agagagaaca aactgcaaat 720

gaccagaaaa tgcttttgga tcagagccca taccctctg actgacttct ccagaaattc 780

tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta 840

ggggtatttc ttattcaaac cactacaatg gggtgggggg agcaatcctc caagtaggca 900

ctacacacag acaaataaaa actctagtaa ctggaatgga ttgacttatt tgaattactt 960

gccagtggag ctacatagag cacaattatt gtatttaaat taccctttat gatcttacaa 1020

aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatggtg 1080

atatctagtt gttcttcaac tggaaacttc atgctttctg cccagcattc atgttgctgg 1140

aaagagcaat gtacactacc agtgtagaaa ttaaatcatc aatcttatca agatgtggat 1200

cctataagtt acaataaaaa ttagcctgat aagatatccc caccagaaga atattcacat 1260

aaatgctatg ggagcaacaa gctattttct aaattagctt taatcctatt ctacaagaga 1320

gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc 1380

caatgggaga tcctaatatt attgttctac aaaatgaaaa taactcctaa tgacttgttg 1440

ctgcagtaat aagttagtat gttgctcaac tctcacaaga gaagttttgt cttacaataa 1500

atggcaatta aagcagcccc acaagattta tatcataccg atctcctcat ggcctatgca 1560

tctagaagct aggaaacaaa gaggacccta agagagacat acatggtccc cctggagaag 1620

gggaaggggg caagacctcc aaagctaatt gggagcatgg gggaggggag agggagttag 1680

aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt 1740

caagagaaga tagaggagag caagaaaaga gataccatag tagagggagc cttgtatgtt 1800

taaatagaaa actggcacta gggaattgtc caaagatcca caaggtccaa ctaataatct 1860

aagcaatagt cgagaggcta ccttaaaagc ctttctctga taatgagatt gatgactacc 1920

ttatatacca tcctagagcc ttcatccagt agctgatgga agcagaagca gacatctaca 1980

gctaaacact gagctagttg cagacaggga ggagtgatga gcaaagtcaa gaccaggctg 2040

gagaaacaca cagaaacagc agacctgaaa aaaatgttgc acatggaccc cagactgata 2100

gctgggagtc cagcatagga cttttctaga aaccctgaat gaggatatca gtttggaggt 2160

ctggttaatc tatggggaca ctggtagtgg atcaatattt atccctagtt catgactgga 2220

atttgggtac ccattccaca tggaggaatt ctctgtcagc ctagacacat gggggaggtt 2280

ctaggtcctg ctccaaataa tgtgttagac tttgaagaac tcccttgaga agactcaccc 2340

tccctgggga gcagaaaggg gatgggatga gggttggtga gggacaggag aggaggggag 2400

ggtgagggaa ctgggattga caagtaaatg atgcttgttt ctaatttaaa tgaataaagg 2460

aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg 2520

actataaaga aacactatta tctaaataaa aacatgtcag aagcacacat gaacttatag 2580

tgtttatgaa agtatgtata ataactacat aatctcaagc caagaaaaaa atatcatctt 2640

tcagtgatga aggtgatttt atttctccca gaattaaagc caaagaccta atgaaagtaa 2700

ttatcttcaa aaggttgaaa atacatactt tgcaatacac agatctgcct agaaatctca 2760

tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca 2820

gtttgtagat aaactcacaa tgtatcattt ctttttattt tttgaccaaa cagcttctca 2880

tctgttattc agaataattc ctcgatggca ggatatccat cccaattggg ggaaggggag 2940

aatttgaaga aaacctagac cacatacata tttgccattg ggaaacaaag tctaaaatga 3000

tgttgttcac atcttctcta ctagtcctct ccccgtccca aagaaccttg gtatatgtgc 3060

ctcattttac agagagagga aagcaggaac tgagcatccc ttacttgcca tcctcaaccc 3120

aaaatttgca tcattgctca gctctgccct tctcatatga cagttacaag tcaaggcttc 3180
caaagtccct ctgtcatgtt tggtgtcaat agtttataca gatgacttca tgtcttcata 3240
tctaatgtct tatatagatt aatattaaac aatgttattt ctctaaccac attttaaatt 3300
aatttaaaaa tccattaatt gtgtctataa aatgcagaca gagtgctgag acacaatata 3360
agcctgatga tctgaatttg aaactcacac ccaccacatg gagaatcaac ttccaaaaat 3420
tttcctatta cttccacact tacaccattg tacaaacaca ataataatga acaaaatgaa 3480
atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc 3540
agcttaagct tgctttatgg ttacacttta ccatcttcca ttaattataa ggacttcaat 3600
catggcagaa ctatgctgtt attgtctcag tgtaacctaa ccaggtgttc cagatgttct 3660
taatgtggac acctaaacta tttgatattt gggttaagat ctttccctct ttcagaagaa 3720
acctcaggac agagggaatc ttgtctttta attttgagtc tgtagacttt ttccatttca 3780
aatatacatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt 3840
aggttcaata ttaagcttca atattatcat ggaatcgcct gttatacact gagtgtttgg 3900
caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agtttagctc 3960
aaataaatct aagcaatcca ctctagaatt aaatagtttc c 4001

