



(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 I753168 B

(45)公告日：中華民國 111 (2022) 年 01 月 21 日

(21)申請案號：107117388

(22)申請日：中華民國 107 (2018) 年 05 月 22 日

(51)Int. Cl.：

A61K31/7088 (2006.01)

C12N15/12 (2006.01)

C12N15/861 (2006.01)

A61P7/04 (2006.01)

(30)優先權：2017/05/22

美國

62/509,616

(71)申請人：日商武田藥品工業股份有限公司 (日本) TAKEDA PHARMACEUTICAL COMPANY LIMITED (JP)

日本

(72)發明人：侯恩林 法郎西斯卡 HORLING, FRANZISKA (AT)；朗格勒 約翰尼斯 LENGLER, JOHANNES (DE)；佛卡納 佛可 岡特 FALKNER, FALKO-GUNTER (DE)；羅斯登能 翰斯彼得 ROTTENSTEINER, HANSPETER (AT)；雪佛玲格 費德瑞奇 SCHEIFLINGER, FRIEDRICH (AT)

(74)代理人：林景郁；閻啓泰

(56)參考文獻：

EP 2438931B1

審查人員：吳思嫻

申請專利範圍項數：29 項 圖式數：25 共 240 頁

(54)名稱

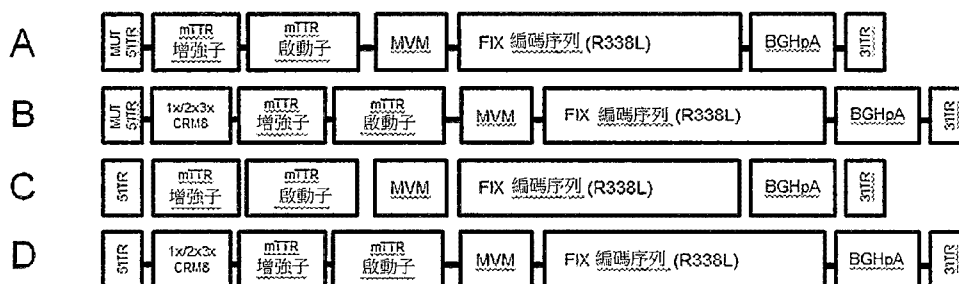
用於血友病 B 基因療法之編碼具有增強表現的重組 F I X 變異體之病毒載體

(57)摘要

本發明提供編碼用於在哺乳動物細胞中表現之因子 IX 變異體的密碼子改變之聚核苷酸以及其他態樣。在一些實施例中，本發明亦提供哺乳動物基因療法載體及治療血友病 B 之方法。

The present disclosure provides, among other aspects, codon-altered polynucleotides encoding Factor IX variants for expression in mammalian cells. In some embodiments, the disclosure also provides mammalian gene therapy vectors and methods for treating hemophilia B.

指定代表圖：



【圖 1】

I753168

【發明摘要】

【中文發明名稱】 用於血友病B基因療法之編碼具有增強表現的重組FIX變異體之病毒載體

【英文發明名稱】 VIRAL VECTORS ENCODING RECOMBINANT FIX VARIANTS WITH INCREASED EXPRESSION FOR GENE THERAPY OF HEMOPHILIA B

【中文】

本發明提供編碼用於在哺乳動物細胞中表現之因子IX變異體的密碼子改變之聚核苷酸以及其他態樣。在一些實施例中，本發明亦提供哺乳動物基因療法載體及治療血友病B之方法。

【英文】

The present disclosure provides, among other aspects, codon-altered polynucleotides encoding Factor IX variants for expression in mammalian cells. In some embodiments, the disclosure also provides mammalian gene therapy vectors and methods for treating hemophilia B.

【指定代表圖】 圖1

【代表圖之符號簡單說明】

無

【特徵化學式】

無

【發明說明書】

【中文發明名稱】 用於血友病B基因療法之編碼具有增強表現的重組FIX變異體之病毒載體

【英文發明名稱】 VIRAL VECTORS ENCODING RECOMBINANT FIX VARIANTS WITH INCREASED EXPRESSION FOR GENE THERAPY OF HEMOPHILIA B

【技術領域】

【0001】 本發明提供編碼用於在哺乳動物細胞中表現之因子IX變異體的密碼子改變之聚核苷酸、哺乳動物基因療法載體及治療血友病B之方法。

【先前技術】

【0002】 血液凝固係通過相互依賴之生化反應的複雜且動態之生物路徑(稱為凝血級聯反應)來進行。凝血因子VIII (FVIII)及因子IX (FIX)為級聯反應中之主要組分。因子VIII被招募至出血位點，且與活化之因子IX及因子X (FX)一起形成X酶複合物。X酶複合物活化FX，FX進而將凝血酶原活化為凝血酶，然後凝血酶活化凝血級聯反應中之其他組分以產生穩定的凝塊(綜述於Saenko等人, *Trends Cardiovasc. Med.*, 9:185-192 (1999)；Lenting等人, *Blood*, 92:3983-3996 (1998)中)。

【0003】 血友病B為以因子IX活性缺乏為特徵之先天性X連鎖出血病症。通常，削弱之因子VIII/因子IX活性抑制凝血級聯反應中之正反饋環路。此引起不完全凝固，其表現為出血事件持續時間增加、大範圍瘀傷、自發性口腔及鼻部出血、關節僵硬及慢性疼痛且在嚴重情況下可能為內部出血及貧血。(Zhang等人, *Clinic. Rev. Allerg. Immunol.*, 37:114-124 (2009))。

【0004】 常規地，藉由因子IX替代療法來治療血友病B，該因子IX替代療法由向患有血友病B之個體投與因子IX蛋白(例如血漿源性或重組製備之因子IX)組成。因子IX係預防性地投與以預防出血事件或降低出血事件之頻率、響應於急性出血事件而投與及/或術前投與以管理手術期間之出血。然而，因子IX替代療法存在若干不需要之特徵。

【0005】 首先，因子IX替代療法用於治療或管理血友病B，但不治癒潛在之因子IX缺乏。由於此原因，患有血友病B之個體在其生命中持續需要因子IX替代療法。連續治療為昂貴的且需要個體維持嚴格順應性，因為僅遺漏幾個預防性劑量患有重度血友病B之個體就可能會具有嚴重後果。

【0006】 其次，因為習知因子IX產品在活體內具有相對短之半衰期，約為24小時，故預防性因子IX替代療法需要每週投與兩次或三次。此對個體構成在其整個生命中維持順應性之負擔。雖然第三代「長效」因子IX藥物可降低投與頻率，但使用此等藥物之預防性因子FIX替代療法仍需要每月、每週或更頻繁地永久投與。舉例而言，使用Nonacog β pegol [聚乙二醇化重組因子IX] (Novo Nordisk, 美國及歐洲管理部門待批准)之預防性治療仍需要每週投與 (Collins P.W.等人, Blood, 124(26):3880-86 (2014))。此外，經化學修飾之生物製品(例如聚乙二醇化多肽)之長期作用尚未完全理解。

【0007】 第三，高至5%之重度血友病B患者因子IX替代療法形成抗因子IX抑制劑抗體，從而使療法效率低(Osooli及Berntorp, J. Intern. Med., 277(1):1-15 (2015))。不同於因子VIII旁路療法可用於治療已產生抗因子VIII抑制抗體之血友病A患者，不存在用於治療血友病B之因子IX旁路療法。

【0008】 第四，因子IX替代療法為昂貴的，視患者體重而定每一劑量在約\$1,000至約\$3,000範圍內(美國血友病聯盟(Hemophilia Federation of America)在線材料)。因此，在每週給藥兩次之情況下，因子IX替代療法每年之成本可高

至\$300,000。

【0009】 基因療法對於血友病B之治療來說有著極大前景，因為其將補救潛在之功能因子IX活性低表現(例如歸因於錯義或無義突變)，而非向個體提供一次性劑量之因子IX活性。由於提供因子IX之機制的差異，與因子IX替代療法相比，單次投與因子IX基因療法載體可為個體提供足夠之因子IX含量達幾年，甚至更久。此降低治療成本且消除對持續患者順應性之需要。

【0010】 已顯示血友病B之因子IX基因療法治療之概念的證據。參見例如Manno C.S.等人, Nat Med., 12(3):342-47 (2006)。然而，仍存在關於治療有效量之因子IX是否可表現足夠之時間的問題。參見例如Giangrande, Semin Thromb Hemost. 42(5):513-17 (2016)。

【0011】 已進行若干嘗試以構築密碼子優化之因子IX。舉例而言，WO 2006/036502揭示一種具有ApoE HCR-1增強子及 α -1抗胰蛋白酶(AAT)啟動子之密碼子優化之因子IX AAV基因療法載體。類似地，WO 2014/064277及WO 2016/146757揭示包括肝特異性SERPIN調控元件之一或多個拷貝之密碼子優化之因子VIII及因子IX AAV基因療法載體。最後，WO 2016/210170揭示具有ApoE HCR-1增強子及 α -1抗胰蛋白酶(AAT)啟動子之密碼子優化之因子IX AAV基因療法載體。

【發明內容】

【0012】 因此，對改良之因子IX基因療法構築體存在需要。舉例而言，對編碼更高效地包裝至基因療法載體中且經由基因療法載體遞送之因子IX的合成之密碼子改變之核酸存在需要。對更高效地表現因子IX之合成之密碼子改變的核酸亦存在需要。對編碼與野生型因子IX相比具有改良之折迭特性、改良之自表現細胞之分泌及/或增加之活性的因子IX多肽的密碼子改變之核酸亦存在需

要。此類編碼密碼子改變之核酸的因子IX允許改良之因子IX缺乏(例如血友病B)的治療。以上缺乏及與因子IX缺乏(例如血友病B)之治療相關的其他問題因所揭示之編碼因子IX蛋白的密碼子改變之核酸而減輕或清除。

【0013】 在一個態樣中，描述編碼因子IX及因子IX變異體之核酸組合物(例如密碼子改變之聚核苷酸)。如本文所描述，核酸組合物包括與編碼因子IX之CS02、CS03、CS04、CS05及CS06序列具有高序列一致性之聚核苷酸。本文所描述之核酸組合物相對於野生型因子IX編碼序列提供增加之因子IX表現。核酸組合物亦允許增加之基於AAV之基因療法病毒體的產生。在一些實施例中，與編碼因子IX之野生型序列相比，本文所描述之核酸組合物具有降低之GC含量及或包括較少之CpG二核苷酸。

【0014】 在一些實施例中，核酸組合物包括編碼因子IX之聚核苷酸，其核苷酸序列與所揭示之選自以下之序列具有至少95%的序列一致性(例如至少95%、96%、97%、98%、99%或100%的序列一致性)：CS02-FL-NA (SEQ ID NO:5)、CS02-MP-NA (SEQ ID NO:13)、CS03-FL-NA (SEQ ID NO:6)、CS03-MP-NA (SEQ ID NO:14)、CS04-FL-NA (SEQ ID NO:7)、CS04-MP-NA (SEQ ID NO:15)、CS05-FL-NA (SEQ ID NO:8)、CS05-MP-NA (SEQ ID NO:16)、CS06-FL-NA (SEQ ID NO:9)及CS06-MP-NA (SEQ ID NO:17)。

【0015】 在一些實施例中，核酸組合物包括編碼因子IX之聚核苷酸，其核苷酸序列與所揭示之編碼因子IX輕鏈之序列(例如CS02-LC-NA (SEQ ID NO:42)、CS03-LC-NA (SEQ ID NO:44)、CS04-LC-NA (SEQ ID NO:46)、CS05-LC-NA (SEQ ID NO:48)或CS06-LC-NA (SEQ ID NO:50))及所揭示之編碼因子IX重鏈之序列(例如CS02-HC-NA (SEQ ID NO:41)、CS03-HC-NA (SEQ ID NO:43)、CS04-HC-NA (SEQ ID NO:45)、CS05-HC-NA (SEQ ID NO:47)或CS06-HC-NA (SEQ ID NO:49))具有至少95%的序列一致性(例如至少95%、96%、

97%、98%、99%或100%的序列一致性)。

【0016】 在一些實施例中，核酸組合物包括編碼具有輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子(例如活化肽)的因子IX多肽的聚核苷酸。因子IX多肽之輕鏈由與以下中之一者具有高序列一致性的第一核苷酸序列編碼：CS02-LC-NA (SEQ ID NO:42)、CS03-LC-NA (SEQ ID NO:44)、CS04-LC-NA (SEQ ID NO:46)、CS05-LC-NA (SEQ ID NO:48)及CS06-LC-NA (SEQ ID NO:50)。因子IX多肽之重鏈由與以下中之一者具有高序列一致性的第二核苷酸序列編碼：CS02-HC-NA (SEQ ID NO:41)、CS03-HC-NA (SEQ ID NO:43)、CS04-HC-NA (SEQ ID NO:45)、CS05-HC-NA (SEQ ID NO:47)及CS06-HC-NA (SEQ ID NO:49)。多肽連接子包含蛋白酶裂解位點(例如兩個因子XIa裂解位點)。

【0017】 在上文所描述之聚核苷酸的一些實施例中，多肽連接子具有與野生型因子IX活化肽FIX-AP-AA (SEQ ID NO:56；FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)具有高序列一致性的胺基酸序列。在一些實施例中，多肽連接子由與以下中之一者具有高序列一致性的第三核苷酸序列編碼：CS02-AP-NA (SEQ ID NO:57)、CS03-AP-NA (SEQ ID NO:58)、CS04-AP-NA (SEQ ID NO:59)、CS05-AP-NA (SEQ ID NO:60)及CS06-AP-NA (SEQ ID NO:61)。

【0018】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸編碼前原因子IX多肽，例如其中所編碼之因子IX蛋白包括信號肽及原肽。在一些實施例中，信號肽、原肽或信號肽與原肽兩者由密碼子改變之序列編碼。在一些實施例中，信號肽、原肽或信號肽與原肽兩者由野生型序列編碼，而編碼成熟因子IX單鏈多肽(例如FIX-MP-AA (SEQ ID NO:10)；FIX-FL-AA (SEQ ID NO:2)之胺基酸47-461)的核酸部分為密碼子改變的。

【0019】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸編碼因



子IX變異體多肽，其例如相對於野生型因子IX胺基酸序列(例如FIX-FL-AA (SEQ ID NO:2)或FIX-MP-AA (SEQ ID NO:10))具有一或多個胺基酸取代。在一些實施例中，因子IX變異體為與野生型因子IX相比具有增加之因子IX活性的高活性因子IX變異體。在一特定實施例中，所編碼之因子IX多肽具有『Padua』R384L胺基酸取代(相對於因子IX前原多肽序列FIX-FL-AA (SEQ ID NO:2)，相對於成熟因子IX單鏈序列FIX-MP-AA (SEQ ID NO:10)為R338L胺基酸取代)。

【0020】 在一個態樣中，描述治療血友病B之方法。該等方法包括向有需要之患者投與本文所描述之核酸組合物(例如密碼子改變之因子IX聚核苷酸構築體)(例如與CS02、CS03、CS04、CS05或CS06因子IX編碼序列具有高序列一致性之聚核苷酸)。在一些實施例中，如本文所描述，因子IX聚核苷酸構築體為哺乳動物基因療法載體。在一特定實施例中，因子IX聚核苷酸構築體為腺相關病毒(AAV)載體。在一些實施例中，基因療法載體包括肝特異性調控控制元件之一或多個拷貝(例如CRM8調控控制元件之1至3個拷貝)。

【0021】 在一個態樣中，描述製備腺相關病毒(AAV)粒子之方法。該方法包括將本文所描述之核酸組合物(例如密碼子改變之因子IX聚核苷酸構築體)(例如與CS02、CS03、CS04、CS05或CS06因子IX編碼序列具有高序列一致性之聚核苷酸)引入哺乳動物宿主細胞中，其中聚核苷酸勝任於在哺乳動物宿主細胞中複製。

【圖式簡單說明】

【0022】 圖1說明根據一些實現方式之示例性因子IX基因療法構築體。自我互補型(A、B)及單股(C、D)載體之序列元件顯示沒有(A、C)且具有(B、D)肝特異性順式調控模組(CRM8)。

【0023】 圖2顯示登錄號CCDS14666.1之野生型因子IX編碼序列(SEQ ID

NO:1) (「FIX-FL-NA」)。

【0024】 圖3顯示在人類中表現之兩種野生型因子IX前原多肽同種型之胺基酸序列。圖3A顯示對應於UniProt登錄號P00740及NCBI登錄號NP_000124.1之第一較長因子IX前原多肽同種型之野生型胺基酸序列(SEQ ID NO:2) (「FIX-FL-AA」)。

圖3B顯示對應於NCBI登錄號NP_001300842.1之第二較短因子IX前原多肽同種型之野生型胺基酸序列(SEQ ID NO:3) (「FIX2-FL-AA」)。

【0025】 圖4顯示Padua (R384L)因子IX胺基酸序列(SEQ ID NO:4 ; 「FIXp-FL-AA」)。

【0026】 圖5顯示根據一些實現方式編碼具有R384L胺基酸取代之因子IX變異體的CS02密碼子改變之核苷酸序列(SEQ ID NO:5) (CS02-FL-NA)。

【0027】 圖6顯示根據一些實現方式編碼具有R384L胺基酸取代之因子IX變異體的CS03密碼子改變之核苷酸序列(SEQ ID NO:6) (CS03-FL-NA)。

【0028】 圖7顯示根據一些實現方式編碼具有R384L胺基酸取代之因子IX變異體的CS04密碼子改變之核苷酸序列(SEQ ID NO:7) (CS04-FL-NA)。

【0029】 圖8顯示根據一些實現方式編碼具有R384L胺基酸取代之因子IX變異體的CS05密碼子改變之核苷酸序列(SEQ ID NO:8) (CS05-FL-NA)。

【0030】 圖9顯示根據一些實現方式編碼具有R384L胺基酸取代之因子IX變異體的CS06密碼子改變之核苷酸序列(SEQ ID NO:9) (CS06-FL-NA)。

【0031】 圖10說明以每公斤體重 2×10^{11} vg之劑量注射具有CRM8肝特異性順式調控控制元件之0、1、2或3個拷貝的CS02基因療法構築體之野生型小鼠中的FIX抗原含量。

【0032】 圖11顯示在人類中表現之兩種單鏈野生型因子IX蛋白同種型(例如缺乏信號肽及原肽)的胺基酸序列。圖11A顯示對應於UniProt登錄號P00740及NCBI登錄號NP_000124.1之第一較長因子IX前原多肽同種型的野生型胺基酸序

列(SEQ ID NO:10) (「FIX-MA-AA」)。圖11B顯示對應於NCBI登錄號NP_001300842.1之第二較短因子IX前原多肽同種型的野生型胺基酸序列(SEQ ID NO:11) (「FIX2-MA-AA」)。

【0033】 圖12顯示單鏈因子IX (R338L)「Padua」胺基酸序列(SEQ ID NO:12；「FIXp-MP-AA」)。

【0034】 圖13顯示根據一些實現方式編碼具有R338L胺基酸取代之單鏈因子IX變異體的CS02密碼子改變之核苷酸序列(SEQ ID NO:13) (CS02-MP-NA)。

【0035】 圖14顯示根據一些實現方式編碼具有R338L胺基酸取代之單鏈因子IX變異體的CS03密碼子改變之核苷酸序列(SEQ ID NO:14) (CS03-MP-NA)。

【0036】 圖15顯示根據一些實現方式編碼具有R338L胺基酸取代之單鏈因子IX變異體的CS04密碼子改變之核苷酸序列(SEQ ID NO:15) (CS04-MP-NA)。

【0037】 圖16顯示根據一些實現方式編碼具有R338L胺基酸取代之單鏈因子IX變異體的CS05密碼子改變之核苷酸序列(SEQ ID NO:16) (CS05-MP-NA)。

【0038】 圖17顯示根據一些實現方式編碼具有R338L胺基酸取代之單鏈因子IX變異體的CS06密碼子改變之核苷酸序列(SEQ ID NO:17) (CS06-MP-NA)。

【0039】 圖18顯示根據一些實現方式編碼本文所描述之許多構築體之前原肽(PPP)的核酸序列(NA)。

【0040】 圖19顯示根據一些實現方式編碼本文所描述之許多構築體之信號肽(SP)的核酸序列(NA)。

【0041】 圖20顯示根據一些實現方式編碼本文所描述之許多構築體之原肽(PP)的核酸序列(NA)。

【0042】 圖21顯示FIX前原肽(PPP)之胺基酸序列(AA)。

【0043】 圖22顯示FIX信號肽(SP)之胺基酸序列(AA)。

【0044】 圖23顯示FIX原肽(PP)之胺基酸序列(AA)。

【0045】 圖24顯示CRM8序列之核酸序列(SEQ ID NO:39)。

【0046】 圖25A及B顯示CS06-CRM8.3-ssV構築體之核酸序列(SEQ ID NO:40)。

【實施方式】

相關申請案之交叉參考

【0047】 本申請案主張於2017年5月22日申請之美國臨時申請案第62/509,616號之權益，該臨時申請案出於所有目的以全文引用之方式明確併入本文中。

I. 引言

【0048】 基於AAV之基因療法對於血友病之治療來說有著極大前景。對於血友病B，第一臨床資料為鼓舞人心的，因為在至少一些患者中可維持約10%之FIX含量超過1年。舉例而言，在初始人類試驗中，證實AVV-FIX構築體之肝動脈插管術引起因子IX在活體內之短暫表現。Kay M.等人, Nat Genet. 24(3):257-61 (2000)。然而，轉導引起免疫系統針對AAV源性衣殼抗原之適度活化。Manno C.S.等人, Nat Med. 12(3):342-47 (2006)及Mingozzi F.等人, Nat Med. 13(4):419-22 (2007)。

【0049】 非病毒載體可具較低免疫原性，因為其係基於與非抗原性載體(例如惰性聚合物、脂質或奈米粒子)相關之裸DNA或DNA的遞送。然而，非病

毒載體之細胞轉染速率低於病毒遞送載體。另外，由非病毒載體產生之長期表現因用於大規模製備構築體之細菌序列之存在而受阻礙。

【0050】 然而，此等挑戰不能簡單地藉由投與更高劑量之基因療法構築體來解決。根據當前之知識，基於AAV之基因療法載體的載體劑量不應增至超過每公斤體重 2×10^{12} vg。此係因為在此類高劑量下T細胞免疫反應被觸發，T細胞免疫反應破壞轉導之細胞，且因此，轉基因表現降低或甚至消除。因此，需要改良FIX表現之策略來使FIX基因療法成為對於血友病B患者來說更可行之治療選擇。

【0051】 因此，支持改良之因子IX表現及活性的改良之因子IX多肽構築體將改良兩種治療方法。舉例而言，藉由降低構築體之初始劑量，由此降低對個體之免疫系統之刺激將改良病毒遞送方法。藉由使用較少之治療性聚核苷酸拷貝支持較大之因子IX活性將改良依賴於投與裸DNA之方法。

【0052】 本發明係關於解決與因子IX基因療法相關之此等及其他問題的密碼子改變之因子IX變異體編碼序列的發現。舉例而言，本文所揭示之聚核苷酸在哺乳動物宿主中提供明顯改良之因子IX表現及活性。在一些實現方式中，藉由使用與密碼子改變之CS02、CS03、CS04、CS05及CS06構築體具有高序列一致性的編碼因子IX之聚核苷酸實現此等優勢。在一些實施例中，如下文更充分描述，與野生型構築體相比，此等序列包括顯著較少之CpG二核苷酸。

【0053】 有利地，本文所描述之CS02、CS03、CS04、CS05及CS06密碼子改變之因子IX序列與同等之野生型序列相比在活體內提供優良因子IX表現。舉例而言，實例1表明相對於運載野生型因子IX編碼序列之自我互補型AAV載體，運載CS02、CS03、CS04、CS05或CS06密碼子改變之因子IX (R384L)編碼序列的自我互補型AAV載體提供20倍至40倍之活體內因子IX活性增加。類似地，觀察到相對於運載野生型因子IX (R384L)編碼序列之自我互補型AAV載體2

倍至4倍之因子IX表現增加(表2)。

【0054】 有利地，可藉由在因子IX編碼序上游列引入肝特異性調控元件之一或多個拷貝使由CS02、CS03、CS04、CS05及CS06密碼子改變之序列產生的改良之因子IX活性進一步增強。舉例而言，如實例2及3中所證實，自我互補型AAV因子IX載體中包括一或多個肝特異性CRM8調控控制元件使小鼠模型中之因子IX表現進一步增加至2倍至3倍且在人類肝細胞中則為2倍至13倍(分別見表3及表4)。同樣地，單股AAV因子IX載體中包括肝特異性CRM8調控控制元件之一或多個拷貝使活體內因子IX活性增加至2倍(小鼠模型；表5)且在人類肝細胞中則高至26倍(表6)。

【0055】 令人驚訝地，雖然缺乏肝特異性CRM8調控控制元件之編碼密碼子改變之因子IX多肽的自我互補型AAV載體與類似單股AAV載體相比提供更大之因子IX表現增加(將表6中由CS06-CRM.0-scV提供之6.2倍FIX活性增加與由CS06-CRM.0-ssV (SEQ ID NO:40)提供之3.9倍因子IX活性增加相比較)，但含有肝特異性CRM8調控控制元件之多個拷貝的單股AAV因子IX載體顯著優於類似自我互補型AAV載體(將表4中由CS02-CRM8.3-scV相對於CS02-CRM8.0-scV提供之12.8倍因子IX活性增加與表6中由CS06-CRM8.3-ssV (SEQ ID NO:40)相對於CS06-CRM8.0-scV提供之16.8倍因子IX活性增加相比較)。

II. 定義

【0056】 除非另外指明，否則如本文所用，以下術語具有歸屬於它們之含義。

【0057】 如本文所用，術語「因子IX」及「FIX」(其中「IX」指意謂「九」之羅馬數字)為可互換使用的，且係指具有因子IX活性之任何蛋白質(例如活性FIX，常常稱為「FIXa」)或具有因子IX活性之蛋白質的蛋白質前驅體(例如原蛋白或前原蛋白，常常稱為pFIX及ppFIX)，所述因子IX活性特定而言為在

存在因子VIII之情況下的例如如使用歐洲藥典9.0之章節2.7.11中所描述之一期因子IX凝結分析所量測的因子X裂解活性，該參考文獻之內容以引用之方式併入本文中。

【0058】 因子IX轉譯為非活性單鏈多肽，其包括信號肽、原肽、輕鏈、活化肽及重鏈，常常稱為因子IX前原多肽。因子IX前原肽經歷轉譯後加工以形成活性因子IX蛋白(例如FIXa)。此加工包括移除(例如藉由裂解)信號肽，隨後移除(例如藉由裂解)原肽，以形成仍為非活性之含有因子IX輕鏈及因子IX重鏈之單鏈成熟因子IX多肽。進一步裂解成熟因子IX多肽以切除因子IX輕鏈與因子IX重鏈之間的活化肽，從而形成活性因子IX蛋白(例如FIXa)。因子IX輕鏈及因子IX重鏈保持通過二硫鍵締合。

【0059】 舉例來說舉例而言，首先裂解野生型人類因子IX前原蛋白以釋放所編碼之信號肽(FIX-FL-AA (SEQ ID NO:2)之胺基酸1-28)，從而形成第一單鏈原蛋白。然後裂解此單鏈原肽以釋放原肽(FIX-FL-AA (SEQ ID NO:2)之胺基酸29-46)，從而形成第二單鏈原蛋白(例如FIX-MP-AA (SEQ ID NO:10)，其中「MP」命名代表「成熟蛋白」)。然後藉由因子XIa在FIX輕鏈與FIX重鏈之間將第二單鏈原蛋白裂解兩次，以釋放活化肽(FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)。此形成由分開的輕鏈及重鏈通過二硫鍵締合組成之活性因子IXa蛋白。關於因子IX之結構、功能及活化之其他資訊參見例如Brandstetter H.等人 P.N.A.S. USA, 92(21):9796-800 (1995)；Hopfner KP等人, Structure, 7(8):989-96 (1999)；及Gailani D.等人, Thromb Res., 133增刊1:S48-51 (2014)，該等參考文獻之內容出於所有目的以全文引用之方式併入本文中。

【0060】 如本文所描述，此活性因子IXa蛋白可包括一或多個變異體，並且R338L變異體在一些實施例中有特定用途。此變異體稱為具有在本文中稱為「FIXp-MP-NA」之核酸序列的「FIXp-MP-AA」(SEQ ID NO:12)；「FIXp」代

表在最終蛋白質中包括Padua R338L變異體。應注意，本文例示之密碼子優化之序列CS02至CS06編碼FIXp蛋白，包括R338L變異體。因此，FIXp尤其包括在FIX之定義中。

【0061】 如本文所用，術語「因子IX多肽」及「FIX多肽」係指在特定條件下例如如使用歐洲藥典9.0之章節2.7.11中所描述之一期因子IX凝結分析所量測具有因子IX絲胺酸蛋白酶活性的多肽。因子IX多肽包括單鏈前驅體多肽(包括因子IX前原多肽、因子IX原肽及成熟單鏈因子IX多肽)，其在藉由上文所描述之轉譯後加工活化時變為具有因子IX絲胺酸蛋白酶活性之活性因子IX蛋白，以及活性因子IX蛋白本身。包括R338L變異體之因子IX多肽尤其包括在因子IX多肽之定義中。在一示例性實施例中，人類因子IX多肽係指包括與野生型人類因子IX多肽中包括輕鏈及重鏈之部分(FIX-MP-AA，SEQ ID NO:10，顯示於圖11A中)或與padua人類因子IX多肽中包括輕鏈及重鏈之部分(FIXp-MP-AA，SEQ ID NO:12，顯示於圖12中)具有高序列一致性(例如至少85%、90%、95%、99%或更高)之胺基酸序列的多肽。

【0062】 如本文所用，術語「因子IX輕鏈」或簡單地「輕鏈」係指活化因子IXa蛋白中源自因子IX單鏈多肽之N端部分的多肽，其含有因子IX之Gla模組、EGF樣1及EGF樣2結構域。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸47-191構成因子IX輕鏈。如本文所用，野生型因子IX輕鏈之胺基酸序列稱為FIX-LC-AA (SEQ ID NO:62)。

【0063】 如本文所用，術語「因子IX重鏈」或簡單地「重鏈」係指活化因子IXa蛋白中源自因子IX單鏈多肽之C端部分的多肽，其含有因子IX之肽酶S1結構域。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸227-461構成因子IX重鏈。如本文所用，野生型因子IX重鏈之胺基酸序列稱為FIX-HC-AA (SEQ ID NO:63)且當包括R338L變異體時稱為FIXp-

HC-AA (SEQ ID NO:64)。

【0064】 通常，因子IX輕鏈及重鏈表現為單一多肽鏈，例如與活化肽一起。然而，在一些實施例中，因子IX輕鏈及因子VIII重鏈表現為分開之多肽鏈(例如共表現)，且重構以形成因子IX蛋白(例如活體內或活體外)。一般而言，出於本發明之目的，即使兩條鏈分開表現，其通常亦位於相同表現載體(例如病毒基因組)上，而非位於不同表現載體上。

【0065】 如本文所用，術語「因子IX活化肽」或簡單地「活化肽」係指在因子IXa蛋白活化後自因子IX單鏈多肽切除之肽。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸192-226構成因子IX活化肽。如本文所用，野生型因子IX活化肽之胺基酸序列稱為FIX-AP-AA (SEQ ID NO:56)。

【0066】 如本文所用，術語「因子IX信號肽」或簡單地「信號肽」係指由信號肽酶自因子IX前原多肽之N端切除之肽。信號肽將新轉譯之因子IX前原蛋白引導至內質網。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸1-28構成因子IX信號肽。如本文所用，野生型因子IX信號肽之胺基酸序列稱為FIX-SP-AA (SEQ ID NO:37)。本發明之許多信號肽顯示於圖19及22中。

【0067】 如本文所用，術語「因子IX原肽」或簡單地「原肽」係指由弗林蛋白酶(Furin)自因子IX原多肽之N端切除(例如在裂解信號肽之後)之肽。原肽包括 γ -羧化識別位點，其招募羧化酶至相鄰Gla模組，由此促進麩醯胺殘基之羧化。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸29-46構成因子IX原肽。如本文所用，野生型因子IX原肽之胺基酸序列稱為FIX-PP-AA (SEQ ID NO:38)。

【0068】 如本文所用，術語「因子IX前原肽」或簡單地「前原肽」係指

因子IX信號肽及原多肽之聚集物。在一示例性實施例中，人類前原因子IX多肽(FIX-FL-AA (SEQ ID NO:2))之胺基酸1-46構成因子IX前原肽。如本文所用，野生型因子IX前原肽之胺基酸序列稱為FIX-PPP-AA (SEQ ID NO:36)，其具有圖18中所示之稱為FIX-PPP-NA (SEQ ID NO:18)的核酸序列(當使用R338L變異體時具有對應之FIXp-PPP-AA及FIXp-PPP-NA)。

