PATENTOVÝ SPIS

(19) ČESKÁ REPUBLIKA

ÚŘAD PRŮMYSLOVÉHO VLASTNICTVÍ

(21) Číslo přihlášky: **1996-2575**
(22) Přihlášeno: **01.03.1995**
(30) Právo přednosti: **03.03.1994 US 1994/206176**
(40) Zveřejněno: **11.06.1997**
(Věstník č. 6/1997)
(47) Uděleno: **23.01.2006**
(24) Oznámení o udělení ve Věstníku: **15.03.2006**
(Věstník č. 3/2006)
(86) PCT číslo: **PCT/US1995/002648**
(87) PCT číslo zveřejnění: **WO 1995/023868**

(11) Číslo dokumentu:

296 382

(13) Druh dokumentu: **B6**

(51) Int. Cl.:

C12N 15/89	(2006.01)
C12N 15/90	(2006.01)
C12N 15/63	(2006.01)
C12N 15/62	(2006.01)
C12N 15/85	(2006.01)
A01K 67/027	(2006.01)
C07K 14/75	(2006.01)

(73) Majitel patentu:
Pharming Intellectual Property BV, Leiden, NL

(72) Původce:
Garner Ian, Edinburgh, GB
Dalrymple Michael, Edinburgh, GB
Prunkard Donna, Seattle, WA, US
Foster Donald C., Seattle, WA, US

(74) Zástupce:
PATENTSERVIS Praha a.s., Jivenská 1273/1, Praha 4, 14021

(54) Název vynálezu:

Způsob produkce fibrinogenu transgenními zvířaty

(57) Anotace:
Předmětem řešení jsou materiály a způsoby produkce fibrinogenu transgenními savci kromě člověka. Do zárodečné linie savce kromě člověka jsou zavedeny segmenty DNA kódující řetězce Aα, Bβ a γ fibrinogenu a savec nebo jeho samičí potomstvo produkuje mléko obsahující fibrinogen exprimovaný zavedenými segmenty DNA. Dále se popisují savčí embrya, kromě člověka, a transgenní savci, kromě člověka, kteří nesou segmenty DNA kódující heterologní polypeptidové řetězce fibrinogenu.

CZ 296382 B6

Způsob produkce fibrinogenu transgenními zvířaty

Oblast techniky

Vynález se týká způsobu produkce fibrinogenu transgenními savci kromě člověka. Samice savce nebo její samičí potomstvo produkuje mléko obsahující fibrinogen exprimovaný segmenty DNA, které jsou zavedeny do jejich zárodečné linie. Tyto segmenty DNA kódují řetězce Aa, Bβ a γ fibrinogenu. Předkládaný vynález dále popisuje embrya savce a produkci transgenních savců kromě člověka, kteří nesou segmenty DNA kódující heterologní polypeptidové řetězce fibrinogenu.

Dosavadní stav techniky

Závěrečným krokem kaskády plazmatické koagulace je konverze v plazmě rozpustného proteinu fibrinogenu v nerozpustný fibrin katalyzovaná trombinem. Trombin odštěpuje malý peptid (fibrinopeptid A) od jednoho ze tří řetězců (řetězec Aa) tvořících fibrinogen. Fibrinové monomery následně polymerizují a aktivovaný faktor XIII vytvořením příčných vazeb stabilizujesraženinu.

Fibrinogen je klíčová složka biologických tkáňových lepidel (viz např. patenty US 4 377 572 a 4 442 655), která napodobují tvoru přirozených krevních sraženin tím, že nastolí hemostázu a spojí poškozenou tkáň. Tkáňová lepidla poskytují doplněk nebo alternativu k šití, svorkám a jiným mechanickým prostředkům k uzavírání ran. Avšak hlavní složky těchto produktů (fibrinogen, faktor XIII a trombin) jsou připravovány z lidské plazmy sloučené od více dárců kryoprecipitací (např. patentu US 4 377 572, 4 362 567, 4 909 251) nebo etanolovou precipitací (např. patent US 4 442 655) nebo z plazmy jednotlivých dárců (např. patent US 4 627 879, spotnitz a kol., Am. surg. 55: 166–168, 1989). Výsledný přípravek fibrinogenu/faktoru XIII se těsně před použitím smíchá s bovinním trombinem, aby se fibrinogen přeměnil na fibrin a zaktivoval se faktor XIII, čímž se koagulace lepidla.

Komerčně dostupná lepidla jsou založena na plazmě sloučené od více dárců. Protože krevní deriváty jsou spojeny s přenosem viru lidské imunodeficience (HIV), viru hepatitidy a jinými etiologickými činiteli, přijatelnost a dostupnost takovýchto lepidel je omezena. V současnosti není jejich používání ve spojených státech povoleno.

Zatímco použití autologní plazmy snižuje riziko přenosu nemocí, autologní lepidla mohou být použita pouze při plánovaném chirurgickém výkonu, kdy pacient je schopen darovat potřebnou krev předem.

Jak bylo zmíněno výše, fibrinogen se skládá za tří polypeptidových řetězců, každý z nichž je v sestavené molekule přítomen ve dvou kopiích. Tyto řetězce, označené Aa, Bβ a gamma, jsou koordinovaně exprimovány, sestavovány a secernovány játry. Ačkoli by se očekávalo, že technologie rekombinantní DNA by mohla poskytnout alternativu k izolaci fibrinogenu z plazmy, zatímco tento cíl stále uniká. Tři řetězce fibrinogenu byly jednotlivě exprimovány v *E. coli* (Lord, DNA 4: 33–38, 1985, Bolyard a Lord, Gene 66: 183–192, 1988, Bolyard a Lord, Blood 73: 1202–1206), ale v prokaryotickém systému nebyl vytvořen funkční fibrinogen. Exprese biologicky způsobilého fibrinogenu kvasinkami nebyla publikována. K expresi biologicky aktivního fibrinogenu byly použity kultury savčích buněk podrobených transfekci (Farrel a kol. Blood 74: 55a, 1989, Hartwig a Danishefsky, J. Biol. Chem. 266: 6578–6585, 1991, Farrell a kol. Biochemistry 30: 9414–9420, 1991), ale expresní hladiny byly tak nízké, že produkce rekombinantního fibrinogenu v komerčních množstvích je neuskutečnitelná. Experimenty naznačují, že nižší rychlost transkripce v buněčných kulturách ve srovnání s játry může způsobovat dosud popisovanou nízkou rychlost exprese, ale zvýšení množství mRNA pro řetězce fibrinogenu v buňkách BHK podrobených transfekci nevyvolalo odpovídající zvýšení v sekreci fibrinoge-

nového proteinu (Prunkard a Foster, XIV. Congress of the International Society on Thrombosis and Haemostasis, 1993). Tyto nedávné výsledky svědčí o tom, že řádné sestavení a zpracování molekuly fibrinogenu vyžaduje tkáňově specifický mechanismus, který není přítomný v běžných laboratorních buněčných liniích.

V oboru přetrvává potřeba najít způsoby, které by produkovaly velká množství vysoce kvalitního fibrinogenu pro použití v tkáňových lepidlech a dalších aplikacích. Dále je potřebné mít fibrinogen, který neobsahuje krví přenášené patogenní zárodky. Předkládaný vynález tyto potřeby uspokojuje a poskytuje další výhody s nimi související.

Podstata vynálezu

Záměrem předkládaného vynálezu je zajistit komerčně využitelná množství rekombinantního lidského fibrinogenu. Dalším cílem vynálezu je poskytnout materiály a způsoby pro expresi fibrinogenu v prsní tkáni transgenních zvířat, zejména hospodářských zvířat jako je dobytek, ovce, prasata a kozy.

Předkládaný vynález poskytuje způsob pro produkci fibrinogenu, kdy se: poskytne první segment DNA kódující sekreční signál funkčně spojený s řetězcem Aα fibrinogenu, druhý segment DNA kódující sekreční signál funkčně spojený s řetězcem Bβ fibrinogenu a třetí segment DNA kódující sekreční signál funkčně spojený s řetězcem γ fibrinogenu, kde každý segment, tj. první, druhý i třetí, je funkčně spojený s dalšími segmenty DNA, které jsou nutné pro jeho expresi v mléčné žláze hostitelské samice savce,

zavedou segmenty DNA do oplodněného vejce savčích druhů kromě člověka, kde molární poměr segmentů DNA, prvního, druhého a třetího, je v rozmezí 0,5 až 1 : 0,5 až 1 : 0,5 až 1,

vloží vejce do vejcovodu savčích druhů kromě člověka, nebo dělohy samice druhu, aby se získalo potomstvo nesoucí konstrukty DNA,

chovají potomci pro produkci samičího potomstva, které exprimuje první, druhý a třetí segment DNA a produkuje mléko obsahující biologicky způsobilý fibrinogen kódovaný segmenty, sbírá mléko od samičího potomstva, a získá fibrinogen z mléka.

V jednom ztělesnění vynálezu je vejce obsahující zavedené segmenty po určitou dobu kultivováno před vložením do dělohy.

Dále vynález poskytuje produkci fibrinogenu obsahující kroky (a) včlenění prvního segmentu DNA kódujícího sekreční signál funkčně spojený s řetězcem Aα fibrinogenu do β–laktoglobulinového genu za vzniku první genové fúze, (b) včlenění druhého segmentu DNA kódujícího sekreční signál funkčně spojený s řetězcem Bβ fibrinogenu do β–laktoglobulinového genu za vzniku druhé genové fúze, (c) včlenění třetího segmentu DNA kódujícího sekreční signál funkčně spojený s řetězcem gamma fibrinogenu do β–laktoglobulinového genu za vzniku třetí genové fúze, (d) zavedení první, druhé a třetí genové fúze do zárodečné linie savce kromě člověka tak, že segmenty DNA jsou exprimovány v mléčné žláze savce nebo jeho samičího potomstva a biologicky způsobilý fibrinogen je secernován do mléka sace nebo jeho samičího potomstva, (e) získání mléka od savce nebo jeho samičího potomstva a (f) získání fibrinogenu z mléka. V preferovaných ztělesněních samce je ovce, prase, koza nebo hovězí dobytek.

Dále vynález poskytuje způsob produkce fibrinogenu skládající se z kroků (a) poskytnutí transgenní samice savců kromě člověka nesoucí ve své zárodečné linii heterologní segmenty DNA kódující řetězce Aα, Bβ a gamma fibrinogenu, kde segmenty DNA je secernován do mléka savce, (b) sběr mléka od savce a (c) získání fibrinogenu z mléka.

Dále vynález poskytuje embrya savců kromě člověka obsahující ve svém jádře segmenty heterologní DNA kódující řetězce Aα, Bβ a gama fibrinogenu. V souvislosti s tím vynález poskytuje transgenní samice savců kromě člověka, které produkují ve svém mléče získatelná množství lidského fibrinogenu.

Dále vynález poskytuje způsob produkce transgenního potomstva savce skládajícího se z kroků (a) poskytnutí prvního segmentu DNA kódujícího řetězec Aα fibrinogenu, druhého segmentu DNA kódujícího řetězec Bβ fibrinogenu a třetího segmentu DNA kódujícího řetězec gamma fibrinogenu, kde každý první a druhý a třetí segment je funkčně spojen s dalšími segmenty DNA nutnými pro jeho expresi v mléčné žláze hostitelské samice savce a sekreci do mléka hostitelské samice savce, (b) zavedení segmentů DNA do oplodněného vejce savců kromě člověka, (c) vložení vejce do vejcovodu nebo dělohy samice savců kromě člověka, aby se získalo potomstvo nesoucí první, druhý a třetí segment DNA. V souvislosti s tím vynález poskytuje savce kromě člověka tvořené podle tohoto postupu.

Dále vynález poskytuje savce kromě člověka nesoucího zárodečnou linii se segmenty DNA kódujícími heterologní řetězce Aα, Bβ a gama fibrinogenu, kde samičí potomstvo savce exprimuje segment DNA v mléčné žláze za tvorby biologicky způsobilého fibrinogenu.

Tyto a další stránky vynálezu budou odborníkovi zjevné z následujícího detailního popisu a přiložených obrázků.

Před detailním vysvětlením vynálezu bude užitečné definovat určité termíny zde užívané:

Zde užívaný termín „biologicky způsobilý fibrinogen“ označuje fibrinogen, který polymerizuje, je-li ošetřen trombinem a vytváří nerozpustný fibrin.

Termín „vejce“ je použit k pojmenování neoplodněného vejce, oplodněného vejce před fúzi haploidních jader (pronukleů) nebo embrya v časném stádiu (oplodněné vejce s fúzovanými pronukley).

„Samice savce, která produkuje mléko obsahující biologicky způsobilý fibrinogen“ je ta, která produkuje následně po těhotenství a porodu během laktačního období mléko obsahující získatelné množství biologicky způsobilého fibrinogenu. Odborníci vědí, že taková zvířata budou produkovat mléko, a tudíž fibrinogen, přerušovaně.

Termín „potomstvo“ je použit ve svém obvyklém smyslu a zahrnuje děti a potomky.

Termín „heterologní“ se používá k označení genetického materiálu, který má původ z odlišného druhu než je druh, do kterého byl zaveden, nebo protein vzniklý z takového genetického materiálu.

V předkládaném vynálezu je technologie transgenních zvířat použita k tomu, aby produkovala fibrinogen v mléčných žlázách hostitelské samice savce. Expres v mléčné žláze a následná sekrece proteinu do mléka překonává mnoho obtíží, na které se naráží při izolaci proteinů z jiných zdrojů. Mléko se snadno sbírá, je dostupné ve velkých množstvích a je biologicky dobře charakterizované. Navíc hlavní mléčné proteiny jsou v mléku přítomné ve vysokých koncentracích (od 1 do 15 g/l).

Z komerčního úhlu pohledu je jasně výhodné jako hostitele použít druh, který má velký výnos mléka. Ačkoliv mohou být použita menší zvířata jako myši a krysy (a jsou preferována ve stadiu ověřování a prokazování), v předkládaném vynálezu se preferuje použít hospodářská zvířata, savce, zahrnující prasata, kozy, ovce a hovězí dobytek, ale neomezující se pouze na ně. Ovcím se obzvláště dává přednost pro takové faktory jako jsou předchozí historie transgeneze u tohoto

druhu, výnos mléka, cena a snadná dostupnost zařízení pro sběr ovčího mléka. Pro srovnání faktorů, které mají vliv na výběr druhu hostitele viz WO 88/00239. Je všeobecně žádoucí vybrat plemeno hostitelského zvířete, které bylo pěstováno pro mléčné použití, jako ovce East Friesland, nebo zavést mléčný dobytek chovem transgenní linie později. V každém případě by měla být použita zvířata o známém dobrém zdravotním stavu.