<210> 2
<211> 1001
<212> DNA
<213> 人工序列

<220>
<223> 合成多核苷酸

<400> 2
tacccttat gatcttacaa aacttgacag taagatcata ttgctaaaga aaccacatat 60
ttgaatcagg gaacatggtg atatctagtt gttcttcaac tggaaacttc atgctttctg 120
cccagcattc atgttgctgg aaagagcaat gtacactacc agtgtagaaa ttaaatcatc 180
aatcttatca agatgtggat cctataagtt acaataaaaa ttagcctgat aagatatccc 240
caccagaaga atattcacat aaatgctatg ggagcaacaa gctattttct aaattagctt 300
taatcctatt ctacaagaga gaatccatat ctagaatagt tatagggatc aagaacccat 360
ggcttgattg gtcataggcc caatgggaga tcctaatatt attgttctac aaaatgaaaa 420
taactcctaa tgacttgttg ctgcagtaat aagttagtat gttgctcaac tctcacaaga 480
gaagttttgt cttacaataa atggcaatta aagcagcccc acaagattta tatcataccg 540
atctcctcat ggcctatgca tctagaagct aggaaacaaa gaggacccta agagagacat 600
acatggtccc cctggagaag gggaaggggg caagacctcc aaagctaatt gggagcatgg 660