【0069】 除非本文中另外指明，否則因子IX胺基酸之編號係指圖3A中以SEQ ID NO:2形式呈現之全長野生型人類因子IX前原多肽序列(FIX-FL-AA)中之對應胺基酸。因而，當提及本文所揭示之因子IX多肽中之胺基酸取代時，所敘述之胺基酸編號係指全長野生型因子IX前原多肽序列中之類似(例如在結構上或功能上等效)及/或同源(例如在進化上在一級胺基酸序列中為保守的)胺基酸。舉例而言，R384L胺基酸取代係指在全長野生型人類因子IX前原肽序列(FIX-FL-AA (SEQ ID NO:2))之位置384處R至L之取代、在成熟野生型因子IX單鏈多肽(FIX-MP-AA (SEQ ID NO:10))之位置338處R至L之取代、在全長野生型人類因子IX前原肽同種型2序列(FIX2-FL-AA (SEQ ID NO:3))之位置346處R至L之取代、在成熟野生型人類因子IX前原肽同種型2序列(FIX2-FL-AA (SEQ ID NO:3))之位置300處R至L之取代及在野生型人類因子IX重鏈序列(FIX-HC-AA (SEQ ID NO:63))之位置158處R至L之取代。因此，所有此等命名法描述不同因子IX構築體中之相同「Padua」胺基酸取代。

【0070】 如本文所描述，因子IX胺基酸編號系統視是否包括因子IX前原肽(例如全長野生型人類因子IX序列之胺基酸1-46，包括信號肽及原肽)而定。在包括前原肽之情況下，編號稱為「包括前原肽的」或「PPI」。在不包括前原肽之情況下，編號稱為「不包括前原肽的」或「PPE」。舉例而言，R384L為與PPE編號中之R338L相同的胺基酸取代之PPI編號。類似地，因子IX胺基酸編號亦視因子IX同種型而定。舉例而言，R384L為與同種型2編號中之R346L相

同之胺基酸取代的同種型1編號。除非另外指出，否則所有胺基酸編號係指圖3A中以SEQ ID NO:2形式呈現之全長野生型人類因子IX同種型1序列(FIX-FL-AA)中之對應胺基酸。此編號對於除R384L「Padua」突變外具有相同胺基酸序列之FIXp-FL-AA (SEQ ID NO:4)而言為相同的。

【0071】 IX多肽之非限制性實例包括人類前原因子IX(例如基因庫登錄號NP_000124.1 (FIX-FL-AA (SEQ ID NO:2))及NP_001300842.1 (FIX2-FL-AA (SEQ ID NO:3))、缺乏信號肽(前原蛋白之胺基酸1-28)及/或原肽(前原蛋白之胺基酸29-46)之對應單鏈因子IX及其天然變異體；豬前原因子IX(例如UniProt登錄號P00741)、缺乏信號肽之對應單鏈因子IX及其天然變異體；鼠前原因子IX(例如UniProt登錄號P16294)、缺乏信號肽之對應單鏈因子IX及其天然變異體；大鼠前原因子IX(例如UniProt登錄號P16296)、缺乏信號肽之對應單鏈因子IX及其天然變異體；及其他哺乳動物因子VIII同系物(例如黑猩猩、猿、倉鼠、豚鼠等)。

【0072】 如本文所用，因子IX多肽包括在存在因子VIII之情況下具有因子X裂解活性之天然變異體及人工構築體。如本發明中所用，因子IX涵蓋任何天然變異體、替代序列、同種型或突變體蛋白質，其保留一定基礎因子IX裂解活性(例如如根據歐洲藥典9.0之章節2.7.11之一期凝結分析中所分析為對應野生型活性之至少5%、10%、25%、50%、75%或更多，該參考文獻關於其在章節2.7.11中人類凝血因子IX分析之教示以引用之方式併入本文中。在人類群體中所發現之因子IX胺基酸變異(相對於FIX-FL-AA (SEQ ID NO:2))之實例包括但不限於I17N、L20S、C28R、C28Y、V30I、R43L、R43Q、R43W、K45N、R46S、R46T、N48I、S49P、L52S、E53A、E54D、E54G、F55C、G58A、G58E、G58R、E66V、E67K、F71S、E73K、E73V、R75Q、E79D、T84R、Y91C、D93G、Q96P、C97S、P101R、C102R、C102R、G106D、G106S、C108S、D110N、I112S、N113K、Y115C、C119F、C119R、E124K、G125E、

G125R、G125V、C134Y、I136T、N138H、G139D、G139S、C155F、G160E、Q167H、S169C、C170F、C178R、C178W、R191C、R191H、R226G、R226Q、R226W、V227D、V227F、V228F、V228L、Q241H、Q241K、C252S、C252Y、G253E、G253R、A265T、C268W、A279T、N283D、E291V、R294G、R294Q、V296M、H302R、N306S、I316F、L318R、L321Q、N328K、N328Y、P333H、P333T、T342K、T342M、I344L、G351D、W356C、G357E、G357R、K362E、G363W、A366D、R379G、R379Q、C382Y、L392F、L383I、R384L、K387E、I390F、M394K、F395I、F395L、C396F、C396S、A397P、R404T、C407R、C407S、D410H、S411G、S411I、G412E、G413R、P414T、V419E、F424V、T426P、S430T、W431G、W431R、G432S、E433A、G433K、C435Y、A436V、G442E、G442R、I443T、R449Q、R449W、Y450C、W453R及I454T。如下文更充分論述，此編號係相對於野生型人類FIX。在人類群體中鑑定出之其他胺基酸變異為已知的且可例如使用國家生物技術資訊中心(National Center for Biotechnology Information，「NCBI」)之變異觀察器找到，登錄號GCF_000001405.25。因子VIII蛋白亦包括含有轉譯後修飾之多肽。

【0073】 在本發明中具有特定用途的為FIX蛋白，其包括所謂的「Padua」突變，該突變為在成熟單股因子IX蛋白之位置338 (R338L)、因子IX前原多肽之位置384 (R384L)處精胺酸變為白胺酸。此突變賦予FIX蛋白高功能活性。舉例而言，已證實「Padua」蛋白(例如含有R338L突變之因子IX)在活體內之活性為野生型因子IX的5倍至10倍。美國專利第6,531,298號；Simioni P.等人, N Engl J Med. 361(17):1671-75 (2009)，其以全文引用之方式併入本文中。因此，本發明提供編碼Padua-FIX蛋白(有時在本文中稱為「FIXp」或「pFIX」)之胺基酸及核酸構築體。

【0074】 如本文所用，術語「因子IX聚核苷酸」及「FIX聚核苷酸」係指編碼在特定條件下例如如使用歐洲藥典9.0之章節2.7.11中所描述之一期因子IX凝結分析所量測具有因子IX絲胺酸蛋白酶活性的因子IX多肽的聚核苷酸。因子IX聚核苷酸包括編碼因子IX單鏈前驅體多肽(包括因子IX前原多肽、因子IX原肽及成熟單鏈因子IX多肽)之聚核苷酸，該等因子IX單鏈前驅體多肽當藉由上文所描述之轉譯後加工活化時變為具有因子IX絲胺酸蛋白酶活性之活性因子IX蛋白。編碼包括R338L變異體之因子IX多肽之聚核苷酸尤其包括於因子IX聚核苷酸之定義中。在一示例性實施例中，人類因子IX聚核苷酸係指編碼包括與野生型人類因子IX多肽中包括輕鏈及重鏈之部分(FIX-MP-AA，SEQ ID NO:10，顯示於圖11A中)或與padua人類因子IX聚核苷酸中包括輕鏈及重鏈之部分(FIXp-MP-AA，SEQ ID NO:12，顯示於圖12中)具有高序列一致性(例如至少85%、90%、95%、99%或更高)之胺基酸序列的多肽的聚核苷酸。

【0075】 如本文所描述，因子IX聚核苷酸可包括調控元件，諸如啟動子、增強子、終止子、聚腺苷酸化序列及內含子；以及病毒包裝元件，諸如反向末端重複(「ITR」)；及/或支持聚核苷酸在非病毒宿主細胞中之複製的其他元件，例如支持聚核苷酸例如在細菌、酵母或哺乳動物宿主細胞中之增殖的複製子。

【0076】 在本發明中具有特定用途的為密碼子改變之因子IX聚核苷酸。如本文所描述，與由天然編碼之因子IX構築體(例如使用野生型人類密碼子編碼相同因子IX胺基酸序列之聚核苷酸)提供之因子IX表現水準相比，密碼子改變之FIX聚核苷酸提供增加之轉基因因子IX活體內表現。如本文所用，術語「增加之表現」係指與投與天然編碼之因子IX構築體之動物的血液中之轉基因因子IX蛋白含量相比，投與編碼因子IX之密碼子改變之聚核苷酸之動物的血液中之轉基因因子IX蛋白含量增加。增加之蛋白質表現引起因子IX活性增加；因此，

增加之表現引起增加之活性。

【0077】 在一些實施例中，增加之表現係指與投與天然編碼之因子IX聚核苷酸之動物的血液中之轉基因因子IX多肽含量相比，投與密碼子改變之因子IX聚核苷酸之動物的血液中之轉基因因子IX多肽多至少25%。出於本發明之目的，增加之表現係指由密碼子序列改變產生之影響，而非由潛在胺基酸取代(例如「Padua」突變)引起之高活性。換句話說，相對於由天然編碼之「Padua」蛋白獲得之表現水準比較由編碼「Padua」因子IX聚核苷酸之密碼子優化之序列獲得的表現水準。在一些實施例中，增加之表現係指與投與天然編碼之因子IX聚核苷酸之動物的血液中之轉基因因子IX多肽含量相比，投與密碼子改變之因子IX聚核苷酸之動物的血液中之轉基因因子IX多肽多至少50%、多至少75%、多至少100%、多至少3倍、多至少4倍、多至少5倍、多至少6倍、多至少7倍、多至少8倍、多至少9倍、多至少10倍、多至少15倍、多至少20倍、多至少25倍、多至少30倍、多至少40倍、多至少50倍、多至少60倍、多至少70倍、多至少80倍、多至少90倍、多至少100倍、多至少125倍、多至少150倍、多至少175倍、多至少200倍、多至少225倍或多至少250倍。動物血液中之因子IX多肽含量可例如使用特定於因子IX多肽之ELISA分析來量測。

【0078】 「因子IX活性」或「因子IX絲胺酸蛋白酶活性」在本文中意謂在存在因子VIIIa輔因子之情況下例如經由野生型因子IX中之Arg194-Ile195肽鍵之水解裂解因子X多肽，因此將因子X活化為因子Xa之能力。活性水準可使用此項技術中已知之任何因子IX活性來量測；適合之分析概述於本文中；用於測定因子IX活性之一種示例性分析為本文所提供之實例中所用的歐洲藥典9.0之章節2.7.11中所描述之一期因子IX凝結分析。在一些實施例中，在一期凝結分析中使用缺乏FIX活性之人類血漿作為對照來測定因子IX特異性。

【0079】 因為某些因子IX變異體與野生型因子IX相比在活體內具有增強

之特定活性，例如人類『Padua』變異體之因子IX絲胺酸蛋白酶活性為天然編碼類型之人類因子IX的5倍至10倍，故在一些實施例中，例如替代增加之因子IX表現或除增加之因子IX表現之外，藉由投與因子IX聚核苷酸之動物的血液中之因子IX活性增加來評估因子IX聚核苷酸組合物之治療潛能。在一些實施例中，如本文所用，增加之因子IX活性係指與投與天然編碼之因子IX聚核苷酸之動物的血液中之因子IX活性相對於在投與天然編碼之因子IX聚核苷酸之前動物血液中之基線因子IX活性的增加相比，投與密碼子改變之因子IX聚核苷酸之動物的血液中之因子IX活性相對於在投與密碼子改變之因子IX聚核苷酸之前動物血液中之基線因子IX活性的增加更大。在一些實施例中，增加之因子IX活性係指與投與天然編碼之因子IX聚核苷酸之動物的血液中之因子IX活性水準相對於在投與天然編碼之因子IX聚核苷酸之前動物中之因子IX活性基線水準的增加相比，投與密碼子改變之因子IX聚核苷酸之動物的血液中之因子IX活性相對於在投與密碼子改變之因子IX聚核苷酸之前動物血液中之因子IX活性基線水準的增加至少大25%。在一些實施例中，增加之因子IX活性係指與投與天然編碼之因子IX聚核苷酸之動物的血液中之因子IX活性水準相對於在投與天然編碼之因子IX聚核苷酸之前動物中之因子IX活性基線水準的增加相比，投與密碼子改變之因子IX聚核苷酸之動物的血液中之因子IX活性相對於在投與密碼子改變之因子IX聚核苷酸之前動物血液中之因子IX活性基線水準的增加大至少50%、大至少75%、大至少100%、大至少3倍、大至少4倍、大至少5倍、大至少6倍、大至少7倍、大至少8倍、大至少9倍、大至少10倍、大至少15倍、大至少20倍、大至少25倍、大至少30倍、大至少40倍、大至少50倍、大至少60倍、大至少70倍、大至少80倍、大至少90倍、大至少100倍、大至少125倍、大至少150倍、大至少175倍、大至少200倍、大至少225倍或大至少250倍。如本文所描述，使用歐洲藥典9.0之章節2.7.11中所描述之一期因子IX凝結分析來量測活性。

【0080】 如本文所用，術語「血友病」係指廣泛地以降低之血液凝結或凝固為特徵之疾病狀態群組。血友病可指A型、B型或C型血友病，或指所有三種疾病類型之複合疾病。A型血友病(血友病A)由因子VIII (FVIII)活性降低或損失引起且為最突出之血友病亞型。B型血友病(血友病B)由因子IX (FIX)凝結功能之損失或降低引起。C型血友病(血友病C)為抗血友病因子XI (FXI)凝結活性損失或降低之結果。血友病A及B為X連鎖疾病，而血友病C為常染色體疾病。血友病之習知治療包括預防性地與按需投與凝血因子，諸如FVIII、FIX (包括Bebulin®-VH及FXI)；以及FEIBA-VH、去胺加壓素及血漿輸注。

【0081】 如本文所用，術語「因子IX基因療法」或「FIX基因療法」包括為患者提供編碼因子IX之核酸以緩和、減少或預防與因子IX缺乏(例如血友病B)相關之一或多種症狀(例如臨床因素)之復發的任何治療方法。該術語涵蓋投與包含編碼因子IX分子，包括任何修飾形式之因子IX (例如因子VIII R384L變異體)之核酸以維持或改良患有因子IX缺乏(例如血友病B)之個體的健康之任何化合物、藥物、程序或方案。熟習此項技術者將瞭解，FIX基因療法之過程或FIX基因療法治療劑之劑量可例如基於根據本發明所獲得之結果而變化。

【0082】 術語「治療有效量或劑量」或「治療充足量或劑量」或「有效或充足量或劑量」係指產生投與其所要獲得之治療作用的劑量。舉例而言，適合用於治療血友病之藥物之治療有效量可為能夠預防或緩和與血友病相關之一或多種症狀的量。

【0083】 在一些實施例中，治療有效性治療使得個體中出血事件之頻率及/或嚴重程度降低。

【0084】 如本文所用，術語「基因」係指編碼多肽鏈之DNA分子區段(例如編碼區)。在一些實施例中，基因按區域緊接在產生多肽鏈時所涉及之編碼區之前、之後及/或插入其中而定位(例如調控元件，諸如啟動子、增強子、聚腺

苷酸化序列、5'-非轉譯區、3'-非轉譯區或內含子)。

【0085】 如本文所用，術語「調控元件」係指提供編碼序列在細胞中之表現的核苷酸序列，諸如啟動子、增強子、終止子、聚腺苷酸化序列、內含子等。

【0086】 如本文所用，術語「啟動子元件」係指幫助控制編碼序列表現之核苷酸序列。通常，啟動子元件定位於基因轉譯起始位點之5'。然而，在某些實施例中，啟動子元件可定位於內含子序列內或編碼序列之3'。在一些實施例中，適合於基因療法載體之啟動子源自目標蛋白之天然基因(例如因子VIII啟動子)。在一些實施例中，適合於基因療法載體之啟動子關於在目標有機體之特定細胞或組織中表現具特異性(例如肝特異性啟動子)。在其他實施例中，將複數個充分表徵之啟動子元件中之一者用於本文所描述之基因療法載體中。充分表徵之啟動子元件之非限制性實例包括CMV早期啟動子、 β -肌動蛋白啟動子及甲基CpG結合蛋白2 (MeCP2)啟動子。在一些實施例中，啟動子為組成性啟動子，其實質上驅動目標蛋白之恆定表現。在其他實施例中，啟動子為誘導型啟動子，其響應於特定刺激(例如暴露於特定治療或藥劑)而驅動目標蛋白之表現。關於設計用於AAV介導之基因療法之啟動子之綜述，參見Gray等人(Human Gene Therapy 22:1143-53 (2011))，其內容出於所有目的以全文引用之方式明確併入本文中。

【0087】 如本文所用，「CRM8」元件係指源自SERPINA1基因(NCBI登錄號NM_000295.4)之順式作用調控模組，其以肝特異性方式增強可操作連接之基因(例如編碼因子IX多肽之序列)的表現，與SEQ ID NO:39具有高序列一致性。如本文所用，CRM8元件係指其調控元件之單個拷貝，在一些實施例中，包括在因子IX聚核苷酸內之一或多個拷貝(例如1、2、3或更多個拷貝)中。關於CRM元件(諸如CRM8)之其他資訊參見Chuah MK等人, Mol Ther., 22(9):1605-13

(2014)，其以引用之方式併入本文中。

【0088】 如本文所用，「MVM內含子」係指源自與SEQ ID NO:53具有高序列一致性之小鼠細小病毒的內含子序列。關於MVM內含子本身之其他資訊，參見Haut及Pintel, *J Virol.* 72(3):1834-43 (1998)，且關於MVM內含子在AAV基因療法載體中之用途，參見Wu Z等人, *Mol Ther.*, 16(2):280-9 (2008)，該兩個參考文獻以引用之方式併入本文中。

【0089】 如本文所用，術語「可操作地連接」係指第一參考核苷酸序列(例如基因)與第二核苷酸序列(例如調控控制元件)之間允許第二核苷酸序列影響與第一參考核苷酸序列相關之一或多種特性(例如轉錄速率)的關係。在本發明之情形中，當調控元件定位於基因療法載體內使得其發揮對因子IX轉基因轉錄之作用(例如促進或組織選擇性作用)時，調控控制元件可操作地連接至因子IX轉基因。

【0090】 如本文所用，術語「載體」係指用於將因子IX核酸轉移至宿主細胞中之任何核酸構築體。在一些實施例中，載體包括複製子，其用以複製核酸構築體。適合於基因療法之載體的非限制性實例包括質粒、噬菌體、黏粒、人工染色體及病毒，其充當活體內複製之自主性單元。在一些實施例中，載體為用於將因子IX核酸引入宿主細胞中之病毒載體。許多適合於基因療法之經修飾真核病毒為此項技術中已知的。舉例而言，腺相關病毒(AAV)特別好地適合於在人類基因療法中使用，因為人類為該病毒之天然寄主，天然病毒已知不會促成任何疾病，且該等病毒引起輕度免疫反應。

【0091】 如本文所用，術語「因子IX病毒載體」係指包含編碼因子IX多肽之因子IX聚核苷酸的重組病毒，其足以使因子IX多肽在引入適合之動物寄主(例如人類)中時表現。其中編碼因子IX多肽之密碼子改變之因子IX聚核苷酸已插入病毒之基因組中的重組病毒尤其包括於因子IX病毒載體之定義內。其中病

毒之天然基因組已經編碼因子IX多肽之因子IX聚核苷酸置換的重組病毒亦尤其包括於因子IX病毒載體之定義內。包含編碼因子IX之「Padua」變異體之因子IX聚核苷酸的重組病毒包括於因子IX病毒載體之定義內。

【0092】 如本文所用，術語「因子IX病毒粒子」係指封裝編碼因子IX多肽之因子IX聚核苷酸的病毒粒子，其關於因子IX多肽在引入適合之動物寄主(例如人類)中時之表現具特異性。封裝其中已插入編碼因子IX多肽之密碼子改變之因子IX聚核苷酸的基因組的重組病毒粒子尤其包括於因子IX病毒粒子之定義內。封裝置換病毒天然基因組之編碼因子IX多肽之因子IX聚核苷酸的重組病毒粒子亦尤其包括於因子IX病毒粒子之定義內。封裝編碼因子IX之「Padua」變異體之因子IX聚核苷酸的重組病毒粒子包括於因子IX病毒粒子之定義內。

【0093】 「AAV」或「腺相關病毒」在本文中意謂細小病毒科病毒屬內之依賴細小病毒。如本文所用，AAV可指源自其中已插入因子IX聚核苷酸之天然存在之「野生型」AAV基因組的病毒、源自使用由天然存在之AAV cap基因編碼之衣殼蛋白包裝至衣殼中的重組因子IX聚核苷酸的重組病毒或源自使用由非天然衣殼cap基因編碼之衣殼蛋白包裝至衣殼中之重組因子IX聚核苷酸的重組病毒。封裝因子IX聚核苷酸之1型AAV (AAV1)、2型AAV (AAV2)、3型AAV (AAV3)、4型AAV (AAV4)、5型AAV (AAV5)、6型AAV (AAV6)、7型AAV (AAV7)、8型AAV (AAV8)及9型AAV (AAV9)病毒及由一或多種封裝因子IX聚核苷酸之變異體AAV衣殼蛋白形成之病毒包括在AAV之定義內。

【0094】 「AAV8」、「AAV-8」或「AAV血清型8」在本文中意謂由封裝因子IX聚核苷酸之AAV8衣殼病毒蛋白形成之病毒。

【0095】 如本文所用，術語「CpG」係指沿DNA之單股的胞嘧啶-鳥嘌呤二核苷酸，其中「p」表示兩者之間的磷酸酯鍵。

【0096】 如本文所用，術語「CpG島」係指聚核苷酸內具有統計學上升

高之CpG二核苷酸密度的區域。如本文所用，若在200-鹼基對窗中：(i)聚核苷酸(例如編碼密碼子改變之因子IX蛋白之聚核苷酸)之區域的GC含量大於50%，並且(ii)如通過以下關係所定義，每單位預期CpG二核苷酸觀測到之CpG二核苷酸之比率為至少0.6：

$$\frac{N[\text{CpG}] * N[\text{窗長度}]}{N[\text{C}] * N[\text{G}]} \geq 0.6, \text{ 那麼該區域為CpG島。}$$

【0097】 關於鑑定CpG島之方法的其他資訊，參見Gardiner-Garden M.等人, J Mol Biol., 196(2):261-82 (1987)，該參考文獻之內容出於所有目的以全文引用之方式明確併入本文中。

【0098】 如本文所用，術語「核酸」係指脫氧核糖核苷酸或核糖核苷酸及其呈單股或雙股形式之聚合物及其互補物。該術語涵蓋含有已知核苷酸類似物或經修飾之主鏈殘基或鍵聯之核酸，其為合成的、天然存在的及非天然存在的，且其具有與參考核酸類似之結合特性，且其以類似於參考核苷酸之方式代謝。此類類似物之實例包括但不限於硫代磷酸酯、胺基磷酸酯、膦酸甲酯、對掌性膦酸甲酯、2-O-甲基核糖核苷酸及肽-核酸(PNA)。然而，本文中特別適合在患者中用於基因療法中之實施例使用磷酸二酯鍵。

【0099】 「核酸組合物」在本文中意謂包括編碼因子IX聚核苷酸之因子IX聚核苷酸之任何分子或分子調配物。因子IX聚核苷酸、因子IX聚核苷酸之水溶液、封裝因子IX聚核苷酸之病毒粒子及封裝因子IX聚核苷酸之病毒粒子的水性調配物包括在核酸組合物之定義內。如本文所揭示之核酸組合物包括編碼FIX多肽之密碼子改變之FIX基因。

【0100】 術語「胺基酸」係指天然存在及非天然之胺基酸，包括以類似於天然存在之胺基酸的方式起作用之胺基酸類似物及胺基酸模擬物。天然存在之胺基酸包括由遺傳密碼編碼之彼等胺基酸，以及隨後經修飾之彼等胺基酸，例如羥基脯胺酸、γ-羧基麩胺酸鹽及O-磷酸絲胺酸。天然存在之胺基酸可包括

例如D-胺基酸及L-胺基酸。關於胺基酸序列，一般熟習此項技術者將認識到對核酸或肽序列之使所編碼序列中之單個胺基酸或小百分比之胺基酸改變、增加或缺失的個別取代、缺失或添加為「經保守修飾之變異體」，其中該改變使得胺基酸經化學上類似之胺基酸取代。能夠提供功能類似之胺基酸的保守取代為此項技術中熟知的。此類經保守修飾之變異體為除本發明之多形變異體、種間同源物及對偶基因之外的且不排除本發明之多形變異體、種間同源物及對偶基因。

【0101】 提供功能類似之胺基酸的保守胺基酸取代為此項技術中熟知的。視特定胺基酸之官能性(例如催化、結構或空間上重要之胺基酸)而定，可將不同胺基酸分組視為彼此之保守取代。**表1**提供基於胺基酸之電荷及極性、胺基酸之疏水性、胺基酸之表面暴露/結構性質及胺基酸之二級結構傾向視為保守取代之胺基酸的分組。

表1. 基於蛋白質中之殘基官能性的保守胺基酸取代分組.

重要特徵	保守分組
電荷/極性	1. H、R及K 2. D及E 3. C、T、S、G、N、Q及Y 4. A、P、M、L、I、V、F及W
疏水性	1. D、E、N、Q、R及K 2. C、S、T、P、G、H及Y 3. A、M、I、L、V、F及W
結構/表面暴露	1. D、E、N、Q、H、R及K 2. C、S、T、P、A、G、W及Y 3. M、I、L、V及F
二級結構傾向	1. A、E、Q、H、K、M、L及R 2. C、T、I、V、F、Y及W 3. S、G、P、D及N
進化保守性	1. D及E 2. H、K及R 3. N及Q 4. S及T 5. L、I及V 6. F、Y及W 7. A及G 8. M及C

【0102】 在兩個或更多個核酸或肽序列之背景下，術語「一致」或「一致性」百分比係指如使用BLAST或BLAST 2.0序列比較算法用下文所描述之預設參數或藉由人工比對及視覺檢查所量測，兩個或更多個序列或子序列為相同的或具有指定百分比之相同氨基酸殘基或核苷酸(亦即當在比較窗或指定區域上比較及比對最大對應性時，在指定區域內具有約60%之一致性，較佳為65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高一致性)。

【0103】 如此項技術中已知，許多不同程式可用於鑑定蛋白質(或如下文所論述之核酸)是否與已知序列具有序列一致性或相似性。序列一致性及/或相似性係使用此項技術中已知之標準技術來測定，包括但不限於Smith及Waterman, *Adv. Appl. Math.*, 2:482 (1981)之局部序列一致性算法、藉由Needleman及Wunsch, *J. Mol. Biol.*, 48:443 (1970)之序列一致性比對算法、藉由Pearson及Lipman, *Proc. Natl. Acad. Sci. U.S.A.*, 85:2444 (1988)之檢索相似性法、藉由此等算法之電腦化實現方式(Wisconsin Genetics軟體包, Genetics Computer Group, 575 Science Drive, Madison, WI中之GAP、BESTFIT、FASTA及TFASTA)、由Devereux等人, *Nucl. Acid Res.*, 12:387-395 (1984)描述之Best Fit序列程式，較佳使用預設設置或通過檢查來進行。較佳地，藉由FastDB基於以下參數來計算一致性百分比：錯配罰分為1；空位罰分為1；空位尺寸罰分為0.33；且連接罰分為30，「Current Methods in Sequence Comparison and Analysis」，*Macromolecule Sequencing and Synthesis, Selected Methods and Applications*, 第127-149頁(1988), Alan R. Liss, Inc，其全部以引用之方式併入本文中。

【0104】 適用算法之實例為PILEUP。PILEUP使用漸進性成對比對自相關序列之群組形成多重序列比對。其亦可繪製顯示用於形成比對之群集關係的

樹形圖。PILEUP使用Feng及Doolittle, J. Mol. Evol. 35:351-360 (1987)之漸進性比對方法的簡化型式；該方法類似於由Higgins及Sharp CABIOS 5:151-153 (1989)描述之方法，兩個參考文獻均以引用之方式併入本文中。適用PILEUP參數包括預設空位權重3.00、預設空位長度權重0.10及加權末端空位。

【0105】 適用算法之另一實例為以下文獻中所描述之BLAST算法：Altschul等人, J. Mol. Biol. 215, 403-410, (1990)；Altschul等人, Nucleic Acids Res. 25:3389-3402 (1997)；及Karlin等人, Proc. Natl. Acad. Sci. U.S.A. 90:5873-5787 (1993)，該等文獻以引用之方式併入本文中。特別適用之BLAST程式為WU-BLAST-2程式，其獲自Altschul等人, Methods in Enzymology, 266:460-480 (1996)；[http://blast.wustl.edu/blast/ README.html](http://blast.wustl.edu/blast/README.html)]。WU-BLAST-2使用若干檢索參數，其中大多數設定為預設值。可調整參數設定為以下值：重迭跨度= 1，重迭分數= 0.125，字語臨限值(T) = 11。HSP S及HSP S2參數為動態值且由程式本身根據特定序列之組成及檢索相關序列之特定資料庫之組成來確定；然而，可調整該等值以增加靈敏性。

【0106】 另一適用算法為如由Altschul等人, Nucl. Acids Res., 25:3389-3402報導之帶空位之BLAST，該參考文獻以引用之方式併入本文中。帶空位之BLAST使用BLOSUM-62取代得分；臨限值T參數設定為9；引發無空位延伸之雙擊法；使空位長度k承擔10+k之代價；Xu設定為16，且對於資料庫檢索階段，Xg設定為40，且對於算法之輸出階段設定為67。藉由對應於約22位元之得分來引發帶空位之比對。

【0107】 胺基酸序列一致性%值係藉由用匹配之一致殘基數目除以比對區域中「較長」序列之殘基總數目來確定。「較長」序列為具有比對區域中之大多數實際殘基的序列(忽略由WU-Blast-2為使比對得分最大化而引入之空位)。以類似方式，將關於所鑑定多肽之編碼序列的「核酸序列一致性百分比

(%)」定義為候選序列中與細胞週期蛋白之編碼序列中之核苷酸殘基一致的核苷酸殘基的百分比。較佳方法利用設定為預設參數之WU-BLAST-2的BLASTN模組，其中重迭跨度及重迭分數分別設定為1及0.125。

【0108】 比對可包括在所要比對之序列中引入空位。此外，對於含有比由圖3A之野生型因子IX序列(SEQ ID NO:2)編碼之蛋白質更多或更少之胺基酸的序列，應瞭解，在一個實施例中，序列一致性百分比將係基於相對於胺基酸或核苷酸總數目之一致胺基酸或核苷酸數目來確定。因此，舉例而言，如下文所論述，在一個實施例中，比圖3A中所示之序列(SEQ ID NO:2)更短之序列的序列一致性將係使用較短序列中之核苷酸數目來確定。在一致性百分比計算中，未為序列變異之諸如插入、缺失、取代等各種表現分配相對權重。

【0109】 在一個實施例中，僅一致性為正評分(+1)且為所有形式之序列變異(包括空位)分配值「0」，此排除了對如下文關於序列相似性計算所描述之加權標度或參數的需要。序列一致性百分比可例如藉由用匹配之一致殘基的數目除以比對區域內「較短」序列之殘基總數目且乘以100來計算。「較長」序列為具有比對區域中之大多數實際殘基的序列。

【0110】 術語「對偶變異體」係指特定基因位點處基因之多形形式以及源自基因之mRNA轉錄物的cDNA及由其編碼之多肽。術語「較佳哺乳動物密碼子」係指如選自以下列表之在哺乳動物細胞中表現之蛋白質中最常用的編碼胺基酸之密碼子集合中的密碼子子集：Gly (GGC, GGG)；Glu (GAG)；Asp (GAC)；Val (GTG, GTC)；Ala (GCC, GCT)；Ser (AGC, TCC)；Lys (AAG)；Asn (AAC)；Met (ATG)；Ile (ATC)；Thr (ACC)；Trp (TGG)；Cys (TGC)；Tyr (TAT, TAC)；Leu (CTG)；Phe (TTC)；Arg (CGC, AGG, AGA)；Gln (CAG)；His (CAC)；及Pro (CCC)。

【0111】 如本文所用，術語密碼子改變的係指編碼多肽(例如因子IX蛋

白)之聚核苷酸序列，其中編碼多肽之天然聚核苷酸之至少一個密碼子已改變以改良聚核苷酸序列之特性。在一些實施例中，改良之特性促使編碼多肽之mRNA之轉錄增加、mRNA穩定性增加(例如mRNA半衰期改良)、多肽轉譯增加及/或聚核苷酸於載體內之包裝增加。可用於達成改良之特性的改變的非限制性實例包括改變用於特定胺基酸之密碼子的使用及/或分佈、調整總體及/或局部GC含量、移除含AT之序列、移除重複之序列元件、調整總體及/或局部CpG二核苷酸含量、移除隱蔽調控元件(例如TATA盒及CCAAT盒元件)、移除內含子/外顯子剪接位點、改良調控序列(例如引入Kozak共同序列)及移除能夠在轉錄之mRNA中形成二級結構(例如莖環)的序列元件。