Fibrinogen produkovaný podle předkládaného vynálezu může být fibrinogen lidský nebo fibrinogen zvířete. Pro léčebná využití se dává přednost pacientovým nativním proteinům. Předkládaný vynález tak poskytuje fibrinogen pro použití jak v lidské, tak veterinární medicíně. Klonované molekuly DNA kódující skladebné řetězce lidského fibrinogenu popsali Rixon a kol. (Biochem 22: 3237, 1983), Chung a kol. (Biochem 22: 3244, 1983), Chung a kol. (Biochem 22: 3250, 1983), Chung a kol. (Adv. Exp. Med. Biol. 281: 39–48, 1990) a Chung a kol. (Ann. NY Acad. Sci. 408: 449–456, 1983). Klony bovinního fibrinogenu popsal Brown a kol. (Nucl. Acids Res. 17: 6397, 1989) a Chung a kol. (Proc. Natl. Acad. sci. USA 78: 1466–1470, 1981). Klony fibrinogenu dalších savců jsou popsány Murakawou a kol. (Thromb. Haemost. 69: 351–360, 1993). Reprezentativní sekvence genů pro lidské řetězce Aα, Bβ a gamma jsou uvedeny pod identifikačními čísly: 1, 3 a 5 v uvedeném pořadí. Odborníci vědí, že existují alelové varianty těchto sekvencí, že další varianty mohou vznikat substitucí aminokyselin, delecí, nebo inzercí, a že takové varianty jsou použitelné v předkládaném vynálezu. Všeobecně se preferuje, že jakékoliv sestrojené varianty obsahují pouze omezený počet substitucí aminokyselin, delecí nebo inzercí a že jakékoliv substituce jsou konzervativní. Dává se přednost produkci takových polypeptidů fibrinogenového řetězce, aby měly přinejmenším z 90 %, výhodně přinejmenším z 95 % a nejvýhodněji z 99 % nebo více procent sekvenci identickou s odpovídajícími nativními řetězci. Termínem „řetězec gamma“ se míní, že zahrnuje i alternativně sestřižený řetězec gamma' fibrinogenu (Chung a kol., Biochem. 23: 4232–4236, 1984). Aminokyselinová sekvence lidského řetězce gamma' je ukázána pod identifikačním číslem 6. Kratší řetězec gamma se vytváří alternativním sestřižením v nukleotidech 9511 a 10054 sekvence identifikačního čísla 5, což má za následek translaci, která je ukončena po nukleotidu 10065 sekvence identifikačního čísla 5.

Aby se dosáhlo exprese v mléčné žláze je použit transkripční promotor z genu mléčného proteinu. Geny mléčných proteinů zahrnující geny, které kódují kaseiny, beta–laktoglobulin (BLG), α–laktalbumin a kyselý protein syrovátky. Preferuje se promotor beta–laktoglobulinu. V případě genu ovčího genu–laktoglobulinu se obvykle používá oblast alespoň proximálních 406 párů bází 5' hraničních sekvence ovčího genu BLG (obsaženém mezi nukleotidy 3844 až 4257 sekvence identifikačního čísla 7). Preferují se větší části 5' hraniční sekvence až do 5 kpb (kilo párů bází). Obzvláště se preferuje větší segment DNA, který obsahuje 5' hraniční oblast promotoru a oblast kódující 5' nekódující část beta–laktoglobulinového genu (obsažené mezi nukleotidy 1 až 4257 sekvence identifikačního čísla 7). Viz Whitelaw a kol., Biochem. J. 286: 31–39, 1992. Vhodné jsou také podobné fragmenty promotorové DNA z jiných druhů.

Do konstruktů mohou být také včleněny jiné oblasti beta–laktoglobulinového genu, aby mohly být exprimovány genomové oblasti genu. Například ne v oboru všeobecně přijímáno, že konstrukty, které neobsahují introny, špatně exprimují ve srovnání s těmi, které takové sekvence DNA obsahují (viz Brinster a kol., Proc. Natl. Acad. sci. USA 85: 836–840, 1988, Palmiter a kol., Proc. Nalt. Acad. sci. USA 88: 478–482, 1991, Whitelaw a kol.., Transgenic Res. 1: 3–13, 1991, WO 89/01342, WO 91/02318). Z tohoto hlediska se všeobecně preferuje použít tam, kde je to možné, genomové sekvence, které obsahují všechny nebo některé z nativních intronů genu kódujícího protein nebo polypeptid. V jistých ztělesněních vynálezu se preferuje další zahrnutí přinejmenším několika intronů beta–laktoglobulinového genu. Jedna taková oblast je segment DNA, který zajišťuje sestřih intronu a polyadenylaci RNA ze 3' nekódující oblasti ovčího beta–laktoglobulinového genu. Je-li nahražen přirozenou 3' nekódující genovou sekvencí, tento segment ovčího beta–laktoglobulinu může jak zvyšovat, tak stabilizovat expresní hladiny proteinu nebo polypeptidu. V jiných ztělesněních je oblast obklopující iniciační ATG jedné či více fibrinogenových sekvencí nahrazena odpovídající sekvencí z genu mléčného specifického proteinu. Taková

náhrada poskytuje předpokládané tkáňové specifické proteiny. Taková náhrada poskytuje předkládané tkáňové specifické iniciační prostředí ke zvýšení exprese. Je příhodné zaměnit celý pre–pro řetězec fibrinogenu a 5' nekódující sekvence např. sekvencemi genu BLG, ačkoliv mohou být zaměněny menší oblasti.

Pro expresi fibrinogenu jsou segmenty DNA kódující každý ze tří dílčích polypeptidových řetězců fibrinogenu funkčně spojen s dalšími segmenty DNA, které jsou nutné pro jejich expresi, a tím se tvoří expresní jednotky. Tyto další segmenty zahrnují výše zmíněný promotor genu mléčného proteinu, jakož i sekvence, které zajišťují terminaci transkripce a polyadenylaci mRNA. Expresní jednotky dále obsahují segment DNA kódující sekreční signál funkčně spojený se segmentem kódujícím polypeptidový řetězec fibrinogenu. Sekreční signál může být nativní sekreční signál fibrinogenu nebo jiného proteinu, např. mléčného proteinu. Termín „sekreční signál" se zde používá k označení té části proteinu, která jej řídí sekreční drahou z buňky do vnějšího prostředí. sekreční signály se obecně nacházejí na amino koncích proteinů. Viz např. von Heinje, Nuc. Acids Res. 14: 4683–4690, 1986 a Meade a ko., U.S. patent č. 4 873 316, zahrnuté zde v odkazech.

Konstrukce expresních jednotek se příhodně provádí vložením sekvence pro řetězec fibrinogenu do plazmidového nebo fágového vektoru, který obsahuje další segmenty DNA, ačkoliv expresní jednotky může být vytvořena v podstatě jakýmkoliv pořadím ligací. Je obzvláště vyhovující obstarat vektor, který obsahuje segment DNA kódující mléčný protein a nahradit kódující sekvenci mléčného proteinu toutéž pro řetězec fibrinogenu, (včetně sekrečního signálu), a tím vytvořit genovou fúzi, která zahrnuje sekvence genu mléčného proteinu pro řízení exprese. V každém případě klonování expresních jednotek v plazmidech nebo jiných vektorech usnadňuje amplifikaci fibrinogenových sekvencí. Amplifikace je pohodlně prováděna v bakteriálních (např. E. coli) hostitelských buňkách, tudíž vektory typicky obsahují počátek replikace a selekční marker funkční v bakteriálních hostitelských buňkách.

Vzhledem k velikosti genů pro řetězce fibrinogenu je nejpraktičtější připravit tři oddělené expresní jednotky, smísit je a zavést směs do hostitele. Odborníci vědí, že mohou být použity také jiné postupy. Například expresní jednotky pro tři řetězce mohou být individuálně zavedeny do odlišných embryí a zkombinovány později v chodu. Ve třetím přístupu tři expresní jednotky mohou být spojeny v jednom vhodném vektor, např. jako je kvasinkový arteficiální chromozom nebo klon fága P1. Kódující sekvence pro dva nebo tři řetězce se mohou kombinovat v polycistronických expresních jednotkách (viz např. Levinson a kol. patent US 4 713 339).

Expresní jednotka(y) je(jsou) potom vložena(y) do oplodněných vajec (včetně embryí v časném stadiu) vybraného hostitelského druhu. Zavedení heterologní DNA může být provedeno jedním z mnoha způsobů včetně mikroinjekce (např. patent US 4 873 191), retrovirové infekce (Jaenisch, science 240: 1468–1474, 1988) nebo místně cílenou integrací za použití embryonálních kmenových (ES) buněk (souhrn Bradley a kol., Bio/Technology 10: 534–539, 1992). Vejce jsou potom implantována do vejcovodů nebo děloh falešně březích samic a je jim umožněno dozrát do porodu. Potomci nesoucí zavedenou DNA ve svých zárodečných liniích mohou předat DNA svému potomstvu normálním mendelovským způsobem umožňujícím vývoj transgenních stád. V oboru jsou známy obecné postupy pro produkci transgenních zvířat. Viz např. Hogan a kol., Manipulating the Mouse Embryo: A Laboratory Manual, Cold spring Harbor Laboratory, 1986, Simon a kol., Bio/Technology 6: 179–183, 1988, Wall a kol., Biol, Reprod. 32: 645–651, 1985, Buhler a kol.., Bio/Technology 8: 140–143, 1990, Ebert a kol., Bio/Technology 9: 835–838, 1991, Krimpenfort a kol, Bio/Technology 9: 844–847, 1991, Wall a kol., J. Cell. Biochem. 49: 113–120, 1992, a publikace WIPO WO 88/00239, WO 90/05188, WO 92/11757, a GB 87/00458, které jsou zde zahrnuty v odkazech. Techniky k zavádění cizích sekvencí DNA do savců a jejich zárodečných buněk byly původně vyvinuty na myších. Viz např. Gordon a kol., Proc. Natl. Acad. Sci. USA 77: 7380–7384, 1980, Gordon a Ruddle, Science 214: 1244–1246, 1981, Palmiter a Brinster, Cell 41: 343–345, 1985, Brinster a kol, Proc., Natl. Acad. Sci. USA 82: 4438–4442, 1985, a Hogan a kol., (tamtéž). Tyto techniky byly následně adaptovány k použití z větších zvířat

včetně druhů hospodářských zvířat (viz např. publikace WIPO WO 88/00239, WO 90/05188 a WO 92/11757 a Simons a kol., Bio./Technology 6: 179–182, 1988). Souhrnem, až dosud nejúčinnějším způsobem dodnes používaným pro vytvoření transgenních myší nebo hospodářských zvířat je injekce několika set lineárních molekul DNA do jednoho z pronukleů oplodněného vejce. Může být také použita injekce DNA do cytoplazmy zygoty.

Preferuje se získání vyvážené exprese jednotlivých řetězců fibrinogenu, aby se umožnilo účinné vytváření hotového proteinu. Ideálně by pro zavádění do vajec měly být všechny tři expresní jednotky na téže molekule DNA. Avšak tento přístup může tvořit technické problémy, např. ve stádiích injekce a manipulace. Například velikost fibrinogenových expresních jednotek může nutně vyžadovat použití kvasinkových arteficiálních chromozomů (YAC) nebo fágu P1 pro amplifikaci a manipulaci DNA před injekcemi. Při tomto přístupu by byly segmenty DNA, které mají být injikovány a obsahují všechny tři expresní jednotky, velmi velké, tudíž by vyžadovaly modifikaci injekční procedury za použití např. jehel většího kalibru. Při mnohem jednodušším přístupu je použita směs všech jednotlivých expresních jednotek. Preferuje se mísit ekvimolární množství tří expresních jednotek, ačkoliv odborníci vědí, že tento poměr může kolísat, aby kompenzoval vlastnosti daných expresních jednotek. Určitá exprese, zpravidla redukované úrovně, se získá, když jsou použita menší molární množství jednoho nebo dvou řetězců a účinnost exprese se snižuje přibližně proporčně k odchylce od preferovaného ekvimolárního poměru. V každém případě se dává přednost použití směsi, která má poměr expresních jednotek Aα:Bβ:gamma v rozmezí 0,5–1:0,5–1:0,5–1. Při poměru jiném než ekvimolárním se preferuje použití relativně více Bβ expresních jednotek. Alternativně se zavádí do jednotlivých vajec jedna nebo směs dvou expresních jednotek. Avšak zvířata získaná tímto způsobem budou exprimovat pouze jeden nebo dva řetězce fibrinogenu. Vytvoření intaktní fibrinogenové molekuly tímto způsobem vyžaduje program pro další chov navržený tak, aby sloučil v jednotlivých skupin zvířat všechny tři expresní jednotky.

Obecně je u samic vyvolaná zvýšená ovulace ošetřením hormonem stimulujícím folikuly a potom jsou samice spářeny. Oplodněná vejce se sbírají a heterologní DNA je injikována do vajec za použití známých způsobů. Viz např. patent US 4 873 191, Gordon a kol., Proc. Natl. Acad. sci. USA 77: 7380–7384, 1980, Gordon a Ruddle, science 214: 1244–1246, 1981, Palmiter a Brinster, Cell 41: 343–345, 1985, Brinster a kol., Proc. Natl, scad. sci. USA 82: 4438–4442, 1985, Hogan a kol., Manipulating the Mouse Embryo: A Laboratory Manual, Cold spring Harbor Laboratory, 1986, simons a kol, Bio/Technology 6: 179–183, 1988, Wall a kol., Biol. Reprod. 32: 645–651, 1985, Buhler a kol., Bio/Technology 8: 140–142, 1990, Ebert a kol. Bio/Technology 9: 835–838, 1991, Krimpenfort a kol., Bio/Technology 9: 844–847, 1991, Wall a kol., J. Cell. Biochem. 49: 113–120, 1992, a publikace WIPO WO 88/00239, WO 90/05118, WO 92/11757, a GB 87/00458, které jsou zde zahrnuty v odkazech.

Pro injekce do oplodněných vajec jsou expresní jednotky odstraněny ze svých příslušných vektorů digescí příslušnými restrikčními enzymy. Pro pohodlnost se preferuje navrhnout vektory tak, že expresní jednotky jsou odstraněny štěpením enzymy, které neštěpí ani v expresních jednotkách ani kdekoli jinde ve vektorech. Expresní jednotky jsou získávány zpět konvenčními metodami, jako je elektroeluce následovaná fenolovou extrakcí a precipitací etanolem, odstředění v sacharózovém hustotním gradientu nebo kombinace těchto postupů.