gggaggggag agggagttag aagaaagaga aggggataaa aggagggaga ggaggacaag 720

agagagaagg aagatctagt caagagaaga tagaggagag caagaaaaga gataccatag 780

tagagggagc cttgtatgtt taaatagaaa actggcacta gggaattgtc caaagatcca 840

caaggtccaa ctaataatct aagcaatagt cgagaggcta ccttaaaagc ctttctctga 900

taatgagatt gatgactacc ttatatacca tcctagagcc ttcatccagt agctgatgga 960

agcagaagca gacatctaca gctaaacact gagctagttg c 1001

<210> 3
<211> 1001
<212> DNA
<213> 人工序列

<220>
<223> 合成多核苷酸

<400> 3
caaagtcaag accaggctgg agaaacacac agaaacagca gacctgaaaa aaatgttgca 60

catggacccc agactgatag ctgggagtcc agcataggac ttttctagaa accctgaatg 120

aggatatcag tttggaggtc tggttaatct atggggacac tggtagtgga tcaatattta 180

tccctagttc atgactggaa tttgggtacc cattccacat ggaggaattc tctgtcagcc 240

tagacacatg ggggaggttc taggtcctgc tccaaataat gtgttagact ttgaagaact 300

cccttgagaa gactcaccct ccctggggag cagaaagggg atgggatgag ggttggtgag 360

ggacaggaga ggaggggagg gtgagggaac tgggattgac aagtaaatga tgcttgtttc 420

taatttaaat gaataaagga aaagtaaaag aagaaaagaa aacaggccaa aagattataa 480

aagacagagg tggtgggtga ctataaagaa acactattat ctaaataaaa atatgtcaga 540

agcacacatg aacttatagt gtttatgaaa gtatgtataa taactacata atctcaagcc 600

aagaaaaaaa tatcatcttt cagtgatgaa ggtgatttta tttctcccag aattaaagcc 660

aaagacctaa tgaaagtaat tatcttcaaa aggttgaaaa tacatacttt gcaatacaca 720

gatctgccta gaaatctcat gttcacaata cacatgatgc tcaattgaat tccattcaat 780

gttacagttt agataaacag tttgtagata aactcacaat gtatcatttc tttttatttt 840

ttgaccaaac agcttctcat ctgttattca gaataattcc tcgatggcag gatatccatc 900

ccaattgggg gaaggggaga atttgaagaa aacctagacc acatacatat ttgccattgg 960

gaaacaaagt ctaaaatgat gttgttcaca tcttctctac t 1001

<210> 4
<211> 14931
<212> DNA
<213> 灰倉鼠

<220>
<221> 各種特徵
<222> (2176)..(2239)
<223> n為a、c、g、t或核苷酸缺失

<400> 4

catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca atttttatct 60

ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttctttc 120

caagtaataa acacatacac tgaaattttg gttcttgtgg ataattttaa tgaaacagga 180

aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac 240

attgatggat catgtagtga aatgtatttt taggtatcta aggaattttg gcttcgtttt 300

gtgcttgttg acactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc 360

ttttaccagt atttgtccat ttgcatttte tttattattc atggctgctg ttctagaaag 420

tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatcagtg tcttaacacc 480

tctctgagta catgttggcc aatgtcgttt ctagacccat ctattcttgc ttgacttatc 540

ctggtacatg cctgccaaga aatttctcct catcctttct gtctcttcac tgatttactt 600

gatgtgtgga tttcacattg atcatatgga aatagaagat acaattttct ttattcacag 660

tttggaagac tttcaatctc atagatcatc attatttttt gctactgttc cctatgctat 720

ggtgaaattt ccatttgaat aattgcttaa acaattaaca agaaagaatc tatttttact 780

tgcaataact tccatttcag aacatttact acactgttac tatatccaaa aactagtttt 840

atatatcatg tgagaaatga ctaattcata atttggccat gacatttttt tcagaaacag 900

aaaaagtgac caatacatac acaatgctat aaatattaag acttcagcaa attaaatatt 960

tattcatgat atcacataaa attcatttat tatgttttat ttaaatgtgt ttttaaaaca 1020

gtggtatcac taaatattaa gttagatgtg tttatgtgct taatgaattt atattttaga 1080

atgttataag ttgtatatag tcaaatatgt aataaatttt atttttttagg tctttctcat 1140

taaggtattt taattttggg tccctttttcc agagtgactc tagctcatga tgagttgaca 1200

taaaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat ctttgcactg 1260

aagtttgagt cagttcatac atttagtact tgggaagtac attaagctaa ctttcattgc 1320

tctggcaaaa tgctcgataa gataagagtc tattgtggaa agccatggca gcaggaaagt 1380

aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tgctggtatc 1440

caacattttt tgctgttcat tttctctaga accctagtcc ataaagatgt atgacttgca 1500

ttcaaaatgc gtccccttca gttgttcaac ttttctgtaa atatcctttc aggcatgtct 1560

agaagattgt ttcgcaaata cttctcaatc cattcaagtt gatagtgcag attaatcact 1620

gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat 1680

caataagtaa atcaaagtgt agcttttgcc tttaacattg ccagacttat gtaatgttct 1740

gcacgttctt cctccatcac tttttattct aatggtgttt ccttgacatt gaatcacgct 1800

gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta 1860

gatttactat tttacttaga atttttttata attgagagaa tataatattt tcacagttat 1920

ctatctgctg taaatagagg attttaaaaa aaatctctat aacttttttt tacaacacac 1980

agtaaaatta agttaaaatt taataaagtc actatgttga tttcaaagtg tgctacgccc 2040

acggtggtca cgcaggtgta gcagaagatg ccactaaggt gggctaaggc cgatgggttg 2100

gggtctgcgc tccctggaga tgagccccag gcggttccct ggcaatcagc tgcgatcatg 2160

atgcccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220

nnnnnnnnnn nnnnnnnnnc tgggtgactt tatggaaaga atttgataga tttcatgatg 2280

tagaagaatt ttattaggct tattttacag gagactaaga ccctgggacc taaagatatc 2340

tgggtcctga gaatcaggaa atgggtagag acgtggttga tggtatgaga cagattttag 2400

agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa 2460

caattatgct gttctttttc tttctgttgt atgatctgat gatgtagccc ccttgccaag 2520

ttccctgatc ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag 2580

catattcaac tacattgagt gtgtctgtct gtcatttcct cgccgattcc tgatttctcc 2640

ttgagccttt tcccttgttc tccctcggtc ggtggtctcc acgagaggcg gtccgtggca 2700

aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaaatga aaaattaaaa 2760

taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaaacaata ctatataata 2820

ttttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattctttc 2880

aataataatt tcaataaaaa aagtttcccc atgatagtag aaaataataa catatgtatc 2940

tatctattta tttaactaca catatatagc atttgtttca actaaaataa atgaatgagc 3000

aaagcaccta agtaattggt gtctattata tttatgaagc caatagtttc aaataaatta 3060

tcatgcataa ggaggtattg caaatgttaa acctttttttg aaacagatat tcccagttac 3120

agaaattata atttctaatc tttcctataa gtagaatgat gataattaat ataggccatt 3180

tgtaaataat gttcagatta aaatattctc tatttcacta gagaagaatg atattaaatg 3240

tattatattt tatttcccat tttgtttgca ccactattct atatccctca gcagtttaaa 3300

tttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggcactaaaa ttagaataat 3360