【0112】 如本文所論述，存在各種命名法來指本文中之本發明組分。「CS-編號」(例如「CS02」、「CS03」、「CS04」、「CS05」、「CS06」等)係指編碼FIX多肽之密碼子改變之聚核苷酸及/或所編碼之多肽，包括變異體。舉例而言，CS02-FL係指全長(Full Length)密碼子改變之CS02聚核苷酸序列或由CS02聚核苷酸序列編碼之胺基酸序列(對於胺基酸(Amino Acid)序列在本文中有時稱為「CS02-FL-AA」，而對於核酸(Nucleic Acid)序列則為「CS02-FL-NA」(SEQ ID NO:5))。類似地，「CS02-LC」係指編碼FIX多肽之輕鏈的密碼子改變之核酸序列(「CS02-LC-NA」(SEQ ID NO:42))或由CS02聚核苷酸序列編碼之FIX輕鏈之胺基酸序列(在本文中有時亦稱為「CS02-LC-AA」)。同樣地，CS02-HC、CS02-HC-AA及CS02-HC-NA (SEQ ID NO:41)對於FIX重鏈而言同樣如此。如熟習此項技術者將瞭解，對於僅密碼子改變(例如與Padua因子IX變異體相比不含其他胺基酸取代)之諸如CS02、CS03、CS04、CS05、CS06等構築體，胺基酸序列將一致的，因為胺基酸序列不因密碼子最佳化而改變。因此，本發明之序列構築體包括但不限於CS02-FL-NA (SEQ ID NO:5)、CS02-FL-AA、CS02-LC-NA (SEQ ID NO:42)、CS02-LC-AA、CS02-HC-AA、CS02-HC-

NA (SEQ ID NO:41)、CS03-FL-NA (SEQ ID NO:6)、CS03-FL-AA、CS03-LC-NA (SEQ ID NO:44)、CS03-LC-AA、CS03-HC-AA、CS03-HC-NA (SEQ ID NO:43)、CS04-FL-NA (SEQ ID NO:7)、CS04-FL-AA、CS04-LC-NA (SEQ ID NO:46)、CS04-LC-AA、CS04-HC-AA、CS04-HC-NA、CS05-FL-NA (SEQ ID NO:8)、CS05-FL-AA、CS05-LC-NA (SEQ ID NO:48)、CS05-LC-AA、CS05-HC-AA、CS05-HC-NA (SEQ ID NO:47)、CS06-FL-NA (SEQ ID NO:9)、CS06-FL-AA、CS06-LC-NA (SEQ ID NO:50)、CS06-LC-AA、CS06-HC-AA及CS06-HC-NA (SEQ ID NO:49)。應注意，本文中之所有「CS」構築體編碼或含有FIXp胺基酸序列，不過彼等編碼或含有人類野生型FIX胺基酸序列之CS構築體包括在CS構築體之定義內。

【0113】 如本文所用，術語「肝特異性表現」係指與在其他組織中相比，特定基因(例如密碼子改變之轉基因因子IX基因)在活體內在肝組織中之優先或主導表現。在一些實施例中，肝特異性表現意謂特定基因之所有表現中至少50%發生在個體之肝組織內。在其他實施例中，肝特異性表現意謂特定基因之所有表現中至少55%、60%、65%、70%、75%、80%、85%、90%、95%、99%或100%發生在個體之肝組織內。因此，肝特異性調控元件為驅動基因在肝組織中之肝特異性表現之調控元件。

【0114】 如本文所用，術語「小於」X及「小於」X%係指0至X之範圍，不包括值X，例如0%至X%，不包括X%。如本文所用，該等術語與以0或0%開始至但不包括X或X%之範圍可互換使用。

【0115】 如本文所用，術語「不超過」X或「不超過」X%係指0至X之範圍，包括值X，例如0%至X%，包括X%。如本文所用，該等術語與以0或0%開始至且包括X或X%之範圍可互換使用。

【0116】 如本文所用，術語「大於」X或「大於」X%係指X至上限之範

圍，不包括值X，例如X%至100%，不包括X%。如本文所用，該等術語在百分比之情形中與以X或X%開始(但不包括X或X%)至100%之上限的範圍可互換使用。

【0117】 如本文所用，術語「至少」X或「至少」X%係指X至上限之範圍，包括值X，例如X%至100%，包括X%。如本文所用，該等術語在百分比之情形中與以X或X%開始(且包括X或X%)至100%之上限的範圍可互換使用。

【0118】 如本文所用，術語「在『X』與『Y』之間」、「在『X』%與『Y』%之間」、「『X』至『Y』」及「『X』%至『Y』%」係指X至Y之範圍，包括值X及Y，例如X%至Y%，包括X%及Y%。如本文所用，該等術語與以X或X%開始至且包括Y或Y%之範圍可互換使用。

III. 密碼子改變之因子IX聚核苷酸

【0119】 在一些實施例中，本發明提供編碼因子IX或因子IX變異體(其中FIXp在特定實施例中獲得應用)之密碼子改變之核酸組合物。此等密碼子改變之聚核苷酸當於基於AAV之基因療法構築體中投與時提供明顯改良之因子IX表現。與習知密碼子優化之構築體相比，該等密碼子改變之聚核苷酸亦展示改良之AAV-病毒體包裝。如實例1中所證實，申請人已通過發現編碼具有過度活性R338L胺基酸取代(基於成熟單鏈因子IX多肽序列；基於因子IX前原蛋白序列為R384L)之因子IXp多肽的若干密碼子改變之聚核苷酸(例如CS02-FL-NA、CS03-FL-NA、CS04-FL-NA、CS05-FL-NA及CS06-FL-NA (分別為SEQ ID NO:5-9))達成此等優點。如實例2及3中所證實，一或多個肝特異性調控控制元件(例如CRM8)併入編碼因子IX分子之基因療法載體中進一步增加因子IX之活體內及活體外表現及因子IX活性。

【0120】 野生型因子IX經編碼具有28胺基酸信號肽(FIX-SP-AA (SEQ ID NO:37))及18胺基酸原肽(FIX-PP-AA (SEQ ID NO:38))，其在因子IXa活化之前自

所編碼之多肽裂解。如熟習此項技術者所瞭解，信號肽及/或原肽可為突變的、經來自其他基因或其他有機體之信號肽及/或原肽置換或完全移除，而不會影響在藉由細胞加工移除信號肽及原肽之後所留下之成熟多肽的序列。

【0121】 因此，在一些實施例中，本文所提供之密碼子改變之聚核苷酸(例如核酸組合物)具有與編碼成熟因子IX單鏈多肽之CS02-FL-NA、CS03-FL-NA、CS04-FL-NA、CS05-FL-NA或CS06-FL-NA(分別為SEQ ID NO:5-9)具有高序列一致性之核苷酸序列，該成熟因子IX單鏈多肽為因子IX輕鏈、活化肽及重鏈(例如由野生型因子IX基因編碼之全長多肽之胺基酸47-461；FIX-FL-AA (SEQ ID NO:2))。

【0122】 另外，如此項技術中已知的，人類野生型因子IX具有定位於因子IX輕鏈與重鏈之間的34胺基酸活化肽，其在蛋白質活化後自單鏈因子IX多肽切除。因為活化肽自活性因子IX多肽移除，肽本身對於最終因子IX活性而言為非必需的。因此，不要求由本文所揭示之密碼子改變之聚核苷酸編碼的因子IX多肽與人類野生型活化肽序列(FIX-AP-AA (SEQ ID NO:56))具有高序列一致性。然而，所編碼之活化肽應為在因子IX多肽活化後可切除的。舉例而言，在一些實施例中，所編碼之活化肽應在其N端及C端包括因子XI裂解位點，該等因子XI裂解位點為由人類因子IX在活體內可識別且可裂解的。

【0123】 因此，在一些實施例中，本文所提供之密碼子改變之聚核苷酸(例如核酸組合物)編碼與人類野生型FIX輕鏈序列(FIX-LC-AA (SEQ ID NO:62))及人類野生型FIX重鏈序列(FIX-HC-AA (SEQ ID NO:63))具有高序列一致性之單鏈因子IX多肽，且另外編碼具有兩個因子XI裂解位點的將輕鏈之C端連接至重鏈之N端的多肽連接子(例如活化肽)。

【0124】 在一些實施例中，由密碼子改變之聚核苷酸編碼之因子IX輕鏈及重鏈分別為人類因子IX輕鏈及重鏈，包括FIXp重鏈。在其他實施例中，由密

碼子改變之聚核苷酸編碼之因子IX輕鏈及重鏈為來自另一哺乳動物(例如豬因子IX)之重鏈及輕鏈序列。在其他實施例中，因子IX輕鏈及重鏈為嵌合輕鏈及重鏈(例如人類及第二哺乳動物序列之組合)。在其他實施例中，因子IX輕鏈及重鏈為來自另一哺乳動物之輕鏈及重鏈之人類化型式，例如其中人類殘基在所選位置經取代以降低所得肽當向人類投與時之免疫原性的來自另一哺乳動物之輕鏈及重鏈序列。

【0125】 人類基因之GC含量變化很大，自小於25%至大於90%。然而，一般而言，具有較高GC含量之人類基因以更高水準表現。舉例而言，Kudla等人(PLoS Biol., 4(6):80 (2006))證實增加基因之GC含量主要藉由增加轉錄及實現較高之mRNA轉錄物穩態含量來增加所編碼多肽之表現。通常，密碼子優化之基因構築體之所需GC含量被認為等於或大於60%。舉例而言，scAAV8.FIXR338L基因療法載體中之因子IX基因為使用GeneOptimizer軟體(Geneart)進行特定密碼子改變的，以使野生型編碼序列之GC含量自41% GC增加至61% GC。參見Wu Z.等人, Mol Ther 16:280-89 (2008)及Monahan PE等人, Hum Gene Ther., 26(2):69-81 (2015)。然而，天然AAV基因組具有約56%之GC含量。

【0126】 因此，在一些實施例中，本文所提供之密碼子改變之聚核苷酸(例如核酸組合物)的CG含量更接近地匹配天然AAV病毒體之GC含量(例如約56% GC)，此低於關於在哺乳動物細胞中之表現習知密碼子優化之聚核苷酸的較佳CG含量(例如等於或高於60% GC)。舉例而言，CS02-FL-NA (SEQ ID NO:5)具有約54%之GC含量，CS03-FL-NA (SEQ ID NO:6)具有約55%之GC含量，CS04-FL-NA (SEQ ID NO:7)具有約54.5%之GC含量，CS05-FL-NA (SEQ ID NO:8)具有約56.6%之GC含量，且CS06-FL-NA (SEQ ID NO:9)具有約55%之GC含量。此等構築體與具有較高GC含量之類似密碼子改變之序列相比將提供改良

之病毒體包裝。

【0127】 因此，在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸(例如與CS02-CS06因子IX編碼序列中之一者具有高序列一致性之聚核苷酸)的總GC含量小於60%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量小於59%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量小於58%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量小於57%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量不超過56%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量不超過55%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量不超過54%。

【0128】 在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為53%至59%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%至59%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為55%至59%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為56%至59%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為53%至58%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%至58%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為55%至58%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為56%至58%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為53%至57%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%至57%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為55%至57%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為56%至57%。在一些實施例中，編碼因

子IX多肽之密碼子改變之聚核苷酸的總GC含量為53%至56%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%至56%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為55%至56%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為53%至55%。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%至55%。

【0129】 在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $54\pm 0.5\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $54\pm 0.4\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $54\pm 0.3\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $54\pm 0.2\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $54\pm 0.1\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為54%。

【0130】 在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $55\pm 0.5\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $55\pm 0.4\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $55\pm 0.3\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $55\pm 0.2\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $55\pm 0.1\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為55%。

【0131】 在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $56\pm 0.5\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $56\pm 0.4\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $56\pm 0.3\%$ 。在一些實施例中，編碼因子IX多肽之密碼

子改變之聚核苷酸的總GC含量為 $56\pm 0.2\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為 $56\pm 0.1\%$ 。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的總GC含量為56%。

【0132】 此等CpG二核苷酸(亦即，胞嘧啶核苷酸隨後鳥苷酸)在活體內經由鐘樣受體誘導免疫反應已經理論化。一些證據表明在某些情況下，CpG耗盡之AAV載體避開小鼠中之免疫偵測(Faust等人, J. Clin. Invest. 2013; 123, 2994-3001)。野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))含有20個CpG二核苷酸。

【0133】 因此，在一些實施例中，本文所提供之核酸組合物(例如密碼子改變之聚核苷酸)為密碼子改變的以減少因子IX編碼序列中之CpG二核苷酸數目。舉例而言，CS02-FL-NA (SEQ ID NO:5)不具有CpG二核苷酸，CS03-FL-NA (SEQ ID NO:6)不具有CpG二核苷酸，CS04-FL-NA (SEQ ID NO:7)不具有CpG二核苷酸，CS05-FL-NA (SEQ ID NO:8)具有11個CpG二核苷酸，且CS06-FL-NA (SEQ ID NO:9)具有3個CpG二核苷酸。此等構築體與野生型因子IX編碼序列及具有較高CpG二核苷酸數目之類似密碼子改變之序列相比會引起較低免疫原性反應。

【0134】 因此，在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸(例如與CS02-CS06因子IX編碼序列中之一者具有高序列一致性之聚核苷酸)的序列具有少於20個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有少於15個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有少於12個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有少於10個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有少於5個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子

改變之聚核苷酸的序列具有少於3個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0135】 在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有超過15個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有超過12個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有超過10個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有超過5個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有超過3個CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。在一些實施例中，編碼因子IX多肽之密碼子改變之聚核苷酸的序列具有不超過18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、1個CpG二核苷酸或沒有CpG二核苷酸。

A. 因子IX胺基酸取代

【0136】 為進一步增加本文所描述之因子IX構築體之基於AAV載體的表現效率，根據一些實現方式，進一步併入已知改良分泌、增加特定活性及/或增強因子IX穩定性之胺基酸取代。許多潛在因子IX變異體在此項技術中已知增加血漿FIX活性水準。此等變異體包括增加因子IX催化活性(例如過度活性突變體)、增加對抗凝血酶III及/或肝素之抗性、增加血清半衰期及產生改變之轉譯後修飾模式的胺基酸取代。

【0137】 舉例來說舉例而言，殘基R338 (PPE)之突變可增加因子IX之凝結活性。關於綜述，參見美國專利第6,531,298號，其內容出於所有目的以全文引用之方式併入本文中。如美國專利第6,531,298號中所揭示，在此位置精胺酸至白胺酸之胺基酸取代增加因子IX之活性。此隨後在活體內得以證實，其中R338L (PPE)突變在活體內使因子IX活性增加至5倍至10倍。關於綜述，參見

Simioni P.等人, N Engl J Med. 361(17):1671-75 (2009), 其以全文引用之方式併入本文中。因此, 在一些實施例中, 本文所描述之密碼子改變之聚核苷酸編碼在精胺酸384 (PPI; 殘基338 (PPE)處具有胺基酸取代之因子IX多肽。在一特定實施例中, 胺基酸取代為R384L (PPI)。在其他實施例中, 在殘基384 (PPI)/338 (PPE)處之胺基酸取代為變成不為白胺酸之殘基。舉例而言, 據報導R384A (PPI)胺基酸取代在小鼠中提供2倍至6倍高之活性。Schuetttrumpf J等人, Blood, 105(6):2316-23 (2005), 其內容出於所有目的以全文引用之方式明確併入本文中。

【0138】 類似地, 殘基Y305、K311、S365及Y391之突變對合成受質引起增加之因子IX活性。特定而言, K311M及K311T單突變對合成裂解受質產生2.8倍及6.7倍之增加的活性。Sichler K.等人, J Biol Chem. 278(6):4121-26 (2003) (使用不同殘基編號)。另外, Y305F/K311T/Y391T三重突變體對合成受質產生7000倍之增加的活性。同上。因此, 在一些實施例中, 本文所描述之密碼子改變之聚核苷酸編碼在酪胺酸305 (PPI)、離胺酸311 (PPI)及酪胺酸391 (PPI)處具有胺基酸取代之因子IX多肽。在一特定實施例中, 胺基酸取代為K311M (PPI)。在一特定實施例中, 胺基酸取代為K311T (PPI)。在另一特定實施例中, 胺基酸取代為Y305F/K311T/Y391T (PPI)。

【0139】 提供改良之特性的其他胺基酸取代為此項技術中已知的且可併入所描述之密碼子改變之因子IX聚核苷酸中。舉例而言, 參見美國專利第8,778,870號, 其內容出於所有目的以全文引用之方式明確併入本文中。

B. 編碼因子IX蛋白之密碼子改變之聚核苷酸

CS02密碼子改變之聚核苷酸

【0140】 在一個實施例中, 本文所提供之核酸組合物包括編碼單鏈因子IX多肽之因子IX聚核苷酸(例如密碼子優化之聚核苷酸), 其中因子IX聚核苷酸

包括與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之核苷酸序列。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0141】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-FL-NA (SEQ ID NO:5)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS02-FL-NA (SEQ ID NO:5)。

【0142】 在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於60%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於59%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於58%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於57%之GC含量。在一些實施例中，與CS02-

FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於56%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於55%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有小於54%之GC含量。

【0143】 在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0144】 在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS02-

FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\%\pm 0.4$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\%\pm 0.3$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\%\pm 0.2$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\%\pm 0.1$ 之GC含量。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0145】 在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列

一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS02-FL-NA (SEQ ID NO:5)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0146】 在一些實施例中，所編碼之因子IX多肽(例如由與CS02-FL-NA (SEQ ID NO:5)具有高序列同源性之聚核苷酸編碼之多肽)與野生型因子IX前原蛋白序列FIX-FL-AA (SEQ ID NO:2)及/或Padua (hFIX (R384L))前原蛋白序列FIXp-FL-AA (SEQ ID NO:4)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除信號肽及原肽，及藉由切除活化多肽)之能力。

【0147】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.9%之一致性。在一個實施例

中，所編碼之因子IX多肽的序列為FIX-FL-AA (SEQ ID NO:2)。

【0148】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-FL-AA (SEQ ID NO:4)。

【0149】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽(例如具有絲胺酸蛋白酶活性)之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之核苷酸序列。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高

序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0150】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS02-MP-NA (SEQ ID NO:13)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS02-MP-NA (SEQ ID NO:13)。

【0151】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)

具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0152】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0153】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID

NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 53.8% 之GC含量。

【0154】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3

個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0155】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-SP-AA (SEQ ID NO:37)之胺基酸序列之因子IX信號肽的因子IX信號聚核苷酸。在一些實施例中，因子IX信號聚核苷酸具有與CS02-SP-NA (SEQ ID NO:25)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS03-SP-NA (SEQ ID NO:26)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS04-SP-NA (SEQ ID NO:27)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS05-SP-NA (SEQ ID NO:28)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS06-SP-NA (SEQ ID NO:29)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0156】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PP-AA (SEQ ID NO:38)之胺基酸序列之因子IX原肽的因子IX原肽聚核苷酸。在一些實施例中，因子IX原肽聚核苷酸具有與CS02-PP-NA (SEQ ID NO:31)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS03-PP-NA (SEQ ID NO:32)至少90%、95%、96%、97%、98%、99%或

100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS04-PP-NA (SEQ ID NO:33)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS05-PP-NA (SEQ ID NO:34)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS06-PP-NA (SEQ ID NO:35)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0157】 在一些實施例中，與CS02-MP-NA (SEQ ID NO:13)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列之因子IX前原肽的因子IX前原肽聚核苷酸。在一些實施例中，因子IX前原肽聚核苷酸具有與CS02-PPP-NA (SEQ ID NO:19)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS03-PPP-NA (SEQ ID NO:20)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS04-PPP-NA (SEQ ID NO:21)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS05-PPP-NA (SEQ ID NO:22)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS06-PPP-NA (SEQ ID NO:23)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0158】 在一些實施例中，所編碼之因子IX多肽(例如由與CS02-FL-NA (SEQ ID NO:5)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活

化多肽)之能力。

【0159】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0160】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括

白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0161】 在一個實施例中，本文所提供之密碼子改變之聚核苷酸編碼單鏈因子IX多肽，該單鏈因子IX多肽包括輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子。因子IX多肽之輕鏈由第一核苷酸序列編碼，該第一核苷酸序列與作為CS02-FL-NA (SEQ ID NO:5)之編碼因子IX輕鏈之部分的CS02-LC-NA (SEQ ID NO:42)具有高序列一致性。因子IX多肽之重鏈由第二核苷酸序列編碼，該第二核苷酸序列與作為CS02-FL-NA (SEQ ID NO:5)之編碼因子IX重鏈之部分的CS02-HC-NA (SEQ ID NO:41)具有高序列一致性。多肽連接子包括因子XI裂解位點，其允許活體內成熟(例如在前驅體單鏈因子IX多肽表現之後)。

【0162】 在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少95%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少96%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少97%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少98%之序列一致性。在一些實施例

中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少99%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少99.5%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)具有至少99.9%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別為CS02-LC-NA及CS02-HC-NA (SEQ ID NO:42及41)。

【0163】 在一些實施例中，因子IX構築體之多肽連接子由第三核苷酸序列編碼，該第三核苷酸序列與作為編碼野生型因子IX活化多肽(例如FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)之密碼子改變之序列的CS02-AP-NA (SEQ ID NO:57)具有高序列一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少80%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少90%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少95%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少96%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少97%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少98%之一致性。在一些實施例中，第三核苷酸序列與CS02-AP-NA (SEQ ID NO:57)具有至少99%之一致性。在一些實施例中，第三核苷酸序列為CS02-AP-NA (SEQ ID NO:57)。

【0164】 在一些實施例中，所編碼之因子IX多肽亦包括信號肽(例如因子IX信號肽)及/或原肽(例如因子IX原肽)。在一些實施例中，信號肽為野生型因子IX信號肽(FIX-SP-AA (SEQ ID NO:37))。在一些實施例中，信號肽由與CS02-SP-NA (SEQ ID NO:25)具有高序列一致性(例如至少95%、96%、97%、98%或

99%)之密碼子改變之聚核苷酸序列編碼。在一些實施例中，原肽為野生型因子IX原肽(FIX-PP-AA (SEQ ID NO:38))。在一些實施例中，原肽由與CS02-PP-NA (SEQ ID NO:31)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。

【0165】 在一些實施例中，所編碼之因子IX多肽(例如由與CS02-LC-NA (SEQ ID NO:42)及CS02-HC-NA (SEQ ID NO:41)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0166】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0167】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA

(SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0168】 在一些實施例中，參考圖1，提供包括結構A之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS02-MP-NA (SEQ ID NO:13)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致

性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28) 及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS02-FL-NA (SEQ ID NO:5)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0169】 在一些實施例中，參考圖1，提供包括結構B之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS02-MP-NA (SEQ ID NO:13)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28) 及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID

NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS02-FL-NA (SEQ ID NO:5)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0170】 在一些實施例中，參考圖1，提供包括結構C之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS02-MP-NA (SEQ ID NO:13)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS02-FL-NA (SEQ ID NO:5)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0171】 在一些實施例中，參考圖1，提供包括結構D之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS02-MP-NA (SEQ ID NO:13)具有至少

95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS02-FL-NA (SEQ ID NO:5)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

CS03密碼子改變之聚核苷酸

【0172】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之核苷酸序列。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0173】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-

NA (SEQ ID NO:6)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-FL-NA (SEQ ID NO:6)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS03-FL-NA (SEQ ID NO:6)。

【0174】 在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0175】 在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例

中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0176】 在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高

序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0177】 在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS03-FL-NA (SEQ ID NO:6)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0178】 在一些實施例中，所編碼之因子IX多肽(例如由與CS03-FL-NA

(SEQ ID NO:6)具有高序列同源性之聚核苷酸編碼之多肽)與野生型因子IX前原蛋白序列FIX-FL-AA (SEQ ID NO:2)及/或Padua (hFIX (R384L))前原蛋白序列FIXp-FL-AA (SEQ ID NO:4)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除信號肽及原肽，及藉由切除活化多肽)之能力。

【0179】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-FL-AA (SEQ ID NO:2)。

【0180】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少95%之一致性且在前

原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-FL-AA (SEQ ID NO:4)。

【0181】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽(例如具有絲胺酸蛋白酶活性)之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之核苷酸序列。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0182】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少96%之一致性。在一

特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS03-MP-NA (SEQ ID NO:14)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS03-MP-NA (SEQ ID NO:14)。

【0183】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0184】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ

ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0185】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0186】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0187】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列

一致性之因子IX聚核苷酸進一步包括編碼具有FIX-SP-AA (SEQ ID NO:37)之胺基酸序列之因子IX信號肽的因子IX信號聚核苷酸。在一些實施例中，因子IX信號聚核苷酸具有與CS02-SP-NA (SEQ ID NO:25)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS03-SP-NA (SEQ ID NO:26)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS04-SP-NA (SEQ ID NO:27)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS05-SP-NA (SEQ ID NO:28)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS06-SP-NA (SEQ ID NO:29)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0188】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PP-AA (SEQ ID NO:38)之胺基酸序列之因子IX原肽的因子IX原肽聚核苷酸。在一些實施例中，因子IX原肽聚核苷酸具有與CS02-PP-NA (SEQ ID NO:31)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS03-PP-NA (SEQ ID NO:32)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS04-PP-NA (SEQ ID NO:33)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS05-PP-NA (SEQ ID NO:34)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS06-PP-NA (SEQ ID NO:35)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0189】 在一些實施例中，與CS03-MP-NA (SEQ ID NO:14)具有高序列

一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列之因子IX前原肽的因子IX前原肽聚核苷酸。在一些實施例中，因子IX前原肽聚核苷酸具有與CS02-PPP-NA (SEQ ID NO:19)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS03-PPP-NA (SEQ ID NO:20)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS04-PPP-NA (SEQ ID NO:21)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS05-PPP-NA (SEQ ID NO:22)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS06-PPP-NA (SEQ ID NO:23)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0190】 在一些實施例中，所編碼之因子IX多肽(例如由與CS03-FL-NA (SEQ ID NO:6)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0191】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列

與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0192】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多

肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0193】 在一個實施例中，本文所提供之密碼子改變之聚核苷酸編碼單鏈因子IX多肽，該單鏈因子IX多肽包括輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子。因子IX多肽之輕鏈由第一核苷酸序列編碼，該第一核苷酸序列與作為CS03-FL-NA (SEQ ID NO:6)之編碼因子IX輕鏈之部分的CS03-LC-NA (SEQ ID NO:44)具有高序列一致性。因子IX多肽之重鏈由第二核苷酸序列編碼，該第二核苷酸序列與作為CS03-FL-NA (SEQ ID NO:6)之編碼因子IX重鏈之部分的CS03-HC-NA (SEQ ID NO:43)具有高序列一致性。多肽連接子包括因子XI裂解位點，其允許活體內成熟(例如在前驅體單鏈因子IX多肽表現之後)。

【0194】 在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少95%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少96%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少97%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少98%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少99%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少99.5%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有至少99.9%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別為CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA

(SEQ ID NO:43)。

【0195】 在一些實施例中，因子IX構築體之多肽連接子由第三核苷酸序列編碼，該第三核苷酸序列與作為編碼野生型因子IX活化多肽(例如FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)之密碼子改變之序列的CS03-AP-NA (SEQ ID NO:58)具有高序列一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少80%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少90%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少95%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少96%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少97%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少98%之一致性。在一些實施例中，第三核苷酸序列與CS03-AP-NA (SEQ ID NO:58)具有至少99%之一致性。在一些實施例中，第三核苷酸序列為CS03-AP-NA (SEQ ID NO:58)。

【0196】 在一些實施例中，所編碼之因子IX多肽亦包括信號肽(例如因子IX信號肽)及/或原肽(例如因子IX原肽)。在一些實施例中，信號肽為野生型因子IX信號肽(FIX-SP-AA (SEQ ID NO:37))。在一些實施例中，信號肽由與CS03-SP-NA (SEQ ID NO:26)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。在一些實施例中，原肽為野生型因子IX原肽(FIX-PP-AA (SEQ ID NO:38))。在一些實施例中，原肽由與CS03-PP-NA (SEQ ID NO:32)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。

【0197】 在一些實施例中，所編碼之因子IX多肽(例如由與CS03-LC-NA (SEQ ID NO:44)及CS03-HC-NA (SEQ ID NO:43)具有高序列同源性之聚核苷酸

編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0198】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0199】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽

的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0200】 在一些實施例中，參考圖1，提供包括結構A之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS03-MP-NA (SEQ ID NO:14)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA

(SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS03-FL-NA (SEQ ID NO:6)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0201】 在一些實施例中，參考圖1，提供包括結構B之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS03-MP-NA (SEQ ID NO:14)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS03-FL-NA (SEQ ID NO:6)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0202】 在一些實施例中，參考圖1，提供包括結構C之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因

子IX多肽之核酸序列，該核酸序列與CS03-MP-NA (SEQ ID NO:14)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS03-FL-NA (SEQ ID NO:6)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0203】 在一些實施例中，參考圖1，提供包括結構D之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS03-MP-NA (SEQ ID NO:14)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼

序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS03-FL-NA (SEQ ID NO:6)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

CS04密碼子改變之聚核苷酸

【0204】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之核苷酸序列。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0205】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少99%之一致性。在一

特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-FL-NA (SEQ ID NO:7)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS04-FL-NA (SEQ ID NO:7)。

【0206】 在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0207】 在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列

具有50%至56%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0208】 在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0209】 在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10

個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS04-FL-NA (SEQ ID NO:7)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0210】 在一些實施例中，所編碼之因子IX多肽(例如由與CS04-FL-NA (SEQ ID NO:7)具有高序列同源性之聚核苷酸編碼之多肽)與野生型因子IX前原蛋白序列FIX-FL-AA (SEQ ID NO:2)及/或Padua (hFIX (R384L))前原蛋白序列FIXp-FL-AA (SEQ ID NO:4)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除信號肽及原肽，及藉由切除活化多肽)之能力。

【0211】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA

(SEQ ID NO:2)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-FL-AA (SEQ ID NO:2)。

【0212】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，

所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-FL-AA (SEQ ID NO:4)。

【0213】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽(例如具有絲胺酸蛋白酶活性)之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之核苷酸序列。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0214】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷

酸的序列與CS04-MP-NA (SEQ ID NO:15)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS04-MP-NA (SEQ ID NO:15)。

【0215】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0216】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC

含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0217】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0218】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具

有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0219】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-SP-AA (SEQ ID NO:37)之胺基酸序列之因子IX信號肽的因子IX信號聚核苷酸。在一些實施例中，因子IX信號聚核苷酸具有與CS02-SP-NA (SEQ ID NO:25)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS03-SP-NA (SEQ ID NO:26)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS04-SP-

NA (SEQ ID NO:27)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS05-SP-NA (SEQ ID NO:28)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS06-SP-NA (SEQ ID NO:29)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0220】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PP-AA (SEQ ID NO:38)之胺基酸序列之因子IX原肽的因子IX原肽聚核苷酸。在一些實施例中，因子IX原肽聚核苷酸具有與CS02-PP-NA (SEQ ID NO:31)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS03-PP-NA (SEQ ID NO:32)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS04-PP-NA (SEQ ID NO:33)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS05-PP-NA (SEQ ID NO:34)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS06-PP-NA (SEQ ID NO:35)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0221】 在一些實施例中，與CS04-MP-NA (SEQ ID NO:15)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列之因子IX前原肽的因子IX前原肽聚核苷酸。在一些實施例中，因子IX前原肽聚核苷酸具有與CS02-PPP-NA (SEQ ID NO:19)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS03-PPP-NA (SEQ ID NO:20)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚

核苷酸具有與CS04-PPP-NA (SEQ ID NO:21)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS05-PPP-NA (SEQ ID NO:22)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS06-PPP-NA (SEQ ID NO:23)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0222】 在一些實施例中，所編碼之因子IX多肽(例如由與CS04-FL-NA (SEQ ID NO:7)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0223】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在

一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0224】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0225】 在一個實施例中，本文所提供之密碼子改變之聚核苷酸編碼單鏈因子IX多肽，該單鏈因子IX多肽包括輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子。因子IX多肽之輕鏈由第一核苷酸序列編碼，該第一核苷酸序列與作為CS04-FL-NA (SEQ ID NO:7)之編碼因子IX輕鏈之部分的CS04-LC-

NA (SEQ ID NO:46)具有高序列一致性。因子IX多肽之重鏈由第二核苷酸序列編碼，該第二核苷酸序列與作為CS04-FL-NA (SEQ ID NO:7)之編碼因子IX重鏈之部分的CS04-HC-NA (SEQ ID NO:45)具有高序列一致性。多肽連接子包括因子XI裂解位點，其允許活體內成熟(例如在前驅體單鏈因子IX多肽表現之後)。

【0226】 在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少95%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少96%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少97%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少98%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少99%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少99.5%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)具有至少99.9%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別為CS04-LC-NA及CS04-HC-NA (SEQ ID NO:46及45)。

【0227】 在一些實施例中，因子IX構築體之多肽連接子由第三核苷酸序列編碼，該第三核苷酸序列與作為編碼野生型因子IX活化多肽(例如FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)之密碼子改變之序列的CS04-AP-NA (SEQ ID NO:59)具有高序列一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少80%之一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少90%之一致性。在一些實施例中，第三

核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少95%之一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少96%之一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少97%之一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少98%之一致性。在一些實施例中，第三核苷酸序列與CS04-AP-NA (SEQ ID NO:59)具有至少99%之一致性。在一些實施例中，第三核苷酸序列為CS04-AP-NA (SEQ ID NO:59)。

【0228】 在一些實施例中，所編碼之因子IX多肽亦包括信號肽(例如因子IX信號肽)及/或原肽(例如因子IX原肽)。在一些實施例中，信號肽為野生型因子IX信號肽(FIX-SP-AA (SEQ ID NO:37))。在一些實施例中，信號肽由與CS04-SP-NA (SEQ ID NO:27)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。在一些實施例中，原肽為野生型因子IX原肽(FIX-PP-AA (SEQ ID NO:38))。在一些實施例中，原肽由與CS04-PP-NA (SEQ ID NO:33)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。