DNA se injikuje do vajec v zásadě tak, jak je popsáno v Hoganovi a kol, tamtéž. Při typické injekci jsou vejce umístěna v nádobě s embryonálním kultivačním médiem za použití stereo mikroskopu s transfokátorem (preferuje se zvětšení 50x nebo 63x) Vhodné média zahrnují Hepes (N–2–hydroxyethylpiperazin–N'–2–ethansulfonová kyselina nebo bikarbonátová pufrovaná média jsou jako M2 nebo M16 (dostupné od sigma Chemical Co., st. Louis, USA) nebo syntetické médium pro vejcovody (popsáno níže). Vejce jsou zadrženy a přeneseny doprostřed podložního sklíčka na injekční zařízení za použití např. kompletní Drummondovy pipety s kapilárou. V tomto stádiu se pozoruje při nižším (např. 4x) zvětšení. Za použití držící pipety injekčního zařízení jsou vejce centrálně umístěny na sklíčko. Jednotlivá vejce jsou před injekcí postupně upev-

něna k držící pipetě. Pro každý injekční proces je komplex držící pipety s vejcem umístěn do centra zorného pole. Injekční jehla je poté nasměrována přesně pod vejce. Za výhodného použití objektivu Nomarského zvětšujícího 40x, jsou nastaveny obě manipulační výšky tak, aby bylo zaměřeno vejce a jehla. Pronuklea jsou lokalizována otáčením vejce a nastavováním držící pipety dle potřeby. Jakmile je lokalizován pronukleas, výška manipulátoru se změní tak, aby se zaostřilo na pronukleární membránu. Injekční jehla se umístí pod vejce tak, že špička jehly je pod středem pronuklea. Poté se změní pozice jehly za použití manipulátoru tak, aby se jehla a pronukleas dostaly do stejné ohniskové roviny. Jehla se přenese řídicí pákou manipulátoru vpravo od vejce. Krátkým plynulým bodavým pohybem se propíchne pronukleární membrána a špička jehly se ponechá uvnitř pronuklea. Ze skleněné stříkačky se injekční jehlou aplikuje tlak tak dlouho, dokud pronukleas nezvětší přibližně dvakrát svůj objem. V této chvíli se jehla pomalu odstraní. Při přechodu zpět na nižší zvětšení (např. 4x) se injikované vejce odstraní od odlišné oblasti sklíčka a proces se opakuje s jiným vejcem.

Poté, co se injikuje DNA, vejce se kultivují, aby se umožnilo pronuklům fúzovat za tvorby embryí jednobuněčných nebo embryí pozdního stádia. Obecně jsou vejce kultivována při přibližně tělesné teplotě použitého druhu v pufrovém médiu, které obsahuje vyvážené množství solí a sérum. Přežívající embrya jsou potom přenesena do falešně březích samic příjemkyň typicky vložením vajec do vejcovodů nebo dělohy a umožní se jim dozrát k porodu. Během embryogeneze se injikovaná DNA náhodně integruje do genomů malého počtu vyvíjejících se embryí.

Možní transgenní potomci se prozkoumávají pomocí krevních vzorků a/nebo tkáňové biopsie. Z těchto vzorků se připravuje DNA a vyšetřuje se na přítomnost injikovaného konstruktu technikami jako je polymerázová řetězová reakce (PCR, viz Mullis, patent US 4 683 202) a Southernův přenos (Southern, J. Mol. Biol., 98: 503, 1975, Maniatis a kol., Molecular Cloning: A Laboratory Manual, Cold spring Harbor Laboratory,1982). Transgenní zvířata zakladatelské generace, neboli G0, mohou být plně transgenní, mající transgeny ve všech svých buňkách, nebo mozaiky, mající trangeny pouze v částí svých buněk, (viz např. Wilkie a kol., Develop, Biol. 118: 9–18, 1986), V druhém případě skupiny zárodečných buněk mohou být plně nebo částečně transgenní. V druhém případě počet transgenního potomstva od zakladatelského zvířete bude menší než očekávaných 50 %, která jsou předpovězena podle mendelovských zákonů. Zakladatelská G0 zvířata vyrostou do pohlaví zralosti a spáří se, aby se získalo potomstvo neboli G1. G1 se také vyšetří na přítomnost transgenů, aby se prokázal přenos ze zakladatelských G0 zvířat. Co se týče G0 samců, ti mohou být spářeni s několika netransgenními samicemi a dát vznik mnoha potomkům. To zvyšuje šance na pozorování přenosu transgenů. Samice G0 zakladatelské generace mohou být spářeny přirozeně, mohou být uměle oplodněny anebo zvýšeně ovulovány, abys e získalo velké množství vajec, které jsou přeneseny do náhradních matek. Druhý postup dává nejlepší šanci k přenosu u zvířat, která mají omezený počet potomků. Výše zmíněné metody chovu se používají, aby se získala zvířata, která mohou předat DNA následným generacím potomků normálním mendelovským způsobem tak, aby se umožnil vývoj např. kolonií (myší) nebo stád (ovcí, prasat, koz a hovězího dobytka) transgenních zvířat.

Mléko od laktujících G0 a G1 samic se vyšetřuje na expresi heterologního proteinu za použití imunologických technik jako je ELISA (viz Harlow a Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, 1988) a Wsternův přenos (Towbin a kol., Proc. Ntl. Acad. sci. USA 76: 4350–4354, 1979). Expresní hladiny heterologního proteinu jsou mezi jednotlivci dle očekávání odlišné z mnoha důvodů v oboru známých.

Vyhovující rodina zvířat by měla splňovat tři kritéria. Měla by být odvozena ze stejného zakladatelského G0 zvířete, měla by projevovat stabilní přenos transgenu a měla by projevovat stabilní expresní hladiny z generace na generaci a od laktace k laktaci jednotlivých zvířat. tyto principy byly demonstrovány a diskutovány (Carver a kol., Bio/Technology 11: 1263–1270, 1993). Zvířata z takové vhodné rodiny se nazývají linie. Nejprve jsou použity samci G0 nebo G1 ke vzniku stáda producentských zvířat přirozených nebo umělým oplozením. Tímto způsobem může

rychle vzniknout velké množství samic, u kterých je integrován tentýž transgen a od kterých se může získávat zásoba mléka.

Fibrinogen se získává z mléka za použití standardních praktik, jako je odstřeďování mléka, srážení, filtrace a techniky proteinové chromatografie.

Fibrinogen produkovaný podle předkládaného vynálezu je využitelný v lidské a veterinární medicíně, jako například součást chirurgických lepidel. Lepidla tohoto typu jsou v oboru známá. Viz např. patenty US 4 377 572, 4 442 655, 4 462 567 a 4 627 879, které jsou zde zahrnuty v odkazech. Obecně se fibrinogen a faktor XIII slučují a vytvářejí první složku, která se smísí těsně před použitím s druhou složkou obsahující trombin. Trombin konvertuje fibrinogen na fibrin, způsobí změnu směsi na gel a aktivuje faktor XIII. Aktivovaný faktor XIII přičemž váže fibrin a posílí a sterilizuje lepivou hmotu. Taková lepivá typicky obsahující od 30 mg/ml do 100 mg/ml fibrinogenu a od 50 μg/ml do 500 μg/ml faktoru XIII. Mohou také obsahovat další složky, jako aprotinin, albumin, fibrinektin, plnidla a rozpouštědla. V oboru jsou známy způsoby produkce faktoru XIII. Viz např. patent US 5 204 447. Fibrinogen je také užitečný pro potahování povrchů polymerovaných předmět, např. syntetických cévních štěpů, jak vyznačeno v patentu US 5 272 074 (zahrnuto zde v odkazech).

Vynález je dále ilustrován následujícími neomezujícími příklady.

Přehled obrázků na výkresech

Obr. 1 ilustruje subklonování sekvence DNA řetězce Aα lidského fibrinogenu.

Obr. 2 je částečná restrikční mapa vektoru Zem228. Použité symboly jsou MT–1p, promotor myšího metalothioneinu, SV40t, terminátor sV40 a sV40p, promotor sV40.

Obr. 3 ilustruje subklonování sekvence DNA řetězce Bβ lidského fibrinogenu.

Obr. 4 ilustruje subklonování sekvence DNA řetězce gamma lidského fibrinogenu.

Obr. 5 je částečná restrikční mapa vektoru Zem 219b. Použité symboly jsou MT–1p, promotor myšího metalothioneinu, hGHt, terminátor lidského růstového hormonu sV40p, promotor SV40, DHFR, gen dihydrofolátreduktázy a SV40t, terminátor sV40.

Příklady provedení vynálezu

Příklad 1

Odstranilo se mnohočetné klonování místo vektoru pUC18 (Yanisch–Perron a kol., Gene 33: 103–119, 1985) a nahradilo se syntetickým dvouvláknovým oligonukleotidem (jehož vlákna jsou ukázána v sekvencích identifikačních čísel 8 a 27), který obsahuje restrikční místa Pvu I/Mlu I/EcoR V/Xba I/Pvu I/Mlu I a byl ohraničen 5′ přesahy kompatibilními s restrikčními místy EcoR I a Hind III. pUC18 se štěpil oběma EcoR I a Hind III, fosfatázou z telecího střeva se odstranily 5′ terminální fosfátové skupiny a oligonukleotid se zaligoval do základu vektoru. Sekvence DNA od jednoho k druhému konci spojení se potvrdila sekvenováním a nový plazmid byl nazván pUCPM.

Sekvence beta–laktoglobulinového genu (BGL) z pSS1tgXS (vyznačen v publikaci WIPO WO 88/00239) byly vystříženy jako fragment Sal I–Xba I a znovu zaklonovány do vektoru pUCPM, který byl štěpen Sal I a XbaI a vytvořil vektor pUCXS. pUCXS je tedy derivát pUC18,

který obsahuje celý gen BLG od místa sal I k místu Xba I z fágu SS1 (Ali a Clark, J. Mol. Bil. 199: 415–426, 1988).

Plazmid pSS1tgSE (vyznačen v publikaci WIPO WO 88/00239) obsahuje fragment BLG dlouhý 1290 párů bází (bp) ohraničený restrikčními místy pro Sph I a EcoR I, oblast pokrývající unikátní místo Not I a osamocené místo Pvu II, která leží v 5′ nepřekládané vedoucí oblasti mRNA genu BLG. Do tohoto místa Pvu II byl zaligován dvouvláknový linker (spojník) o 8 bp DNA (5′–GGATATCC–3′), který kóduje rozeznávací místo pro enzym EcoR V. Tento plazmid byl nazván pSS1tgSE/RV. Sekvence DNA vymezená restrikčními místy sph I a Not I v pSS1tgSE/RV byla vystřižena enzymatickou digescí a použita k nahrazení ekvivalentního fragmentu v pUCXS. Výsledný plazmid byl nazván pUCXSRV. sekvence inzertu BLG v pUCXSRV je ukázán v sekvenci identifikačního čísla 7, s unikátním místem EcoR V v nukleotidu 4245 v 5′ nepřekládané vedoucí oblasti genu BLG. Toto místo umožňuje inzerci jakýchkoliv dalších sekvencí DNA pod řízením promotoru BLG 3′ k transkripčnímu iniciačnímu místu.

Za použití primerů BLGAMP3 (5′–TGG ATC CCC TGC CGG TGC CTC TGG–3′, sekvence identifikačního čísla 9) a BLGAMP4 (5′–AAC GCG TCA TCC TCT GTG AGC CAG–3′, sekvence identifikačního čísla 10) byl tvořen fragment PCR o přibližné velikosti 650 bp ze sekvencí bezprostředně 3′ ke stopovacímu kodonu genu BLG v pUCXSRV. Fragment PCR byl navržen tak, aby měl na svém 5′ konci místo BamH I a na svém 3′ konci místi Mlu I a byl jako takový klonován do pGEM7zf(+) (Promega) naštěpeným BamH I a Mlu I za vzniku pDAM200(+).

pUCXSRV byl štěpen Kpn I a nejdelší pás, obsahující vektor, byl přečištěn na gelu. Tento pás obsahoval sekvence celého plazmidu pUC a některé 3′ nekódující sekvence z genu BLG. Do tohoto základu byl zaligován malý fragment Kpn I z pDAM200 (+), který ve správné orientaci účinně vytvořil místo BamH I na vzdáleném 5′konci 3′ hraniční oblasti BLG dlouhé 2,6 kbp. Tento plazmid byl nazván pBLAC200. Fragment Cla I–Xba I dlouhý 2,6 kbp z pBLAC200 byl zaligován do vektoru pSP72 (Promega) štěpeném Cla I–Xba I, tak se umístilo místo EcoR V bezprostředně v protisměru sekvencí BLG. Tento plazmid byl nazván pBLAC210.

Fragment EcoR V–Xba I dlouhý 2,6 kbp z pBLAC210 byl zaligován do pUCXSRV štěpeném EcoR V–XbaI za tvorby pMAD6. Toto vlastně vystřihlo z pUCXSRV všechny kódující a intronové sekvence a vytvořilo malý gen BLG skládající se z 5′ promotoru o délce 4,3 kbp a sekvencí o délce 2,6 kbp 3′ po směru ohraničujících unikátní místo EcoR V. Do místa EcoR vektoru pMAD6 byl vložen oligonukleotidový linker (ZC6839: ACTOCGTAGT, sekvence identifikačního čísla 11). Tato modifikace zrušila místo EcoR V a vytvořila místo SnaB I k použití pro klonovací účely. Vektor byl označen pMAD6–Sna. Messenger–RNA (mRNA) začíná v protisměru od místa SnaB I a končí po směru od místa snaB I. Prekurzorový transkript bude kódovat jednotlivý intron pocházející z BLG, intron 6, který je plně v 3′ nepředkládané oblasti genu.

Příklad 2

Klony kódující jednotlivé řetězce fibrinogenu byly získány z laboratoře Dr. Earla W. Davieho, University of Washintgon, seattle. Genomický klon řetězce Aα fibrinogenu (Chung a kol., 1990, tamtéž) byl získán z plazmidu BS4. Tento plazmid obsahuje klon Aα vložený do míst Sal I a BamH I vektoru pUC18, ale postrádá kódující sekvenci pro první čtyři aminokyseliny řetězce Aα. Genomická DNA řetězce Bβ (Chung a kol., tamtéž) byla izolována z klonu fága lambda Charon 4A (označeného βλ4) jako dva fragmenty, každý o délce přibližně 5,6 kbp. Tyto dva fragmenty byly klonovány odděleně v pUC19, který byl naštěpen EcoR I a ošetřen fosfatázou z telecího střeva. Restrikční klony se prozkoumávaly digescí s restrikčním enzymem Pvu II, aby se rozlišily plazmidy s 5′ a 3′ inzerty Bβ (označené Bata5′Ri/puc a Beta3′RI/puc, v příslušném pořadí). Genomické klony řetězce gamma byly izolovány, jak popsáno Rixonem a kol.

(Biochemistry 24: 2077–2086, 1985). Klon pgamma12A9 zahrnuje 5′ nekódující sekvence a přibližně 4535 bp kódující sekvence pro řetězec gamma. Klon pgamma12F3 zahrnuje zbývající kódující sekvenci a 3′ nekódující nukleotidy. Oba klony jsou založené na plazmidu pBR322 s fibrinogenovými sekvencemi vloženými do místa EcoR I. Tyto plazmidy byly použity jako templáty pro příslušné reakce PCR.