ttaatataaa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct 3420

tttaattgtc caagaatact ttttcttcta ttttaggact gatcagaccc aggactaata 3480

ttttatatgt actaattcta tgtaccaaaa tatgttatta tctcatgaat tctgtctcaa 3540
tattgaggta ataaaaatag tccatcatga actttaaaat taaaataatg attaattaat 3600
ttttattcat attttgtttg tatgaatggt tatacatcac atgtgtgcct ggtgactgtg 3660
aatgtcagga gaaggtatga aagccactgg aattggaata agagataata tttgagatgt 3720
tatgtgggtg ctgagaatta gacgcaagcc atcttcaaga atagccagca tactatacca 3780
ctgagtaatc cattcatccc tcaataatta tctttgtaga cagtaaatat atttctaaac 3840
tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac 3900
ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca 3960
cttactattt ctaaagctgt ccaatggggg ccagtcccca gtcagcaccc agtgagataa 4020
tccatgaatg catttatatc ttaggaaaaa ttcttatcta tgtagtattt agaacatttt 4080
catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctctttaatt 4140
catctagaaa aaaaaaacct ctcaggaaaa tctctcttgc tctcctccca atgctctatt 4200
cagcatcttc tccctactta attctagatc tttttctcta tgcctccttg ctgctgccct 4260
gctggctctg ctctatgcct ccccatgtca cttttctttg ctatctcacc gttaccttct 4320
ctgcctcact ctctgccttc ttctctgctt ctcacatggc caggctctgg acaattatag 4380
ttatatgtta cattctcata acacatgata tgtcacatag tttctctcag gctagggata 4440
tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggttc 4500
ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat 4560
tttcttcaaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat 4620
ttgtctttag aggcaaccac atattgtctg cagggcctgc tttgaatttg ctctgaagtt 4680
agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaaa atatcctagc 4740
atttcaagtc atcaaaatta tatggagagt ataaatcatc cttctgacta ttcatagtca 4800
tatttgtgtc caccaagtat aaaacacact accaaagggc tgtggaaaaa atcgccataa 4860
ctgttcttat tagggaggca tagcagtggt acctgaggaa gttacagcaa caaccagtca 4920
tccagtcaat aaccccatgg ctttgccact tggaggtacc caataatgtt tggctttgcc 4980
gagtaggact ccaacaaatt cagagggtca atttttaaat gctggttgtc actgctgaac 5040
agtcccattg ccctctgcat aattccacaa tggaaagctt tttacactga ttgccaatca 5100
ttaaacagcc tactcagcat aaacaggtat gatattattc tgcattttgt tacattacta 5160
gatgaattcc tatttcttcc tacaatagtg gaactgaaaa aagatacaca atcatactac 5220
ccctctacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg 5280
accaaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa 5340