【0229】 在一些實施例中，所編碼之因子IX多肽(例如由與CS04-LC-NA (SEQ ID NO:46)及CS04-HC-NA (SEQ ID NO:45)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0230】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例

中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0231】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所

編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0232】 在一些實施例中，參考圖1，提供包括結構A之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，其與CS04-MP-NA (SEQ ID NO:15)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS04-FL-NA (SEQ ID NO:7)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0233】 在一些實施例中，參考圖1，提供包括結構B之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多

肽之核酸序列，該核酸序列與CS04-MP-NA (SEQ ID NO:15)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS04-FL-NA (SEQ ID NO:7)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0234】 在一些實施例中，參考圖1，提供包括結構C之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS04-MP-NA (SEQ ID NO:15)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼

序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS04-FL-NA (SEQ ID NO:7)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0235】 在一些實施例中，參考圖1，提供包括結構D之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS04-MP-NA (SEQ ID NO:15)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS04-FL-NA (SEQ ID NO:7)具有至少95%、96%、97%、

98%、99%、99.5%、99.9%或100%之一致性。

CS05密碼子改變之聚核苷酸

【0236】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之核苷酸序列。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0237】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-FL-NA (SEQ ID NO:8)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS05-FL-NA (SEQ ID NO:8)。

【0238】 在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸

的序列具有少於59%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0239】 在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0240】 在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸

的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 53.8% 之GC含量。

【0241】 在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實

施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS05-FL-NA (SEQ ID NO:8)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0242】 在一些實施例中，所編碼之因子IX多肽(例如由與CS05-FL-NA (SEQ ID NO:8)具有高序列同源性之聚核苷酸編碼之多肽)與野生型因子IX前原蛋白序列FIX-FL-AA (SEQ ID NO:2)及/或Padua (hFIX (R384L))前原蛋白序列FIXp-FL-AA (SEQ ID NO:4)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除信號肽及原肽，及藉由切除活化多肽)之能力。

【0243】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少98%之一致性。在一

個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-FL-AA (SEQ ID NO:2)。

【0244】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-FL-AA (SEQ ID NO:4)。

【0245】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽(例如具有絲胺酸蛋白酶活性)之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之核苷酸序列。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0246】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS05-MP-NA (SEQ ID NO:16)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS05-MP-NA (SEQ ID NO:16)。

【0247】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含

量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0248】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0249】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之

GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 53.8% 之GC含量。

【0250】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致

性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0251】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-SP-AA (SEQ ID NO:37)之胺基酸序列之因子IX信號肽的因子IX信號聚核苷酸。在一些實施例中，因子IX信號聚核苷酸具有與CS02-SP-NA (SEQ ID NO:25)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS03-SP-NA (SEQ ID NO:26)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS04-SP-NA (SEQ ID NO:27)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS05-SP-NA (SEQ ID NO:28)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS06-SP-NA (SEQ ID NO:29)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0252】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PP-AA (SEQ ID NO:38)之胺

基酸序列之因子IX原肽的因子IX原肽聚核苷酸。在一些實施例中，因子IX原肽聚核苷酸具有與CS02-PP-NA (SEQ ID NO:31)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS03-PP-NA (SEQ ID NO:32)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS04-PP-NA (SEQ ID NO:33)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS05-PP-NA (SEQ ID NO:34)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS06-PP-NA (SEQ ID NO:35)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0253】 在一些實施例中，與CS05-MP-NA (SEQ ID NO:16)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列之因子IX前原肽的因子IX前原肽聚核苷酸。在一些實施例中，因子IX前原肽聚核苷酸具有與CS02-PPP-NA (SEQ ID NO:19)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS03-PPP-NA (SEQ ID NO:20)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS04-PPP-NA (SEQ ID NO:21)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS05-PPP-NA (SEQ ID NO:22)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS06-PPP-NA (SEQ ID NO:23)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0254】 在一些實施例中，所編碼之因子IX多肽(例如由與CS05-FL-NA

(SEQ ID NO:8)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0255】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0256】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且

在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0257】 在一個實施例中，本文所提供之密碼子改變之聚核苷酸編碼單鏈因子IX多肽，該單鏈因子IX多肽包括輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子。因子IX多肽之輕鏈由第一核苷酸序列編碼，該第一核苷酸序列與作為CS05-FL-NA (SEQ ID NO:8)之編碼因子IX輕鏈之部分的CS05-LC-NA (SEQ ID NO:48)具有高序列一致性。因子IX多肽之重鏈由第二核苷酸序列編碼，該第二核苷酸序列與作為CS05-FL-NA (SEQ ID NO:8)之編碼因子IX重鏈之部分的CS05-HC-NA (SEQ ID NO:47)具有高序列一致性。多肽連接子包括因子XI裂解位點，其允許活體內成熟(例如在前驅體單鏈因子IX多肽表現之後)。

【0258】 在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少95%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID

NO:48及47)具有至少96%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少97%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少98%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少99%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少99.5%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)具有至少99.9%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別為CS05-LC-NA及CS05-HC-NA (SEQ ID NO:48及47)。

【0259】 在一些實施例中，因子IX構築體之多肽連接子由第三核苷酸序列編碼，該第三核苷酸序列與作為編碼野生型因子IX活化多肽(例如FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)之密碼子改變之序列的CS05-AP-NA (SEQ ID NO:60)具有高序列一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少80%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少90%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少95%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少96%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少97%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少98%之一致性。在一些實施例中，第三核苷酸序列與CS05-AP-NA (SEQ ID NO:60)具有至少99%之一致性。在一些實施例中，第三核苷酸序列為CS05-AP-NA (SEQ ID NO:60)。

【0260】 在一些實施例中，所編碼之因子IX多肽亦包括信號肽(例如因子IX信號肽)及/或原肽(例如因子IX原肽)。在一些實施例中，信號肽為野生型因子IX信號肽(FIX-SP-AA (SEQ ID NO:37))。在一些實施例中，信號肽由與CS05-SP-NA (SEQ ID NO:28)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。在一些實施例中，原肽為野生型因子IX原肽(FIX-PP-AA (SEQ ID NO:38))。在一些實施例中，原肽由與CS05-PP-NA (SEQ ID NO:34)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。

【0261】 在一些實施例中，所編碼之因子IX多肽(例如由與CS05-LC-NA (SEQ ID NO:48)及CS05-HC-NA (SEQ ID NO:47)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0262】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-

MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0263】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0264】 在一些實施例中，參考圖1，提供包括結構A之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多

肽之核酸序列，其與CS03-MP-NA (SEQ ID NO:14)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS05-FL-NA (SEQ ID NO:8)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0265】 在一些實施例中，參考圖1，提供包括結構B之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS05-MP-NA (SEQ ID NO:16)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部

分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS05-FL-NA (SEQ ID NO:8)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0266】 在一些實施例中，參考圖1，提供包括結構C之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS05-MP-NA (SEQ ID NO:16)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS05-FL-NA (SEQ ID NO:8)具有至少95%、96%、97%、

98%、99%、99.5%、99.9%或100%之一致性。

【0267】 在一些實施例中，參考圖1，提供包括結構D之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS05-MP-NA (SEQ ID NO:16)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS05-FL-NA (SEQ ID NO:8)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

CS06密碼子改變之聚核苷酸

【0268】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之核苷酸序列。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低

之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0269】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-FL-NA (SEQ ID NO:9)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS06-FL-NA (SEQ ID NO:9)。

【0270】 在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施

例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0271】 在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0272】 在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些

實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0273】 在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS06-FL-NA

(SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS06-FL-NA (SEQ ID NO:9)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0274】 在一些實施例中，所編碼之因子IX多肽(例如由與CS06-FL-NA (SEQ ID NO:9)具有高序列同源性之聚核苷酸編碼之多肽)與野生型因子IX前原蛋白序列FIX-FL-AA (SEQ ID NO:2)及/或Padua (hFIX (R384L))前原蛋白序列FIXp-FL-AA (SEQ ID NO:4)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除信號肽及原肽，及藉由切除活化多肽)之能力。

【0275】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-FL-AA (SEQ ID NO:2)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-FL-AA (SEQ ID NO:2)。

【0276】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX

單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-FL-AA (SEQ ID NO:4)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-FL-AA (SEQ ID NO:4)。

【0277】 在一個實施例中，本文所提供之核酸組合物包括編碼單鏈因子IX多肽(例如具有絲胺酸蛋白酶活性)之因子IX聚核苷酸(例如密碼子改變之聚核苷酸)，其中因子IX聚核苷酸包括與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之核苷酸序列。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有降低之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之因子IX聚核苷酸之核苷酸序列與野生型因

子IX編碼序列(FIX-FL-NA (SEQ ID NO:1))相比具有減少之CpG二核苷酸數目。

【0278】 在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少95%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少96%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少97%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少98%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少99%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少99.5%之一致性。在一特定實施例中，密碼子改變之聚核苷酸的序列與CS06-MP-NA (SEQ ID NO:17)具有至少99.9%之一致性。在另一特定實施例中，密碼子改變之聚核苷酸的序列為CS06-MP-NA (SEQ ID NO:17)。

【0279】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於60%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於59%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於58%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於57%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於56%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於55%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有少於54%之GC含量。

【0280】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至60%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至59%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至58%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至57%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至56%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至55%之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有50%至54%之GC含量。

【0281】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 1.0$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.8$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.6$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.5$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.4$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.3$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.2$ 之GC含量。在一些實施例中，與

CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有 $53.8\% \pm 0.1$ 之GC含量。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有53.8%之GC含量。

【0282】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過15個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過12個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過10個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過9個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過8個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過7個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過6個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過5個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過4個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過3個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列具有不超過2個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之

聚核苷酸的序列具有不超過1個CpG二核苷酸。在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之密碼子改變之聚核苷酸的序列不具有CpG二核苷酸。

【0283】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-SP-AA (SEQ ID NO:37)之胺基酸序列之因子IX信號肽的因子IX信號聚核苷酸。在一些實施例中，因子IX信號聚核苷酸具有與CS02-SP-NA (SEQ ID NO:25)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS03-SP-NA (SEQ ID NO:26)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS04-SP-NA (SEQ ID NO:27)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS05-SP-NA (SEQ ID NO:28)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX信號聚核苷酸具有與CS06-SP-NA (SEQ ID NO:29)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0284】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PP-AA (SEQ ID NO:38)之胺基酸序列之因子IX原肽的因子IX原肽聚核苷酸。在一些實施例中，因子IX原肽聚核苷酸具有與CS02-PP-NA (SEQ ID NO:31)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS03-PP-NA (SEQ ID NO:32)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS04-PP-NA (SEQ ID NO:33)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS05-PP-NA (SEQ ID

NO:34)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX原肽聚核苷酸具有與CS06-PP-NA (SEQ ID NO:35)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0285】 在一些實施例中，與CS06-MP-NA (SEQ ID NO:17)具有高序列一致性之因子IX聚核苷酸進一步包括編碼具有FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列之因子IX前原肽的因子IX前原肽聚核苷酸。在一些實施例中，因子IX前原肽聚核苷酸具有與CS02-PPP-NA (SEQ ID NO:19)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS03-PPP-NA (SEQ ID NO:20)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS04-PPP-NA (SEQ ID NO:21)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS05-PPP-NA (SEQ ID NO:22)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。在一些實施例中，因子IX前原肽聚核苷酸具有與CS06-PPP-NA (SEQ ID NO:23)至少90%、95%、96%、97%、98%、99%或100%一致之核酸序列。

【0286】 在一些實施例中，所編碼之因子IX多肽(例如由與CS06-FL-NA (SEQ ID NO:9)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0287】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多

肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0288】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有

至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0289】 在一個實施例中，本文所提供之密碼子改變之聚核苷酸編碼單鏈因子IX多肽，該單鏈因子IX多肽包括輕鏈、重鏈及將輕鏈之C端連接至重鏈之N端的多肽連接子。因子IX多肽之輕鏈由第一核苷酸序列編碼，該第一核苷酸序列與作為CS06-FL-NA (SEQ ID NO:9)之編碼因子IX輕鏈之部分的CS06-LC-NA (SEQ ID NO:50)具有高序列一致性。因子IX多肽之重鏈由第二核苷酸序列編碼，該第二核苷酸序列與作為CS06-FL-NA (SEQ ID NO:9)之編碼因子IX重鏈之部分的CS06-HC-NA (SEQ ID NO:49)具有高序列一致性。多肽連接子包括因子XI裂解位點，其允許活體內成熟(例如在前驅體單鏈因子IX多肽表現之後)。

【0290】 在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)。具有至少95%之序列一致性 在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少96%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少97%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少98%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少99%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少99.5%之序列一

致性。在一些實施例中，第一及第二核苷酸序列分別與CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)具有至少99.9%之序列一致性。在一些實施例中，第一及第二核苷酸序列分別為CS06-LC-NA及CS06-HC-NA (SEQ ID NO:50及49)。

【0291】 在一些實施例中，因子IX構築體之多肽連接子由第三核苷酸序列編碼，該第三核苷酸序列與作為編碼野生型因子IX活化多肽(例如FIX-FL-AA (SEQ ID NO:2)之胺基酸192-226)之密碼子改變之序列的CS06-AP-NA (SEQ ID NO:61)具有高序列一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少80%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少90%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少95%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少96%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少97%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少98%之一致性。在一些實施例中，第三核苷酸序列與CS06-AP-NA (SEQ ID NO:61)具有至少99%之一致性。在一些實施例中，第三核苷酸序列為CS06-AP-NA (SEQ ID NO:61)。

【0292】 在一些實施例中，所編碼之因子IX多肽亦包括信號肽(例如因子IX信號肽)及/或原肽(例如因子IX原肽)。在一些實施例中，信號肽為野生型因子IX信號肽(FIX-SP-AA (SEQ ID NO:37))。在一些實施例中，信號肽由與CS06-SP-NA (SEQ ID NO:29)具有高序列一致性(例如至少95%、96%、97%、98%或99%)之密碼子改變之聚核苷酸序列編碼。在一些實施例中，原肽為野生型因子IX原肽(FIX-PP-AA (SEQ ID NO:38))。在一些實施例中，原肽由與CS06-PP-NA (SEQ ID NO:35)具有高序列一致性(例如至少95%、96%、97%、98%或99%)

之密碼子改變之聚核苷酸序列編碼。

【0293】 在一些實施例中，所編碼之因子IX多肽(例如由與CS06-LC-NA (SEQ ID NO:50)及CS06-HC-NA (SEQ ID NO:49)具有高序列同源性之聚核苷酸編碼之多肽)與野生型成熟因子IX單鏈多肽序列FIX-MP-AA (SEQ ID NO:10)及/或成熟Padua (hFIX (R384L))單鏈序列FIXp-MP-AA (SEQ ID NO:12)具有高序列一致性。所編碼之因子IX多肽應保留變得活化為功能因子IXa蛋白(例如藉由移除任何信號肽及原肽，及藉由切除活化多肽)之能力。

【0294】 在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少85%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少90%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少95%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少96%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少97%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少98%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.5%之一致性。在一個實施例中，所編碼之因子IX多肽的序列與FIX-MP-AA (SEQ ID NO:10)具有至少99.9%之一致性。在一個實施例中，所編碼之因子IX多肽的序列為FIX-MP-AA (SEQ ID NO:10)。

【0295】 在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少85%之一致性且在前原多肽之位置384 (例如成熟因子IX單鏈多肽FIXp-MP-AA (SEQ ID NO:12)之位置338)處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少

90%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少95%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少96%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少97%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少98%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.5%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列與FIXp-MP-AA (SEQ ID NO:12)具有至少99.9%之一致性且在前原多肽之位置384處包括白胺酸。在一個實施例中，所編碼之因子IX多肽的序列為FIXp-MP-AA (SEQ ID NO:12)。

【0296】 在一些實施例中，參考圖1，提供包括結構A之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，其與CS06-MP-NA (SEQ ID NO:17)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包

括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS06-FL-NA (SEQ ID NO:9)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0297】 在一些實施例中，參考圖1，提供包括結構B之自我互補型聚核苷酸的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS06-MP-NA (SEQ ID NO:17)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS06-FL-NA (SEQ ID NO:9)具有至少95%、96%、97%、

98%、99%、99.5%、99.9%或100%之一致性。

【0298】 在一些實施例中，參考圖1，提供包括結構C之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS06-MP-NA (SEQ ID NO:17)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS06-FL-NA (SEQ ID NO:9)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

【0299】 在一些實施例中，參考圖1，提供包括結構D之聚核苷酸(例如單股聚核苷酸)的核酸組合物，其中聚核苷酸之FIX編碼序列部分包括編碼成熟因子IX多肽之核酸序列，該核酸序列與CS06-MP-NA (SEQ ID NO:17)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX信號肽之核酸序列，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%

之一致性：FIX-SP-NA (SEQ ID NO:24)、CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)及CS06-SP-NA (SEQ ID NO:29)。在一些實施例中，聚核苷酸之FIX編碼序列部分亦包括編碼因子IX原肽之核酸序列(視情況與如上文所描述之因子IX信號肽之核酸序列組合)，該核酸序列與以下中之一者具有至少90%、95%、96%、97%、98%、99%或100%之一致性：FIX-PP-NA (SEQ ID NO:30)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID NO:33)、CS05-PP-NA (SEQ ID NO:34)及CS06-PP-NA (SEQ ID NO:35)。在一些實施例中，聚核苷酸之FIX編碼序列部分包括編碼前原因子IX多肽之核酸序列，該核酸序列與CS06-FL-NA (SEQ ID NO:9)具有至少95%、96%、97%、98%、99%、99.5%、99.9%或100%之一致性。

C. 密碼子改變之因子IX信號肽及原肽

【0300】 在一個態樣中，本發明提供編碼因子IX信號肽、因子IX原肽及兩者(例如因子IX前原肽)之密碼子改變之聚核苷酸。此等密碼子改變之聚核苷酸改良因子IX表現且可位於例如密碼子改變的或以其他方式編碼因子IX單鏈多肽(例如因子IX輕鏈、活化肽及重鏈)之聚核苷酸的上游。通常，所編碼之肽為野生型因子IX信號肽(例如FIX-SP-AA (SEQ ID NO:37))、原肽(例如FIX-PP-AA (SEQ ID NO:38)及前原肽(FIX-PPP-AA (SEQ ID NO:36))。

【0301】 在某些實施例中，編碼因子IX信號肽、原肽及前原肽之密碼子改變之聚核苷酸具有與以下中之一者具有高一致性(例如至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%)的序列：CS02-SP-NA (SEQ ID NO:25)、CS03-SP-NA (SEQ ID NO:26)、CS04-SP-NA (SEQ ID NO:27)、CS05-SP-NA (SEQ ID NO:28)、CS06-SP-NA (SEQ ID NO:29)、CS02-PP-NA (SEQ ID NO:31)、CS03-PP-NA (SEQ ID NO:32)、CS04-PP-NA (SEQ ID

NO:33)、CS05-PP-NA (SEQ ID NO:34)、CS06-PP-NA (SEQ ID NO:35)、CS02-PPP-NA (SEQ ID NO:19)、CS03-PPP-NA (SEQ ID NO:20)、CS04-PPP-NA (SEQ ID NO:21)、CS05-PPP-NA (SEQ ID NO:22)及CS06-PPP-NA (SEQ ID NO:23)。

CS02信號肽及原肽

【0302】 在一個實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS02-SP-NA (SEQ ID NO:25)具有至少95%之序列一致性。在其他實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS02-SP-NA (SEQ ID NO:25)具有至少96%、97%、98%、99%或100%之一致性。

【0303】 在一個實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS02-PP-NA (SEQ ID NO:31)具有至少95%之序列一致性。在其他實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS02-PP-NA (SEQ ID NO:31)具有至少96%、97%、98%、99%或100%之一致性。

【0304】 在一個實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS02-PPP-NA (SEQ ID NO:19)具有至少95%之序列一致性。在其他實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS02-PPP-NA (SEQ ID NO:19)具有至少96%、97%、98%、99%或100%之一致性。

CS03信號肽及原肽

【0305】 在一個實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS03-SP-NA (SEQ ID NO:26)具有至少95%之序列一致性。在其他實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS03-SP-NA (SEQ ID NO:26)具有至少96%、97%、98%、99%或100%之一致性。

【0306】 在一個實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS03-PP-NA (SEQ ID NO:32)具有至少95%之序列一致性。在其他實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS03-PP-NA (SEQ ID NO:32)具有至

少96%、97%、98%、99%或100%之一致性。

【0307】 在一個實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS03-PPP-NA (SEQ ID NO:20)具有至少95%之序列一致性。在其他實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS03-PPP-NA (SEQ ID NO:20)具有至少96%、97%、98%、99%或100%之一致性。

CS04信號肽及原肽

【0308】 在一個實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS04-SP-NA (SEQ ID NO:27)具有至少95%之序列一致性。在其他實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS04-SP-NA (SEQ ID NO:27)具有至少96%、97%、98%、99%或100%之一致性。

【0309】 在一個實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS04-PP-NA (SEQ ID NO:33)具有至少95%之序列一致性。在其他實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS04-PP-NA (SEQ ID NO:33)具有至少96%、97%、98%、99%或100%之一致性。

【0310】 在一個實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS04-PPP-NA (SEQ ID NO:21)具有至少95%之序列一致性。在其他實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS04-PPP-NA (SEQ ID NO:21)具有至少96%、97%、98%、99%或100%之一致性。

CS05信號肽及原肽

【0311】 在一個實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS05-SP-NA (SEQ ID NO:28)具有至少95%之序列一致性。在其他實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS05-SP-NA (SEQ ID NO:28)具有至少96%、97%、98%、99%或100%之一致性。

【0312】 在一個實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與

CS05-PP-NA (SEQ ID NO:34)具有至少95%之序列一致性。在其他實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS05-PP-NA (SEQ ID NO:34)具有至少96%、97%、98%、99%或100%之一致性。

【0313】 在一個實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS05-PPP-NA (SEQ ID NO:22)具有至少95%之序列一致性。在其他實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS05-PPP-NA (SEQ ID NO:22)具有至少96%、97%、98%、99%或100%之一致性。

CS06信號肽及原肽

【0314】 在一個實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS06-SP-NA (SEQ ID NO:29)具有至少95%之序列一致性。在其他實施例中，編碼因子IX信號肽之密碼子改變之聚核苷酸與CS06-SP-NA (SEQ ID NO:29)具有至少96%、97%、98%、99%或100%之一致性。

【0315】 在一個實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS06-PP-NA (SEQ ID NO:35)具有至少95%之序列一致性。在其他實施例中，編碼因子IX原肽之密碼子改變之聚核苷酸與CS06-PP-NA (SEQ ID NO:35)具有至少96%、97%、98%、99%或100%之一致性。

【0316】 在一個實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS06-PPP-NA (SEQ ID NO:23)具有至少95%之序列一致性。在其他實施例中，編碼因子IX前原肽之密碼子改變之聚核苷酸與CS06-PPP-NA (SEQ ID NO:23)具有至少96%、97%、98%、99%或100%之一致性。

IV. 因子IX表現載體

【0317】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸整合至表現載體中。如熟習此項技術者將瞭解，許多形式之載體可用於使用本文所揭示之密碼子改變之因子IX聚核苷酸序列實現因子IX基因療法。表現載體之非限

制性實例包括病毒載體(例如適合用於基因療法之載體)、質粒載體、噬菌體載體、黏粒、噬菌粒、人工染色體及類似載體。

【0318】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸整合至病毒基因療法載體中。病毒載體之非限制性實例包括：逆轉錄病毒，例如莫洛尼鼠白血病毒(Moloney murine leukemia virus, MMLV)、哈維鼠肉瘤病毒(Harvey murine sarcoma virus)、鼠乳腺腫瘤病毒(murine mammary tumor virus)及勞斯氏肉瘤病毒(Rous sarcoma virus)；腺病毒、腺相關病毒；SV40型病毒；多瘤病毒；愛潑斯坦-巴爾病毒(Epstein-Barr virus)；乳頭瘤病毒；疱疹病毒；痘苗病毒；及脊髓灰質炎病毒。

【0319】 在活體內，因子IX主要在肝中合成。因而，已靶向肝細胞作為適合於因子IX基因療法構築體之宿主細胞。若干類別之病毒載體已顯示有能力進行基因療法構築體之肝靶向遞送，包括逆轉錄病毒載體(參見例如Axelrod等人, 1990；Kay等人, 1992；Van den Driessche等人, 1999；及Xu等人, 2003, 2005，其揭示內容出於所有目的以全文引用之方式明確地併入本文中)、慢病毒(參見例如Ward等人, 2011；Brown等人, 2007；及Matrai等人, 2011，其揭示內容出於所有目的以全文引用之方式明確地併入本文中)、腺相關病毒(AAV) (參見例如Herzog等人, 1999，其揭示內容出於所有目的以全文引用之方式明確地併入本文中)及腺病毒載體(參見例如Brown等人, 2004及Ehrhardt及Kay, 2002，其揭示內容出於所有目的以全文引用之方式明確地併入本文中)。

【0320】 在一些實施例中，基因療法載體為逆轉錄病毒，且特定而言為複製缺陷型逆轉錄病毒。用於製備複製缺陷型逆轉錄病毒之方案為此項技術中已知的。關於綜述，參見Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co., New York (1990)及Murry, E. J., Methods in Molecular Biology, 第7卷, Humana Press, Inc., Clifton, N.J. (1991)。

【0321】 在一個實施例中，基因療法載體為基於腺相關病毒(AAV)之基因療法載體。AAV系統先前已有描述且為此項技術中普遍熟知的(Kelleher及Vos, *Biotechniques*, 17(6):1110-17 (1994)；Cotten等 人, *P.N.A.S. U.S.A.*, 89(13):6094-98 (1992)；Curiel, *Nat Immun*, 13(2-3):141-64 (1994)；Muzyczka, *Curr Top Microbiol Immunol*, 158:97-129 (1992)；及Asokan A等人, *Mol. Ther.*, 20(4):699-708 (2012)，各自出於所有目的以全文引用之方式併入本文中)。關於rAAV載體之產生及使用的細節描述於例如美國專利第5,139,941號及第4,797,368號中，該等專利各自出於所有目的以全文引用之方式併入本文中。在一特定實施例中，AAV載體為AAV-8載體。

【0322】 用於肝特異性因子IX表現之示例性AAV遞送載體描述於WO 2009/130208中，該專利之內容出於所有目的以全文引用之方式明確併入本文中。該載體為編碼人類因子IX之單股AAV載體，且包括驅動因子cDNA之TTR *Serp*調控序列。該載體亦包括人類因子IX基因之內含子I及聚腺苷酸化信號。

【0323】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸整合至逆轉錄病毒表現載體中。此等系統先前已有描述，且為此項技術中普遍熟知的(Mann等人, *Cell*, 33:153-159, 1983；Nicolas及Rubinstein, *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez及Denhardt編, Stoneham: Butterworth, 第494-513頁, 1988；Temin, *Gene Transfer*, Kucherlapati (編), New York: Plenum Press, 第149-188頁, 1986)。在一特定實施例中，逆轉錄病毒載體為慢病毒載體(參見例如Naldini等人, *Science*, 272(5259):263-267, 1996；Zufferey等人, *Nat Biotechnol*, 15(9):871-875, 1997；Blomer等人, *J Virol.*, 71(9):6641-6649, 1997；美國專利第6,013,516號及第5,994,136號)。

【0324】 在一些實施例中，本文所描述之密碼子改變之聚核苷酸可藉由非病毒方法向個體投與。舉例而言，裸DNA可藉由電穿孔、聲孔作用、粒子轟

擊或水動力遞送投與至細胞中。亦可將DNA與聚合物(例如脂質體、聚核糖體、polyplexes、樹狀聚體)一起封裝或偶合，且作為複合物投與個體。同樣地，DNA可偶合至無機奈米粒子，例如金、二氧化矽、氧化鐵或磷酸鈣粒子，或附接至細胞穿透肽以活體內遞送至細胞。

【0325】 密碼子改變之因子IX編碼聚核苷酸亦可併入人工染色體中，諸如人工染色體表現(ACE) (參見例如Lindenbaum等人, *Nucleic Acids Res.*, 32(21):e172 (2004))及哺乳動物人工染色體(MACS)。關於綜述，參見例如Pérez-Luz及Díaz-Nido, *J Biomed Biotechnol.* 2010: Article ID 642804 (2010)。

【0326】 多種載體可用於在細胞培養物中由密碼子改變之多肽表現因子IX多肽，包括真核及原核表現載體。在某些實施例中，涵蓋用於在細胞培養物中表現因子IX多肽之質粒載體。一般而言，結合此等宿主使用含有複製子及源自與宿主細胞相容之物種的控制序列之質粒載體。該載體可帶有複製位點，以及能夠在轉型之細胞中提供表型選擇的標誌序列。質粒將包括編碼因子IX多肽之密碼子改變之聚核苷酸，其可操作地連接至一或多個控制序列，例如啟動子。

【0327】 用於原核表現之載體的非限制性實例包括質粒，諸如pRSET、pET、pBAD等，其中用於原核表現載體中之啟動子包括lac、trc、trp、recA、araBAD等。用於真核表現之載體的實例包括：(i)對於在酵母中表現，載體為諸如pAO、pPIC、pYES、pMET，使用諸如AOX1、GAP、GAL1、AUG1之啟動子；(ii)對於在昆蟲細胞中表現，載體為諸如pMT、pAc5、pIB、pMIB、pBAC等，使用諸如PH、p10、MT、Ac5、OpIE2、gp64、polh之啟動子，以及(iii)對於在哺乳動物細胞中表現，載體為諸如pSVL、pCMV、pRc/RSV、pcDNA3、pBPV等，以及源自諸如痘苗病毒、腺相關病毒、疱疹病毒、逆轉錄病毒等病毒系統之載體，使用諸如CMV、SV40、EF-1、UbC、RSV、ADV、BPV以及β-

肌動蛋白之啟動子。

【0328】 在一些實施例中，本發明提供AAV基因療法載體，其包括如本文所描述之密碼子改變之因子IX聚核苷酸、載體5'及3'末端之內部末端重複(ITR)序列、可操作地連接至因子IX聚核苷酸之一或多個啟動子及/或增強子序列及在因子IX聚核苷酸序列3'末端之後的聚腺苷酸化信號。在一些實施例中，一或多個啟動子及/或增強子序列包括肝特異性調控控制元件之一或多個拷貝。

【0329】 圖1說明根據一些實現方式之因子IX基因療法載體的若干示例性架構。圖1A說明具有突變之5' ITR、截短之TTR增強子/啟動子序列、MVM病毒內含子序列、密碼子改變之因子IX編碼序列、聚腺苷酸化序列及3'-ITR之自我互補型AAV載體。圖1B說明類似於圖1A之編碼因子IX多肽之自我互補型AAV載體，但進一步包括一或多個(例如一個、兩個、三個或更多個)肝特異性調控控制元件。圖1C說明具有與圖1A相同之元件的單股載體，除了5'-ITR未突變，從而防止自我互補性。圖1D說明類似於圖1A之編碼因子IX多肽之單股AAV載體，但進一步包括一或多個(例如一個、兩個、三個或更多個)肝特異性調控控制元件。雖然參考圖1中包括R384L『Padua』胺基酸取代之因子IX蛋白進行了說明，但在一些實施例中，具有如1圖中所描繪之一般結構(例如結構A、B、C或D)之因子IX核苷酸構築體編碼不包括R384L『Padua』胺基酸取代之因子IX蛋白。

【0330】 圖25顯示AAV因子IX基因療法載體CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸序列，其例示圖1D中所說明之基因療法載體架構。CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸1-145為AAV2 5'-ITR序列(SEQ ID NO:51)。5'-ITR序列隨後為位於核苷酸165-236、238-309及311-382處之肝特異性CRM8調控控制元件CRM8 (SEQ ID NO:39)之三個拷貝。在CRM8序列之後為位於核苷酸383-712處之截短之TTR增強子/啟動子序列(SEQ ID NO:52)。接著，

該載體包括位於核苷酸724-800處之小鼠(MVM)細小病毒內含子(SEQ ID NO:53)。該載體之核苷酸814-2199為CS06密碼子改變之因子IX (R384L)編碼序列(CS06-FL-NA (SEQ ID NO:9))。因子IX聚核苷酸序列隨後為位於核苷酸2208-2441處之BGH聚腺苷酸化信號，且最後為位於核苷酸2458-2602處之AAV2 3'-ITR序列(SEQ ID NO:55)。

【0331】 已表徵若干AAV血清型，包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8及AAV9。通常，任何AAV血清型均可用於本文所描述之因子IX基因療法構築體。然而，該等血清型具有不同趨向性，例如其優先感染不同組織。在一個實施例中，因為因子IX主要在肝中產生，基於至少在血清型AAV7、AAV8及AAV9中發現之肝趨向性來選擇用於所揭示之基因療法構築體的AAV血清型。因此，在一個實施例中，因子IX基因療法構築體為AAV7血清型載體。在另一實施例中，因子IX基因療法構築體為AAV8血清型載體。在另一實施例中，因子IX基因療法構築體為AAV9血清型載體。