Kódující sekvence pro řetězce fibrinogenu byly upraveny k inzerci do expresních vektorů za použití polymerázové řetězcové reakce (PCR), jak obecně popsáno Mullisem (patent US 4 683 202). Tento postup odstranil nativní 5′ a 3′ nepřekládané sekvence, přidal sekvenci 9 bází (CCT GCA GCC) v protisměru prvního ATG každé kódující sekvence, poskytl první čtyři kodony sekvence řetězce Aα, odstranil vnitřní místi Mlu I v sekvenci Aα a přidal restrikční místa k usnadnění následujících klonovacích kroků.

Odvolávaje se na obr. 1, 5′ konec kódující sekvence Aα byl upraven reakcí PCR, která obsahovala 20 pmol každého z primerů ZC6632 (sekvence identifikačního čísla 12) a ZC6627 (sekvence identifikačního čísla 13), přibližně 10 ng plazmidu BS4 jako templátovou DNA, 10 µl směsi obsahující 2,5 mM každého z dNTP, 7,5 µl 10x DNA polymerázového pufru *Pyrococcus furiosus* (Pfu) č. 1 (200 mM Tris–HCl, pH 8,2, 100 mM KCl, 60 mM $(NH_4)_2SO_4$, 20 mM $MgCl_2$, 1% Triton X–100, 100 µg/ml bovinního sérového albuminu bez přítomnosti nukleáz) (Stratagene, La Jolla, CA) a vodu do objemu 75 µl. směs byla zahřáta na 94 °C v terminálním cykleru pro DNA (Perkin–Elmer Corp., Norwalk, CT). K zahřáté směsi se přidalo 25 µl směsi obsahující 2,5 µl 10x pufru Pfu č. 1, 22 µl H_2O a 1 µl DNA–polymerázy Pfu o koncentraci 2,5 jednotek/µl (Stratagene). Reakce probíhaly v termálním cykleru pro DNA (Perkin–Elmer) po 5 cyklů v 94 °C, 45 s, 40 °C, 90 s, 72 °C, 120 s, po 20 cyklů 94 °C, 45 s, 45 °C, 90 s, 72 °C, 120 s, pak byly inkubovány v 72 °C po 7 minut. 5′ fragment vytvořený PCR byl štěpen BamH I a Hind III a fragment BamH I–Hind III byl potom spojen s vnitřním fragmentem Hind III–Xba I o délce 2,91 kbp a s vektorem pUC18 štěpeným BamH I, Xba I. Exonové sekvence vzniklé PCR byly sekvenovány.

Odvolávaje se opět na obr. 1, 3′ konec kódující sekvence Aα byl upraven sérií kroků, ve kterých bylo místo Mlu I 563 bp v protisměru od stopovacího kodonu sekvence Aα změněno za použití PCR s překrývajícím se prodloužením řetězců (Ho a ko., Gene 77: 51-59, 1989). V první reakci bylo v reakční směsi, jak je popsána výše, smícháno 40 pmol každého z primerů ZC6521 (sekvence identifikačního čísla 14) ZC6520 (sekvence identifikačního čísla 15) a přibližně 10 ng plazmidu BS4 s templátovou DNA. Reakce probíhala 5 cyklů v 94 °C, 45 s, 40 °C, 60 s, 72 °C, 120 s, 15 cyklů v 94 °C, 45 s, 45 °C, 60 s, 72 °C, 120 s a pak inkubace v 72 °C po 7 minut. Druhá reakce byla uskutečněna stejným způsobem za použití 40 pmol každého z primerů ZC6519 (sekvence identifikačního čísla 16) a ZC6518 (sekvence identifikačního čísla 17) a BS4 jako templátu. DNA fragmenty vzniklé první a druhou reakcí PCR byly izolovány gelovou elektroforézo a elucí z gelu. Přibližně 1/10 každého znovu získaného reakčního produktu byla smíchána se 40 pmol každého z primerů ZC6521 (sekvence identifikačního čísla 14) a ZC6518 (sekvence identifikačního čísla 17) v reakci PCR, ve které se komplementární 3′ konce každého fragmentu (obsahující změnu jedné fáze) spojily a složily jako primer k extenzi komplementárního vlákna ve směru 3′. PCR se prováděla za použití stelných reakčních podmínek jako v prvním a druhém kroku 3′ PCR. Reakční produkt byl potom naštěpen Xba I a BamH I a tento fragment Xba I–BamH I byl klonován do vektoru pUC18 štěpeného Xba I a BamH I. Exony vzniklé PCR byly sekvenovány.

Jak je ukázáno na obr. 1, 5′ fragment BamH I–Xba I (3,9 kbp) a 3′ fragment Xba I–BamH I (1,3 kbp) byly vloženy do místa BamH I vektoru Zem228. Zem228 je odvozen od pUC18 obsahujícího klonovací místo BamH I mezi myším promotorem MT–1 a terminátorem SV40 a marker pro neomycinovou rezistenci ohraničený promotorem SV40 a terminačními sekvencemi. Viz European patent Office Publication EP 119 944 a obr. 2. Celá kódující sekvence Aα byla izo-

lována z vektoru Zem228 jako fragment SnaB I, který byl vložen do místa SnaB I plazmidu pMAD6–Sna.

Odvolávaje se na obr. 3, 5′ konec řetězce Bβ byl upraven PCR za použití oligonukleotidů ZC6629 (sekvence identifikačního čísla 18), ZC6630 (sekvence identifikačního čísla 19) a ZC6625 (sekvence identifikačního čísla 19) a ZC6625 (sekvence identifikačního čísla 20). Tyto primery byly použity v párových kombinacích (ZC6629 + ZC6625 nebo ZC6630 + ZC6625), aby vznikly kódující sekvence Bβ začínající v prvním ATG kodonu (pozice 470 v sekvenci identifikačního čísla 3) (označené N1–Beta) nebo ve třetím ATG kodonu (pozice 512 v sekvenci identifikačního čísla 3) (označené N3–Beta). Přibližně 5 ng templátové DNA Beta5′RI/puc se smíchalo s 20 pmol každého primeru (N1–Beta. ZC6629, sekvence identifikačního čísla 18 + ZC6625, sekvence identifikačního čísla 2 nebo N3–Beta: ZC6630, sekvence identifikačního čísla 19 + ZC6625, sekvence identifikačního čísla 20) v reakční směsi, jak popsáno výše. směsi byly inkubovány po 5 cyklů 94 °C, 45 s, 40 °C, 120 s, (N1–Beta) nebo 90 s (N3–Beta), 72 °C, 120 s, 20 cyklů 94 °C, 45 s, 45 °C, 120 s, (N1–Beta) nebo 90 s (N3–Beta), 72 °C, 120 s, pak inkubovány v 72 °C po 7 minut. Tyto dva reakční produkty (N1, 555 bp, nebo N3, 510 bp) byly štěpeny Eco RI a Bgl II a fragmenty byly zaligovány do vnitřního fragmentu Bgl II–Xba I a vektoru pUC19 naštěpeného EcoR I + Xba I. 3′ konec sekvence Bβ byl upraven v reakční směsi, jak je popsána výše, za použití oligonukleotidových primerů ZC6626 (sekvence identifikačního čísla 21) a ZC6624 (sekvence identifikačního čísla 22) a přibližně 5 ng templátu Beta3′RI/puc. Směsi byly inkubovány po 5 cyklů 94 °C, 45 s, 40 °C, 90 s, 72 °C, 120 s, 15 cyklů 94 °C, 45 s, 45 °C, 90 s, 72 °C, 120 s, pak inkubovány v 72 °C po 7 minut. Byl izolován fragment Bgl II–EcoR I dlouhý 990 bp. Tento 3′ fragmenty byl spojen se sousedním kódujícím fragmentem (340 bp, sph I–Bgl II) a s vektorem pUC19 naštěpeným Sph I + EcoR I. 3′ a 5′ epoxy vzniklé PCR byly sekvenovány. Třetí přechodný vektor byl vytvořen spojením dvou vnitřních fragmentů (Xba I–EcoR I, dlouhý 4285 bp a EcoR I–Sph I, dlouhý 383 bp) ve vzniku pUC19 naštěpeném Xba I + Sph I. Veškerá kódující sekvence Bβ (dvě formy) byla potom sestavena ligací jednoho z 5′ fragmentů EcoR I–Xba I, vnitřního fragmentu Xba I–Sph I, z 3′ fragmentu sph I–EcoR I vektoru pUC19 naštěpeného EcoR I. sekvence Bβ byla potom izolována jako fragment Sna BI dlouhý 7,6 kbp a vložena do místa SnaB I vektoru pMAD6–Sna.

Odvolávaje se na obr. 4, 5′ konec sekvence pro řetězec gamma byl upraveny PCR za použití oligonukleotidových primerů ZC6514 (sekvence identifikačního čísla 23) a ZC6517 (sekvence identifikačního čísla 24) a přibližně 50 ng plazmidu pgamma 12A9 jako templátu. Reakce probíhala za použití 40 pM každého primeru, jak je popsáno výše. Reakce probíhala 5 cyklů v 94 °C, 45 s, 40 °C, 60 s, 72 °C, 120 s, následováno 15 cykly v 94 °C, 45 s, 45 °C, 60 s, 72 °C, 120 s. Výsledný fragment dlouhý 213 bp byl štěpen BamH I a spe I a výsledný restrikční fragment byl spojen se sousedním po směru fragmentem spe I–EcoR I, dlouhým 4,4 kbp a vektorem pUC19 naštěpením BamH I + EcoR I. 3′ konec sekvence řetězce gamma byl upraven za použití oligonukleotidových primerů ZC6516 (sekvence identifikačního čísla 25) a ZC6515 (sekvence identifikačního čísla 26), 40 pmol každého primeru, přibližně 50 ng templátu pgamma12F3 a stejného schématu pro termální cykly, jak bylo použito pro 5′ fragment. Výsledný fragment dlouhý 500 bp byl štěpen Spe I a BamH I a výsledný restrikční fragment byl spojen s protisměrným fragmentem EcoR I–Spe I dlouhý, 2,77 kbp a vektorem pUC19 naštěpeným EcoR I + BamH I. Všechny exony vzniklé PCR byly sekvenována. Celá kódující sekvence řetězce gamma′ byla potom sestavena ligací 5′ fragmentu BamH I–EcoR I o délce 4,5 kbp, vnitřního fragmentu EcoR I–Pst I o délce 1,1 kbp a 3′ fragmentu Pst I–Xba I o délce 2,14 kbp s vektorem Zem219b naštěpeným BamH I + Xba I. Zem 219b je vektor odvozený od pUC18 obsahující promotor myšího metalothioneinu a selektivní marker DHFR funkčně spojený s promotorem sV40 (obr. 5). Plazmid Zem 219b byl uložen v American Type Culture Collection jako *E. coli* XL1–modrá transformanta pod přístupovým číslem 68979. Celá kódující sekvence řetězce gamma′ byla poté izolována jako fragment SnaB I dlouhý 7,8 kbp a vložena do místa SnaB I vektoru pMAD6–Sna.

Příklad 3

Myši pro počáteční chovné kmeny (C57BL6J, CBACA) byly získány od Harlan Olac Ltd. (Bicester, UK). Byly spárovány a produkovaly F1 hybridní chov (B6CBAF1) na příjemkyně, superovulovaná samice, chovné samce a vazektomované samce. Všechna zvířata byla držena při cyklu 14 hodin světlo/10 hodin tma a krmena vodou a jídlem (Special Diet services, RM3, Edinburgh, scotland) *ad libitum*.

Transgenní myši vznikaly v podstatě jak popsáno v Hoganovi a kol., Manipulating the Mouse Embryo: A Laboratory Manual, Cold spring Harbor Laboratory, 1986, což je ve své celistvosti zahrnuto zde v odkazech. U samic B6CBAF1 byla vyvolána zvýšená ovulace ve 4 až 5 týdnech věku i.p. (intraperitoneální) injekcí sérového gonadotropinu od březí klisny (FOLLIGON, Vet–Drug, Falkirk, scotliand) (5 m.j.), která byla o 45 hodin později následována i.p. injekcí lidského choriogonadotropinu (CHORULON), Vet–Drug, Falkirk, scotland) (5 m.j.). Poté byly spářeny přes noc s chovným samcem. Samice byly potom vyšetřeny na přítomnost kopulačních zátek. Ty, které zabřezly, byly usmrceny a jejich vejce byla shromážděna pro mikroinjekce.

Do oplodněných vajec byla injikována DNA, jak je popsáno v Hoganovi a kol. (taktéž). Ve stručnosti, každý z vektorů obsahující expresní jednotky Aα, Bβ a gamma byl naštěpen Mlu I a expresní jednotky byly izolovány odstředěním v sacharózovém hustotním gradientu. Všechny použité chemikálie byly reagenčního stupně čistoty „reagent grade" (Sigma Chemicals Cl., st. Louis, MO, U.S.A) a všechny roztoky byly sterilní a neobsahovaly nukleázy. Roztoky 20 % a 40 % sacharózy v 1 m NaCl, 20 mM Tris pH 8,0, 5 mM EDTA byly připraveny za použití vody o ultravysokém stupni čistoty a sterilizovány filtrací. 30% roztok sacharózy byl připraven smísením stejných objemů 20% a 40% roztoků. Gradient byl připraven vrstvením 0,5 ml stupňů 40 %, 30 % a 20 % roztoků sacharózy do 2 ml polyalomerové kyvety a ponechán stát po dobu jedné hodiny. 100 µl roztoku DNA (max. 8 µg DNA) bylo navrstveno na vrchol gradientu a gradient se odstřeďoval 17 až 20 hodin ve 26 000 rpm, 15 °C v ultracentrifuze Beckman TL100 za použití rotoru TLS–55 (Beckman Instruments, Fullerton, CA, USA). Gradienty byly rozděleny do frakcí propíchnutím dna zkumavky jehlou č. 20 a sbíráním kapek do mikrotitrační destičky s 96 jamkami. Poměrné objemy 3 µl se analyzovaly na 1% agarózovém minigelu. Frakce obsahující požadovaný fragment DNA byly sloučeniny a precipitovány etanolem v 0,3 M Na–acetátu přes noc v –20 °C. Pelety DNA byly resuspendovány v 50 až 100 µl vody o ultravysokém stupni čistoty a kvantifikovány fluorimetrií. Expresní jednotky byly naředěny v Dulbeccově fosfátovém pufrovaném fyziologickém roztoku bez kalcia a magnézia (obsahující na litr 0,2 g KCl, 0,2 g KH_2PO_4, 8,0 g NaCl, 1,15 g Na_2HPO_4), smíchány (za použití buď N1–Beta, nebo N3–Beta expresní jednotky) v molárním poměru 1:1:1, koncentrace byla nastavena na 6 µg/ml a injikovaly se do vajec (~2 pl roztoku celkové DNA na vejce).