ggcggactta ttttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata 5400

taataaaggg acaataaatc ttcaagaaac ttacccctac tgaattaaaa tattaaagaa 5460

ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt 5520

aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa 5580

ccaagtggga atcatatcag aaagcagtca aaagcttata ttacaagtaa cagatgcttg 5640

gttatatgac ctcccagagc ttgactgtct atacacaaaa agtggtgtta ataaaactgt 5700

aatttgggct atgtttttt aaatggcttc accaacatga aaggaaggga atgagcatgt 5760

catggatgct tagagattat gcttccagca agaagaattg agctttggct cttattacag 5820

aaacatgaca aggtgtgagt tttatttatt agaaattata taatatttta agctggggac 5880

taaaaatttt attgaaacaa acaggcaagg gataggcatg tactagaagc aaaaatagga 5940

tgtcaatgct gtaatgttat tttttggacc aaaatagtat ttcctataga aatgacaatg 6000

atcttaggtt attattcttc ataaagatga caagttcaca agatatccta gttcattaaa 6060

atcgttttag tcatttaata gagtgctgtg atagattaca caaaggaaag cacttacgat 6120

gagaaataat gatatccaca attattttct taattcttag aaacattcta ttgttatatc 6180

tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata 6240

ttgtctatgg tgacatttca aagaacatgt gacgtacagt gtagcacaga taaagaacat 6300

aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt 6360

gtgtgcacta ccaaaggatg tactaatgct cacgacactc ccctatgtca ccctttgttc 6420

atcattacat cataggtcta ttttgtttgc ttttgaaatc tagaccaagt cttttgtgtc 6480

tttccaagca cagagctcat taatttacct catagacttg ttaaacttct tctggttcat 6540

caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg 6600

gataattttt acataaaaca tgcatacaag taagattatt cagactgaac atgaatttta 6660

gagaaatcag gaaggagtat atgggagtgg ttggagtgag actagagaaa tgtaattaaa 6720

ctataatctc aatacaaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga 6780

aacatcagtc ttcaaattgg aaagatattt ttactaggaa aatgtctggt agatggttat 6840

tatctagaaa acacaaaaat tagaaaacgg taaactttaa taaaaagaat aatacaatga 6900

gactacatga aaagttctta actaatgaaa caaatatctt gaaacttttt tcttaaaagt 6960

ttaatatcaa taaccatcat ggaaattcaa attaaaacta tttacatatt acccctgaaa 7020

taataactaa tacccaataa aaataatata aacaaaaaat ggcaatgcat gccatcatgg 7080

atttgggaga gagaatgttc attgcagttc tgaatggata ctggtgccac cacggtgaaa 7140

atctctgtat aggtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat 7200

ttttcaagca aatacccaaa ggactctacc tgactgcaga gacactttct cataaaatat 7260
tattgttgat ctattcataa tatctggaaa atagaaacag ccaagatgcc catcaactga 7320
ttaatagatg ataaaattat tgtacatttc agtgtaatat tattcagttt ttaagaaaaa 7380
tgaaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat 7440
tacctaaaca ttgaaagtcc aaaatcatat gatcttttta gtggatctac taatcttttg 7500
ctatatgtat tttattgaac tacccatgga tgtgagataa ttggtaacaa cagcacatgg 7560
gagagcatgg gatcattcaa ggaagattag agagaatgca ttttttagga gataatggag 7620
gagcaataga aaggattaaa tgaggttact gatgaaagtg atggttagag aaggcaatat 7680
gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta 7740
gaaacttcct ataattggtg tatagttata tacaccaaag agctcagatg gagttaccct 7800
ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat 7860
agattgggaa gcatttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag 7920
aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc 7980
ttgaagtgta agaatgtagc agagagaaca aactgcaaat gaccagaaaa tgcttttgga 8040
tcagagccca taccctctg actgacttct ccagaaattc tgaacaaata aaactcccca 8100
aacagagcca taactgaagg tccagtgtct gagactacta ggggtatttc ttattcaaac 8160
cactacaatg gggtgggggg agcaatcctc caagtaggca ctacacacag acaaataaaa 8220
actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtggag ctacatagag 8280
cacaattatt gtatttaaat taccctttat gatcttacaa aacttgacag taagatcata 8340
ttgctaaaga aaccacatat ttgaatcagg gaacatggtg atatctagtt gttcttcaac 8400
tggaaacttc atgctttctg cccagcattc atgttgctgg aaagagcaat gtacactacc 8460
agtgtagaaa ttaaatcatc aatcttatca agatgtggat cctataagtt acaataaaaa 8520
ttagcctgat aagatatccc caccagaaga atattcacat aaatgctatg ggagcaacaa 8580
gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt 8640
tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaatatt 8700
attgttctac aaaatgaaaa taactcctaa tgacttgttg ctgcagtaat aagttagtat 8760
gttgctcaac tctcacaaga gaagttttgt cttacaataa atggcaatta aagcagcccc 8820
acaagattta tatcataccg atctcctcat ggcctatgca tctagaagct aggaaacaaa 8880
gaggacccta agagagacat acatggtccc cctggagaag gggaaggggg caagacctcc 8940
aaagctaatt gggagcatgg gggaggggag agggagttag aagaaagaga aggggataaa 9000
aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag 9060

caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta 9120
gggaattgtc caaagatcca caaggtccaa ctaataatct aagcaatagt cgagaggcta 9180
ccttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcctagagcc 9240
ttcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagttg 9300
cagacaggga ggagtgatga gcaaagtcaa gaccaggctg gagaaacaca cagaaacagc 9360
agacctgaaa aaaatgttgc acatggaccc cagactgata gctgggagtc cagcatagga 9420
cttttctaga aaccctgaat gaggatatca gtttggaggt ctggttaatc tatggggaca 9480
ctggtagtgg atcaatattt atccctagtt catgactgga atttgggtac ccattccaca 9540
tggaggaatt ctctgtcagc ctagacacat gggggaggtt ctaggtcctg ctccaaataa 9600
tgtgttagac tttgaagaac tcccttgaga agactcaccc tccctgggga gcagaaaggg 9660
gatgggatga gggttggtga gggacaggag aggaggggag ggtgagggaa ctgggattga 9720
caagtaaatg atgcttgttt ctaatttaaa tgaataaagg aaaagtaaaa gaagaaaaga 9780
aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta 9840
tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata 9900
ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgatga aggtgatttt 9960
atttctccca gaattaaagc caaagaccta atgaaagtaa ttatcttcaa aaggttgaaa 10020
atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg 10080
ctcaattgaa ttccattcaa tgttacagtt tagataaaca gtttgtagat aaactcacaa 10140
tgtatcattt ctttttattt tttgaccaaa cagcttctca tctgttattc agaataattc 10200
ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac 10260
cacatacata tttgccattg ggaaacaaag tctaaaatga tgttgttcac atcttctcta 10320
ctagtcctct ccccgtccca aagaaccttg gtatatgtgc ctcattttac agagagagga 10380
aagcaggaac tgagcatccc ttacttgcca tcctcaaccc aaaatttgca tcattgctca 10440
gctctgccct tctcatatga cagttacaag tcaaggcttc caaagtccct ctgtcatgtt 10500
tggtgtcaat agtttataca gatgacttca tgtcttcata tctaatgtct tatatagatt 10560
aatattaaac aatgttattt ctctaaccac attttaaatt aatttaaaaa tccattaatt 10620
gtgtctataa aatgcagaca gagtgctgag acacaatata agcctgatga tctgaatttg 10680
aaactcacac ccaccacatg gagaatcaac ttccaaaaat tttcctatta cttccacact 10740
tacaccattg tacaaacaca ataataatga acaaaatgaa atgaaataaa aaattaagtc 10800
tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg 10860
ttacacttta ccatcttcca ttaattataa ggacttcaat catggcagaa ctatgctgtt 10920