【0332】 本文所描述之因子IX基因療法構築體可為單股的(例如ssAAV載體，如圖1C及1D中所說明)或自我互補的(例如scAAV載體，如圖1A及1B中所說明)。雖然研究及理論已表明自我互補型AAV載體將促成更佳之轉基因表現，但藉由回避在轉譯之前對第二股合成之需要，如實例5中所報導，單股AAV載體經鑑定促成類似於自我互補型載體之更佳因子IX表現。

啟動子及增強子

【0333】 本文所描述之因子IX基因療法構築體通常包括驅動活體內基因表現之一或多個啟動子及/或增強子元件，例如調控元件。在一些實施例中，啟動子或增強子元件以組織依賴性方式例如主要在特定組織中驅動表現。因為因子IX主要在肝中合成，故在一些實施例中，本文所描述之基因療法載體包括肝特異性調控元件，其實質上將基因療法載體之表現限制於肝細胞。

【0334】 通常，肝特異性調控元件可源自已知排他地在肝中表現之任何基因。WO 2009/130208鑑定出若干以肝特異性方式表現之基因，包括serpin肽酶抑制劑A支成員1 (亦稱為 α -抗胰蛋白酶) (SERPINA1；GeneID 5265)、脂蛋白元C-I (APOC1；GeneID 341)、脂蛋白元C-IV (APOC4；GeneID 346)、脂蛋白元H (APOH；GeneID 350)、甲狀腺素轉運蛋白(TTR；GeneID 7276)、白蛋白 (ALB；GeneID 213)、醛縮酶B (ALDOB；GeneID 229)、細胞色素P450家族2亞家族E多肽1 (CYP2E1；GeneID 1571)、纖維原蛋白 α 鏈(FGA；GeneID 2243)、轉鐵蛋白(TF；GeneID 7018)、結合球蛋白相關蛋白(HPR；GeneID 3250)。在一些實施例中，本文所描述之因子IX基因療法構築體包括源自此等蛋白質中之一者或多者之基因組基因座的肝特異性調控元件。此類元件之若干實例描述於WO 2009/130208中，該專利之內容出於所有目的以全文引用之方式明確併入本文中。

【0335】 肝臟特異性肝特異性調控元件之一個實例為來自甲狀腺素轉運蛋白(TTR)基因，通常稱為「TTR_e」或「TTR_{Enh}」。Hsieh J.L.等人, *Cancer Sci.*, 100(3):537-45 (2009)，其內容出於所有目的以全文引用之方式明確併入本文中。在一些實施例中，本文所描述之因子IX基因療法構築體包括截短之TTR增強子及啟動子元件。此等元件之實例提供於如圖25所提供之CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸383-712處。在一些實施例中，截短之TTR增強子及啟動子元件與CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸383-712具有至少85%的序列一致性。在其他實施例中，截短之TTR增強子及啟動子元件與CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸383-712具有至少90%、95%、96%、97%、98%、99%、99.5%或100%的序列一致性。

【0336】 肝特異性調控元件之另一實例為如PCT公開案第WO 2016/146757號中所描述來自SERPINA1基因，該公開案之內容出於所有目的以

全文引用之方式明確併入本文中。此類元件之實例為提供於CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸165-236處的CRM8調控控制元件(SEQ ID NO:39)。在一些實施例中，SERPINA1源性調控控制元件與CRM8 (SEQ ID NO:39)具有至少85%之序列一致性。在其他實施例中，截短之SERPINA1源性調控控制元件與CRM8 (SEQ ID NO:39)具有至少90%、95%、96%、97%、98%、99%或100%之序列一致性。

【0337】 在一些實施例中，如由圖1B及1D中說明之構築體所例示，因子IX基因療法構築體包括一或多個SERPINA1源性調控控制元件。在一個實施例中，構築體包括一個SERPINA1源性調控控制元件(例如CRM8)。在另一實施例中，構築體包括兩個SERPINA1源性調控控制元件(例如CRM8)。在另一實施例中，構築體包括三個SERPINA1源性調控控制元件(例如CRM8)。在其他實施例中，構築體包括4、5、6或更多個SERPINA1源性調控控制元件(例如CRM8)。

【0338】 在一個實施例中，如圖1B、1D及25中所例示，因子IX基因療法構築體包括一或多個SERPINA1源性調控控制元件(例如CRM8)及截短之TTR增強子及啟動子元件。

內含子

【0339】 在一些實施例中，本文所描述之因子IX基因療法構築體包括內含子，例如病毒源性內含子，以增加因子IX基因之表現。適合於基因療法構築體之表現的內含子為此項技術中已知的。典型地，如圖1及圖25中所示之因子IX構築體中所例示，內含子定位於轉基因編碼序列之5'。然而，在一些實施例中，內含子可例如位於天然因子IX內含子接點或以其他方式定位於轉基因編碼序列內，或定位於轉基因編碼序列之3'。可用於本文所描述之因子IX基因療法構築體中之內含子的非限制性實例包括源自以下各項之內含子：小鼠細小病毒(MVM)內含子、 β -球蛋白內含子(β IVS-II)、因子IX (FIX)內含子A、猿猴病毒40

(SV40)小T內含子及 β -肌動蛋白內含子。

【0340】 在一個實施例中，本文所描述之因子IX基因療法構築體包括例如如圖1中所說明及由位於圖25中之CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸724-800處的MVM內含子(SEQ ID NO:53)所例示的MVM源性內含子。在一些實施例中，用於本文所描述之基因療法構築體中之內含子與MVM (SEQ ID NO:53)具有至少85%的序列一致性。在其他實施例中，用於本文所描述之基因療法構築體中之內含子與MVM (SEQ ID NO:53)具有至少90%、95%、96%、97%、98%、99%或100%的序列一致性。

聚腺苷酸化信號

【0341】 在一些實施例中，例如如圖1中所說明，本文所描述之因子IX基因療法構築體包括聚腺苷酸化信號。聚腺苷酸化信號在由因子IX轉基因產生之mRNA轉錄物之3'末端引導poly-A尾之合成。因此，聚腺苷酸化信號定位於因子IX編碼序列之3'。可用於本文所描述之因子IX基因療法構築體中之聚腺苷酸化信號的非限制性實例包括源自以下各項之聚腺苷酸化信號：猿猴病毒40 (SV40)後期基因、牛生長激素(BGH)聚腺苷酸化信號及極小兔 β -球蛋白(mRBG)基因。

【0342】 在一個實施例中，本文所描述之因子IX基因療法構築體包括例如如圖1中所說明及由位於圖25中之CS06-CRM8.3-ssV (SEQ ID NO:40)之核苷酸2208-2441處之BGHpA信號(SEQ ID NO:54)所例示的源自牛生長激素(BGH)聚腺苷酸化信號之聚腺苷酸化信號。在一些實施例中，用於本文所描述之基因療法構築體中之聚腺苷酸化信號與BGHpA信號(SEQ ID NO:54)具有至少85%的序列一致性。在其他實施例中，用於本文所描述之基因療法構築體中之聚腺苷酸化信號與BGHpA信號(SEQ ID NO:54)具有至少90%、95%、96%、97%、98%、99%或100%的序列一致性。

V. 方法

製備

【0343】 本文所描述之密碼子改變之因子IX聚核苷酸及病毒載體(例如核酸組合物)係根據用於核酸擴增及載體製備之習知方法來製備。已開發兩種主要平台用於大規模製備重組AAV載體。第一平台係基於在哺乳動物細胞中複製，而第二平台係基於在無脊椎動物細胞中複製。關於綜述，參見Kotin R.M., Hum. Mol. Genet., 20(R1):R2-6 (2011)，其內容出於所有目的以全文引用之方式明確併入本文中。

【0344】 因此，本發明提供用於製備腺相關病毒(AAV)粒子之方法。在一些實施例中，該等方法包括將如本文所描述與CS02、CS03、CS04、CS05或CS06序列中之一者具有高核苷酸序列一致性(例如至少95%、96%、97%、98%、99%、99.5%、99.9%或100%)之密碼子改變之因子IX聚核苷酸構築體引入宿主細胞中，其中聚核苷酸構築體勝任於在宿主細胞中複製。

【0345】 在一些實施例中，宿主細胞為哺乳動物宿主細胞，例如HEK、CHO或BHK細胞。在一特定實施例中，宿主細胞為HEK 293細胞。在一些實施例中，宿主細胞為無脊椎動物細胞，例如昆蟲細胞。在一特定實施例中，宿主細胞為SF9細胞。

調配物

【0346】 本文提供用於治療出血病症之組合物。此類組合物含有治療有效量之密碼子改變之因子IX聚核苷酸，例如如本文所描述包括密碼子改變之聚核苷酸編碼因子IX之AAV基因療法載體。治療有效量之密碼子改變之FIX聚核苷酸(例如包括密碼子改變之因子IX編碼序列之AAV基因療法載體)與適合用於全身、表面或局部投與之藥物載體或媒劑混合。本文所揭示之密碼子改變之因子IX聚核苷酸之最終調配物將在熟習此項技術者之能力範圍內。

劑量

【0347】 本發明之核酸組合物係投與有需要之患者。投與之治療性基因治療劑之量或劑量取決於諸如以下因素：特定密碼子改變之FIX聚核苷酸構築體、所用之遞送載體、疾病嚴重程度及個體之一般特徵。確切劑量將視治療目的而定，且將為熟習此項技術者使用已知技術可確定的(參見例如Lieberman, *Pharmaceutical Dosage Forms* (第1-3卷, 1992)；Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999)；Pickar, *Dosage Calculations* (1999)；及Remington: *The Science and Practice of Pharmacy*, 第20版, 2003, Gennaro編, Lippincott, Williams & Wilkins)。確定用於治療特定個體之特定劑量及給藥方案在熟練醫師之能力範圍內。

【0348】 在一些實施例中，以治療有效劑量向有需要之個體(例如患有輕度、中度或重度血友病B之個體)靜脈內投與具有密碼子改變之因子IX聚核苷酸之基因療法載體(例如AAV基因療法載體粒子)。在一些實施例中，治療有效劑量在每公斤個體體重約 2×10^{11} 與 2×10^{14} 個載體基因組之間。在一特定實施例中，治療有效劑量在每公斤個體體重約 2×10^{12} 與 2×10^{13} 個載體基因組之間。在一些實施例中，為個體投與每公斤個體體重約 2×10^{11} 、 3×10^{11} 、 4×10^{11} 、 5×10^{11} 、 6×10^{11} 、 7×10^{11} 、 8×10^{11} 、 9×10^{11} 、 1×10^{12} 、 2×10^{12} 、 3×10^{12} 、 4×10^{12} 、 5×10^{12} 、 6×10^{12} 、 7×10^{12} 、 8×10^{12} 、 9×10^{12} 、 1×10^{13} 、 2×10^{13} 、 3×10^{13} 、 4×10^{13} 、 5×10^{13} 、 6×10^{13} 、 7×10^{13} 、 8×10^{13} 、 9×10^{13} 、 1×10^{14} 或 2×10^{14} 個載體基因組。

【0349】 因此，本發明提供用於治療因子IX缺乏(例如血友病B)之方法。在一些實施例中，該等方法包括向有需要之患者投與如本文所描述與CS02、CS03、CS04、CS05或CS06序列至中之一者具有高核苷酸序列一致性(例如至少95%、96%、97%、98%、99%、99.5%、99.9%或100%)的密碼子改變之因子IX聚核苷酸構築體。在一些實施例中，密碼子改變之因子聚核苷酸與密碼子改變

之因子IX前原多肽編碼序列具有高序列一致性，例如與CS02-FL-NA (SEQ ID NO:5)、CS03-FL-NA (SEQ ID NO:6)、CS04-FL-NA (SEQ ID NO:7)、CS05-FL-NA (SEQ ID NO:8)或CS06-FL-NA (SEQ ID NO:9)中之一者具有高序列一致性。在一些實施例中，密碼子改變之因子聚核苷酸與密碼子改變之成熟因子IX單鏈多肽編碼序列具有高序列一致性，例如與CS02-MP-NA (SEQ ID NO:13)、CS03-MP-NA (SEQ ID NO:14)、CS04-MP-NA (SEQ ID NO:15)、CS05-MP-NA (SEQ ID NO:16)或CS06-MP-NA (SEQ ID NO:17)中之一者具有高序列一致性。

【0350】 在一些實施例中，治療包括向有需要之患者投與包括如本文所描述與CS02、CS03、CS04、CS05或CS06序列中之一者具有高核苷酸序列一致性(例如至少95%、96%、97%、98%、99%、99.5%、99.9%或100%)的密碼子改變之因子IX聚核苷酸構築體的基因療法載體。在一個實施例中，基因療法載體為哺乳動物基因療法載體。在一特定實施例中，哺乳動物基因療法載體為病毒載體，例如慢病毒、逆轉錄病毒、腺病毒或腺相關病毒載體。

【0351】 在一個實施例中，基因療法載體為具有編碼密碼子改變之因子IX編碼序列之病毒載體的腺相關病毒(AAV)粒子。通常，病毒載體包括各末端之反向末端重複(ITR)、一或多個表現調控元件、密碼子改變之因子IX編碼序列及poly-A信號序列。在一特定實施例中，基因療法載體包括肝特異性調控控制元件(例如CRM8元件之一或多個拷貝)。

製備

【0352】 本文所描述之密碼子改變之因子IX聚核苷酸及病毒載體(例如核酸組合物)係根據用於核酸擴增及載體製備之習知方法來製備。已開發兩種主要平台用於大規模製備重組AAV載體。第一平台係基於在哺乳動物細胞中複製，而第二平台係基於在無脊椎動物細胞中複製。關於綜述，參見Kotin R.M., Hum. Mol. Genet., 20(R1):R2-6 (2011)，其內容出於所有目的以全文引用之方式明確併

入本文中。

【0353】 因此，本發明提供用於製備腺相關病毒(AAV)粒子之方法。在一些實施例中，該等方法包括將如本文所描述與CS02、CS03、CS04、CS05或CS06序列中之一者具有高核苷酸序列一致性(例如至少95%、96%、97%、98%、99%、99.5%、99.9%或100%)之密碼子改變之因子IX聚核苷酸構築體引入宿主細胞中，其中聚核苷酸構築體勝任於在宿主細胞中複製。

【0354】 在一些實施例中，宿主細胞為哺乳動物宿主細胞，例如HEK、CHO或BHK細胞。在一特定實施例中，宿主細胞為HEK 293細胞。在一些實施例中，宿主細胞為無脊椎動物細胞，例如昆蟲細胞。在一特定實施例中，宿主細胞為SF9細胞。

治療

【0355】 在一些實施例中，根據已知投與方法向有需要之個體投與本文所描述之核酸組合物(例如密碼子改變之聚核苷酸)。用於投與基因療法載體之方法為此項技術中熟知的。此等方法包括但不限於靜脈內投與、肌肉內注射、間質內注射及肝內投與(例如肝內動脈或靜脈)。舉例而言，參見Chuah MK等人, Hum Gene Ther., 23(6):557-65 (2012)；Chuah MK等人, J Thromb Haemost., 10(8):1566-69 (2012)；Chuah MK等人, J Thromb Haemost. 11 Suppl 1:99-110 (2013)；VandenDriessche等人, Hum Gene Ther. 23(1):4-6 (2012)；High KA, Blood, 120(23):4482-87 (2012)；Matrai等人, Mol Ther., 18(3):477-90 (2010)；及Matrai等人, Curr Opin Hematol., 17(5):387-92 (2010)，其各自關於綜述以引用之方式併入本文中。

評估治療功效

【0356】 血友病B治療之治療功效可例如藉由量測所治療個體之血液的因子IX依賴性凝結潛能來評估。用於評估凝結潛能之量度包括但不限於活體外

活化部分凝血活酶時間分析(APPT)、因子IX生色活性分析、血液凝結時間及因子IX抗原含量(例如使用因子IX特異性ELISA)。應注意，治療劑量不需要在患者中產生野生型含量之FIX；確切地說，足以以有意義或可量測方式減輕症狀之表現出於本發明之目的被視為治療性的。

【0357】 根據美國血友病基金會(National Hemophilia Foundation)，當個體之血漿含有正常人類血漿因子IX活性之6%至49%時，將其歸類為具有輕度血友病B。患有輕度血友病B之個體典型地僅在嚴重損傷、外傷或手術之後經歷出血。在許多情況下，直至損傷、手術或拔牙產生延長之出血才診斷為輕度血友病。可能直至成年才發生第一事件。患有輕度血友病之女性常常經歷痛經、月經過多且可能會產後出血。

【0358】 根據美國血友病基金會，當個體之血漿含有正常人類血漿之因子IX活性的1%至5%時，將其歸類為具有中度血友病B。患有中度血友病B之個體傾向於在損傷之後具有出血事件。沒有明顯原因發生之出血稱為自發出血事件。

【0359】 根據美國血友病基金會，當個體之血漿含有不到正常人類血漿因子IX活性之1%時，將其歸類為具有重度血友病B。患有嚴重血友病B之個體在損傷後經歷出血且可能會具有頻繁自發出血事件，常常進入其關節及肌肉。

【0360】 在一些實施例中，正常人類血漿定義為每毫升含有1 IU之因子IX活性。因此，在一些實施例中，歸類為患有輕度血友病B之個體的血漿含有每毫升0.06與0.49 IU之間的因子IX活性。在一些實施例中，歸類為患有中度血友病B之個體的血漿含有每毫升0.01與0.05 IU之間的因子IX活性。在一些實施例中，歸類為患有重度血友病B之個體的血漿含有每毫升0.01與0.05 IU之間的因子IX活性。

【0361】 因此，在一些實施例中，當血友病B療法使個體之血液/血漿中

之平均因子IX活性水準升高時，其為治療有效的。在一些實施例中，治療有效之治療使個體之血液/血漿中之平均因子IX活性水準升高至少1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%或更多。在一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少5%。在另一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少10%。在另一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少15%。在另一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少20%。在另一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少25%。在另一特定實施例中，治療有效之血友病療法使個體血液/血漿中之平均因子IX活性增加至少30%。

【0362】 在一些實施例中，治療有效之治療使個體血液中之平均因子IX活性水準升高，使得個體歸類為具有較不嚴重形式之血友病B。舉例而言，在一個實施例中，最初歸類為患有重度血友病B之個體在經歷治療有效之治療之後重新歸類為患有中度血友病B或輕度血友病B。在另一實施例中，最初歸類為患有中度血友病B之個體在經歷治療有效之治療之後重新歸類為患有輕度血友病B。

VI. 實例

實例1-密碼子改變之因子IX表現序列使FIX表現水準增加

【0363】 為產生提供改良之異源因子IX活體內表現之基因療法構築體，構築一組編碼具有R384L胺基酸取代之全長因子IX前原蛋白(FIXp-FL-AA SEQ ID NO:4))的自我互補型基於AAV8之載體。通過若干步驟改變各構築體之因子IX編碼序列以改良在人類中之表現。根據經設計用於解釋較佳/不利序列基元且使密碼子使用偏向於較佳人類密碼子之算法對各因子IX編碼序列進行修飾。在

此第一步驟中使用如表2中所報導之若干算法。然後將由應用表2中所報導之算法產生之中間密碼子改變之序列進一步人工修飾以減少或消除CpG二核苷酸、調整最終GC含量、調整以允許較佳密碼子對、調整以避免不利密碼子對且調整最終密碼子用法。關於此等考慮之其他資訊，參見例如Fath S.等人, PLoS. One., 6, e17596 (2011)；Haas J.等人, Curr. Biol., 6, 315-324 (1996)；Tats A., BMC Genomics. 9:463 (2008)；Grote A.等人, Nucleic Acids Research, 33 (網站伺服器發行), W526-W531 (2005)；Mirsafian H.等人, Scientific World Journal., 639682 (2014)；及Pechmann S.等人, Nat Struct Mol Biol. 20(2):237-43 (2013)，其內容出於所有目的尤其關於其密碼子改變考慮之教示以全文引用之方式明確併入本文中。

【0364】 各個所產生之密碼子改變之編碼序列(例如分別顯示於圖5至9中之CS02、CS03、CS04、CS05及CS06)編碼一致FIX (R384L)蛋白(FIXp -FL-AA (SEQ ID NO:4))。CS02、CS03及CS04構築體不含CpG基元，而CS05及CS06分別含有11及3個CpG。

【0365】 為用作對照，亦產生在存在及不存在R384L Padua胺基酸取代之情況下併有野生型FIX編碼序列之載體構築體。WH01構築體編碼不具有R384L Padua突變之野生型FIX前原蛋白，且包括20個CpG二核苷酸。WH02構築體編碼具有R384L Padua突變之野生型FIX前原蛋白，且包括19個CpG二核苷酸。

【0366】 WH01及WH02構築體分別在其編碼序列中包括20及19個CpG。相比之下，CS02、CS03及CS04構築體不含CpG基元，而CS05及CS06構築體分別含有11及3個CpG。

【0367】 如圖1A中所示，將密碼子改變之因子IX編碼序列插入腺相關病毒(「AAV」)轉基因盒中，該腺相關病毒轉基因盒含有小鼠甲狀腺素轉運蛋白增強子/啟動子(SEQ ID NO:52)、小鼠細小病毒(「MVM」)內含子(SEQ ID

NO:53)、包括R384L「Padua」胺基酸取代之密碼子改變之FIX構築體(美國專利第6,531,298號; Simione等人, NEJM 361:1671-75 (2009); 參考野生型精胺酸在缺乏信號肽及原肽之人類單鏈FIX蛋白中之位置, R384L突變通常報導為R338L突變), 隨後為牛生長激素polyA元件(SEQ ID NO:54)。基因盒由AAV2反向末端重複(「ITR」)(SEQ ID NO:51及55)側接。左側ITR重複在末端多聯體拆分位點中包括突變, 從而產生載體之自我互補型(sc)表型。基本載體設計詳細描述於Wu等人, Mol. Ther. 16:280-89 (2008)中及PCT公開案第WO 2014/064277 A1號中, 該等參考文獻出於所有目的以全文引用之方式併入本文中。

【0368】 向B6/129P2- F9tm1Dws FIX敲出小鼠(描述於Lin等人, Blood, 90:3962-66 (1997)中, 該參考文獻之內容出於所有目的以全文引用之方式併入本文中)投與CS及WH因子IX AAV構築體。基於個別動物體重(4×10^{11} 個載體基因組/公斤(vg/kg)體重)經由外側尾部靜脈將AAV載體稀釋液注入動物(每組4-8個動物)中。在給藥之後根據已知程序使用玻璃毛細管藉由眶後穿刺按所規定之時間間隔收集血液樣品。然後將血液轉移至事先填充有檸檬酸鈉抗凝劑之管中且藉由標準程序獲得血漿並在 -20°C 下冷凍。

【0369】 測定各種因子IX構築體之表現且如表2中所報導, 使用在第14天在FIX敲出小鼠中之血漿FIX含量來判斷將載體尾部靜脈注射至小鼠中之後的構築體效能。到第14天, 敲出小鼠模型中之表現水準幾乎達至最大FIX表現。如表2中所示, 在投與 4×10^{11} 個載體基因組/公斤(vg/kg)體重之後第14天時WH02 FIX (R384L)對照構築體以1.03個單位之FIX進行表現。使用此表現水準作為基線來測定密碼子改變之因子IX構築體的富集倍數。如表2中所報導, 與由野生型聚核苷酸序列編碼之WH02對照構築體相比, CS密碼子改變之構築體提供約2倍與4倍之間的增加之表現。最值得注意地, CS06密碼子改變之構築體提供為WH02對照構築體之4.2倍大之因子IX活性及為WH01 (野生型因子IX)對照構築

體之21.6倍大之因子IX活性。

表2. 來自具有野生型密碼子序列(WH01-wtFIX；WH02-FIX (R384L)及密碼子改變之序列(CS02-CS06)之構築體的因子IX表現.

AAV 構築體	載體基因組之修飾	CpG之 數目	第14天表 現水準 [人 類 FIX%]	與 WH01 相 之 現 數 比 表 倍	與 WH02 相 比 之 表 現 數 之 倍
WH01	不 具 有R338L (Padua)突變之人類FIX野生型序列 (GeneBank NM000133.3)	20	0.20	1.0	0.19
WH02	具 有R338L (Padua)突 變之人類FIX野生型序列	19	1.03	5.2	1.0
CS02	具 有R338L突 變 之 人 類FIX序列；針對人類血清白蛋白密碼子進一步優化之Geneart基本算法。	0	2.12	10.6	2.1
CS03	具 有R338L突 變 之 人 類FIX序列；針對最常用之人類密碼子進一步優化的Geneart基本算法(Haas等人, 1996. Curr Biol. 6, 315-324)。	0	1.98	9.9	1.9
CS04	具 有R338L突 變 之 人 類FIX序列；如Uhlen等人, 2015 Science 347, 6220中所描述針對肝密碼子使用進一步優化之Geneart基本算法。	0	2.77	13.9	2.7
CS05	具 有R338L突 變 之 人 類FIX序列；經改良以減少CpG之JCAT算法；(Grote等人, 2005. Nucleic Acids Res 1,33)。	11	3.93	19.7	3.8
CS06	具 有R338L突 變 之 人 類FIX序列；針對最常用之人類密碼子進一步優化的Geneart基本算法(Haas等人, 1996. Curr Biol. 6, 315-324)。	3	4.32	21.6	4.2

實例2--肝特異性CRM8元件增強FIX在小鼠中之表現

【0370】 為進一步增加來自密碼子改變之構築體的因子IX表現及活性，將如Nair等人, Blood 123:3195-99 (2014)中所報導之肝特異性順式調控控制元件(CRM8 (SEQ ID NO:39))之一至三個拷貝併入基因盒中，形成在圖1B中作圖之構築體。藉由尾部靜脈途徑將具有零個(CS02-CRM8.0-V)、一個(CS02-CRM8.1-V)、兩個(CS02-CRM8.2-V)或三個(CS02-CRM8.3-V) CRM8對照元件之具有CS02密碼子改變之FIX編碼序列之AAV載體注入野生型小鼠中。然後使用人類FIX特異性ELISA分析隨時間推移量測小鼠血漿中之人類FIX抗原。

【0371】 如表3中所報導，在感染後21天時，與來自缺乏CRM8元件之對照構築體的表現相比，CRM8調控元件之使用使活體內因子IX表現增加至約2倍及4倍。舉例而言，含有單個CRM8元件之CS02-CRM8.1-V載體提供兩倍於CS02-CRM8.0-V對照載體之FIX表現。包括CRM8元件之多個拷貝進一步改良此表現。舉例而言，相對於對照載體，含有CRM8元件之2個拷貝的載體提供三倍之表現，且含有CRM8元件之3個拷貝的載體提供3.4倍之表現。

表3. 注射具有CRM8調控控制元件之0-3個拷貝的密碼子改變之AAV載體的野生型小鼠血漿中之因子IX表現水準.

#	AAV構築體	CRM8元件之數目	FIX (ng/ml) 第4天	FIX (ng/ml) 第11天	FIX (ng/ml) 第21天	增加倍數 第21天
1	CS02-CRM8.0-scV	0	65.8	133.4	239.2	1.0
2	CS02-CRM8.1-scV	1	120.7	250.7	442.8	1.9
3	CS02-CRM8.2-scV	2	152.9	417.3	713.8	3.0
4	CS02-CRM8.3-scV	3	130.9	432.6	800.9	3.4

實例3-肝特異性CRM8元件增強FIX在人類肝細胞中之表現

【0372】 HepG2進行之活體外生物效能分析對如實例2中所描述含有

CRM8肝特異性調控控制元件之0-3個拷貝的CS02因子IX基因療法構築體進一步測試。簡單來說，如實例2中所描述使HepG2細胞感染CS02-CRM8-V AAV載體中之一者，且在感染之後三天時藉由生色受質分析來量測FIX活性。與實例2中所報導之結果一致，如表4中所報導，含有CRM8調控控制元件之所有載體提供較高之FIX表現。明顯地，與在小鼠模型中相比，在人類HepG2細胞中使用多個CRM8元件之效果甚至更顯著。舉例而言，相對於對照載體，含有CRM8元件之2個拷貝的載體提供6.7倍之表現且含有CRM8元件之3個拷貝的載體提供12.8倍之表現。此證實在此等載體中CRM8調控控制元件對FIX表現具有之積極作用。

表4. 注射具有CRM8調控控制元件之0-3個拷貝的密碼子改變之AAV載體的人類肝HepG2細胞中之因子IX表現水準.

AAV構築體	CRM8元件之數目	FIX活性[生物效能單位]	增加倍數
CS02-CRM8.0-scV	0	0.35	1
CS02-CRM8.1-scV	1	0.82	2.3
CS02-CRM8.2-scV	2	2.36	6.7
CS02-CRM8.3-scV	3	4.48	12.8

實例4-單股FIX AAV8載體提供與類似之自我互補型載體類似的活體內表現

【0373】 在一些情況下，自我互補型(sc) AAV載體比類似單股(ss) AAV載體更高效地表現轉基因盒。此大概係歸因於在細胞核中自我互補型載體基因組去殼之後更快速地形成雙股。關於綜述，參見McCarty D., Mol. Ther., (16):1648-56 (2008)，其內容出於所有目的以全文引用之方式併入本文中。

【0374】 當前研究使用EGFP載體證實此影響。Bell等人, Hum. Gene Ther. Methods, (27):228-37 (2016)。然而，研究亦表明此影響為轉基因及劑量依賴性的。舉例而言，與對應單股載體相比，自我互補型AAV8載體中之人類鳥胺酸胺甲醯基轉移酶(hOTC)基因盒在低劑量下在小鼠肝中顯示更佳之表現。然而，在高劑量下未能展示此影響，表明至少在所研究之非分泌基因中該影響為轉基

因及劑量依賴性的。同上。

【0375】 為探索所揭示之密碼子改變之FIX基因構築體在單股及自我互補型設計之情形中的特性，分別如圖1D及1C中所作圖，在存在及不存在CRM8調控控制元件之情況下構築具有CS06密碼子改變之FIX (R338L)基因及兩個完整ITR之單股構築體。在HEK293細胞系統中製備單股(ss)載體，且將因子IX表現與實例1-3中所報導之自我互補型構築體的表現相比較。

【0376】 第一首先，在如上文所描述將自我互補型(sc)及單股(ss) CS06-CRM8.0-V構築體注射至B6/129P2-F9tm1Dws FIX敲出小鼠中之後，對其進行活體內測試。令人驚訝地，如表5中所報導，自我互補型(sc)及單股(ss) CS06載體構築體顯示非常類似之血漿FIX活性水準，表明對於本文所描述之密碼子改變之因子IX構築體，所報導之sc載體與ss載體相比之優點不成立。表現取決於許多參數，包括轉基因構築體、轉錄物之穩定性、構築體中所用之啟動子、時間及劑量。如表5中所示，在經選擇以校正出血且在FIX ko小鼠中獲得長期表現之條件下，對應sc及ss載體提供實質上類似之表現水準。

【0377】 亦在單股載體背景中研究肝特異性CRM8調控控制元件對FIX表現之影響。如表5中所報導，單股載體中包括CRM8元件改良B6/129P2-F9tm1Dws FIX敲出小鼠中之FIX表現。包括三個CRM8元件使來自單股CS06構築體之FIX表現進一步改良至超過缺乏CRM8元件之自我互補型CS06對照略高於2倍之水準。與野生型WH02構築體相比，當與三個CRM8調控控制元件配對時，單股CS06載體提供高至7倍大之表現。

表5. 注射各種單股(ss)及自我互補型(sc) AAV因子IX載體之FIX敲出小鼠中之因子IX表現水準.

AAV構築體	CRM 8元件之數目	表現水準 ¹ (第7天)	表現水準 (第14天)	表現水準 (第28天)	相較於CS06之增加倍數 (第7天/第14天/第28天)	相較於WH02之增加倍數 (第7天/第14天/第28天)
CS06-CRM8.0-ssV	0	1.38	2.73	2.99	0.7 / 0.9 / 0.9	1.5 / 2.7 / 2.7
CS06-CRM8.1-ssV	1	1.92	3.57	3.47	1.0 / 1.1 / 1.0	2.1 / 3.5 / 3.2
CS06-CRM8.3-ssV (SEQ ID NO:40)	3	4.43	6.65	7.78	2.3 / 2.1 / 2.2	4.9 / 6.5 / 7.1
CS06-CRM8.0-scV	0	1.89	3.17	3.50	1.0 / 1.0 / 1.0	2.1 / 3.1 / 3.2
WH02-CRM8.0-scV	0	0.90	1.03	1.10	0.5 / 0.3 / 0.3	1.0

¹ 以國際單位計之FIX活性(7-8隻小鼠之平均值)；d，天；

實例5-單股FIX AAV8載體在人類肝細胞中提供比類似自我互補型載體更佳之FIX表現

【0378】 然後在人類HepG2細胞中研究來自實例4中所描述之單股CS06載體的因子IX表現且與類似自我互補型載體構築體相比較。與實例4中所報導之活體內結果一致，在HepG2細胞中不具有CRM8元件之單股CS06載體以比比較自我互補型載體略微較低之水準提供FIX表現。然而，如表6中所報導，包括單個CRM8元件使來自單股CS06載體的FIX表現增加至為來自自我互補型CS06載體之表現的2.6倍大之水準。

【0379】 然而，最令人驚訝地，單股CS06載體中包括三個CRM8元件使FIX表現增加至為來自自我互補型CS06載體之表現的16.8倍大的水準。增加之FIX表現為來自WH02控制載體之FIX表現的超過100倍大。概括地說，在活體內與活體外生物效能分析兩者中含有三個CRM8元件之單股CS06載體提供最高表

現水準。

表6. 在人類肝細胞中來自單股(ss)及自我互補型載體之因子IX表現水準.