Samice příjemkyně staré 6 až 8 týdnů se připravují spářením samic B6CBAF1 v přirozeném estru s vazektomovanými samci. Samice, u kterých se vyskytnou kopulační zátky, jsou poté chovány pro přenos vejce.

Potenciálním transgenním zvířatům se čtyři týdny po narození odebírají biopsie z ocasů v celkové anestezii. Vzorky tkáně se umístí do 2 ml pufru pro ocasní biopsie (0,3 M Na acetát, 50 mM HCl, 1,5 mM $MgCl_2$, 10 mM Tris–HCl, pH 8,5, 0,5 % NP40, 0,5 % Tween 20), který obsahuje 200 µg/ml proteinázy K (Boehringer Mannheim, Mannheim, SRN), a promíchají se (vortex). Vzorky se protřepávají (250 rpm) v 55 až 60 °C po dobu od 3 hodin až přes noc. DNA připravená z bioptických vzorků se vyšetřuje na přítomnost injikovaných konstruktů metodami PCR a Southernova přenosu. Natrávená tkáň je důkladně promíchána a poměrné objemy po 5 µl jsou přeneseny do mikrocentrifugačních zkumavek o objemu 0,5 ml. Jako kontroly jsou zahrnuty pozitivní a negativní vzorky z ocasní biopsie. Do každé zkumavky se přidá 40 µl silikonového oleje (BDH, Poole, UK) a zkumavky se krátce stočí. Zkumavky se inkubují v zahřívacím bloku

termálního cykleru (např. Omni–gene, Hybaid, Teddington, UK) na 95 °C po 10 minut. Poté jsou do každé zkumavky přidány poměrné objemy po 45μl směsi PCR tak, že konečná koncentrace každé reakční směsi je: 50 mM KCl, 2 mM $MgCl_2$, 10 mM Tris–HCl (pH 8,3), 0,01% želatina, 0,01 % NP40, 10% DMSO, 500 nM každého primeru, 200 μM každého z dNTP, 0,02 U/μl polymerázy Taq (Boehringer Mannheim, Mannheim, SRN). Zkumavky potom probíhají cykly s 30 opakovanými teplotními změnami, jak je vyžadují konkrétní použité primery. Primery se mohou různit, ale ve všech případech musí cílit na promotorovou oblast BLG. Tato je specifická pro fragmenty injikované DNA, neboť myš nemá gen BLG. Poté je do každé zkumavky přidáno 12 μl 5x nanášecího pufru obsahujícího značkovací barvu Orange G (0,25% Orange G [Sigma] 15% Ficcol typ 400 [Pharmacia Biosystems Ltd., Milton Keynes, VB]) a reakční směsi jsou děleny elektroforézou v 1,6% agarózovém gelu obsahujícím ethidien bromid (Sigma) dokud značkovací barva nedorazí do 2/3 délky gelu. Gel je vizualizován zdrojem UV světla vysílajícím o vlnové délce 254 nm. Tímto přístupem jsou identifikovány transgenní myši, které mají jeden nebo více fragmentů injikované DNA.

Pozitivní vzorky ocasní biopsie jsou dále zpracovávány pro získání čisté DNA. Vzorky DNA jsou prozkoumávány metodou southernova přenosu za použití sondy pro promotor genu GLB (nukleotidy 2523–4253 sekvence identifikačního čísla 7). Specifická štěpení s příslušnými restrikčními enzymy (např. EcoR I) umožní rozlišení tří konstruktů obsahujících sekvence Aα, Bβ a gamma.

Analýzy transgenních myší Southernovým přenosem připravená v podstatě tak, jak popsána výše, dokázala, že více než 50 % potomstva obsahovalo všechny tři sekvence fibrinogenu. Vyšetření mléka pozitivních zvířat elektroforézou v redukujícím polyakrylamidovém gelu s SDS ukázalo přítomnosti všech tří proteinových řetězců v koncentracích až 1 mg/ml. Množství plně sestaveného fibrinogenu bylo ve vztahu k poměru jednotlivých podjednotek přítomných v mléce. S vysokými koncentracemi lidského fibrinogenu v myším mléce nebyl spojen žádný zjevný fenotyp.

Příklad 4

Bahnice dárkyně jsou v den 0 ošetřeny intravaginální houbou napuštěnou protesteronem (CHRONOGEST Goat Sponge, Intervet, Cambridge, VB). Houby jsou ponechány *in situ* deset nebo dvanáct dní.

Zvýšená ovulace je navozena ošetřením bahnic dárkyň celkem jednou jednotkou ovčího hormonu stimulujícího folikuly (OFSH) (OVAGEN, Horizon Animal Reproduction Technology Pty. Ltd., Nový Zéland) podané v osmi intramuskulárních injekcích o 0,125 jednotky na injekci, začínající v 17,00 dne –4 a končící v 8,0 dne 0. Dárkyně jsou injikovány intramuskulárně 0,5 ml luteolytického činidla (ESTRUMATE, Vet–Drug) dne –4, aby došlo k regresi corpus luteum a umožnil se návrat k esteru a ovulaci. Aby se ovulace synchronizovala, jsou zvířata dárkyně v den 0 v 17,00 intramuskulárně injikována 2,0 ml syntetického analogu uvolňujícího hormonu (RECEPTAL, Vet–Drug).

Dárkyně hladoví a žízní alespoň dvanáct hodin před umělým oplodněním (U.O.). Zvířata jsou uměle oplodněna v den 1 intrauterinní laparoskopií při sedativech a lokální anestezii. Pro zajištění zklidnění se intramuskulárně injikuje přibližně patnáct minut před umělým oplodněním buď xylazin (ROMPUN, Vet–Drug) v dávce 0,05 až 0,1 ml na 10 kg tělesné váhy, nebo injekce ACP 10 mg/ml (Vet–Drug) v dávce 0,1 ml na 10 kg tělesné váhy. Umělé oplodnění je provedené za použití čerstvě odebraného semene od berana Poll Dorset. Semeno je naředěno stejným objemem filtrovaného fyziologického roztoku pufrovaného fosfáty a do každého děložního rohu se injikuje po 0,2 ml naředěného semene. Okamžitě před nebo po umělém oplodnění je dárkyním podána intramuskulární injekce AMOXYPENU (Vet–Drug).

Oplodněná vejce jsou získána v den 2, poté co dárkyně od 17,00 den 1 hladověly a žíznily. Získání vajec je provedeno v celkové anestezii navozené intravenózní injekcí 5% thiopenton natria (INTAVAL SODIUM, Vet–Drug) v dávce 3 ml na 10 kg tělesné váhy. Anestezie je udržována inhalací směsi 1 až 2% halotanu/O_2/N_2O po intubaci. Aby se získala oplodněná vejce, je provedena laparotomická incize a děloha je vyňata ven z dutiny břišní. Vejce se získají retrográdním výplachem vejcovodů kultivačním médiem pro vejce (Ovum Culture Medium, Advanced Protein Products, Brierly Hill, West Midlands, VB) doplněným bovinním sérovým albuminem novozélandského původu. Po výplachu se děloha vrátila do břicha a incize byla uzavřena. Dárkyním bylo po operaci buď dovoleno zotavit se, nebo na nich byla provedena eutanazie. Dárkyním, které se zotavovaly, byla okamžitě před nebo po operaci podána intramuskulární injekce Amoxypenu L.A. v dávce doporučené výrobcem.

Plazmidy obsahující tři expresní jednotky řetězců fibrinogenu jsou štěpeny s Mlu I a fragmenty expresních jednotek jsou získány nazpět a čistí se na sacharózových hustotních gradientech. Koncentrace fragmentů se určují fluorimetricky a ředí se Dulbeccovým fyziologickým roztokem pufrovaným fosfáty bez kalcia a magnézia, jak popsáno výše. Koncentrace je upravena na 6 μg/ml a přibližně 2 pl směsi jsou mikroinjikovány do jednoho pronuklea každého oplodněného vejce s viditelnými pronukley.

Všechna oplodněná vejce, která přežijí pronukleární mikroinjekci jsou kultivována *in vitro* v 38,5 °C v atmosféře 5% CO_2: 5% O_2: 90% N_2 a ~100% vlhkosti v syntetickém vejcovodovém médiu pufrovaném bikarbonáty (viz tabulka) doplněném 20% v/v sérem vazektomovaného berana. Sérum může být inaktivováno teplem v 56 °C po 30 minut a před použitím uskladněno zmražené v –20 °C. Oplodněná vejce jsou kultivována po dobu přiměřenou k tomu, aby se mohla projevit časná mortalita embryí (způsobená manipulačními technikami). Tato mrtvá nebo zastavená embrya jsou vyřazena. Embrya projevující 5 nebo 6 buněčných dělení jsou přenesena do synchronizovaných bahnic příjemkyň.

Tabulka

Syntetické vejcovodové médium

Zásobní roztok A (trvanlivost 3 měsíce)

NaCl	6,29 g
KCl	0,534 g
KH_2PO_4	0,162 g
$MgSO_4.H_2O$	0,182 g
Penicilin	0,06 g
60% sirup natrium laktátu	0,6 ml
H_2O (v čistotě pro TK)	99,4 ml

Zásobní roztok B (trvanlivost 2 týdny)

$NaHCO_3$	0,21 g
Fenolová červeň	0,001 g
H_2O (v čistotě pro TK)	10 ml

Zásobní roztok C (trvanlivost 2 týdny)

Natrium pyruvát	0,051 g
H_2O (v čistotě pro TK)	10 ml

Zásobní roztok D (trvanlivost 3 měsíce)

$CaCl_2.2H_2O$	0,262 g
H_2O (v čistotě pro TK)	10 ml

Zásobní roztok E (trvanlivost 3 měsíce)

Hepes	0,651 g
Fenolová červeň	0,001 g
H_2O (v čistotě pro TK)	10 ml

10 ml média pufrovaného bikarbonáty

Zásobní roztok A	1 ml
Zásobní roztok B	1 ml
Zásobní roztok C	0,07 ml
Zásobní roztok D	0,1 ml
H_2O (v čistotě pro TK)	7,83 ml

Osmolarita by měla být 265–285 mOsm.
Přidej 2,5 ml teplem inaktivovaného ovčího séra a sterilizuj filtrací.

10 ml média pufrovaného HEPES

Zásobní roztok A	1 ml
Zásobní roztok B	0,2 ml
Zásobní roztok C	0,07 ml
Zásobní roztok D	0,1 ml
Zásobní roztok E	0,8 ml
H_2O (v čistotě pro TK)	7,83 ml

Osmolarita byl měla být 265–285 mOsm.
Přidej 2,5 ml teplem inaktivovaného ovčího séra a sterilizuj filtrací.

Bahnice příjemkyně jsou ošetřeny intravaginální houbou napuštěnou progesteronem (Chronogest Ewe Sponge nebo Chronogest Ewe–Lamb Sponge, Intervet) ponechanou *in situ* deset nebo dvanáct dní. Bahnicím je v den odstranění houby (den –1) intramuskulárně injikováno 1,5 ml (300 m.j.) substituentu hormonu stimulujícího folikuly (P.M.S.G., Intervet) a 0,5 ml Luteolytického činidla (Estrumate, Coopers Pitman–Moore). Bahnice jsou testovány na estrus vazektomovaným beranem ve dnech 0 a 1 mezi 8. a 17. hodinou.

Embrya přežívající kultivaci *in vitro* jsou vrácena příjemkyním (hladovějícím od 17. hodiny den 5 nebo 6) v den 6 nebo 7. Přes embrya je prováděn v celkové anestezii, jak je popsáno výše. Děloha je vyňata z dutiny břišní laparotomickou incizí s nebo bez laparoskopie. Embrya jsou vrácena do jednoho nebo obou rohů děložních pouze bahnicím s alespoň jedním vhodným žlutým tělískem. Po vrácení dělohy na místo je břicho zavřeno a příjemkyním je umožněno se zotavit. Zvířatům jsou okamžitě před nebo po operaci podány intramuskulární injekce Amoxypenu L.A. v dávce doporučené výrobcem.

Jehňata jsou označena ušními visačkami a ponechána matkám na výchovu. Bahnice a jehňata jsou buď ustájena a krmena kompletními dietními koncentráty a dalšími doplňky anebo senem *ad libitum*, nebo jsou ponechána venku na pastvě.

Během prvního týdne života (nebo tak brzy, jak je možné bez ublížení na zdraví) je každé jehně testováno na přítomnost heterologní DNA dvěma odběry vzorků. Z jugulární žíly se odebírá 10 ml krve do podtlakové odběrové zkumavky s EDTA. Jsou-li zdravá, mají jehňata také druhý odběr 10 ml krve, který je odebraný během jednoho týdne od prvního. Vzorky tkáně se odebírají biopsií ocasu co nejdříve poté, co byl ocas znecitlivěn gumovým kroužkem v jeho proximální třetině (obvykle během 200 minut po „ocasení"). Tkáň je okamžitě umístěna v roztoku pufru pro ocasní biopsie. Vzorky ocasku jsou drženy v teplotě místnosti a analyzovány v den odběru. Všem jehňatům byla podána intramuskulární injekce Amoxypenu L.A. v dávce doporučené výrobcem a naříznutý konec ocasu byl ošetřen antibiotickým sprejem.

Při extrakci DNA z ovčí krve se nejdříve oddělí bílé krevní buňky. 10 ml vzorek krve se naředí ve 20 ml Hankova pufrovaného fyziologického roztoku (HBS, získaný na sigma Chemical Co.). Deset ml naředěné krve se navrství přes 5 ml Histopaque (Sigma) do dvou 15 ml zkumavek se šroubovacím uzávěrem. Zkumavky se odstřeďují v 3000 rpm (max. 2000 x g), s pomalým brzděním po 15 minut při teplotě místnosti. Mezi vrstvy bílých buněk jsou přeneseny do čisté 15 ml zkumavky a naředěny na 15 ml HBS. Naředěné buňky se stačí 10 minut ve 3000 rpm při teplotě místnosti a získaná peleta buněk je resuspendována ve 2 až 5 ml pufru pro ocasní biopsie.

Aby se z bílých buněk extrahovala DNA, přidá se k resuspendovaným buňkám 10% SDS do konečné 1% koncentrace a zkumavka se převrací, aby se roztok promíchal. Přidá se 1 mg čerstvého roztoku proteinázy K a směs se inkubuje přes noc ve 45 °C. DNA se extrahuje za použití roztoku fenolu a chloroformu (3x) o stejných objemech a roztokem chloroformu/izoamylalkoholu (1x). Poté se DNA precipituje přidáním 0,1 objemu 3 M NaAc a 2 objemů atanolu a zkumavka se převrací, aby se roztok promíchal. Precipitovaná DNA je navinuta na čistou skleněnou tyčinku se zataveným koncem. Navinutá DNA je omyta v 70% etanolu, ponechá s částečně uschnout a pak je znovu rozpuštěna v TE (10 mM Tris–HCl, 1 mM EDTA, pH 7,4).