attgtctcag tgtaacctaa ccaggtgttc cagatgttct taatgtggac acctaaacta 10980
tttgatattt gggttaagat ctttccctct ttcagaagaa acctcaggac agagggaatc 11040
ttgtctttta attttgagtc tgtagacttt ttccatttca aatatacatg aaacaagtga 11100
tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca 11160
atattatcat ggaatcgcct gttatacact gagtgtttgg caataaggga tttttagaag 11220
aaggagtttt tattctcaac aggttcctta agtttagctc aaataaatct aagcaatcca 11280
ctctagaatt aaatagtttc ctaagggcac agctatgaat agagctcaat ttacatataa 11340
aattttgttc accatttatg tcattccagt tttcattagt acaaggaaaa tacaaaatat 11400
ttagatgtca atatcaagtg aatagttcat ctccttttt aatatatatc acctaaatca 11460
ccattttctc agaaaaatct ggcctgaagt tctgtctgga acttcaacat gaaaaatatg 11520
cacagcttgc tattataaat cctagttgat ttttaagatt catgtctggt gtctgactca 11580
gaggggccag aggctagaca aatatttttt gaatcttcat tgtgaagatt tttaatgatt 11640
attttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg 11700
aactactttg tgcttaaaat gttagttccc tatcataaat gataggtgat aagtgtatgt 11760
ttaatacttt ccctctgagc tatattcatg tactagagaa ttattttaaa catgaaaaga 11820
ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga 11880
gcaacgtttt tcttctacag aggatgcttt gctgccaagc aacctggttg tgtggaaatg 11940
ttcctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg 12000
ttagagcttg gtccttggag tattatcttt taccagaaaa ttcatagtat tctttcaata 12060
actaacaact aaacttttcg ataaaaaaga attggaattt caattttaaa gcctgagtaa 12120
aattcttgtg aatcaggata ttttatttta agtcttatct tttaaaaagt tattttattt 12180
tttaaaaaat tataatatac tttcataatt tccctccttc acttttcttt acaaacactt 12240
ctatagatca ccatgtgttt tttttttac atttatggcc tctttctgtt cattgttatt 12300
acatacaaat agtcttgcct atagaagaac accacaattt gttacctgat aacaaattat 12360
caacccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa 12420
gatatatgtg tgtgcacata tatagataca catatatgta ggatttttaa ttttagattt 12480
tagacatcaa aattatttat atgactgaga aactagacac tataaatgag cattcagtat 12540
tcaacaccgt gattttagat attgtcacaa tgacagaaaa ttttcttata gaaaatttta 12600
agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt 12660
tagtttataa taaaaagtac atataattaa aatggttggc acaaaacaac atttgagcat 12720
ttttcctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa 12780