AAV構築體	CRM8 元件之 數目	FIX活性[生物 效能單位]	相較於CS06- CRM8.0-scV之 增加倍數	相較於WH02- CRM8.0-scV之 增加倍數
CS06-CRM8.0- ssV	0	0.24	0.6	3.9
CS06-CRM8.1- ssV	1	0.95	2.6	15.9
CS06-CRM8.3- ssV (SEQ ID NO: 40)	3	6.20	16.8	103.3
CS06-CRM8.0- scV	0	0.37	1.0	6.2
WH02- CRM8.0-scV	0	0.06	0.2	1.0

實例1-5之材料及方法

【0380】 動物實驗. 對於FIX敲出模型中之實驗，使用FIX ko小鼠品系 B6/129P2-F9tm1Dws (由Lin等人, 1997. Blood 90 :3962-6開發)。在野生型小鼠模型中，使用4-5週齡雄性C57BL6-J Bl6小鼠。兩種品系均獲自商業育種者。基於個別動物體重經由外側尾部靜脈將AAV載體稀釋液注入動物(每組4-8個動物)中。在給藥之後根據已知程序使用玻璃毛細管藉由眶後穿刺按所規定之時間間隔進行血液採樣。然後將血液轉移至事先填充有檸檬酸鈉抗凝劑之管中且藉由標準程序獲得血漿並在-20℃下冷凍。

【0381】 HepG2細胞中進行之包括FIX生色受質分析之活體外生物效能分析. 在人類肝細胞株HepG2 (ATCC HB-8065)中進行對基因療法載體製劑之活體外生物效能分析。在用羥基脲處理之後，使細胞感染AAV8FIX載體且孵育約96 h。在孵育時間期間，FIX表現且釋放至細胞上清液中且藉由生色終點量測 (Rossix AB, Sweden)測定FIX活性。各分析操作包括經純化之AAV-FIX載體材料之標準曲線，使用在700與7000之間的範圍內的MOI。將標準物在MOI 3270下

之FIX活性設定為生物效能單位(BPU) 1。

【0382】 在小鼠血漿中之人類FIX定量. 為定量基因敲出小鼠血漿中之人類FIX，使用標準FIX凝結分析方法進行FIX凝結分析。為定量野生型小鼠血漿中之人類FIX抗原，使用特定地偵測人類FIX之市售ELISA套組 (ASSERACHROM IX:AG (產品目錄號00943 Stago BNL)。

實例6-因併入CRM8元件產生之改良之轉錄功效

【0383】 為解釋含有CRM8之載體的改良之生物效能是否由增加之轉錄功效引起，將人類肝細胞株(HepG2)及小鼠肝細胞(FIX敲出小鼠)用含有0、1或3個CRM8元件之單股CS06載體轉導。測定FIX mRNA及DNA含量且呈現為經校正之FIX mRNA與DNA含量之間的比率。

【0384】 在活體外模型中，包括一個CRM8元件(CS06-CRM8.1-ssV)或三個CRM8元件(CS06-CRM8.3-ssV (SEQ ID NO:40))在轉導之人類肝細胞中分別產生為在用缺乏CRM8元件之載體(CS06-CRM8.0-ssV)轉導之細胞中的5倍及23倍高之人類FIX mRNA含量(表6)。類似地，在活體內模型中，在鼠肝中來自含有一個或三個CRM8元件之載體的FIX表現分別為來自缺乏CRM8元件之載體的FIX表現的2.0倍及2.8倍高(表6)。兩種模型均支持CRM8元件在改良FIX構築體之轉錄活性中提供有益影響。

表7. 人類肝細胞株或小鼠肝之在AAV8-FIX轉導之後的FIX mRNA含量.

AAV構築體	活體外肝細胞株(HepG2)		活體內小鼠肝組織	
	經校正之FIX 比率： mRNA/DNA	相較於CS06- CRM8.0-ssV 之增加倍數	經校正之FIX 比率： mRNA/DNA	相較於CS06- CRM8.0-ssV 之增加倍數
CS06- CRM8.0-ssV	0.025	1	0.38	1
CS06- CRM8.1-ssV	0.14	5.4	0.78	2.0
CS06- CRM8.3-ssV	0.58	23.4	1.09	2.8

實例6之方法

【0385】 包括RNA及DNA提取之定量實時聚合酶鏈反應，藉由標準程序自冷凍肝(參見動物實驗)或HepG2細胞(參見HepG2細胞中之活體外生物效能分析)萃取基因組DNA及總RNA。對於活體內實驗之分析學，每個處理組選擇在第14天接近於各別組之平均FIX活性(在平均值 $\pm$ SD內)的三個動物之子集。使用寡核苷酸(dT20)引物、SuperScript III逆轉錄酶(RT)及DNA酶處理之總RNA(DNeasy血液及組織套組, Qiagen, Germany；RNeasy微型套組, Qiagen)根據手冊合成cDNA。

【0386】 藉由基於螢光之定量實時聚合酶鏈反應(qPCR)擴增FIX外顯子6之96 bp序列來測定gDNA與cDNA樣品兩者中之FIX轉基因拷貝數。鼠 β -肌動蛋白充當內源性對照且使用市售TaqMan分析定量。使用特定裝置之軟體進行qPCR資料分析，基於標準曲線之線性回歸參數計算每一反應之FIX或 β -肌動蛋白拷貝。另外，將結果校正至1 μ g之RNA或DNA且計算mRNA:DNA比率。

【0387】 應瞭解，本文所描述之實例及實施例僅用於說明性目的，且將向熟習此項技術者建議根據其進行之各種修改或變化，且該等修改或變化應包括在本發明之精神及權限以及隨附申請專利範圍之範疇內。本文引用之所有公開案、專利及專利申請案出於所有目的以全文引用之方式併入本文中。

【符號說明】

無

【序列表】

- <110> 日商武田藥品工業股份有限公司(TAKEDA PHARMACEUTICAL COMPANY LIMITED)
- <120> 用於血友病 B 基因療法之編碼具有增強表現的重組 FIX 變異體之病毒載體
- <130> 008073-5117-TW
- <140> TW 107117388
- <141> 2018-05-22
- <150> US 62/509,616
- <151> 2017-05-22
- <160> 64
- <170> PatentIn 版本 3.5
- <210> 1
- <211> 1386
- <212> DNA
- <213> 人工序列
- <220>
- <223> FIX-FL-NA
- <400> 1
- atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta60

ggatatctac tcagtgctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt120

ctgaatcggc caaagaggta taattcaggt aaattggaag agtttgttca agggaacctt180

gagagagaat gtatggaaga aaagtgtagt ttggaagaag cacgagaagt ttttgaaaac240

actgaaagaa caactgaatt ttggaagcag tatgttgatg gagatcagtg tgagtccaat300

ccatgtttta atggcggcag ttgcaaggat gacattaatt cctatgaatg ttggtgtccc360

tttgatttg aaggaaagaa ctgtgaatta gatgtaacat gtaacattaa gaatggcaga420

tgcgagcagt tttgtaaaa tagtgctgat aacaaggtagg ttgctcctg tactgaggga	480
tatcgacttg cagaaaacca gaagtcctgt gaaccagcag tgccatttcc atgtggaaga	540
gtttctgttt cacaaacttc taagctcacc cgtgctgaga ctgtttttcc tgatgtggac	600
tatgtaaatt ctactgaagc tgaaacatt ttggataaca tcaactcaaag cacccaatca	660
tttaatgact tcactcgggt tgttgggtga gaagatgcc aaccaggtca attcccttgg	720
caggttgttt tgaatggtaa agttgatgca ttctgtggag gctctatcgt taatgaaaaa	780
tggattgtaa ctgctgcccc ctgtgttgaa actgggtgta aaattacagt tgcgcaggt	840
gaacataata ttgaggagac agaacataca gagcaaaagc gaaatgtgat tcgaattatt	900
cctcaccaca actacaatgc agctattaat aagtacaacc atgacattgc cttcttgga	960
ctggacgaac ccttagtgct aaacagctac gttacaccta ttgacattgc tgacaaggaa	1020
tacacgaaca tcttcctcaa atttgatct ggctatgtaa gtggctgggg aagagtcttc	1080
cacaaagga gatcagcttt agttcttcag taccttagag ttccacttgt tgaccgagcc	1140
acatgtcttc gatctacaaa gttcaccatc tataacaaca tgttctgtgc tggtttccat	1200
gaaggaggta gagattcatg tcaaggagat agtgggggac cccatgttac tgaagtggaa	1260
gggaccagtt tcttaactgg aattattagc tggggtgaag agtgtgcaat gaaaggcaaa	1320
tatggaatat ataccaaggt atcccgttat gtcaactgga ttaaggaaaa aacaaagctc	1380
acttaa	1386

- <210> 2
- <211> 461
- <212> PRT
- <213> 人工序列

- <220>
- <223> FIX-FL-AA

- <400> 2

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg Tyr Asn
35 40 45

Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg Glu Cys
50 55 60

Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe Glu Asn
65 70 75 80

Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly Asp Gln
85 90 95

Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp Asp Ile
100 105 110

Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys Asn Cys
115 120 125

Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe
130 135 140

Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly
145 150 155 160

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

Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala
180 185 190

Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala Glu
195 200 205

Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp Phe
210 215 220

Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp
225 230 235 240

Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile
245 250 255

Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly
260 265 270

Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu
275 280 285

His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn
290 295 300

Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu
305 310 315 320

Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile
325 330 335

Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr
340 345 350

Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val
355 360 365

Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg
370 375 380

Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His
385 390 395 400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val
405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly
420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser
435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
450 455 460

- <210> 3
- <211> 423
- <212> PRT
- <213> 人工序列

- <220>
- <223> FIX2-FL-AA

<400> 3

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu

195 200 205

Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn Glu Lys Trp Ile
210 215 220

Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys Ile Thr Val Val
225 230 235 240

Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr Glu Gln Lys Arg
245 250 255

Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn Ala Ala Ile Asn
260 265 270

Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp Glu Pro Leu Val
275 280 285

Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp Lys Glu Tyr Thr
290 295 300

Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser Gly Trp Gly Arg
305 310 315 320

Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln Tyr Leu Arg Val
325 330 335

Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr Lys Phe Thr Ile
340 345 350

Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly Gly Arg Asp Ser
355 360 365

Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu Val Glu Gly Thr

—

—

—

—

—

—

—

—

—

—

—

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

Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly
260 265 270

Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu
275 280 285

His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn
290 295 300

Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu
305 310 315 320

Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile
325 330 335

Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr
340 345 350

Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val
355 360 365

Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Leu
370 375 380

Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His
385 390 395 400

Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val
405 410 415

Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly
420 425 430

Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser
435 440 445

Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
450 455 460

<210> 5
<211> 1385
<212> DNA
<213> 人工序列

<220>
<223> CS02-FL-NA

<400> 5
atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtcagcaga gtgcacagtg ttcctggacc atgagaatgc caacaagatc 120

ctgaacaggc ccaagagata caactcaggc aagctggagg agtttgtgca gggcaacctg 180

gagagggagt gcatggagga gaagtgcagc tttagaggagg ccagagaggt gtttgagaac 240

acagagagga ccacagagtt ctggaagcag tatgtggatg gagaccagtg tgagagcaac 300

ccttgccatga atggaggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgccct 360

tttggttttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg 420

tgtgagcagt tctgcaagaa ctcagctgac acaaagtgg tgtgtagctg cacagagggc 480

tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga 540

gtgtcagtgt cccagaccag caagctgacc agagctgaga cagtgttccc tgatgtggac 600

tatgtgaata gcacagaggc tgagaccatc ctggacaaca tcaccagag caccagttcc 660

ttcaatgact tcaccagagt tgtgggagga gaggatgcca agcctggcca gttcccctgg 720

caggtgggtgc tgaatggcaa agtggatgcc ttctgtggag gcagcattgt gaatgagaag 780

tggattgtga cagctgcccc ctgtgtggag acaggagtga agatcacagt ggtggctgga	840
gaacacaata ttgaggagac agagcacaca gagcagaaga ggaatgtcat caggattatc	900
ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag	960
ctggatgagc ctctgggtgct gaatagctat gtgaccccc tctgcattgc tgacaaggag	1020
tacaccaaca tcttcctgaa gtttggctca ggctatgtgt caggctgggg cagagtgttc	1080
cacaagggca gatcagccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc	1140
acctgcctgt tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggtttccat	1200
gagggaggca gagacagctg ccaggagagac tcaggaggac cccatgtgac agaagtggag	1260
ggcaccagct tcttgacagg catcatcagc tggggagagg agtgtgccat gaagggaag	1320
tatggcatct acaccaaagt gagcagatat gtgaactgga tcaaggagaa aaccaagctg	1380
acctg	1385

- <210> 6
- <211> 1386
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS03-FL-NA

<400> 6	
atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg	60
ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc	120
ctgaacaggc ccaagagata caactctggc aagctggagg agtttgtgca gggcaacctg	180
gagagggagt gcatggagga gaagtgcagc tttgaggagg ccagggaagt gtttgagaac	240
actgagagga ccactgagtt ctggaagcag tatgtggatg gggaccagtg tgagagcaac	300
ccttgccctga atgggggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgcct	360

tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg	420
tgtgagcagt tctgcaagaa ctctgctgac aacaaagtgg tgtgtagctg cactgagggc	480
tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga	540
gtgtctgtgt cccagaccag caagctgacc agagctgaga ctgtgttccc tgatgtggac	600
tatgtgaata gcactgaggc tgagaccatc ctggacaaca tcaccagag caccagtc	660
ttcaatgact tcaccagagt ggtggggggg gaggatgcca agcctggcca gttcccctgg	720
caggtgggtgc tgaatggcaa agtggatgcc ttctgtgggg gcagcattgt gaatgagaag	780
tggatttgta ctgctgcca ctgtgtggag actgggggta agatcactgt ggtggctggg	840
gaacacaata ttgaggagac tgagcacact gagcagaaga ggaatgtcat caggattatc	900
ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag	960
ctggatgagc ctctgggtgt gaatagctat gtgaccccca tctgcattgc tgacaaggag	1020
tacaccaaca tcttctgaa gtttggctct ggctatgtgt ctggctgggg cagagtgttc	1080
cacaagggca ggtctgccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc	1140
acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggettccat	1200
gaggggggca gagacagctg ccagggggac tctgggggcc cccatgtgac tgaagtggag	1260
ggcaccagct tctgactgg catcatcagc tggggggagg agtgtgcat gaagggaag	1320
tatggcatct acaccaaagt gagcaggat gtgaactgga tcaaggagaa aaccaagctg	1380
acctga	1386

- <210> 7
- <211> 1386
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS04-FL-NA

<400> 7

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg	60
ggctacctgc tgtctgctga gtgcacagtg ttccctggacc atgagaatgc caacaagatc	120
ctgaacaggc ccaagagata caactctggc aagctggagg agttttgtgca gggcaacctg	180
gagagggagt gcatggagga gaagtgcagc tttagaggagg ccaggagagt gtttgagaac	240
acagagagga ccacagagtt ctggaagcag tatgtggatg gtgaccagtg tgagagcaac	300
ccttgccctga atggaggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgccct	360
tttggctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcagg	420
tgtgagcagt tctgcaagaa ctctgctgac aacaaggctgg tgtgtagctg cacagagggc	480
tacagactgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcaga	540
gtgtctgtgt cccagaccag caagctgacc agagctgaga cagtgttccc tgatgtggac	600
tatgtgaaca gcacagaggc tgagaccatc ctggacaaca tcaccagag caccagtc	660
ttcaatgact tcaccagagt ggtgggagga gaggatgcca agcctggcca gttcccctgg	720
caggtggtgc tgaatggcaa ggtggatgcc ttctgtggag gcagcattgt gaatgagaag	780
tggattgtga cagctgcccc ctgtgtggag acaggagtga agatcacagt ggtggctgga	840
gagcacaaca ttgaggagac agagcacaca gagcagaaga ggaatgtgat caggatcatc	900
cctcaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag	960
ctggatgagc ctctgggtgt gaacagctat gtgacccta tctgcattgc tgacaaggag	1020
tacaccaaca tcttctgaa gtttggctct ggctatgtgt ctggctgggg cagagtgttc	1080
cacaagggca ggtctgccct ggtgctgcag tacctgagag tgcccctggt ggacagagcc	1140
acctgcctgt tgagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggttccat	1200
gagggaggca gagacagctg ccagggtgac tctggaggac cccatgtgac agaggtggag	1260

ggcaccagct tcctgacagg catcatcagc tggggagagg agtgtgccat gaagggcaag 1320
tatggcatct acaccaaagt gagcagatat gtgaactgga tcaaggagaa gaccaagctg 1380
acctga 1386

<210> 8
<211> 1386
<212> DNA
<213> 人工序列

<220>
<223> CS05-FL-NA

<400> 8
atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60
ggctacctgc tgtctgtctga gtgcactgtg ttcttgacc atgagaatgc caacaagatc 120
ctgaaccgcc ccaagcgcta caactctggc aagctggagg agtttgtgca gggcaacctg 180
gagagggagt gcatggagga gaagtgcagc tttagaggagg ccaggagggt gtttgagaac 240
actgagcgca ccactgagtt ctggaagcag tatgtggatg gggaccagtg tgagagcaac 300
ccctgcctga atggggggag ctgcaaggat gacatcaaca gctatgagtg ctggtgcccc 360
tttgctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggccgc 420
tgtgagcagt tctgcaagaa ctctgtctgac aacaaggtagg tgtgtctctg cactgagggc 480
taccgcctgg ctgagaacca gaagagctgt gagcctgctg tgcccttccc ctgtggcagg 540
gtgtctgtga gccagaccag caagctgacc agggctgaga ctgtgttccc tgacgtggac 600
tatgtgaaca gcactgaggc tgagaccatc ctggacaaca tcaccagag caccagagc 660
ttcaatgact tcaccaggtt ggtgggagga gaggatgcca agcctggcca gtccccctgg 720
caggtggtgc tgaatggcaa ggtggatgcc ttctgtggag gcagcattgt gaatgagaag 780
tggattgtga ccgctgcccc ctgtgtggag actggagtga agatcactgt ggtggctggg 840

gagcacaaca ttgaggagac agagcacaca gagcagaagc gcaatgtgat caggatcatc	900
ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag	960
ctggatgagc ccctgggtgct gaacagctac gtgaccccca tctgcattgc agacaaggag	1020
tacaccaaca tcttcctgaa gtttggctct ggctatgtgt ctggctgggg caggggtgtc	1080
cacaagggca ggtctgccct ggtgctgcag tacctgaggg tgcccctggt ggacagggcc	1140
acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgcgc tggcttccat	1200
gagggaggaa gggacagctg ccaggagac tctggaggcc cccatgtgac agaggtggag	1260
ggcaccagct tcctgacagg catcatcagc tggggggagg agtgtgccat gaagggaag	1320
tatggcatct acaccaaagt gtcccgctat gtgaactgga tcaaggagaa gaccaagctg	1380
acctga	1386

- <210> 9
- <211> 1386
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS06-FL-NA

<400> 9	
atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg	60
ggctacctgc tgtctgctga gtgcactgtc ttcttgacc atgagaatgc caacaagatc	120
ctcaacaggc ccaagagata caactctggc aaactggagg agtttgtcca gggcaacctg	180
gagagggagt gcatggagga gaagtgtcc tttgaggagg ccaggaggt ctttgagaac	240
actgagcgca ccaactgagtt ctggaaacag tatgtggatg gggaccagtg tgagtccaac	300
ccctgcctga atgggggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgcccc	360
tttgctttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcaga	420

tgtgagcagt tctgcaagaa ctctgtctgac aacaaggtgg tgtgctcctg cactgagggc	480
taccgcctgg ctgagaacca gaagagctgt gagcctgttg tgccattccc atgtggcaga	540
gtctctgtga gccagaccag caagctcacc agggctgaga ctgtgttccc tgatgtggac	600
tatgtgaaca gcactgaggc tgaaaccatc ctggacaaca tcaccagag caccagagc	660
ttcaatgact tcaccagagt ggtgggagga gaggatgcca agcctggcca gttcccctgg	720
caagtgggtgc tcaatggcaa ggtggatgcc ttctgtgggg gctccattgt gaatgagaag	780
tggattgtca ctgctgcccc ctgtgtggag actgggggtca agatcactgt ggtggctggg	840
gagcacaaca ttgaggagac tgagcacact gagcagaagc gcaatgtgat caggatcatc	900
ccccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag	960
ctggatgagc ccctggctct caacagctat gtgaccccca tctgcattgc tgacaaggag	1020
tacaccaaca tcttctcaa gtttggctct ggctatgtct ctggttgggg cagagtgttc	1080
cacaaaggca ggtctgccct ggtgctccag tacctgagag tgcccctggt ggacagggcc	1140
acctgcctct tgagacacaa gttcaccatc tacaacaaca tgttctgtgc tggtttccat	1200
gagggaggaa gagacagctg ccagggggac tctggaggac cccatgtcac tgaggtggag	1260
ggcacctcct tcctcactgg catcatctcc tggggagagg agtgtgcat gaaaggcaaa	1320
tatggcatct acaccaaagt ctccagatat gtcaactgga tcaaggagaa gaccaagctg	1380
acctga	1386

- <210> 10
- <211> 415
- <212> PRT
- <213> 人工序列

- <220>
- <223> FIX-MP-AA

- <400> 10

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg
1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly
35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp
50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys
65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu
85 90 95

Gln Phe Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr
100 105 110

Glu Gly Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val
115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr
130 135 140

Arg Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu
145 150 155 160

Ala Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn
165 170 175

Asp Phe Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe
180 185 190

Pro Trp Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly
195 200 205

Ser Ile Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu
210 215 220

Thr Gly Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu
225 230 235 240

Thr Glu His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His
245 250 255

His Asn Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu
260 265 270

Leu Glu Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile
275 280 285

Cys Ile Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser
290 295 300

Gly Tyr Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala
305 310 315 320

Leu Val Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys
325 330 335

Leu Arg Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly
340 345 350

Phe His Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro
355 360 365

His Val Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser
370 375 380

Trp Gly Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys
385 390 395 400

Val Ser Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
405 410 415

<210> 11

<211> 377

<212> PRT

<213> 人工序列

<220>

<223> FIX2-MP-AA

<400> 11

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg
1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Val
35 40 45

Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu Gln Phe Cys Lys Asn Ser
50 55 60

Ala Asp Asn Lys Val Val Cys Ser Cys Thr Glu Gly Tyr Arg Leu Ala

I753168

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

245

250

255

Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser Gly Trp

260

265

270

Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln Tyr Leu

275

280

285

Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr Lys Phe

290

295

300

Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly Gly Arg

305

310

315

320

Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu Val Glu

325

330

335

Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu Cys Ala

340

345

350

Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr Val Asn

355

360

365

Trp Ile Lys Glu Lys Thr Lys Leu Thr

370

375

<210> 12
<211> 415
<212> PRT
<213> 人工序列

<220>
<223> FIXp-MP-AA

<400> 12

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg
1 5 10 15

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly
35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp
50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys
65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu
85 90 95

Gln Phe Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr
100 105 110

Glu Gly Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val
115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr
130 135 140

Arg Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu
145 150 155 160

Ala Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn
165 170 175

Asp Phe Thr Arg Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe
180 185 190

Pro Trp Gln Val Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly
195 200 205

Ser Ile Val Asn Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu
210 215 220

Thr Gly Val Lys Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu
225 230 235 240

Thr Glu His Thr Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His
245 250 255

His Asn Tyr Asn Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu
260 265 270

Leu Glu Leu Asp Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile
275 280 285

Cys Ile Ala Asp Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser
290 295 300

Gly Tyr Val Ser Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala
305 310 315 320

Leu Val Leu Gln Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys
325 330 335

Leu Leu Ser Thr Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly
340 345 350

Phe His Glu Gly Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro
355 360 365

His Val Thr Glu Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser
370 375 380

Trp Gly Glu Glu Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys
385 390 395 400

Val Ser Arg Tyr Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
405 410 415

- <210> 13
- <211> 1248
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS02-MP-NA

<400> 13	
tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag	60
gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag	120
ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgctt gaatggaggc	180
agctgcaagg atgacatcaa cagctatgag tgcgtgtgcc cttttggctt tgagggcaag	240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag	300
aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac	360
cagaagagct gtgagcctgc tgtgcccttc cctgtggca gagtgtcagt gtcccagacc	420
agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa tagcacagag	480
gctgagacca tcttggaaca catcaccagc agcaccagc cttcaatga cttcaccaga	540

gttgtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 600
aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc 660
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag 720
acagagcaca cagagcagaa gaggaatgtc atcaggatta tccccacca caactacaat 780
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 840
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg 900
aagtttggt caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc 960
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc 1020
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc 1080
tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca 1140
ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaa 1200
gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tgacctga 1248

<210> 14
<211> 1248
<212> DNA
<213> 人工序列

<220>
<223> CS03-MP-NA

<400> 14
tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60
gagaagtgca gctttgagga ggccaggga gtgtttgaga aactgagag gaccactgag 120
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgctt gaatgggggc 180
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300

aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac	360
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc	420
agcaagctga ccagagctga gactgtgttc cctgatgtgg actatgtgaa tagcactgag	480
gctgagacca tcctggacaa catcaccag agcaccagc cttcaatga cttcaccaga	540
gtggtggggg gggaggatgc caagcctggc cagttcccct ggcagggtgt gctgaatggc	600
aaagtggatg cttctgttg gggcagcatt gtgaatgaga agtggattgt gactgctgcc	660
cactgtgtgg agactggggt gaagatcact gtggtggctg gggaacacaa tattgaggag	720
actgagcaca ctgagcagaa gaggaatgtc atcaggatta tccccacca caactacaat	780
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctgtg	840
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttctg	900
aagtttggt ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc	960
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gctgagcacc	1020
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc	1080
tgccaggggg actctggggg ccccatgtg actgaagtgg agggcaccag cttcctgact	1140
ggcatcatca gctgggggga ggagtgtgcc atgaaggga agtatggcat ctacaccaaa	1200
gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacctga	1248

- <210> 15
- <211> 1250
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS04-MP-NA

<400> 15	
cstacaactc tggcaagctg gaggagttag tgcagggcaa cctggagagg gagtgcattg	60

aggagaagtg cagctttgag gaggccaggg aggtgtttga gaacacagag aggaccacag	120
agttctggaa gcagtatgtg gatggtgacc agtgtgagag caacccttgc ctgaatggag	180
gcagctgcaa ggatgacatc aacagctatg agtgctgggtg cccttttggc tttgagggca	240
agaactgtga gctggatgtg acctgcaaca tcaagaatgg caggtgtgag cagttctgca	300
agaactctgc tgacaacaag gtggtgtgta gctgcacaga gggctacaga ctggctgaga	360
accagaagag ctgtgagcct gctgtgccct tccccgtgg cagagtgtct gtgtcccaga	420
ccagcaagct gaccagagct gagacagtgt tccctgatgt ggactatgtg aacagcacag	480
aggctgagac catcctggac aacatcacc agagcaccca gtccttcaat gacttcacca	540
gagtgggtggg aggagaggat gccaagcctg gccagttccc ctggcaggtg gtgctgaatg	600
gcaaggtgga tgccttctgt ggaggcagca ttgtgaatga gaagtggatt gtgacagctg	660
cccactgtgt ggagacagga gtgaagatca cagtgggtggc tggagagcac aacattgagg	720
agacagagca cacagagcag aagaggaatg tgatcaggat catccctcac cacaactaca	780
atgtgtccat caacaagtac aaccatgaca ttgccctgct ggagctggat gaggcctctgg	840
tgctgaacag ctatgtgacc cctatctgca ttgtgacaa ggagtacacc aacatcttcc	900
tgaagtttgg ctctggctat gtgtctggct ggggcagagt gttccacaag ggcaggtctg	960
ccctgggtgct gcagtacctg agagtgtccc tgggtggacag agccacctgc ctgttgagca	1020
ccaagttcac catctacaac aacatgttct gtgtctggctt ccatgaggga ggcagagaca	1080
gctgccaggg tgactctgga ggaccccatg tgacagaggt ggagggcacc agcttcctga	1140
caggcatcat cagctgggga gaggagtgtg ccatgaaggg caagtatggc atctacacca	1200
aagtgagcag atatgtgaac tggatcaagg agaagaccaa gctgacctga	1250

<210> 16
<211> 1248
<212> DNA

<213> 人工序列

<220>

<223> CS05-MP-NA

<400> 16

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag	60
gagaagtgca gctttgagga ggccaggag gtgtttgaga aactgagcg caccactgag	120
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatggggg	180
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggcaag	240
aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag	300
aactctgctg acaacaaggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac	360
cagaagagct gtgagcctgc tgtgcccttc cctgtggca ggggtgtctgt gagccagacc	420
agcaagctga ccagggtga gactgtgttc cctgacgtgg actatgtgaa cagcactgag	480
gctgagacca tcctggacaa catcaccag agcaccaga gttcaatga cttcaccagg	540
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcagggtgt gctgaatggc	600
aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc	660
cactgtgtgg agactggagt gaagatcact gtggtggctg gggagcaca cattgaggag	720
acagagcaca cagagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat	780
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gccctggtg	840
ctgaacagct acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttctg	900
aagtttggct ctggctatgt gtctggctgg ggcagggtgt tccacaaggg caggtctgcc	960
ctggtgctgc agtacctgag ggtgccccctg gtggacaggg ccacctgcct gctgagcacc	1020
aagttcacca tctacaacaa catgttctgc gctggcttcc atgagggagg aagggacagc	1080
tgccagggag actctggagg ccccatgtg acagaggtgg agggcaccag cttcctgaca	1140

ggcatcatca gctgggggga ggagtgtgcc atgaaggga agtatggcat ctacacaaaa 1200
gtgtcccgt atgtgaactg gatcaaggag aagaccaagc tgacctga 1248

<210> 17
<211> 1248
<212> DNA
<213> 人工序列

<220>
<223> CS06-MP-NA

<400> 17
tacaactctg gcaaactgga ggagtttgc cagggaacc tggagaggga gtgcatggag 60
gagaagtgtc cttttgagga ggccaggag gtctttgaga aactgagcg caccactgag 120
ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatgggggc 180
agctgcaagg atgacatcaa cagctatgag tgtgtgtgcc cttttggctt tgagggaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag 300
aactctgtg acaacaaggt ggtgtgtcc tgcactgagg gctaccgctt ggctgagaac 360
cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc 420
agcaagctca ccagggtga gactgtgttc cctgatgtgg actatgtgaa cagcactgag 480
gctgaaacca tcctggacaa catcaccag agcaccaga gcttcaatga cttcaccaga 540
gtgtgtggag gagaggatgc caagcctggc cagttcccct ggcaagtgtt gctcaatggc 600
aaggtggatg ctttctgtgg gggctccatt gtgaatgaga agtggattgt cactgctgcc 660
cactgtgtgg agactggggt caagatcact gtgtgtgtg gggagcaca cattgaggag 720
actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat 780
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gccctgtgc 840
ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacacaa catcttctc 900

aagtttggct ctggctatgt ctctggctgg ggcagagtgt tccacaaagg caggtctgcc 960
ctggtgctcc agtacctgag agtgccctg gtggacaggg ccacctgcct cttgagcacc 1020
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg aagagacagc 1080
tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttcctcact 1140
ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacaccaa 1200
gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacctga 1248

<210> 18
<211> 138
<212> DNA
<213> 人工序列

<220>
<223> FIX-PPP-NA

<400> 18
atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta 60
ggatatctac tcagtctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt 120
ctgaatcggc caaagagg 138

<210> 19
<211> 138
<212> DNA
<213> 人工序列

<220>
<223> CS02-PPP-NA

<400> 19
atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60
ggctacctgc tgtcagcaga gtgcacagtg ttcttgacc atgagaatgc caacaagatc 120
ctgaacaggc ccaagaga 138

<210>	20	
<211>	138	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS03-PPP-NA	
<400>	20	
atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg		60
ggctacctgc tgtctgctga gtgcactgtg ttctctggacc atgagaatgc caacaagatc		120
ctgaacaggc ccaagaga		138
<210>	21	
<211>	138	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS04-PPP-NA	
<400>	21	
atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg		60
ggctacctgc tgtctgctga gtgcacagtg ttctctggacc atgagaatgc caacaagatc		120
ctgaacaggc ccaagaga		138
<210>	22	
<211>	138	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS05-PPP-NA	
<400>	22	
atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg		60

ggctacctgc tgtctgctga gtgcactgtg ttcctggacc atgagaatgc caacaagatc 120
ctgaaccgcc ccaagcgc 138

<210> 23
<211> 138
<212> DNA
<213> 人工序列

<220>
<223> CS06-PPP-NA

<400> 23
atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg 60
ggctacctgc tgtctgctga gtgcactgtc ttcctggacc atgagaatgc caacaagatc 120
ctcaacaggc ccaagaga 138

<210> 24
<211> 84
<212> DNA
<213> 人工序列

<220>
<223> FIX-SP-NA

<400> 24
atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta 60
ggatatctac tcagtgtga atgt 84

<210> 25
<211> 84
<212> DNA
<213> 人工序列

<220>
<223> CS02-SP-NA

<400> 25

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtcagcaga gtgc 84

<210> 26

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> CS03-SP-NA

<400> 26

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 27

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> CS04-SP-NA

<400> 27

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 28

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> CS05-SP-NA

<400> 28

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 29
<211> 84
<212> DNA
<213> 人工序列