Vzorky DNA z krve a ocasu jsou analyzovány metodou Southernova přenosu za použití sond na promotorovou oblast BLG a oblasti kódující řetězce fibrinogenu.

Z předchozího vyplývá, že ačkoliv zde po ilustrativní účely byla popsána specifická ztělesnění vynálezu, mohou být učiněny různé modifikace bez odchylky od ducha a rámce vynálezu. Proto vynález není omezen jinak než připojenými patentovými nároky.

Průmyslová využitelnost

V oblasti medicíny, lidské i veterinární trvá potřeba najít způsoby, které by produkovaly velká množství vysoce kvalitního fibrinogenu pro použití v tkáňových lepidlech a dalších aplikacích. Dále je potřebné mít fibrinogen, který neobsahuje krví přenášené patogenní zárodky. Záměrem předkládaného vynálezu je zajistit komerčně využitelná množství rekombinantního lidského fibrinogenu. Dalším cílem vynálezu je poskytnout materiály a způsoby pro expresi fibrinogenu v první tkáni transgenních zvířat, zejména hospodářských zvířat jak oje dobytek, ovce, prasata a kozy.

SEZNAM SEKVENCÍ

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 1:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 5943 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: dvojité
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: DNA (genomová)

(vii) PŘÍMÝ ZDROJ:

(B) KLON: Řetězce Aα lidského fibrinogenu

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: CDS
(B) POZICE: spoj. (31..84, 1154..1279, 1739..1922, 3055..3200, 3786..5210)

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 1:

```
GTCTAGGAGC CAGCCCCACC CTTAGAAAAG ATG TTT TCC ATG AGG ATC GTC TGC      54
                                 Met Phe Ser Met Arg Ile Val Cys
                                   1               5

CTA GTT CTA AGT GTG GTG GGC ACA GCA TGG GTATGGCCCT TTTCATTTTT      104
Leu Val Leu Ser Val Val Gly Thr Ala Trp
     10                  15

TCTTCTTGCT TTCTCTCTGG TGTTTATTCC ACAAAGAGCC TGGAGGTCAG AGTCTACCTG  164

CTCTATGTCC TGACACACTC TTAGCTTTAT GACCCCAGGC CTGGGAGGAA ATTTCCTGGG  224

TGGGCTTGAC ACCTCAAGAA TACAGGGTAA TATGACACCA AGAGGAAGAT CTTAGATGGA  284

TGAGAGTGTA CAACTACAAG GGAAACTTTA GCATCTGTCA TTCAGTCTTA CCACATTTTG  344

TTTTGTTTTG TTTTAAAAAG GGCAAGAATT ATTTGCCATC CTTGTACCTA TAAAGCCTTG  404

GTGCATTATA ATGCTAGTTA ATGGAATAAA ACATTTTATG GTAAGATTTG TTTTCTTTAG  464

TTATTAATTT CTTGCTACTT GTCCATAATA AGCAGAACTT TTAGTGTTAG TACAGTTTTG  524

CTGAAAGGTT ATTGTTGTGT TTGTCAAGAC AGAAGAAAAA GCAAACGAAT TATCTTTGGA  584

AATATCTTTG CAGTATCAGA AGAGATTAGT TAGTAAGGCA ATACGCTTTT CCGCAGTAAT  644
```

```
GGTATTCTTT TAAATTATGA ATCCATCTCT AAAGGTTACA TAGAAACTTG AAGGAGAGAG      704

GAACATTCAG TTAAGATAGT CTAGGTTTTT CTACTGAAGC AGCAATTACA GGAGAAAGAG      764

CTCTACAGTA GTTTTCAACT TTCTGTCTGC AGTCATTAGT AAAAATGAAA AGGTAAAATT      824

TAACTGATTT TATAGATTCA AATAATTTTC CTTTTAGGAT GGATTCTTTA AAACTCCTAA      884

TATTTATCAA ATGCTTATTT AAGTGTCACA CACAGTTAAG AAATTTGTAC ACCTTGTCTC      944

CTTTAATTCT CATAACAACT CCATAAAATG GGTCCTAGGA TTTCCATTTG AAGATAAGAA     1004

ACCTGAAGCT TGCCGAAGCC CTGTGTCTGC TCTCCTTAAT CTCTGTGAGA GTGCCATCTC     1064

TTCCTGGGGA CTTGTAGGCA TGCCACTGTC TCCTCTTCTG GCTAACATTG CTGTTGCTCT     1124

CTTTTGTGTA TGTGAATGAA TCTTTAAAG ACT GCA GAT AGT GGT GAA GGT GAC        1177
                               Thr Ala Asp Ser Gly Glu Gly Asp
                                   20                  25

TTT CTA GCT GAA GGA GGA GGC GTG CGT GGC CCA AGG GTT GTG GAA AGA        1225
Phe Leu Ala Glu Gly Gly Gly Val Arg Gly Pro Arg Val Val Glu Arg
            30                  35                  40

CAT CAA TCT GCC TGC AAA GAT TCA GAC TGG CCC TTC TGC TCT GAT GAA        1273
His Gln Ser Ala Cys Lys Asp Ser Asp Trp Pro Phe Cys Ser Asp Glu
        45                  50                  55

GAC TGG GTAAGCAGTC AGCGGGGGAA GCAGGAGATT CCTTCCCTCT GATGCTAGAG         1329
Asp Trp
    60

GGGCTCACAG GCTGACCTGA TTGGTCCCAG AAACTTTTTT AAATAGAAAA TAATTGAATA     1389

GTTACCTACA TAGCAAATAA AGAAAAGGAA CCTACTCCCA AGAGCACTGT TTATTTACCT     1449

CCCCAACTCT GGATCATTAG TGGGTGAACA GACAGGATTT CAGTTGCATG CTCAGGCAAA     1509

ACCAGGCTCC TGAGTATTGT GGCCTCAATT TCCTGGCACC TATTTATGGC TAAGTGGACC     1569

CTCATTCCAG AGTTTCTCTG CGACCTCTAA CTAGTCCTCT TACCTACTTT TAAGCCAACT     1629

TATCTGGAAG AGAAAGGGTA GGAAGAAATG GGGGCTGCAT GGAAACATGC AAAATTATTC     1689

TGAATCTGAG AGATAGATCC TTACTGTAAT TTTCTCCCTT CACTTTCAG AAC TAC         1744
                                                    Asn Tyr
```

```
AAA TGC CCT TCT GGC TGC AGG ATG AAA GGG TTG ATT GAT GAA GTC AAT      1792
Lys Cys Pro Ser Gly Cys Arg Met Lys Gly Leu Ile Asp Glu Val Asn
        65                  70                  75

CAA GAT TTT ACA AAC AGA ATA AAT AAG CTC AAA AAT TCA CTA TTT GAA      1840
Gln Asp Phe Thr Asn Arg Ile Asn Lys Leu Lys Asn Ser Leu Phe Glu
    80                  85                  90

TAT CAG AAG AAC AAT AAG GAT TCT CAT TCG TTG ACC ACT AAT ATA ATG      1888
Tyr Gln Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met
 95                 100                 105                 110

GAA ATT TTG AGA GGC GAT TTT TCC TCA GCC AAT  A GTAAGTATTA            1932
Glu Ile Leu Arg Gly Asp Phe Ser Ser Ala Asn
               115                 120

CATATTTACT TCTTTGACTT TATAACAGAA ACAACAAAAA TCCTAAATAA ATATGATATC    1992

CGCTTATATC TATGACAATT TCATCCCAAA GTACTTAGTG TAGAAACACA TACCTTCATA    2052

ATATCCCTGA AAATTTTAAG AGGGAGCTTT TGTTTTCGTT ATTTTTTCAA AGTAAAAGAT    2112

GTTAACTGAG ATTGTTTAAG GTCACAAAAT AAGTCAGAAT TTTGGATTAA AACAAGAATT    2172

TAAATGTGTT CTTTTCAACA GTATATACTG AAAGTAGGAT GGGTCAGACT CTTTGAGTTG    2232

ATATTTTTGT TTCTGCTTTG TAAAGGTGAA AACTGAGAGG TCAAGGAACT TGTTCAAAGA    2292

CACAGAGCTG GGAATTCAAC TCCCAGACTC CACTGAGCTG ATTAGGTAGA TTTTTAAATT    2352

TAAAATATAG GGTCAAGCTA CGTCATTCTC ACAGTCTACT CATTAGGGTT AGGAAACATT    2412

GCATTCACTC TGGGCATGGA CAGCGAGTCT AGGGAGTCCT CAGTTTCTCA AGTTTTGCTT    2472

TGCCTTTTTA CACCTTCACA AACACTTGAC ATTTAAAATC AGTGATGCCA ACACTAGCTG    2532

GCAAGTGAGT GATCCTGTTG ACCCAAAACA GCTTAGGAAC CATTTCAAAT CTATAGAGTT    2592

AAAAAGAAAA GCTCATCAGT AAGAAAATCC AATATGTTCA AGTCCCTTGA TTAAGGATGT    2652

TATAAAATAA TTGAAATGCA ATCAAACCAA CTATTTTAAC TCCAAATTAC ACCTTTAAAA    2712

TTCCAAAGAA AGTTCTTCTT CTATATTTCT TTGGGATTAC TAATTGCTAT TAGGACATCT    2772

TAACTGGCAT TCATGGAAGG CTGCAGGGCA TAACATTATC CAAAAGTCAA ATGCCCCATA    2832
```

```
GGTTTTGAAC TCACAGATTA AACTGTAACC AAAATAAAAT TAGGCATATT TACAAGCTAG   2892

TTTCTTTCTT TCTTTTTTCT CTTTCTTTCT TTCTTTCTTT CTTTCTTTCT TTCTTTCTTT   2952

CTTTCTTTCT TTCTCCTTCC TTCCTTTCTT CCTTTCTTTT TTGCTGGCAA TTACAGACAA   3012

ATCACTCAGC AGCTACTTCA ATAACCATAT TTTCGATTTC AG  AC CGT GAT AAT       3065
                                                  Asn Arg Asp Asn
                                                      125

ACC TAC AAC CGA GTG TCA GAG GAT CTG AGA AGC AGA ATT GAA GTC CTG      3113
Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu
                130                 135                 140

AAG CGC AAA GTC ATA GAA AAA GTA CAG CAT ATC CAG CTT CTG CAG AAA      3161
Lys Arg Lys Val Ile Glu Lys Val Gln His Ile Gln Leu Leu Gln Lys
            145                 150                 155

AAT GTT AGA GCT CAG TTG GTT GAT ATG AAA CGA CTG GAG GTAAGTATGT      3210
Asn Val Arg Ala Gln Leu Val Asp Met Lys Arg Leu Glu
        160                 165                 170

GGCTGTGGTC CCGAGTGTCC TTGTTTTTGA GTAGAGGGAA AAGGAAGGCG ATAGTTATGC   3270

ACTGAGTGTC TACTATATGC AGAGAAAAGT GTTATATCCA TCATCTACCT AAAAGTAGGT   3330

ATTATTTTCC TCACTCCACA GTTGAAGAAA AAAAAATTCA GAGATATTAA GTAAATTTTC   3390

CAACGTACAT AGATAGTAAT TCAAAGCAAT GTTCAGTCCC TGTCTATTCC AAGCCATTAC   3450

ATCACCACAC CTCTGAGCCC TCAGCCTGAG TTCACCAAGG ATCATTTAAT TAGCGTTTCC   3510

TTTGAGAGGG AATAGCACCT TACTCTTGAT CCATTCTGAG GCTAAGATGA ATTAAACAGC   3570

ATCCATTGCT TATCCTGGCT AGCCCTGCAA TACCCAACAT CTCTTCCACT GAGGGTGCTC   3630

GATAGGCAGA AAACAGAGAA TATTAAGTGG TAGGTCTCCG AGTCAAAAAA AATGAAACCA   3690

GTTTCCAGAA GGAAAATTAA CTACCAGGAA CTCAATAGAC GTAGTTTATG TATTTGTATC   3750

TACATTTTCT CTTTATTTTT CTCCCCTCTC TCTAG GTG GAC ATT GAT ATT AAG      3803
                                        Val Asp Ile Asp Ile Lys
                                                175
```

```
ATC CGA TCT TGT CGA GGG TCA TGC AGT AGG GCT TTA GCT CGT GAA GTA      3851
Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val
            180                 185                 190

GAT CTG AAG GAC TAT GAA GAT CAG CAG AAG CAA CTT GAA CAG GTC ATT      3899
Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile
        195                 200                 205

GCC AAA GAC TTA CTT CCC TCT AGA GAT AGG CAA CAC TTA CCA CTG ATA      3947
Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile
    210                 215                 220

AAA ATG AAA CCA GTT CCA GAC TTG GTT CCC GGA AAT TTT AAG AGC CAG      3995
Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln
225                 230                 235                 240

CTT CAG AAG GTA CCC CCA GAG TGG AAG GCA TTA ACA GAC ATG CCG CAG      4043
Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln
                245                 250                 255

ATG AGA ATG GAG TTA GAG AGA CCT GGT GGA AAT GAG ATT ACT CGA GGA      4091
Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly
            260                 265                 270

GGC TCC ACC TCT TAT GGA ACC GGA TCA GAG ACG GAA AGC CCC AGG AAC      4139
Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn
        275                 280                 285

CCT AGC AGT GCT GGA AGC TGG AAC TCT GGG AGC TCT GGA CCT GGA AGT      4187
Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser
    290                 295                 300

ACT GGA AAC CGA AAC CCT GGG AGC TCT GGG ACT GGA GGG ACT GCA ACC      4235
Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr
305                 310                 315                 320

TGG AAA CCT GGG AGC TCT GGA CCT GGA AGT GCT GGA AGC TGG AAC TCT      4283
Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser
                325                 330                 335

GGG AGC TCT GGA ACT GGA AGT ACT GGA AAC CAA AAC CCT GGG AGC CCT      4331
Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro
            340                 345                 350

AGA CCT GGT AGT ACC GGA ACC TGG AAT CCT GGC AGC TCT GAA CGC GGA      4379
Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly
        355                 360                 365
```

```
AGT GCT GGG CAC TGG ACC TCT GAG AGC TCT GTA TCT GGT AGT ACT GGA      4427
Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly
    370                 375                 380

CAA TGG CAC TCT GAA TCT GGA AGT TTT AGG CCA GAT AGC CCA GGC TCT      4475
Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser
385                 390                 395                 400

GGG AAC GCG AGG CCT AAC AAC CCA GAC TGG GGC ACA TTT GAA GAG GTG      4523
Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val
                405                 410                 415

TCA GGA AAT GTA AGT CCA GGG ACA AGG AGA GAG TAC CAC ACA GAA AAA      4571
Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys
            420                 425                 430

CTG GTC ACT TCT AAA GGA GAT AAA GAG CTC AGG ACT GGT AAA GAG AAG      4619
Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys
        435                 440                 445

GTC ACC TCT GGT AGC ACA ACC ACC ACG CGT CGT TCA TGC TCT AAA ACC      4667
Val Thr Ser Gly Ser Thr Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr
    450                 455                 460

GTT ACT AAG ACT GTT ATT GGT CCT GAT GGT CAC AAA GAA GTT ACC AAA      4715
Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys
465                 470                 475                 480

GAA GTG GTG ACC TCC GAA GAT GGT TCT GAC TGT CCC GAG GCA ATG GAT      4763
Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
                485                 490                 495

TTA GGC ACA TTG TCT GGC ATA GGT ACT CTG GAT GGG TTC CGC CAT AGG      4811
Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
            500                 505                 510

CAC CCT GAT GAA GCT GCC TTC TTC GAC ACT GCC TCA ACT GGA AAA ACA      4859
His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
        515                 520                 525

TTC CCA GGT TTC TTC TCA CCT ATG TTA GGA GAG TTT GTC AGT GAG ACT      4907
Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
    530                 535                 540
```

```
GAG TCT AGG GGC TCA GAA TCT GGC ATC TTC ACA AAT ACA AAG GAA TCC      4955
Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
545                 550                 555                 560

AGT TCT CAT CAC CCT GGG ATA GCT GAA TTC CCT TCC CGT GGT AAA TCT      5003
Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
                565                 570                 575

TCA AGT TAC AGC AAA CAA TTT ACT AGT AGC ACG AGT TAC AAC AGA GGA      5051
Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
            580                 585                 590

GAC TCC ACA TTT GAA AGC AAG AGC TAT AAA ATG GCA GAT GAG GCC GGA      5099
Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
        595                 600                 605

AGT GAA GCC GAT CAT GAA GGA ACA CAT AGC ACC AAG AGA GGC CAT GCT      5147
Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
    610                 615                 620

AAA TCT CGC CCT GTC AGA GGT ATC CAC ACT TCT CCT TTG GGG AAG CCT      5195
Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
625                 630                 635                 640

TCC CTG TCC CCC TAGACTAAGT TAAATATTTC TGCACAGTGT TCCCATGGCC          5247
Ser Leu Ser Pro
                645

CCTTGCATTT CCTTCTTAAC TCTCTGTTAC ACGTCATTGA AACTACACTT TTTTGGTCTG    5307

TTTTTGTGCT AGACTGTAAG TTCCTTGGGG GCAGGGCCTT TGTCTGTCTC ATCTCTGTAT    5367

TCCCAAATGC CTAACAGTAC AGAGCCATGA CTCAATAAAT ACATGTTAAA TGGATGAATG    5427

AATTCCTCTG AAACTCTATT TGAGCTTATT TAGTCAAATT CTTTCACTAT TCAAAGTGTG    5487

TGCTATTAGA ATTGTCACCC AACTGATTAA TCACATTTTT AGTATGTGTC TCAGTTGACA    5547

TTTAGGTCAG GCTAAATACA AGTTGTGTTA GTATTAAGTG AGCTTAGCTA CCTGTACTGG    5607

TTACTTGCTA TTAGTTTGTG CAAGTAAAAT TCCAAATACA TTTGAGGAAA ATCCCCTTTG    5667

CAATTTGTAG GTATAAATAA CCGCTTATTT GCATAAGTTC TATCCCACTG TAAGTGCATC    5727

CTTTCCCTAT GGAGGGAAGG AAAGGAGGAA GAAAGAAAGG AAGGGAAAGA AACAGTATTT    5787

GCCTTATTTA ATCTGAGCCG TGCCTATCTT TGTAAAGTTA AATGAGAATA ACTTCTTCCA    5847

ACCAGCTTAA TTTTTTTTTT AGACTGTGAT GATGTCCTCC AAACACATCC TTCAGGTACC    5907

CAAAGTGGCA TTTTCAATAT CAAGCTATCC GGATCC                              5943
```