tgaaaacaaa atttaaacta aatgcctaat ctagcctgat aacattttta tgaatgaaat 12840

tattcaatag tgttatcaat taggggccca aaacttttcc taaaataaaa cttttaattt 12900

ttttccattt ttatttaaat tagaaacaaa attgttttac atgtaaatca gagtttcctc 12960

accctccect tctccctgtc cctcactaac accctacttg tcccatacca tttctgctcc 13020

ccagggaggg tgaggccttc catggggaaa cttcagagtc tgtctatcct ttcggatagg 13080

gcctaggccc tcacccattt gtctaggcta aggctcacaa agtttactcc tatgctagtg 13140

ataagtactg atctactaca agagacacca tagatttcct aggcttcctc actgacaccc 13200

atgttcatgg ggtctggaac aatcatatgc tagtttccta ggtatcagtc tggggaccat 13260

gagctccccc ttgttcaggt caactgtttc tgtgggtttc accaccctgg tcttgactgc 13320

tttgctcatc actcctccct ttctgtaact gggttccagt acaattccgt gtttagctgt 13380

gggtgtctac ttctacttic atcagcttct gggatggagc ctctaggata gcatacaatt 13440

agtcatcatc tcattatcag ggaagggcat ttaaagtagc ctctccattg ttgcttggat 13500

tgttagttgg tgtcatcttt gtagatctct ggacatttcc ctagtgccag atatctcttt 13560

aaacctacaa gactacctct attatggtat ctcttttctt gctctcgtct attcttccag 13620

acaaaatctt cctgctccct tatattttcc tctcccctcc tcttctcccc ttctcattct 13680

cctagatcca tcttcccttc ccccatgctc ccaagagaga tgttgctcag gagatcttgt 13740

tccttaaccc ttttcttggg gatctgtctc tcttagggtt gtccttgttt cctagcttct 13800

ctggaagtgt ggattgtaag ctggtaatca tttgctccat gtctaaaatc catatatgag 13860

tgatgtttgt ctttttgtga ctgggttacc tcactcaaaa tggtttcttc catatgtctg 13920

tggatttcaa tagcacaaac aacatacagt atcttggggc aacactaacc aaacaagtga 13980

aagaccagta tagcaagaac tttgagttta aagaaagaaa ttaaagaaga taccagaaaa 14040

tggaaagatc tcccatgctc tttgataggc agaatcaaca tagtaaaaat ggcaatcttg 14100

ccaaaatcca tctacagact caatgcaatc cccattaaat accagcacac ttcttcacag 14160

acctgaaaga ataatactta actttatatg gagaaacaaa agacccagga taggccaaac 14220

aaccctgtac aatgaaggca cttccagagg catccccatc cctgacttca agctctatta 14280

tagagtaata atcctgaaaa cagcttggta atggcacaaa aatagacagg tagaccaatg 14340

gaattgagtt gaaaaccctg atattaaccc acatatctat gaacacctga ctttgacaaa 14400

gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatcgtg ctggcataac 14460

tggatgctgg catgtagaag actgcagata gatccatgtc taatgccatg cacaaaactt 14520

aagtccaaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag 14580

aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca 14640

gtagcacaga caatcagatc aataatcaat aaatgggacc tcctgaaact gagaagcttc 14700

tgtaaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat 14760

tcaccaatcc tatatctgac agagggctgc tctctatttg caaagaacac aataagctag 14820

tttttaaaac accaattaat ccgattataa agttgggtag agaactaaat aaagaattgt 14880

taacagagca atctaacttg gcagaaagac acataagaaa gtgctcacca t 14931

<210> 5
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成多核苷酸

<400> 5
tgagctagtt gcagacaggg 20

<210> 6
<211> 79
<212> RNA
<213> 人工序列

<220>
<223> 合成多核苷酸

<400> 6
guuuuagagc uagaauagca aguuaaaaua aggcuagucc guuaucaacu ugaaaaagug 60

gcaccgaguc ggugcuuuu 79

申請專利範圍

1. 一種細胞，其包含併入該細胞基因組之基因座內的一外源性核酸，其中該基因座包含與SEQ ID NO:1或SEQ ID NO:4至少98%相同之核苷酸序列，且其中該外源性核酸包含在一啟動子控制下之第一可篩選標記。
2. 如請求項1之細胞，其中該外源性核酸進一步包含在一啟動子控制下之第二可篩選標記，且其中該第二可篩選標記與該第一可篩選標記不同。
3. 如請求項2之細胞，其中該外源性核酸進一步包含在一啟動子控制下之第三可篩選標記，且其中該第三可篩選標記與該第一可篩選標記及該第二可篩選標記不同。
4. 如請求項1之細胞，其中該第一可篩選標記為藥物抗性基因。
5. 如請求項4之細胞，其中該藥物抗性基因係選自由新黴素抗性基因、潮黴素抗性基因及康黴素抗性基因組成之群。
6. 如請求項2之細胞，其中該第二可篩選標記為螢光蛋白質基因。
7. 如請求項6之細胞，其中該螢光蛋白質基因係選自由下列組成之群：珊瑚(Discosoma coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質。
8. 如請求項3之細胞，其中該第三可篩選標記為螢光蛋白質基因。
9. 如請求項8之細胞，其中該螢光蛋白質基因係選自由下列組成之群：珊瑚(Discosoma coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質。

10. 如請求項1之細胞，其進一步包含在一啟動子控制下之感興趣基因。

11. 如請求項1之細胞，其中該細胞為中國倉鼠卵巢(CHO)細胞。

12. 如請求項1之細胞，其中該細胞為小鼠細胞。

13. 一種細胞，其包含併入該細胞基因組之基因座內的一外源性核酸，其中該基因座包含與SEQ ID NO:1或SEQ ID NO:4至少98%相同之核苷酸序列，且其中該外源性核酸包含

(1) 在一啟動子控制下之第一可篩選標記，

(2) 在一啟動子控制下之第二可篩選標記，

其中該第一可篩選標記為藥物抗性基因，其係選自由新黴素抗性基因、潮黴素抗性基因及康黴素抗性基因組成之群，且該第二可篩選標記為螢光蛋白質基因，其係選自由下列組成之群：珊瑚(Discosoma coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質；及

(3) 在一啟動子控制下之感興趣基因。

14. 如請求項13之細胞，其進一步包含第三可篩選標記。

15. 如請求項14之細胞，其中該第三可篩選標記為螢光蛋白質基因，其係選自由下列組成之群：珊瑚(Discosoma coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質。

16. 如請求項14之細胞，其包含超過一個將被表現之感興趣基因。

17. 如請求項13之細胞，其中該細胞為CHO細胞。

18. 如請求項13之細胞，其中該細胞為小鼠細胞。

19. 一種製造感興趣蛋白質的方法，其中該方法包含以下步驟：

A) 提供一種細胞，該細胞包含併入該細胞基因組之基因座內的一外源性核酸，其中該基因座包含與SEQ ID NO:1或SEQ ID NO:4至少98%相同之核苷酸序列，且其中該外源性核酸包含

(1) 在一啟動子控制下之第一可篩選標記，

(2) 在一啟動子控制下之第二可篩選標記，

其中該第一可篩選標記為藥物抗性基因，其係選自由新黴素抗性基因、潮黴素抗性基因及康黴素抗性基因組成之群，且該第二可篩選標記為螢光蛋白質基因，其係選自由下列組成之群：珊瑚(Discosoma coral)(DsRed)、綠色螢光蛋白質(GFP)、增高綠色螢光蛋白質(eGFP)、氰基螢光蛋白質(CFP)、增高氰基螢光蛋白質(eCFP)、黃色螢光蛋白質(YFP)、增高黃色螢光蛋白質(eYFP)及遠紅外光螢光蛋白質；及