<220>
<223> CS06--SP-NA

<400> 29
atgcagaggg tcaacatgat catggctgag tcccctggcc tcatcaccat ctgcctgctg 60

ggctacctgc tgtctgctga gtgc 84

<210> 30
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> FIX-PP-NA

<400> 30
acagtttttc ttgatcatga aaacgccaac aaaattctga atcggccaaa gagg 54

<210> 31
<211> 54
<212> DNA
<213> 人工序列

<220>
<223> CS02-PP-NA

<400> 31
acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga 54

<210> 32
<211> 54
<212> DNA

<213>	人工序列	
<220>		
<223>	CS03-PP-NA	
<400>	32	
actgtgttcc	tggaccatga gaatgccaac aagatcctga acaggcccaa gaga	54
<210>	33	
<211>	54	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS04-PP-NA	
<400>	33	
acagtgttcc	tggaccatga gaatgccaac aagatcctga acaggcccaa gaga	54
<210>	34	
<211>	54	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS05-PP-NA	
<400>	34	
actgtgttcc	tggaccatga gaatgccaac aagatcctga accgccccaa gcgc	54
<210>	35	
<211>	54	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS06-PP-NA	
<400>	35	
actgtcttcc	tggaccatga gaatgccaac aagatcctca acaggcccaa gaga	54

<210> 36
<211> 46
<212> PRT
<213> 人工序列

<220>
<223> FIX-PPP-AA

<400> 36

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys Thr Val Phe Leu
20 25 30

Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro Lys Arg
35 40 45

<210> 37
<211> 28
<212> PRT
<213> 人工序列

<220>
<223> FIX-SP-AA

<400> 37

Met Gln Arg Val Asn Met Ile Met Ala Glu Ser Pro Gly Leu Ile Thr
1 5 10 15

Ile Cys Leu Leu Gly Tyr Leu Leu Ser Ala Glu Cys
20 25

<210> 38
<211> 18
<212> PRT

<213> 人工序列

<220>

<223> FIX-PP-AA

<400> 38

Thr Val Phe Leu Asp His Glu Asn Ala Asn Lys Ile Leu Asn Arg Pro
1 5 10 15

Lys Arg

<210> 39

<211> 72

<212> DNA

<213> 人工序列

<220>

<223> CRM8

<400> 39

gggggaggct gctggtgaat attaaccaag gtcaccccag ttatcggagg agcaaacagg 60

ggctaagtcc ac 72

<210> 40

<211> 5276

<212> DNA

<213> 人工序列

<220>

<223> CS06-CRM8.3-ssV

<400> 40

ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggtcgcc 60

cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tcactagggg ttcttgagtt taaacttcgt cgacggggga ggctgctggt 180

gaatattaac caaggtcacc ccagttatcg gaggagcaaa caggggctaa gtccaccggg	240
ggaggctgct ggtgaatatt aaccaaggtc accccagtta tcggaggagc aaacaggggc	300
taagtccacc gggggaggct gctggtgaat attaaccaag gtcacccag ttatcggagg	360
agcaaacagg ggctaagtcc accgagggca ctgggaggat gttgagtaag atggaaaact	420
actgatgacc cttgcagaga cagagtatta ggacatgttt gaacaggggc cgggcgatca	480
gcaggtagct ctagaggatc cccgtctgtc tgcacatttc gtagagcgag tgttccgata	540
ctctaatactc cctaggcaag gtccataatt gtgtaggtta cttattctcc ttttgttgac	600
taagtcaata atcagaatca gcaggtttgg agtcagcttg gcagggatca gcagcctggg	660
ttggaaggag ggggtataaa agccccttca ccaggagaag ccgtcacaca gactaggcgc	720
gccctaaggt aagtiggcgc cgtttaaggg atggttggtt ggtggggtat taatgtttaa	780
ttaccttttt tacaggcctg aagatctgcc accatgcaga gggtaacat gatcatggct	840
gagtcacctg gcctcatcac catctgcctg ctgggctacc tgcgtctctc tgagtgcact	900
gtcttctctg accatgagaa tgccaacaag atcctcaaca ggccaagag atacaactct	960
ggcaaactgg aggagtttgt ccagggaac ctggagaggg agtgcattga ggagaagtgc	1020
tcctttgagg aggccaggga ggtctttgag aacctgagc gcaccactga gttctgaaa	1080
cagtatgttg atggggacca gtgtgagtc aacctctgc tgaatggggg cagctgcaag	1140
gatgacatca acagctatga gtgctggtgc ccctttggct ttgagggcaa gaactgtgag	1200
ctggatgtga cctgcaacat caagaatggc agatgtgagc agttctgcaa gaactctgct	1260
gacaacaagg tgggtgtgct ctgcaactgag ggctaccgcc tggctgagaa ccagaagagc	1320
tgtgagcctg ctgtgccatt cccatgtggc agagtctctg tgagccagac cagcaagctc	1380
accagggctg agactgtgtt ccctgatgtg gactatgtga acagcactga ggctgaaacc	1440
atcctggaca acatcaccca gagcaccag agcttcaatg acttcaccag agtggtggga	1500

ggagaggatg ccaagcctgg ccagttcccc tggcaagtgg tgctcaatgg caaggtggat	1560
gccttctgtg ggggctccat tgtgaatgag aagtggattg tctactgtgc ccactgtgtg	1620
gagactgggg tcaagatcac tgtggtggct ggggagcaca acattgagga gactgagcac	1680
actgagcaga agcgcaatgt gatcaggatc atccccacc acaactacaa tgctgccatc	1740
aacaagtaca accatgacat tgccctgtctg gagctggatg agccccctggt cctcaacagc	1800
tatgtgacct ccatctgcat tgctgacaag gagtacacca acatcttctt caagtttggc	1860
tctggctatg tctctggctg gggcagagtg ttccacaaag gcaggctctgc cctggtgctc	1920
cagtacctga gagtgcacct ggtggacagg gccacctgcc tcttgagcac caagttcacc	1980
atctacaaca acatgttctg tgctggcttc catgaggagag gaagagacag ctgccagggg	2040
gactctggag gaccccatgt cactgagggt gagggcacct ccttcctcac tggcatcatc	2100
tcctggggag aggagtgtgc catgaaaggc aaatatggca tctacaccaa agtctccaga	2160
tatgtcaact ggatcaagga gaagaccaag ctgacctgat gagcatgcct agagctcgct	2220
gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc tccccgtgc	2280
cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg	2340
catcgattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca	2400
agggggagga ttgggaagac aatagcaggc atgctgggga attaattaag ctgcggaagg	2460
aacccttagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg	2520
ggcgacaaaa ggtcgcccga cgccgggct ttgccgggc gccctcagtg agcgagcgag	2580
cgcgagaga gggagtggcc aagacgattt aaatgacaag cttagcgtaa tcatggtcat	2640
agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa	2700
gcataaagtg taaagcctgg ggtgcctaata gaggagcta actcacatta attgcgttgc	2760
gctcactgcc cgctttccag tcgggaaacc tgtcgtgcc gctgcattaa tgaatcggcc	2820

aacgcgcggg gagaggcggg ttgcgtatig ggcgtcttc cgttctctcg ctactgact	2880
cgctgcgctc ggtcgttcgg ctgcggcgag cggatcagc tcaactaaag gcgtaatac	2940
ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa	3000
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg	3060
acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa	3120
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tctgttccg accctgccgc	3180
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac	3240
gctgtaggta tctcagttcg gtgtaggtcg ttgcctcaa gctgggctgt gtgcacgaac	3300
cccccgttca gcccgaccgc tgcgccttat ccgtaacta tcgtcttgag tccaaccgg	3360
taagacacga cttatcgcca ctggcagcag ccaactggtaa caggattagc agagcgaggt	3420
atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa	3480
cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct	3540
cttgatccgg caaacaacc accgctggta gcggtggttt tttgtttgc aagcagcaga	3600
ttacgcgcag aaaaaagga tctcaagaag atcctttgat cttttctacg gggctctgacg	3660
ctcagtggaa cgaaaactca cgtaaggga ttttggtcat gagattatca aaaaggatct	3720
tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt	3780
aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc	3840
tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg	3900
gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggtccag	3960
atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc cctgcaactt	4020
tatccgcctc catccagtct attaatgtt gccgggaagc tagagtaagt agttcgccag	4080
ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt	4140

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccc	4200
tgttgtgcaa aaaagcgggt agctccttcg gtcctccgat cggtgtcaga agtaagttgg	4260
ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat	4320
ccgtaagatg cttttctgtg actgggtgagt actcaaccaa gtcattctga gaatagtgt	4380
tgcggcgacc gagttgtct tgcgcggcgt caatacggga taataccgcg ccacatagca	4440
gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct	4500
taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat	4560
cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaat gccgcaaaa	4620
agggataag ggcgacacgg aaatgttgaa tactcatact cticcttttt caatattatt	4680
gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa	4740
ataaacaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa	4800
ccattattat catgacatta acctataaaa ataggcgtat caccaggccc ttctgtctcg	4860
cgcgtttcgg tgatgacggt gaaaacctct gacacatgca gctcccggag acggtcacag	4920
cttgtctgta agcggatgcc gggagcagac aagcccgta gggcgcgtca gcgggtgtg	4980
gcgggtgtcg gggctggctt aactatgcgg catcagagca gattgtactg agagtgcacc	5040
atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc aggcgccatt	5100
cgccattcag gctgcgcaac tgttggaag ggcgatcggg gcgggcctct tcgtattac	5160
gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt	5220
ccagtcacg acgttgtaaa acgacggcca gtgaattcct cgagatttaa atgacg	5276

- <210> 41
- <211> 702
- <212> DNA
- <213> 人工序列

<220>

<223> CS02-HC-NA

<400> 41

gttgtgggag gagaggatgc caagcctggc cagttcccct ggcagggtgt gctgaatggc	60
aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc	120
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag	180
acagagcaca cagagcagaa gaggaatgtc atcaggatta tccccacca caactacaat	240
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg	300
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg	360
aagtttggct caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc	420
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc	480
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc	540
tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca	600
ggcatcatca gctggggaga ggagtgtgcc atgaaggga agtatggcat ctacacaaaa	660
gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tg	702

<210> 42

<211> 435

<212> DNA

<213> 人工序列

<220>

<223> CS02-LC-NA

<400> 42

tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag	60
gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag	120
ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgcct gaatggaggc	180

agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300
aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac 360
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtcagt gtcccagacc 420
agcaagctga ccaga 435

<210> 43
<211> 705
<212> DNA
<213> 人工序列

<220>
<223> CS03-HC-NA

<400> 43
gtggtggggg gggaggatgc caagcctggc cagttccctt ggcagggtgt gctgaatggc 60
aaagtggatg ctttctgtgg gggcagcatt gtgaatgaga agtggattgt gactgctgcc 120
cactgtgtgg agactggggt gaagatcact gtggtggctg gggaacacaa tattgaggag 180
actgagcaca ctgagcagaa gaggaatgtc atcaggatta tccccacca caactacaat 240
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 300
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttctg 360
aagtttggtc ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc 420
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gctgagcacc 480
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc 540
tgccaggggg actctggggg ccccatgtg actgaagtgg agggcaccag cttcctgact 600
ggcatcatca gctgggggga ggagtgtgcc atgaaggga agtatggcat ctacacaaa 660
gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacc 705

<210> 44
<211> 435
<212> DNA
<213> 人工序列

<220>
<223> CS03-LC-NA

<400> 44
tacaactctg gcaagctgga ggagtttgag cagggcaacc tggagaggga gtgcatggag 60
gagaagtgca gctttgagga ggccagggaa gtgtttgaga aactgagag gaccactgag 120
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgccct gaatgggggc 180
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggcct tgagggaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300
aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac 360
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc 420
agcaagctga ccaga 435

<210> 45
<211> 705
<212> DNA
<213> 人工序列

<220>
<223> CS04-HC-NA

<400> 45
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc 60
aaggtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc 120
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagagcaca cattgaggag 180
acagagcaca cagagcagaa gaggaatgtg atcaggatca tccctcacca caactacaat 240

gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg 300
ctgaacagct atgtgacccc tatctgcatt gctgacaagg agtacaccaa catcttcctg 360
aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc 420
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc 480
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc 540
tgccagggtg actctggagg accccatgtg acagaggtgg agggcaccag cttcctgaca 600
ggcatcatca gctggggaga ggagtgtgcc atgaaggga agtatggcat ctacacaaa 660
gtgagcagat atgtgaactg gatcaaggag aagaccaagc tgacc 705

<210> 46
<211> 435
<212> DNA
<213> 人工序列

<220>
<223> CS04-LC-NA

<400> 46~
tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag 60
gagaagtgca gctttgagga ggcagggag gtgtttgaga acacagagag gaccacagag 120
ttctggaagc agtatgtgga tggtgaccag tgtgagagca acccttgcct gaatggaggc 180
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cttttggctt tgagggcaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag 300
aactctgctg acaacaaggt ggtgtgtagc tgcacagagg gctacagact ggctgagaac 360
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc 420
agcaagctga ccaga 435

<210> 47

<211>	705	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS05-HC-NA	
<400>	47	
gtggtgggag	gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc	60
aaggtggatg	ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc	120
cactgtgtgg	agactggagt gaagatcact gtggtggctg gggagcacia cattgaggag	180
acagagcaca	cagagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat	240
gctgccatca	acaagtacia ccatgacatt gccctgctgg agctggatga gcccctggtg	300
ctgaacagct	acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttctg	360
aagtttggt	ctggctatgt gtctggctgg ggcaggggtg tccacaaggg caggtctgcc	420
ctggtgtgc	agtacctgag ggtgccccctg gtggacaggg ccacctgcct gctgagcacc	480
aagttcacca	tctacaacia catgttctgc gctggcttcc atgaggagg aagggacagc	540
tgccagggag	actctggagg ccccatgtg acagaggtgg agggcaccag cttcctgaca	600
ggcatcatca	gctgggggga ggagtgtgcc atgaaggga agtatggcat ctacacaaaa	660
gtgtcccgt	atgtgaactg gatcaaggag aagaccaagc tgacc	705

<210>	48	
<211>	435	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CS05-LC-NA	
<400>	48	
tacaactctg	gcaagctgga ggagtttgtg cagggaacc tggagaggga gtgcatggag	60

gagaagtgca gctttgagga ggccaggag gtgtttgaga acactgagcg caccactgag	120
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatggggg	180
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cctttggctt tgagggaag	240
aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag	300
aactctgctg acaacaaggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac	360
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca ggggtgtctgt gagccagacc	420
agcaagctga ccagg	435

- <210> 49
- <211> 705
- <212> DNA
- <213> 人工序列

- <220>
- <223> CS06-HC-NA

<400> 49	
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaagtggg gctcaatggc	60
aaggtggatg ccttctgtgg gggtccatt gtgaatgaga agtggattgt cactgctgcc	120
cactgtgtgg agactggggt caagatcact gtggtggctg gggagcaca cattgaggag	180
actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat	240
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcccttggtc	300
ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttctc	360
aagtttggt ctggctatgt ctctggctgg ggcagagtgt tccacaaagg caggtctgcc	420
ctggtgctcc agtacctgag agtgcccctg gtggacaggg ccacctgcct cttgagcacc	480
aagttcacca tctacaacaa catgttctgt gctggtctcc atgaggagg aagagacagc	540
tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttcctcact	600

ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacacaaaa 660
gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacc 705

<210> 50
<211> 435
<212> DNA
<213> 人工序列

<220>
<223> CS06-LC-NA

<400> 50
tacaactctg gcaaactgga ggagtttgtc cagggcaacc tggagaggga gtgcatggag 60
gagaagtgct cctttgagga ggccaggag gtctttgaga aactgagcg caccactgag 120
ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatgggggc 180
agctgcaagg atgacatcaa cagctatgag tgctggtgcc cctttggctt tgagggaag 240
aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag 300
aactctgctg acaacaaggt ggtgtgctcc tgcactgagg gctaccgcct ggctgagaac 360
cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc 420
agcaagctca ccagg 435

<210> 51
<211> 145
<212> DNA
<213> 人工序列

<220>
<223> AAV2_5'-ITR

<400> 51
ttggccactc cctctctgcg cgctcgcctg ctactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg 120

gccaaactcca tcactagggg ttcct 145

<210> 52
<211> 330
<212> DNA
<213> 人工序列

<220>
<223> 截短之 TTR 增強子/啟動子

<400> 52
cgagggcact gggaggatgt tgagtaagat ggaaaactac tgatgaccct tgcagagaca 60

gagtattagg acatgtttga acaggggccg ggcgatcagc aggtagctct agaggatccc 120

cgtctgtctg cacatttcgt agagcgagtg ttccgatact ctaatctccc taggcaaggt 180

tcataattgt gtaggttact tattctcctt ttgttgacta agtcaataat cagaatcagc 240

aggtttggag tcagcttggc agggatcagc agcctgggtt ggaaggaggg ggtataaaag 300

ccccttcacc aggagaagcc gtcacacaga 330

<210> 53
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> MVM 內含子

<400> 53
ctaaggtaag ttggcgccgt ttaagggatg gttggttggt ggggtattaa tgtttaatta 60

ccttttttac aggcctg 77

<210> 54
<211> 234
<212> DNA
<213> 人工序列

<220>
<223> BGH 聚腺苷酸化信號

<400> 54
cctagagctc gctgatcagc ctcgactgtg ccttctagtt gccagccatc tgttgtttgc 60

ccctcccccg tgccttcctt gaccctggaa ggtgccactc ccactgtcct ttctaataa 120

aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg gggtggggtg 180

gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg ggaa 234

<210> 55
<211> 145
<212> DNA
<213> 人工序列

<220>
<223> AAV2 3'-ITR

<400> 55
aggaaccctt agtgatggag ttggccactc cctctctgcg cgtcgcctcg ctactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaa 145

<210> 56
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> FIX-AP-AA

<400> 56

Ala Glu Thr Val Phe Pro Asp Val Asp Tyr Val Asn Ser Thr Glu Ala
1 5 10 15

Glu Thr Ile Leu Asp Asn Ile Thr Gln Ser Thr Gln Ser Phe Asn Asp

	20	25	30	
Phe Thr Arg				
35				
<210>	57			
<211>	105			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	CS02-AP-NA			
<400>	57			
gctgagacag tgttcctga tgtggactat gtgaatagca cagaggctga gaccatcctg				60
gacaacatca cccagagcac ccagtccttc aatgacttca ccaga				105
<210>	58			
<211>	105			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	CS03-AP-NA			
<400>	58			
gctgagactg tgttcctga tgtggactat gtgaatagca ctgaggctga gaccatcctg				60
gacaacatca cccagagcac ccagtccttc aatgacttca ccaga				105
<210>	59			
<211>	105			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	CS04-AP-NA			
<400>	59			

gctgagacag tgttcctga tgtggactat gtgaacagca cagaggctga gaccatcctg 60
gacaacatca cccagagcac ccagtccttc aatgacttca ccaga 105

<210> 60
<211> 105
<212> DNA
<213> 人工序列

<220>
<223> CS05-AP-NA

<400> 60
gctgagactg tgttcctga cgtggactat gtgaacagca ctgaggctga gaccatcctg 60
gacaacatca cccagagcac ccagagcttc aatgacttca ccagg 105

<210> 61
<211> 105
<212> DNA
<213> 人工序列

<220>
<223> CS06-AP-NA

<400> 61
gctgagactg tgttcctga tgtggactat gtgaacagca ctgaggctga aaccatcctg 60
gacaacatca cccagagcac ccagagcttc aatgacttca ccaga 105

<210> 62
<211> 145
<212> PRT
<213> 人工序列

<220>
<223> FIX-LC-AA

<400> 62

Tyr Asn Ser Gly Lys Leu Glu Glu Phe Val Gln Gly Asn Leu Glu Arg

15

10

5

1

Glu Cys Met Glu Glu Lys Cys Ser Phe Glu Glu Ala Arg Glu Val Phe
20 25 30

Glu Asn Thr Glu Arg Thr Thr Glu Phe Trp Lys Gln Tyr Val Asp Gly
35 40 45

Asp Gln Cys Glu Ser Asn Pro Cys Leu Asn Gly Gly Ser Cys Lys Asp
50 55 60

Asp Ile Asn Ser Tyr Glu Cys Trp Cys Pro Phe Gly Phe Glu Gly Lys
65 70 75 80

Asn Cys Glu Leu Asp Val Thr Cys Asn Ile Lys Asn Gly Arg Cys Glu
85 90 95

Gln Phe Cys Lys Asn Ser Ala Asp Asn Lys Val Val Cys Ser Cys Thr
100 105 110

Glu Gly Tyr Arg Leu Ala Glu Asn Gln Lys Ser Cys Glu Pro Ala Val
115 120 125

Pro Phe Pro Cys Gly Arg Val Ser Val Ser Gln Thr Ser Lys Leu Thr
130 135 140

Arg
145

- <210> 63
- <211> 235
- <212> PRT
- <213> 人工序列

<220>
<223> FIX-HC-AA

<400> 63

Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val
1 5 10 15

Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn
 20 25 30

Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys
 35 40 45

Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr
50 55 60

Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn
65 70 75 80

Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp
 85 90 95

Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp
100 105 110

Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser
115 120 125

Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln
130 135 140

Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Arg Ser Thr
145 150 155 160

Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly
165 170 175

Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu
180 185 190

Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu
195 200 205

Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr
210 215 220

Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr
225 230 235

<210> 64
<211> 235
<212> PRT
<213> 人工序列

<220>
<223> FIXp-HC-AA

<400> 64

Val Val Gly Gly Glu Asp Ala Lys Pro Gly Gln Phe Pro Trp Gln Val
1 5 10 15

Val Leu Asn Gly Lys Val Asp Ala Phe Cys Gly Gly Ser Ile Val Asn
20 25 30

Glu Lys Trp Ile Val Thr Ala Ala His Cys Val Glu Thr Gly Val Lys
35 40 45

Ile Thr Val Val Ala Gly Glu His Asn Ile Glu Glu Thr Glu His Thr

505560

Glu Gln Lys Arg Asn Val Ile Arg Ile Ile Pro His His Asn Tyr Asn
65707580

Ala Ala Ile Asn Lys Tyr Asn His Asp Ile Ala Leu Leu Glu Leu Asp
859095

Glu Pro Leu Val Leu Asn Ser Tyr Val Thr Pro Ile Cys Ile Ala Asp
100105110

Lys Glu Tyr Thr Asn Ile Phe Leu Lys Phe Gly Ser Gly Tyr Val Ser
115120125

Gly Trp Gly Arg Val Phe His Lys Gly Arg Ser Ala Leu Val Leu Gln
130135140

Tyr Leu Arg Val Pro Leu Val Asp Arg Ala Thr Cys Leu Leu Ser Thr
145150155160

Lys Phe Thr Ile Tyr Asn Asn Met Phe Cys Ala Gly Phe His Glu Gly
165170175

Gly Arg Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Glu
180185190

Val Glu Gly Thr Ser Phe Leu Thr Gly Ile Ile Ser Trp Gly Glu Glu
195200205

Cys Ala Met Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Ser Arg Tyr
210215220

Val Asn Trp Ile Lys Glu Lys Thr Lys Leu Thr

225

230

235

【發明申請專利範圍】

【第1項】 一種核酸組合物，其包含編碼因子IX蛋白之因子IX聚核苷酸，該因子IX聚核苷酸包含與CS06-MP-NA (SEQ ID NO:17)之核酸序列至少97%一致的核酸序列。

【第2項】 如申請專利範圍第1項之核酸組合物，其中該因子IX聚核苷酸包含與CS06-MP-NA (SEQ ID NO:17)之核酸序列至少99%一致的核酸序列。

【第3項】 如申請專利範圍第1項之核酸組合物，其中該因子IX聚核苷酸具有不超過10個CpG二核苷酸。

【第4項】 如申請專利範圍第2項之核酸組合物，其中該因子IX聚核苷酸具有不超過10個CpG二核苷酸。

【第5項】 如申請專利範圍第1項之核酸組合物，其中該因子IX聚核苷酸具有不超過3個CpG二核苷酸。

【第6項】 如申請專利範圍第2項之核酸組合物，其中該因子IX聚核苷酸具有不超過3個CpG二核苷酸。

【第7項】 如申請專利範圍第1項之核酸組合物，其中該因子IX聚核苷酸包含CS06-MP-NA (SEQ ID NO:17)之核酸序列。

【第8項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中該因子IX聚核苷酸在對應於全長野生型因子IX多肽(SEQ ID NO: 2)之殘基384的胺基酸位置編碼白胺酸。

【第9項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中由該因子IX聚核苷酸編碼之該因子IX蛋白與FIXp-MP-AA (SEQ ID NO:12)相比具有1至10個胺基酸取代。

【第10項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中由該因子IX聚核苷酸編碼之該因子IX蛋白具有FIXp-MP-AA (SEQ ID NO:12)

之胺基酸序列。

【第11項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中該因子IX聚核苷酸進一步包含編碼前原前導肽之前原前導聚核苷酸，該前原前導肽包含FIX-PPP-AA (SEQ ID NO:36)之胺基酸序列。

【第12項】 如申請專利範圍第11項之核酸組合物，其中該前原前導聚核苷酸具有CS06-PPP-NA (SEQ ID NO:23)之核酸序列。

【第13項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中該因子IX聚核苷酸具有與CS06-FL-NA (SEQ ID NO:9)之核酸序列至少99%一致的核酸序列。

【第14項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其中該因子IX聚核苷酸具有CS06-FL-NA (SEQ ID NO:9)之核酸序列。

【第15項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其進一步包含可操作地連接至該因子IX聚核苷酸之肝特異性啟動子元件。

【第16項】 如申請專利範圍第15項之核酸組合物，其中該肝特異性啟動子元件包含啟動子聚核苷酸之一個拷貝，該啟動子聚核苷酸包含與CRM8 (SEQ ID NO:39)至少95%一致之核酸序列。

【第17項】 如申請專利範圍第15項之核酸組合物，其中該肝特異性啟動子元件包含啟動子聚核苷酸之三個拷貝，該啟動子聚核苷酸包含與CRM8 (SEQ ID NO:39)至少95%一致之核酸序列。

【第18項】 如申請專利範圍第14項之核酸組合物，其中該啟動子聚核苷酸包含CRM8 (SEQ ID NO:39)之核酸序列。

【第19項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其進一步包含可操作地連接至該因子IX聚核苷酸之內含子。

【第20項】 如申請專利範圍第19項之核酸組合物，其中該内含子包含

含有與MVMI (SEQ ID NO:53)至少95%一致之核酸序列的MVM內含子聚核苷酸。

【第21項】 如申請專利範圍第20項之核酸組合物，其中該MVM內含子聚核苷酸包含MVMI (SEQ ID NO:53)之核酸序列。

【第22項】 如申請專利範圍第18項之核酸組合物，其中該內含子定位於啟動子元件與編碼因子IX多肽之核苷酸序列的轉譯起始位點之間。

【第23項】 如申請專利範圍第1項之核酸組合物，其包含CS06-CRM8.3-ssV (SEQ ID NO:40)之核酸序列。

【第24項】 如申請專利範圍第1項至第6項中任一項之核酸組合物，其包含哺乳動物基因療法載體。

【第25項】 如申請專利範圍第24項之核酸組合物，其中該哺乳動物基因療法載體為腺相關病毒(AAV)載體。

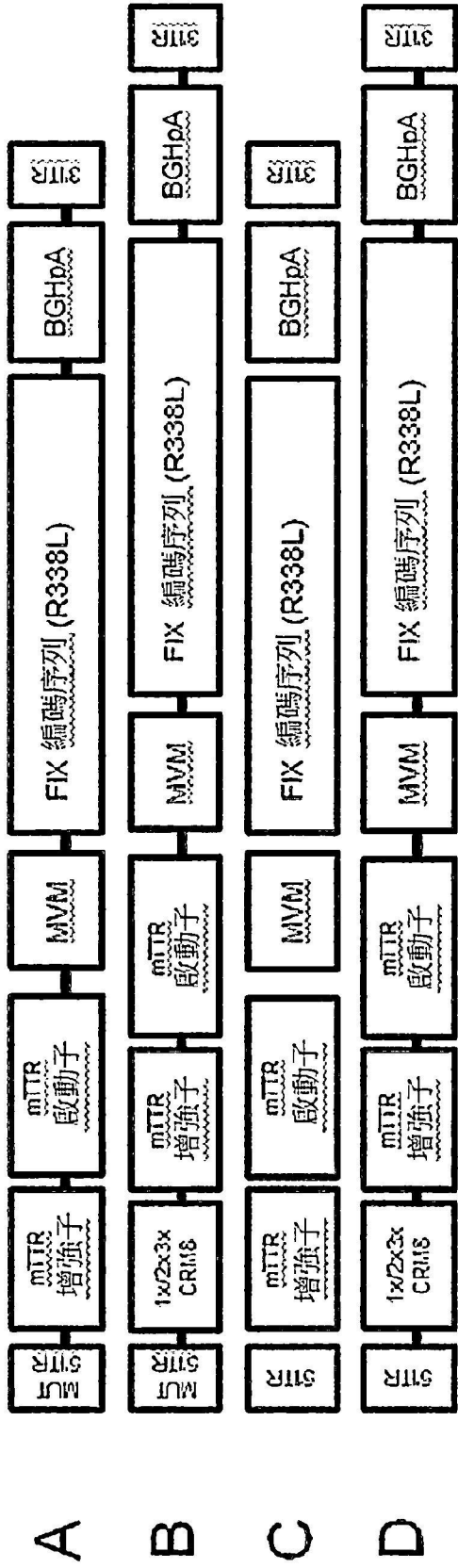
【第26項】 如申請專利範圍第25項之核酸組合物，其中該腺相關病毒載體為血清型8腺相關病毒(AAV-8)載體。

【第27項】 如申請專利範圍第24項之核酸組合物，其中該哺乳動物基因療法載體包含編碼該因子IX蛋白之單鏈聚核苷酸。

【第28項】 一種如申請專利範圍第1項至第27項中任一項之核酸組合物的用途，其用於製造用於治療血友病B之藥劑。

【第29項】 一種製備腺相關病毒(AAV)粒子之方法，其包括將如申請專利範圍第1項至第27項中任一項之核酸組合物引入哺乳動物宿主細胞中，其中該核酸組合物勝任於在該哺乳動物宿主細胞中複製。

【發明圖式】



【圖 1】

FIX-FL-NA

ATGCAGCGCGTGAACATGATCATGGCAGAATCACCAGGCCTCATCACCATCTGCCTTTTAGGATATC
TACTCAGTGCTGAATGTACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAA
GAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAA
AAGTGTAGTTTTTGAAGAAGCACGAGAAGTTTTTGA AAACACTGAAAGAACAAC TGAATTTTGAAGC
AGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACAT
TAATTCCTATGAATGTTGGTGTCCCTTTGGATT TGAAGGAAAGAACTGTGAATTAGATGTAACATGT
AACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGCTCCT
GTACTGAGGGATATCGACTTGCAGAAAACCAGAAGT CCTGTGAACCAGCAGTGCCATTTCCATGTGG
AAGAGTTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCTGATGTGGACTAT
GTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACT
TCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGG
TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTA ACTGCTGCCCCACTGT
GTTGAAACTGGTGTTAAAATTACAGTTGTGCGCAGGTGAACATAATATTGAGGAGACAGAACATACAG
AGCAAAAGCGAAAATGTGATTGCAATTATTCCTCACCACAAC TACAATGCAGCTATTAATAAGTACAA
CCATGACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGC
ATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAA
GAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGC
CACATGTCTTCGATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGA
GGTAGAGATT CATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCT
TAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGT
ATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAA (SEQ ID NO:1)

【圖2】

FIX-FL-AA

MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKRYNSG
 KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
 PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
 NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
 YVNSTEAETI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
 FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIRII
 PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIIDKE YTNIFLKFGS
 GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLRSTKFTI YNNMFCAGFH
 EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRV
 VNWIKEKTKL T (SEQ ID NO:2)

【圖3A】

FIX2-FL-AA

mqrvmimae spgliticll gyllsaectv fldhenanki lnrpkrnsg kleefvqgnl
 erecmeeckcs feearevfen terttefwkq yvdtcnikn grceqfckns adnkvvcsct
 egyrlaenqk scepavpfpc grvsvsqtsk ltraetvfpd vdyvnsteae tildnitqst
 qsfndftrvv ggedakpggf pwqvvlngkv dafcggsivn ekwivtaahc vetgvkitvv
 agehnieete hteqkrnvir iiphhnynaa inkynhdial leldeplvln syvtpiciad
 keytniflkf gsgyvsgwgr vfhkgrsalv lqylrvplvd ratclrstkf tiynnmfcag
 fheggrdscq gdsggphvte vegtsfltgi iswgeecamk gkygiytkvs ryvnwikekt
 klt (SEQ ID NO:3)

【圖3B】

FIXp-FL-AA

MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKRYNSG
KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVD
YVNSTEAE TI LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRN VIRII
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLLSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSR Y
VNWIKEKTKL T (SEQ ID NO:4)