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 2:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 644 aminokyselin
(B) TYP: aminokyselina
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: protein

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 2:

```
Met Phe Ser Met Arg Ile Val Cys Leu Val Leu Ser Val Val Gly Thr
 1               5                  10                  15

Ala Trp Thr Ala Asp Ser Gly Glu Gly Asp Phe Leu Ala Glu Gly Gly
            20                  25                  30

Gly Val Arg Gly Pro Arg Val Val Glu Arg His Gln Ser Ala Cys Lys
        35                  40                  45

Asp Ser Asp Trp Pro Phe Cys Ser Asp Glu Asp Trp Asn Tyr Lys Cys
    50                  55                  60

Pro Ser Gly Cys Arg Met Lys Gly Leu Ile Asp Glu Val Asn Gln Asp
65                  70                  75                  80

Phe Thr Asn Arg Ile Asn Lys Leu Lys Asn Ser Leu Phe Glu Tyr Gln
                85                  90                  95

Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met Glu Ile
            100                 105                 110

Leu Arg Gly Asp Phe Ser Ser Ala Asn Asn Arg Asp Asn Thr Tyr Asn
        115                 120                 125

Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu Lys Arg Lys
    130                 135                 140

Val Ile Glu Lys Val Gln His Ile Gln Leu Leu Gln Lys Asn Val Arg
145                 150                 155                 160
```

Ala Gln Leu Val Asp Met Lys Arg Leu Glu Val Asp Ile Asp Ile Lys
165 170 175

Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val
180 185 190

Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile
195 200 205

Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile
210 215 220

Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln
225 230 235 240

Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln
245 250 255

Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly
260 265 270

Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn
275 280 285

Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser
290 295 300

Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr
305 310 315 320

Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser
325 330 335

Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro
340 345 350

Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly
355 360 365

Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly
370 375 380

Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser
385 390 395 400

Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val
405 410 415

Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys
420 425 430

Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys
435 440 445

Val Thr Ser Gly Ser Thr Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr
450 455 460

Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys
465 470 475 480

Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
485 490 495

Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
500 505 510

His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
515 520 525

Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
530 535 540

Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
545 550 555 560

Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
565 570 575

Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
580 585 590

Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
595 600 605

Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
610 615 620

Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
625 630 635 640

Ser Leu Ser Pro

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 3:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 8878 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: dvojité
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: DNA (genomová)

(vii) PŘÍMÝ ZDROJ:

(B) KLON: Řetězec Bβ lidského fibrinogenu

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: misc–RNA
(B) POZICE: 1...469

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 470...583

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 584...3257

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 3258...3449

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 3450...3938

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 3939...4122

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 4123...5042

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 5043...5270

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 5271...5830

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 5831...5944

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 5945...6632

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 6633...6758

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 6759...6966

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 6967...7252

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: intron
(B) POZICE: 7253...7870

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: exon
(B) POZICE: 7871...8102

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: 3'UTR
(B) POZICE: 8103...8537

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: misc-RNA
(B) POZICE: 8538...8878

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: CDS
(B) POZICE: spoj (470...583, 3258...3449, 3939..4122, 5043...5270, 5831..5944, 6633..6758, 6967..7252, 7871..8102).

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 3:

```
GAATTCATGC CCCTTTTGAA ATAGACTTAT GTCATTGTCA GAAAACATAA GCATTTATGG      60

TATATCATTA ATGAGTCACG ATTTTAGTGG TTGCCTTGTG AGTAGGTCAA ATTTACTAAG     120

CTTAGATTTG TTTTCTCACA TATTCTTTCG GAGCTTGTGT AGTTTCCACA TTAATTTACC     180

AGAAACAAGA TACACACTCT CTTTGAGGAG TGCCCTAACT TCCCATCATT TTGTCCAATT     240

AAATGAATTG AAGAAATTTA ATGTTTCTAA ACTAGACCAA CAAAGAATAA TAGTTGTATG     300

ACAAGTAAAT AAGCTTTGCT GGGAAGATGT TGCTTAAATG ATAAAATGGT TCAGCCAACA     360

AGTGAACCAA AAATTAAATA TTAACTAAGG AAAGGTAACC ATTTCTGAAG TCATTCCTAG     420

CAGAGGACTC AGATATATAT AGGATTGAAG ATCTCTCAGT TAAGTCTAC ATG AAA          475
                                                        Met Lys
                                                          1

AGG ATG GTT TCT TGG AGC TTC CAC AAA CTT AAA ACC ATG AAA CAT CTA        523
Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys His Leu
          5                  10                  15

TTA TTG CTA CTA TTG TGT GTT TTT CTA GTT AAG TCC CAA GGT GTC AAC        571
Leu Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Gln Gly Val Asn
     20                  25                  30

GAC AAT GAG GAG GTGAATTTTT TAAAGCATTA TTATATTATT AGTAGTATTA            623
Asp Asn Glu Glu
 35

TTAATATAAG ATGTAACATA ATCATATTAT GTGCTTATTT TAATGAAATT AGCATTGCTT     683

ATAGTTATGA AATGGAATTG TTAACCTCTG ACTTATTGTA TTTAAAGAAT GTTTCATAGT     743

ATTTCTTATA TAAAAACAAA GTAATTTCTT GTTTTCTAGT TTATCACCTT TGTTTTCTTA     803

AGATGAGGAT GGCTTAGCTA ATGTAAGATG TGTTTTTCTC ACTTGCTATT CTGAGTACTG     863

TGATTTTCAT TTACTTCTAG CAATACAGGA TTACAATTAA GAGGACAAGA TCTGAAAATC     923

TCACAAACTA TAAAATAATA AAAGAGCAGA ATTTTAAGAT AAAAGAAACT GGTGGTAGGT     983

AGATTGTTCT TTGGTGAAGG AAGGTAATAT ATATTGTTAC TGAGATTACT ATTTATAAAA     1043

ATTATAACTA AGCCTAAAAG CAAAATACAT CAAGTGTAAT GATAGAAAAT GAAATATTGC     1103
```

TTTTTTCAGA TGAAAAGTTC AAATTAGAGT TAGTGTGTAT TGTTATTATT AATAGTTATG 1163

AAACACGGTT CAGTCTAATT TATTTATTTG TAGAACAGTT TGTCCTCAAC TATTATTTTT 1223

GCTGACTTAT TGCTGTTAAT TTGCAGTTAC TAAAAATACA GAAATGCATT TAGGACAATG 1283

GATATTTAAG AAATTTAAAT TTTATCATCA AACGTATCAT GGCCAAATTT CTTACATATA 1343

GCATAGTATC ATTAAACTAG AAATAAGAAT ACACAATAAT ATTTAAATGA AGTGATTCAT 1403

TTCGGATCAT TATTGAGTTT CAAGGGAACT TGAGTGTTGT ACTTATCAGA CTCTACATGT 1463

AAGAACATAT AGTTAATCTG GTTGTGTGTG TAAAAACATA TGGTTAATCT GGTTAAGTCT 1523

GGTTAATCAT ATTAGGTAAG AAAAATGTAA AGAATGTGTA AGACGAAATT TTTGTAAAGT 1583

ACTCTGCAAA GCACTTTCAC ATTTCTGCTT ATCAACTAAA CCTCACAGAG ATAGTTTAAT 1643

AGTTTAGGCT TTAAAATGGA TTTTGATTAT TCAACAAGTG GCCTTCATAA TTTCTTTAAG 1703

TGTTTTTCTT TAAGTATATA CTTTCTTTAA ATATTTTTTA AAATTTCCTT TTCTCTAGTA 1763

AAGCCAGACC ATCCATGCTA CCTCTCTAGT GGCACTCTGA AATAAAAAGA AAATAGTTTT 1823

CTCTGTTATA ATTGTATTTG TAATAAGCAG ATGAATCACA TTTCTTAAAA TTTGTTTTAG 1883

AGAGGGTAAG CTCTGACTAG GACCATGACT TCAATGTGAA ATATGTATAT ATCCTCCGAA 1943

TCTTTACATA TTAAGAATGT ATATAGTCAA CTGGTTAAAC AGGAAAATCT GGAACAGCCT 2003

GGCTGGGTTT TAATCTTAGC ACCATCCTAC TAAATGTTAA ATAATATTAT AATCTAATGA 2063

ATAAATGACA ATGCAATTCC AAATAGAGTT CATCTGATGA CTTCTAGACT CACAAAATTG 2123

CAAGAGAGCT CAGTTGTTGC TCAGTTGTTC CAAATCATGT CGTTTGTTAA TTTGTAATTA 2183

AGCTCCAAAG GATGTATAGC TACTGACAAA AAAAAAAATG AGAATGTAGT TAATCCAAAT 2243

CAAAACTTTC CTATTGCAAT GCGTATTTTC TGCTTCATTA TCCTTTAATA TAATATTTTA 2303

AGTTAGCAAG TAATTTTAAT TACAATGCAC AAGCCTTGAG AATTATTTTA AATATAAGAA 2363

AATCATAATG TTTGATAAAG AAATCATGTA AGAAATTTCA AGATAATGGT TTAACAAATA 2423

ATTTTGTTGA TAGAAGATAA GACTAAAAGT GAAATTCGAA GTGGAGAGGA CACTTAAACT 2483

GTAGTACTTG TTATGTGTGA TTCCAGTAAA AATAGTAATG AGCACTTATT ATTGCCAAGT 2543

```
ACTGTTCTGA GGGTACCATA TGCAATAAGT TATTTAATCC TTACAATAAT CTTGTAAGGC    2603

AGATTCAAAC TATCATTACA CTTATTTTAC AGATGAGAAA ACTGGGGCAC AGATAAAGCA    2663

ACTTGCCCAA GGTCTCATAG CTGTAAGTCA ACCCTACGGT CAAGACCTAC AAGTAGCCGA    2723

GCTCCAGAGT ACATTATGAG GGTCAAAGAT TGTCTTATTA CAAATAAATT CCAAGTAGAA    2783

TCAACCTTTA ATAAGTCTTT AATGTCTCTT AAATATGTTT ATATAGGAGT CTAATCACCA    2843

ATTCACAAAA ATGAAAGTAG GGAAATGATT AACAATAATC ATAGGAATCT AACAATCCAA    2903

GTGGCTTGAG AATATTCATT CTTCTTGACA GTATAGATTC TTTACAATTT CGTAAGTTCC    2963

AATGTATGTT TTAGGAATAT GAGGTCATTA CTATTCATAA TCTGATACAG CTTTATCCTA    3023

AGGCCTCTCT TTAAAAACTA CACTGCATCA TAGCTTTTTT GTGCAGTTGG TCTTTCTACT    3083

GTTACTGAAC AGTAAGCAAC CTACAGATTC ACTATCACCA ACCAGCCAGT TGATGGATCT    3143

TAAGCAAATT ATCAAGCTTG TGATAACCTA AATTATAAAA TGAGGGTGTT GGAATAGTTA    3203

CATTCCAAAT CTTCTATAAC ACTCTGTATT ATATTTCTGC CTCATTCCTT GTAG GGT      3260
                                                              Gly

TTC TTC AGT GCC CGT GGT CAT CGA CCC CTT GAC AAG AAG AGA GAA GAG      3308
Phe Phe Ser Ala Arg Gly His Arg Pro Leu Asp Lys Lys Arg Glu Glu
 40                  45                  50                  55

GCT CCC AGC CTG AGG CCT GCC CCA CCG CCC ATC AGT GGA GGT GGC TAT      3356
Ala Pro Ser Leu Arg Pro Ala Pro Pro Pro Ile Ser Gly Gly Gly Tyr
                60                  65                  70

CGG GCT CGT CCA GCC AAA GCA GCT GCC ACT CAA AAG AAA GTA GAA AGA      3404
Arg Ala Arg Pro Ala Lys Ala Ala Ala Thr Gln Lys Lys Val Glu Arg
            75                  80                  85

AAA GCC CCT GAT GCT GGA GGC TGT CTT CAC GCT GAC CCA GAC CTG          3449
Lys Ala Pro Asp Ala Gly Gly Cys Leu His Ala Asp Pro Asp Leu
        90                  95                  100

GTGGGTGCAC TGATGTTTCT TGCAGTGGTG GCTCTCTCAT GCAGAGAAAG CCTGTAGTCA    3509

TGGCAGTCTG CTAATGTTTC ACTGACCCAC ATTACCATCA CTGTTATTTT GTTTGTTTAT    3569
```