(3) 在一啟動子控制下之感興趣基因，其中該感興趣基因編碼該感興趣蛋白質；及

B) 在容許該感興趣基因表現之條件下培養該細胞，藉此製造該感興趣蛋白質。

20. 如請求項19之方法，其中該細胞為CHO細胞。

21. 如請求項19之方法，其中該細胞為小鼠細胞。

圖式

CHO基因組
HC
Hyg
LC
Hyg
LC
Hyg
CHO基因組

圖1A

5'同源性臂
lox
Neo
IRES
mKate
pA
lox
3'同源性臂

圖1B

圖2A

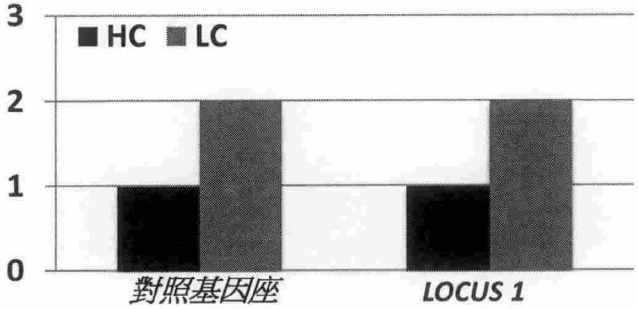

3
2
1
0
HC
LC
對照基因座
LOCUS 1

圖2B

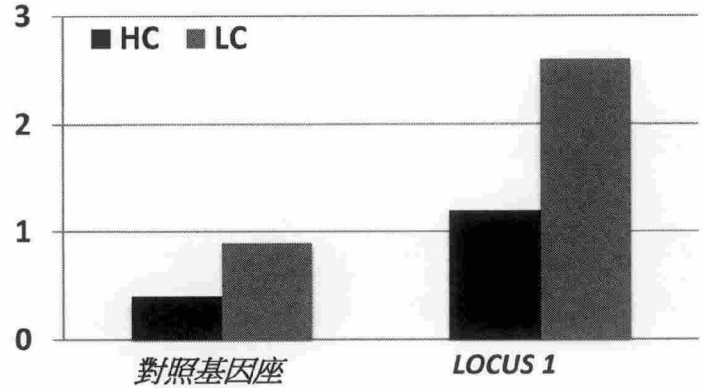

3
2
1
0
HC
LC
對照基因座
LOCUS 1

圖2C

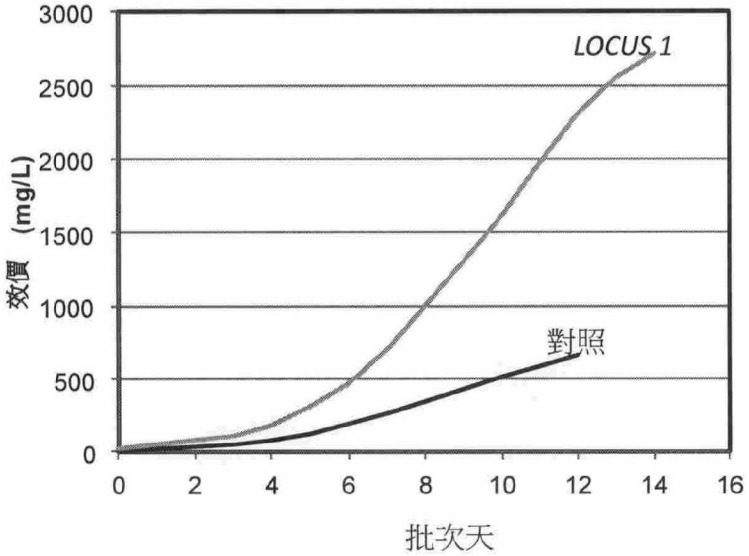
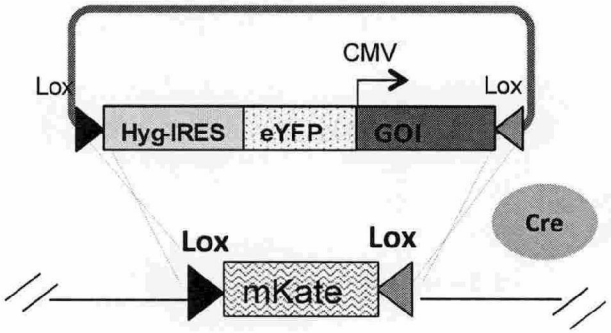
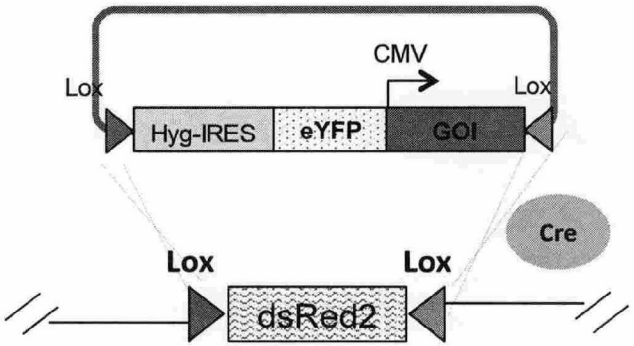

3000
2500
2000
1500
1000
500
0
效價 (mg/L)
LOCUS 1
對照
0
2
4
6
8
10
12
14
16
批次天

圖3A



LOCUS 1

圖3B



對照基因座

圖4