【圖4】

CS02-FL-NA

atgcagaggggtgaacatgatcatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtcagcagagtgcacagtgttcctggaccatgagaatgccacaagatcctgaacaggccc
aagagatacaactcaggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagagaggtgtttgagaacacagagaggaccacagagtctctgg
aagcagtatgtggatggagaccagtgtgagagcaacccttgccctgaatggaggcagctgcaaggat
gacatcaacagctatgagtgtggtgcccttttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactcagctgacaacaaagtgggtg
tgtagctgcacagaggggctacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgtcagtgtcccagaccagcaagctgaccagagctgagacagtgttccctgat
gtggactatgtgaatagcacagaggctgagaccatcctggacaacatcaccagagcaccagtc
ttcaatgacttcaccagagtgtgtgggaggagaggtgccaaagcctggccagttcccctggcaggtg
gtgctgaatggcaaagtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgaca
gctgcccactgtgtggagacaggagtgaagatcacagtgggtggctggagaacacaatattgaggag
acagagcacacagagcagaagaggaatgtcatcaggattatccccaccacaactacaatgctgcc
atcaacaagtacaacatgacattgccctgctggagctggatgagcctctgggtgctgaatagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttcctgaagtttggctcaggctat
gtgtcaggctggggcagagtgttcacaaagggcagatcagccctgggtgctgcagtacctgagagt
ccctgggtggacagagccacctgcctgttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggcagagacagctgccagggagactcaggaggaccccatgtgacagaa
gtggagggcaccagcttccctgacaggcatcatcagctggggagaggagtgtgccatgaagggcaag
tatggcatctacaccaaagtgagcagatatgtgaactggatcaaggagaaaaccaagctgacctga
(SEQ ID NO:5)

【圖5】

CS03-FL-NA

atgcagaggggtgaacatgatcatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcaactgtgttcctggaccatgagaatgccacaagatcctgaacaggccc
aagagatacaactctggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccaggggaagtgtttgagaacactgagaggaccactgagttctgg
aagcagtatgtggatggggaccagtgtgagagcaacccttgccctgaatgggggagctgcaaggat
gacatcaacagctatgagtgctggtgcccttttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactctgctgacaacaaagtgggtg
tgtagctgcaactgagggctacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgtctgtgtgccagaccagcaagctgaccagagctgagactgtgttccctgat
gtggactatgtgaatagcactgaggctgagaccatcctggacaacatcacccagagcaccagtc
ttcaatgacttcaccagagtgggtggggggggaggatgccaaagcctggccagttcccctggcaggtg
gtgctgaatggcaaagtggatgccttctgtgggggcagcattgtgaatgagaagtggattgtgact
gctgcccactgtgtggagactgggggtgaagatcactgtggtggctgggggaacacaatattgaggag
actgagcacactgagcagaagaggaatgtcatcaggattatccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcctctggtgctgaatagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttcctgaagtttggctctggctat
gtgtctggctggggcagagtgttcacaagggcaggtctgccctggtgctgcagtacctgagagt
cccctggtggacagagccacctgcctgctgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggggggcagagacagctgccaggggggactctgggggcccccatgtgactgaa
gtggaGggcaccagcttcctgactggcatcatcagctgggggggaggagtgtgccatgaagggaag
tatggcatctacaccaaagtgagcaggtatgtgaactggatcaaggagaaaaccaagctgacctga
(SEQ ID NO:6)

【圖6】

CS04-FL-NA

atgcagaggggtgaacatgattatgggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcacagtgttcctggaccatgagaatgccacaagatcctgaacaggccc
aagagatacaactctggcaagctggaggagtttgtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccaggagggtgttgagaacacagagaggaccacagagttctgg
aagcagtatgtggatgggtgaccagtgtgagagcaacccttgccctgaatggaggcagctgcaaggat
gacatcaacagctatgagtgtggtgccccttttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcaggtgtgagcagttctgcaagaactctgctgacaacaagggtggtg
tgtagctgcacagaggggtacagactggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcagagtgctgtgtcccagaccagcaagctgaccagagctgagacagtgttcctgat
gtggactatgtgaacagcacagaggctgagaccatcctggacaacatcaccagagcaccagtc
ttcaatgacttcaccagagtggtgggaggagaggatgccaaagcctggccagttcccctggcaggtg
gtgctgaatggcaagggtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgaca
gctgcccactgtgtggagacaggagtgaagatcacagtgggtggctggagagcacacattgaggag
acagagcacacagagcagaagaggaatgtgatcaggatcatccctcaccacaactacaatgctgcc
atcaacaagtacaacatgacattgccctgctggagctggatgagcctctggtgctgaacagctat
gtgacccttatctgcattgctgacaaggagtagaccaacatcttcctgaagtttggctctggctat
gtgtctggctggggcagagtgttccacaagggcaggtctgccctgggtgctgcagtacctgagagt
cccctgggtggacagagccacctgcctgttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggcagagacagctgccagggtgactctggaggaccccatgtgacagag
gtggaGggcaccagcttcctgacaggcatcatcagctggggagaggagtgtgccatgaagggaag
tatggcatctacaccaaagtgagcagatatgtgaactggatcaaggagaagaccaagctgacctga
(SEQ ID NO:7)

【圖7】

CS05-FL-NA

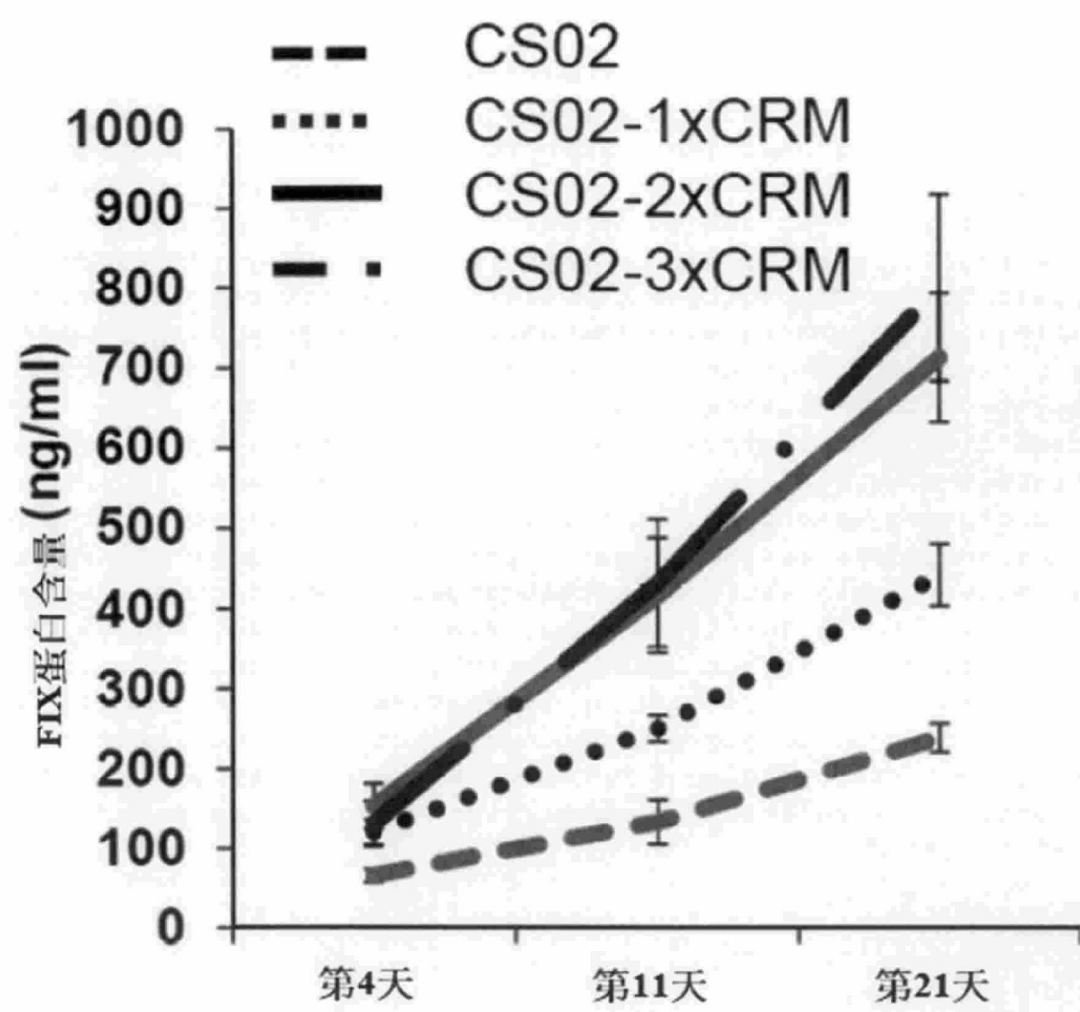
atgcagaggggtgaacatgattatggctgagagccctggcctgatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcactgtgttcctggaccatgagaatgccacaagatcctgaaccgcccc
aagcgctacaactctggcaagctggaggagtttggtgcagggcaacctggagagggagtgcatggag
gagaagtgcagctttgaggaggccagggaggtgtttgagaacactgagcgcaaccactgagttctgg
aagcagtatgtggatggggaccagtggtgagagcaacccctgcctgaatgggggggagctgcaaggat
gacatcaacagctatgagtgcctgggtgcccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggccgctgtgagcagttctgcaagaactctgctgacaacaagggtgggtg
tgctcttgactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgcccttc
ccctgtggcaggggtgtctgtgagccagaccagcaagctgaccagggctgagactgtgttcctgac
gtggactatgtgaacagcactgaggctgagaccatcctggacaacatcaccagagcaccagagc
ttcaatgacttcaccaggggtgggtgggaggagaggatgccaaagcctggccagttcccctggcaggtg
gtgctgaatggcaagggtggatgccttctgtggaggcagcattgtgaatgagaagtggattgtgacc
gctgcccactgtgtggagactggagtgaagatcactgtggtggctggggagcacaacattgaggag
acagagcacacagagcagaagcgcaatgtgatcaggatcatccccaccacaactacaatgctgcc
atcaacaagtacaacctgacattgccctgctggagctggatgagcccctgggtgctgaacagctac
gtgacccccatctgcattgcagacaaggagtacaccaacatcttcctgaagtttggtctctggctat
gtgtctggctggggcaggggtgttcacaaagggcaggtctgccctgggtgctgcagtacctgaggggtg
cccctgggtggacagggccacctgcctgctgagcaccaagttcaccatctacaacaacatgttctgc
gctggcttccatgagggaggaagggacagctgccagggagactctggaggcccccatgtgacagag
gtggagggcaccagcttcctgacaggcatcatcagctggggggaggagtggtgcaatgaagggcaag
tatggcatctacaocaaagtgtcccgtatgtgaactggatcaaggagaagaccaagctgacctga
(SEQ ID NO:8)

【圖8】

CS06-FL-NA

atgcagaggggtcaacatgatcatggctgagtcacctggcctcatcaccatctgcctgctgggctac
ctgctgtctgctgagtgcaactgtcttctctggaccatgagaatgccacaagatcctcaacaggccc
aagagatacaactctggcaaactggaggagtttgtccagggaacctggagagggagtgcatggag
gagaagtgtcctttgaggaggccaggaggtctttgagaacactgagcgcaccactgagttctgg
aaacagtatgtggatggggaccagtgtgagtccaacctgcctgaatgggggcagctgcaaggat
gacatcaacagctatgagtgctggtgcccctttggctttgagggcaagaactgtgagctggatgtg
acctgcaacatcaagaatggcagatgtgagcagttctgcaagaactctgctgacaacaagggtgtg
tgctcctgcactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgccattc
ccatgtggcagagtctctgtgagccagaccagcaagctcaccagggtgagactgtgttccctgat
gtggactatgtgaacagcactgaggctgaaaccaatcctggacaacatcaccagagcaccagagc
ttcaatgacttcaccagagtgttggggaggagaggatgccaaacctggccagttcccctggcaagtg
gtgctcaatggcaagggtggatgccttctgtgggggctccattgtgaatgagaagtggattgtcact
gctgcccactgtgtggagactgggggtcaagatcactgtggtggctggggagcacaacattgaggag
actgagcacactgagcagaagcgcaatgtgatcaggatcatccccaccacaactacaatgctgcc
atcaacaagtacaaccatgacattgccctgctggagctggatgagcccctggtcctcaacagctat
gtgacccccatctgcattgctgacaaggagtacaccaacatcttccctcaagtttggctctggctat
gtctctggctggggcagagtgttccacaaaggcaggtctgcctgggtgctccagtacctgagagtg
cccctgggtggacagggccacctgcctcttgagcaccaagttcaccatctacaacaacatgttctgt
gctggcttccatgagggaggaagagacagctgccagggggactctggaggaccccatgtcactgag
gtggagggcacctccttccctcactggcatcatctcctggggagaggagtgtgccatgaaaggcaaa
tatggcatctacaccaaagtctccagatatgtcaactggatcaaggagaagaccaagctgacctga
(SEQ ID NO: 9)

【圖9】



【圖10】

FIX-MP-AA

YNSG KLEEFVQGNL ERECMEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGSCKD DINSYECWCP FGFEGKNCEL DVTCTNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVS
YVNSTEAEIT LDNITQSTQS FNDFTTRVVG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIIRI
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TCLRSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIEKTKL T (SEQ ID NO:10)

【圖11A】

FIX2-MP-AA

ynsg kleefvqgnl erecmeecks feearevfen terttefwkq yvdtctnikn
grceqfckns adnkvvcsct egyrlaenqk scepavpfpc grvsvsqtsk ltraetvfpd
vdyvnsteae tildnitqst qsfndftrvv ggedakpggf pwqvvlngkv dafcggsivn
ekwivtaahc vetgvkitvv agehnieete hteqkrnvir iiphhnynaa inkynhdial
leldeplvln syvtpiciad keytniflkf gsgyvsgwgr vfhkgrsalv lqylrvplvd
ratclrstkf tiynnmfcag fheggrdscq gdsggphvte vegtsfltgi iswgeecamk
gkygiytkvs ryvnwikekt klt (SEQ ID NO:11)

【圖11B】

FIXp-MP-AA

YNSG KLEEFVQGNL ERECMEEEKCS FEEAREVFEN TERTTEFWKQ YVDGDQCESN
PCLNGGCKD DINSYECWCP FGFEGKNCEL DVTCNIKNGR CEQFCKNSAD
NKVVCSCTEG YRLAENQKSC EPAVPFPCGR VSVSQTSKLT RAETVFPDVG
YVNSTEAEIT LDNITQSTQS FNDFTRVVGG EDAKPGQFPW QVVLNGKVDA
FCGGSIVNEK WIVTAAHCVE TGVKITVVAG EHNIEETEHT EQKRNVIIRI
PHHNYNAAIN KYNHDIALLE LDEPLVLNSY VTPICIADKE YTNIFLKFGS
GYVSGWGRVF HKGRSALVLQ YLRVPLVDRA TLLSTKFTI YNNMFCAGFH
EGGRDSCQGD SGGPHVTEVE GTSFLTGIIS WGEECAMKGK YGIYTKVSRY
VNWIKEKTKL T (SEQ ID NO:12)

【圖12】

CS02-MP-NA

tacaactcag gcaagctgga ggagtttgtg cagggcaacc tggagagggga gtgcatggag
gagaagtgca gctttgagga ggccagagag gtgtttgaga acacagagag gaccacagag
ttctggaagc agtatgtgga tggagaccag tgtgagagca acccttgccct gaatggaggc
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cttttggcct tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactcagctg acaacaaagt ggtgtgtagc tgcacagagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtcagt gtcccagacc
agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa tagcacagag
gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga
gttgtgggag gagaggatgc caagcctggc cagttcccct ggcagggtgt gctgaatggc
aaagtggatg ccttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagaacacaa tattgaggag
acagagcaca cagagcagaa gaggaatgtc atcaggatta tccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct caggctatgt gtcaggctgg ggcagagtgt tccacaaggg cagatcagcc
ctgggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc
tgccagggag actcaggagg accccatgtg acagaagtgg agggcaccag cttcctgaca
ggcatcatca gctggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcagat atgtgaactg gatcaaggag aaaaccaagc tgacctga

(SEQ ID NO:13)

【圖13】

CS03-MP-NA

```
tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagaggga gtgcatggag
gagaagtgca gctttgagga ggccagggaa gtgtttgaga acactgagag gaccactgag
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccttgccct gaatgggggc
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cttttggctt tgaggggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactctgctg acaacaaagt ggtgtgtagc tgcactgagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc
agcaagctga ccagagctga gactgtgttc cctgatgtgg actatgtgaa tagcactgag
gctgagacca tcctggacaa catcacccag agcacccagt cettcaatga cttcaccaga
gtgggtggggg gggaggatgc caagcctggc cagttcccct ggcagggtgg gctgaatggc
aaagtggatg ccttctgtgg gggcagcatt gtgaatgaga agtggattgt gactgctgcc
cactgtgtgg agactggggg gaagatcact gtgggtggctg gggaacacaa tattgaggag
actgagcaca ctgagcagaa gaggaatgtc atcaggatta tccccacca caactacaat
gctgccatca acaagtacaa ccattgacatt gccctgctgg agctggatga gcctctgggtg
ctgaatagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc
ctgggtgctgc agtacctgag agtgccctg gtggacagag ccacctgcct gctgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggggg cagagacagc
tgccaggggg actctggggg ccccatgtg actgaagtgg agggcaccag cttcctgact
ggcatcatca gctgggggga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcaggt atgtgaactg gatcaaggag aaaaccaagc tgacctga
(SEQ ID NO:14)
```

【圖14】

CS04-MP-NA

```

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagagggg gtgcatggag
gagaagtgca gctttgagga ggccagggag gtgtttgaga acacagagag gaccacagag
ttctggaagc agtatgtgga tggtgaccag tgtgagagca acccttgcct gaatggaggc
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cttttggcct tgaggggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca ggtgtgagca gttctgcaag
aactctgctg acaacaagggt ggtgtgtagc tgcacagagg gctacagact ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca gagtgtctgt gtcccagacc
agcaagctga ccagagctga gacagtgttc cctgatgtgg actatgtgaa cagcacagag
gctgagacca tcctggacaa catcacccag agcacccagt ccttcaatga cttcaccaga
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaggtggt gctgaatggc
aaggtggatg ctttctgtgg aggcagcatt gtgaatgaga agtggattgt gacagctgcc
cactgtgtgg agacaggagt gaagatcaca gtggtggctg gagagcaca cattgaggag
acagagcaca cagagcagaa gaggaatgtg atcaggatca tccctcacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gcctctggtg
ctgaacagct atgtgacccc tatctgcatt gctgacaagg agtacaccaa catcttcctg
aagtttggct ctggctatgt gtctggctgg ggcagagtgt tccacaaggg caggtctgcc
ctggtgctgc agtacctgag agtgcccctg gtggacagag ccacctgcct gttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg cagagacagc
tgccaggggtg actctggagg accccatgtg acagaggtgg agggcaccag cttcctgaca
ggcatcatca gctgggggaga ggagtgtgcc atgaagggca agtatggcat ctacaccaaa
gtgagcagat atgtgaactg gatcaaggag aagaccaagc tgacctga
(SEQ ID NO:15)

```

【圖15】

CS05-MP-NA

```

tacaactctg gcaagctgga ggagtttgtg cagggcaacc tggagagga gtgcatggag
gagaagtgca gctttgagga ggccaggag gtgtttgaga aactgagcg caccactgag
ttctggaagc agtatgtgga tggggaccag tgtgagagca acccctgcct gaatgggggg
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cttttggctt tgagggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggcc gctgtgagca gttctgcaag
aactctgctg acaacaagggt ggtgtgctct tgcactgagg gctaccgcct ggctgagaac
cagaagagct gtgagcctgc tgtgcccttc ccctgtggca ggggtgtctgt gagccagacc
agcaagctga ccagggctga gactgtgttc cctgacgtgg actatgtgaa cagcactgag
gctgagacca tcctggacaa catcaccagc agcaccaga gcttcaatga cttcaccagg
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcagggtgg gctgaatggc
aagggtgatg cttctgtgtg aggcagcatt gtgaatgaga agtggattgt gaccgctgcc
cactgtgtgg agactggagt gaagatcact gtggtggctg gggagcacia cattgaggag
acagagcaca cagagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gccctggtg
ctgaacagct acgtgacccc catctgcatt gcagacaagg agtacaccaa catcttctg
aagtttggct ctggctatgt gtctggctgg ggcagggtgt tccacaaggg caggtctgcc
ctggtgctgc agtacctgag ggtgccctg gtggacaggg ccacctgcct gctgagcacc
aagttcacca tctacaacaa catgttctgc gctggcttcc atgaggagg aaggacagc
tgccaggag actctggagg ccccatgtg acagagggtg agggcaccag cttcctgaca
ggcatcatca gctgggggga ggagtgtgcc atgaaggga agtatggcat ctacacaaa
gtgtcccgt atgtgaactg gatcaaggag aagaccaagc tgacctga
(SEQ ID NO:16)

```

【圖16】

CS06-MP-NA

```

tacaactctg gcaaactgga ggagtttgtc cagggcaacc tggagaggga gtgcatggag
gagaagtgct cctttgagga ggccaggag gtctttgaga acactgagcg caccactgag
ttctggaaac agtatgtgga tggggaccag tgtgagtcca acccctgcct gaatgggggc
agctgcaagg atgacatcaa cagctatgag tgctgggtgcc cctttggctt tgaggggcaag
aactgtgagc tggatgtgac ctgcaacatc aagaatggca gatgtgagca gttctgcaag
aactctgctg acaacaagg ggtgtgctcc tgcactgagg gctacgcct ggctgagaac
cagaagagct gtgagcctgc tgtgccattc ccatgtggca gagtctctgt gagccagacc
agcaagctca ccagggtga gactgtgttc cctgatgtgg actatgtgaa cagcactgag
gctgaaacca tcctggacaa catcaccag agcaccaga gcttcaatga cttcaccaga
gtggtgggag gagaggatgc caagcctggc cagttcccct ggcaagtgg gctcaatggc
aaggtggatg ccttctgtgg gggctccatt gtgaatgaga agtggattgt cactgctgcc
cactgtgtgg agactggggt caagatcact gtggtggctg gggagcacia cattgaggag
actgagcaca ctgagcagaa gcgcaatgtg atcaggatca tccccacca caactacaat
gctgccatca acaagtacaa ccatgacatt gccctgctgg agctggatga gccctgggtc
ctcaacagct atgtgacccc catctgcatt gctgacaagg agtacaccaa catcttctc
aagtttggct ctggctatgt ctctggtgg ggcagagtgt tccacaaagg caggtctgcc
ctggtgctcc agtacctgag agtgcccctg gtggacaggg ccacctgcct cttgagcacc
aagttcacca tctacaacaa catgttctgt gctggcttcc atgagggagg aagagacagc
tgccaggggg actctggagg accccatgtc actgaggtgg agggcacctc cttctcact
ggcatcatct cctggggaga ggagtgtgcc atgaaaggca aatatggcat ctacaccaaa
gtctccagat atgtcaactg gatcaaggag aagaccaagc tgacctga

```

(SEQ ID NO:17)

【圖17】

FIX-PPP-NA

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta
ggatatctac tcagtgtctga atgtacagtt tttcttgatc atgaaaacgc caacaaaatt
ctgaatcggc caaagagg (SEQ ID NO:18)

CS02-PPP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtcagcaga gtgcacagtgt ttcttgacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (SEQ ID NO:19)

CS03-PPP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgtctga gtgcactgtg ttcttgacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (SEQ ID NO:20)

CS04-PPP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgtctga gtgcacagtgt ttcttgacc atgagaatgc caacaagatc
ctgaacaggc ccaagaga (SEQ ID NO:21)

CS05-PPP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgtctga gtgcactgtg ttcttgacc atgagaatgc caacaagatc
ctgaacggcc ccaagcgc (SEQ ID NO:22)

CS06-PPP-NA

atgcagaggg tcaacatgat catggctgag tccctggcc tcatcaccat ctgcctgctg
ggctacctgc tgtctgtctga gtgcactgtc ttcttgacc atgagaatgc caacaagatc
ctcaacaggc ccaagaga (SEQ ID NO:23)

【圖18】

FIX-SP-NA

atgcagcgcg tgaacatgat catggcagaa tcaccaggcc tcatcaccat ctgcctttta
ggatatctac tcagtgtga atgt (SEQ ID NO:24)

CS02-SP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtcagcaga gtgc (SEQ ID NO:25)

CS03-SP-NA

atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (SEQ ID NO:26)

CS04-SP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (SEQ ID NO:27)

CS05-SP-NA

atgcagaggg tgaacatgat tatggctgag agccctggcc tgatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (SEQ ID NO:28)

CS06-SP-NA

atgcagaggg tcaacatgat catggctgag tccctggcc tcatcaccat ctgcctgctg
ggctacctgc tgtctgctga gtgc (SEQ ID NO:29)

【圖19】

FIX-PP-NA

acagtttttc ttgatcatga aaacgccaac aaaattctga atcggccaaa gagg
(SEQ ID NO:30)

CS02-PP-NA

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(SEQ ID NO:31)

CS03-PP-NA

actgtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(SEQ ID NO:32)

CS04-PP-NA

acagtgttcc tggaccatga gaatgccaac aagatcctga acaggcccaa gaga
(SEQ ID NO:33)

CS05-PP-NA

actgtgttcc tggaccatga gaatgccaac aagatcctga accgccccaa gcgc
(SEQ ID NO:34)

CS06-PP-NA

actgtcttcc tggaccatga gaatgccaac aagatcctca acaggcccaa gaga
(SEQ ID NO:35)

【圖20】

FIX-PPP-AA

MQRVNMIMAE SPGLITICLL GYLLSAECTV FLDHENANKI LNRPKR (SEQ ID NO:36)

【圖21】

FIX-SP-AA

MQRVNMIMAE SPGLITICLL GYLLSAEC (SEQ ID NO:37)

【圖22】

FIX-PP-AA

TV FLDHENANKI LNRPKR (SEQ ID NO:38)

【圖23】

CRM8

gggggaggctgctggtgaatattaaccaaggtcaccccagttatcggaggagcaaacaggggctaag
tccac (SEQ ID NO:39)

【圖24】

tgggcactccctctctgcgcgctcgctcgctcactgagcgccgggcgaccaaaggtcgcccgacgcgcgggctttgcccgggcggcctcagtgagcgagcgagcgagagggagtgggccaactccatcac taggggttcctgagtttaaaccttcgctcgacgggggagggctgctgggtgaatattaaccaaggtcacc cagttatcggaggagcaaacaggggctaagtccaccgggggagggctgctgggtgaatattaaccaagg tcaccccagttatcggaggagcaaacaggggctaagtccaccgggggagggctgctgggtgaatattaa ccaaggtcaccccagttatcggaggagcaaacaggggctaagtccaccgaggggcaactgggaggatgt tgagtaagatggaaaactactgatgacccttgcagagacagagtattaggacatgtttgaacagggg ccgggcgatcagcaggtagctctagaggatccccgtctgctgacatttcgtagagcgagtggtcc gatactctaattctccctagggcaaggttcataatttggtgtaggttacttattctccttttggtgactaa gtcaataatcagaatcagcaggtttggagtcagcttggcaggggatcagcagcctgggttggaaggag ggggtataaaaagcccttcaccaggagaagccgtcacacagactaggcgcgccctaaggtaagttgg cgccgtttaagggatgggttggttggtgggtattaatgtttaattaccttttttacaggcctgaaga tctgccaccatgcagaggggtcaacatgatcatggctgagtcacctggcctcatcaccatctgcctgc tgggctacctgctgctgctgctgagtgcaactgtcttcctggaccatgagaatgccacaagatccctca caggcccaagagatacaactctggcaaactggaggagtttgtccaggggcaacctggagaggggagtg c atggaggagaagtgtcctttgaggaggccaggggaggtctttgagaacactgagcgccaccactgagt tctggaaacagtatgtggatggggaccagtgtagtccaacccctgctgaatgggggcagctgcaa ggatgacatcaacagctatgagtgctggtgcccttttgctttgagggcaagaactgtgagctggat gtgacctgcaacatcaagaatggcagatgtgagcagttctgcaagaactctgctgacaacaaggtagg tgtgtcctgcactgagggctaccgcctggctgagaaccagaagagctgtgagcctgctgtgccatt cccatgtggcagagtcctctgtgagccagaccagcaagctcaccagggtgtagactgtgttcctgat gtggactatgtgaacagcactgaggctgaaaccatcctggacaacatcaccagagcaccagagct tcaatgacttcaccagagtggtggggaggagaggatgccaaagcctggccagttcccctggcaagtggg gctcaatggcaaggtggatgccttctgtgggggctccattgtgaatgagaagtggttgtcactgct gccactgtgtggagactgggggtcaagatcactgtggtggctgggggagcacaacattgaggagactg agcacactgagcagaagcgcaatgtgatcaggatcatccccaccacaactacaatgctgccatcaa caagtacaaccatgacattgccctgctggagctggatgagccctggtcctcaacagctatgtgacc cccatctgcattgctgacaaggagtacaccaacatcttcctcaagtttggtctctggctatgtctctg gctggggcagagtgttccacaaaggcaggtctgccctgggtgctccagtaacctgagagtgccccctggt ggacaggggccacctgcctcttgagcaccagttcaccatctacaacaacatgttctgtgctggcttc catgaggggaggaagagacagctgccagggggactctggaggacccccatgtcactgaggtggagggca cctccttcctcactggcatcatctcctgggggagaggagtggtgccatgaaaggcaaatatggcatcta caccaaagtctccagatatgtcaactggatcaaggagaagaccaagctgacctgatgagcatgccta gagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttggttgccccctccccctg gccttccttgaccttggaaggtgccactcccactgtcctttcctaataaaaatgaggaaattgcatcg cattgtctgagtaggtgtcattctattctgggggggtgggggtgggggcaggacagcaagggggaggatt gggaagacaatatgcaggcatgctgggggaattaattaagctcgcggaaggaaaccctagtgtatggagtt

【圖25A】

(續)

ggccaactccctctctgcgcgctcgcctcgcctcactgagggccgggaccaaagggtcgcccgacgcccg
ggctttgcccggggcgccctcagtgagcgagcgagcgcgagagaggagtgccaagacgatttaaa
tgacaagcttggcgtaatcatggtcataagctgtttcctgtgtgaaattgttatccgctcacaattcc
acacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaataagtgagtgagcctaactcaca
ttaattgcgttgcgcctcactgcccgtttccagtcgggaaacctgtcgtgcccagctgcattaatgaa
tcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttccctcgcctcactgactc
gctgcgctcgggtcgttcgggtgcggcgagcggtatcagctcactcaaaggcggtataacgggttatcc
acagaatcaggggataacgcaggaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgta
aaaaggccgcggttgcgtggcggttttccataggtccgccccctgacgagcatcacaaaaatcgacg
ctcaagtcagaggtggcgaaacccgacaggactataaagataaccaggcggtttccccctggaagctcc
ctcgtgcgctctcctgttccgacctgcgcgttacccgataacctgtccgcctttctcccttcgggaa
gcggtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagct
gggctgtgtgcacgaaccccccggtcagcccgacgctgcgccttatccggttaactatcgtcttgag
tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcga
ggtatgtagggcggtgctacagagttcttgaagtgggtggcctaactacgggtacactagaagaacagt
atgttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggc
aaacaaaccacgctggtagcggtgggttttttggtttgcaagcagcagattacgcgcagaaaaaaag
gatctcaagaagatcctttgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgtta
agggattttgggtcatgagattatcaaaaaggatcttcacctagatccttttaaatataaaatgaagt
tttaaatcaatctaaagtatatatgagtaaaacttgggtctgacagttaccaatgcttaatcagtgagg
cacctatctcagegatctgtctatttcgttcacccatagttgcctgactccccgtcgtgtagataac
tacgatacgggaggggttaccatctggccccagtgctgcaatgataccgcgagaccacgctcacccg
gctccagatttatcagcaataaaccagccagccggaaggccgagcgcagaagtgggtcctgcaactt
tatccgcctccatccagctctattaattgttgccgggaagctagagtaagtagttccgacagttaatag
tttgcgcaacggttgggtgacattgctacaggcatcgtgggtgtcacgctcgtcgtttgggtatggcttca
ttcagctccgggttcccaacgatcaaggcgagttacatgatcccccatgttggtgcaaaaaagcggtta
gctccttcgggtcctccgatcgttgtcagaagtaagttggccgcagtggttatcactcatgggttatggc
agcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactca
accaagtcattctgagaatagtgtatgcccgcagccaggttgctcttgcccggcgctcaatacgggata
ataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacggttcttcggggcgaaaact
ctcaaggatcttaccgctgttgagatccagttcgatgttaacccactcgtgcacccaactgatcttca
gcactttttactttcaccagcggtttctgggtgagcaaaaaacaggaaggcaaaatgccgcaaaaaagg
gaataagggcgacacggaaatggtgaataactcatactcttctttttcaatattattgaagcattta
tcagggttattgtctcatgagcggatacataatttgaatgtatttagaaaaataaacaatatgggggtt
ccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacct
ataaaaaataggcgatatcacgagggcctttcgtctcgcgcgtttcgggtgatgacggtgaaaacctctg
acacatgcagctcccggagacgggtcacagcttgtctgtaagcggatgccgggagcagacaagcccggt
cagggcgcgctcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattg
tactgagagtgaccatgatgcgggtgtgaaataccgcacagatgcgtaaggagaaaaataccgcacatcag
gcgccatttcgccattcaggctgcgcaactgttgggaaggcgatcggtgcggggcctcTtcgctatta
cgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagct
cacgacgttgtaaaacgacggccagtggaattcctcgagatttaaatgacg

(SEQ ID NO: 40)

【圖25B】