TTTGGAAATA AAATTCAAAA CATAAACATA TTGGGCCTTT GGTTTAGGCT TTCTTTCTTG 3629

TTTTCTTTGG TCTGGGCCCA AAATTTCAAA TTAGGATATG TGGGTGCCAC CTTTCCATTT 3689

GTATTTTGCC ACTGCCTTTG TTTAGTTGGT AAAATTTTCA TAGCCCAATT ATATTTTTTC 3749

TGGGGTAAGT AATATTTTAA ATCTCTATGA GAGTATGATG ATGACTTTCG AATTTCTGGT 3809

CTTACAGAAA ACCAAATAAT AAATTTTTAT GTTGGCTAAT CGTATCGCTG AATTTTCCTA 3869

TGTGCTATTT TAACAAATGT CCATGACCCA AATCCTTCAT CTAATGCCTG CTATTTTCTT 3929

TGTTTTTAG GGG GTG TTG TGT CCT ACA GGA TGT CAG TTG CAA GAG GCT 3977
Gly Val Leu Cys Pro Thr Gly Cys Gln Leu Gln Glu Ala
105 110 115

TTG CTA CAA CAG GAA AGG CCA ATC AGA AAT AGT GTT GAT GAG TTA AAT 4025
Leu Leu Gln Gln Glu Arg Pro Ile Arg Asn Ser Val Asp Glu Leu Asn
120 125 130

AAC AAT GTG GAA GCT GTT TCC CAG ACC TCC TCT TCT TCC TTT CAG TAC 4073
Asn Asn Val Glu Ala Val Ser Gln Thr Ser Ser Ser Ser Phe Gln Tyr
135 140 145

ATG TAT TTG CTG AAA GAC CTG TGG CAA AAG AGG CAG AAG CAA GTA AAA G 4122
Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys Gln Val Lys
150 155 160

GTAGATATCC TTGTGCTTTC CATTCGATTT TCAGCTATAA AATTGGAACC GTTAGACTGC 4182

CACGAGAATG CATGGTTGTG AGAAGATTAA CATTTCTGGG TTAGTGAATA GCATTCATAC 4242

GCTTTTGGGC ACCTTCCCCT GCAACTTGCC AGATAAGCAC TATTCAGCTC TTATTCCCAG 4302

TCTGACATCA GCAAGTGTGA TTTTCTATGA AAAATTCTAC TATGACTCCT TATTTTAAGT 4362

ATACAAGAAA CTTGTGACTC AGAAGATAAT ATTTACAGAG TGGAAAAAAA CCCCTAGCAT 4422

TTATAGTTTT AACATTTGAG GTTTTGAATG AGAGAGTTAT CCATAATATA TTCAATTGTG 4482

TTGTGGATAA TGACACCTAA CCTGTGAATC TTGAGGTCAG AATGTTGAGT GCTGTTGACT 4542

TGGTGGTCAG GAAACAGCTA GTGCGTGAGC CTGGCACAGG CATCTCAGTG AGTAGCATAC 4602

CCACAGTTGG AAATTTTTCA AAGAAATCAA AGGAATCATG ACATCTTATA AATTTCAAGG 4662

TTCTGCTATA CTTATGTGAA ATGGATAAAT AAATCAAGCA TATCCACTCT GTAAGATTGA 4722

```
ACTTCTCAGA TGGAAGACCC CAATACTGCT TTCTCCTCTT TTCCCTCACC AAAGAAATAA    4782

ACAACCTATT TCATTTATTA CTGGACACAA TCTTTAGCGT ATACCTATGG TAAATTACTA    4842

GTATGGTGGT TAGGATTTAT GTTAATTTGT ATATGTCATG CGCCAAATCA TTTCCACTAA    4902

ATATGACTAT ATATCATAAC TGCTTGGTGA TAGCTCAGTG TTTAATAGTT TATTCTCAGA    4962

AAATCAAAAT TGTATAGTTA AATACATTAG TTTTATGAGG CAAAAATGCT AACTATTTCT    5022

ACATAATTTC ATTTTTCCAG  AT AAT GAA AAT GTA GTC AAT GAG TAC TCC          5071
                        Asp Asn Glu Asn Val Val Asn Glu Tyr Ser
                            165                 170

TCA GAA CTG GAA AAG CAC CAA TTA TAT ATA GAT GAG ACT GTG AAT AGC      5119
Ser Glu Leu Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser
    175                 180                 185

AAT ATC CCA ACT AAC CTT CGT GTG CTT CGT TCA ATC CTG GAA AAC CTG      5167
Asn Ile Pro Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu
190                 195                 200                 205

AGA AGC AAA ATA CAA AAG TTA GAA TCT GAT GTC TCA GCT CAA ATG GAA      5215
Arg Ser Lys Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu
                210                 215                 220

TAT TGT CGC ACC CCA TGC ACT GTC AGT TGC AAT ATT CCT GTG GTG TCT      5263
Tyr Cys Arg Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser
            225                 230                 235

GGC AAA  G GTAACTGATT CATAAACATA TTTTTAGAGA GTTCCAGAAG AACTCACACA    5320
Gly Lys

CCAAAAATAA GAGAACAACA ACAACAACAA AAATGCTAAG TGGATTTTCC CAACAGATCA    5380

TAATGACATT ACAGTACATC ATAAAAATAT CCTTAGCCAG TTGTGTTTTG GACTGGCCTG    5440

GTGCATTTGC TGGTTTTGAT GAGCAGGATG GGGCACAGGT AGTCCCAGGG GTGGCTGATG    5500

TGTGCATCTG CGTACTGGCT TGAACAGATG GCAGAACCAC AGATAGATGT AGAAGTTTCT    5560

CCATTTTGTG TGTTCTGGGA GCTCATGGAT ATTCCAGGAC ACAAAAGGTG GAGAAGAGCT    5620

TTGTTCATCC TCTTAGCAGA TAAACGTCCT CAAAACTGGG TTGGACTTAC TAAAGTAAAA    5680
```

```
TGAAAATCTA ATATTTGTTA TATTATTTTC AAAGGTCTAT AATAACACAC TCCTTAGTAA    5740

CTTATGTAAT GTTATTTTAA AGAATTGGTG ACTAAATACA AAGTAATTAT GTCATAAACC    5800

CCTGAACATA ATGTTGTCTT ACATTTGCAG  AA TGT GAG GAA ATT ATC AGG AAA      5853
                                    Glu Cys Glu Glu Ile Ile Arg Lys
                                    240                 245

GGA GGT GAA ACA TCT GAA ATG TAT CTC ATT CAA CCT GAC AGT TCT GTC      5901
Gly Gly Glu Thr Ser Glu Met Tyr Leu Ile Gln Pro Asp Ser Ser Val
        250                 255                 260

AAA CCG TAT AGA GTA TAC TGT GAC ATG AAT ACA GAA AAT GGA  G           5944
Lys Pro Tyr Arg Val Tyr Cys Asp Met Asn Thr Glu Asn Gly
    265                 270                 275

GTAAGCTTTC GACAGTTGTT GACCTGTTGA TCTGTAATTA TTTGGATACC GTAAAATGCC    6004

AGGAAACAAG GCCAGGTGTG GTGGCTCATA CCTGTAATTC CAGCACCTTG GGAGGCCAAA    6064

GTGGGCTGAT AGCTTGAGCC TAGGAGTTTG AAACTAGCCT GGGCAACATA ATGAGACCCT    6124

AACTCTACAA AAAAAAAAAA AATACCAAAA AAAAAAAAAA AATCAGCTGT GTTGGTAGTA    6184

TGTGCCTGTA GTCCCAGCTA TCCAGGAGGC TGAGATGGGA GATCACCTGA GCCCACAACC    6244

TGGAGTCTTG ATCATGCTAC TGAACTGTAG CCTGGGCAAC AGAGGATAGT GAGATCCTGT    6304

CTCAAAAAAA AAAATTAATT AAAAAGCCAG GAAACAAGAC TTAGCTCTAA CATCTAACAT    6364

AGCTGACAAA GGAGTAATTT GATGTGGAAT TCAACCTGAT ATTTAAAAGT TATAAAATAT    6424

CTATAATTCA CAATTTGGGG TAAGATAAAG CACTTGCAGT TTCCAAAGAT TTTACAAGTT    6484

TACCTCTCAT ATTTATTTCC TTATTGTGTC TATTTTAGAG CACCAAATAT ATACTAAATG    6544

GAATGGACAG GGGATTCAGA TATTATTTTC AAAGTGACAT TATTTGCTGT TGGTTAATAT    6604

ATGCTCTTTT TGTTTCTGTC AACCAAAG  GA TGG ACA GTG ATT CAG AAC CGT        6655
                                  Gly Trp Thr Val Ile Gln Asn Arg
                                  280                 285

CAA GAC GGT AGT GTT GAC TTT GGC AGG AAA TGG GAT CCA TAT AAA CAG      6703
Gln Asp Gly Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln
            290                 295                 300
```

```
GGA TTT GGA AAT GTT GCA ACC AAC ACA GAT GGG AAG AAT TAC TGT GGC      6751
Gly Phe Gly Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly
            305                 310                 315

CTA CCA  G GTAACGAACA GGCATGCAAA ATAAAATCAT TCTATTTGAA ATGGGATTTT   6808
Leu Pro

TTTTAATTAA AAAACATTCA TTGTTGGAAG CCTGTTTTAG GCAGTTAAGA GGAGTTTCCT   6868

GACAAAAATG TGGAAGCTAA AGATAAGGGA AGAAAGGCAG TTTTTAGTTT CCCAAAATTT   6928

TATTTTTGGT GAGAGATTTT ATTTTGTTTT TCTTTTAG  GT GAA TAT TGG CTT       6980
                                          Gly Glu Tyr Trp Leu
                                          320

GGA AAT GAT AAA ATT AGC CAG CTT ACC AGG ATG GGA CCC ACA GAA CTT      7028
Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly Pro Thr Glu Leu
325                 330                 335                 340

TTG ATA GAA ATG GAG GAC TGG AAA GGA GAC AAA GTA AAG GCT CAC TAT      7076
Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val Lys Ala His Tyr
                345                 350                 355

GGA GGA TTC ACT GTA CAG AAT GAA GCC AAC AAA TAC CAG ATC TCA GTG      7124
Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr Gln Ile Ser Val
            360                 365                 370

AAC AAA TAC AGA GGA ACA GCC GGT AAT GCC CTC ATG GAT GGA GCA TCT      7172
Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met Asp Gly Ala Ser
        375                 380                 385

CAG CTG ATG GGA GAA AAC AGG ACC ATG ACC ATT CAC AAC GGC ATG TTC      7220
Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His Asn Gly Met Phe
    390                 395                 400

TTC AGC ACG TAT GAC AGA GAC AAT GAC GGC  TG  GTATGTGTGG              7262
Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly  Trp
405                 410                 415

CACTCTTTGC TCCTGCTTTA AAAATCACAC TAATATCATT ACTCAGAATC ATTAACAATA   7322

TTTTTAATAG CTACCACTTC CTGGGCACTT ACTGTCAGCC ACTGTCCTAA GCTCTTTATG   7382

CATCACTCGA AAGCATTTCA ACTATAAGGT AGACATTCTT ATTCTCATTT TACAGATGAG   7442

ATTTAGAGAG ATTACGTGAT TTGTCCAATG TCACACAACT ACCCAGAGAT AAAACTAGAA   7502
```

```
TTTGAGCACA GTTACTTTCT GAATAATGAG CATTTAGATA AATACCTATA TCTCTATATT    7562

CTAAAGTGTG TGTGAAAACT TTCATTTTCA TTTCCAGGGT TCTCTGATAC TAAGGGTTGT    7622

AAAAGCTATT ATTCCAGTAT AAAGTAACAA ACACAGTCCC TAGATGGATT GCCACAAAGG    7682

CCCAGTTATC TCTCTTTCTT GCTATAGGGC ACAGGAGGTC TTTGGTGTAT TAGTGTGACT    7742

CTATGTATAG CACCCAAAGG AAAGACTACT GTGCACACGA GTGTAGCAGT CTTTTATGGG    7802

TAATCTGCAA AACGTAACTT GACCACCGTA GTTCTGTTTC TAATAACGCC AAACACATTT    7862

TCTTTCAG G TTA ACA TCA GAT CCC AGA AAA CAG TGT TCT AAA GAA GAC        7910
          Leu Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp
                          420                 425

GGT GGT GGA TGG TGG TAT AAT AGA TGT CAT GCA GCC AAT CCA AAC GGC      7958
Gly Gly Gly Trp Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly
    430                 435                 440

AGA TAC TAC TGG GGT GGA CAG TAC ACC TGG GAC ATG GCA AAG CAT GGC      8006
Arg Tyr Tyr Trp Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly
445                 450                 455                 460

ACA GAT GAT GGT GTA GTA TGG ATG AAT TGG AAG GGG TCA TGG TAC TCA      8054
Thr Asp Asp Gly Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser
                465                 470                 475

ATG AGG AAG ATG AGT ATG AAG ATC AGG CCC TTC TTC CCA CAG CAA TAGTCCCCAA
8109
Met Arg Lys Met Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln
            480                 485                 490

TACGTAGATT TTTGCTCTTC TGTATGTGAC AACATTTTTG TACATTATGT TATTGGAATT    8169

TTCTTTCATA CATTATATTC CTCTAAAACT CTCAAGCAGA CGTGAGTGTG ACTTTTTGAA    8229

AAAAGTATAG GATAAATTAC ATTAAAATAG CACATGATTT TCTTTTGTTT TCTTCATTTC    8289

TCTTGCTCAC CCAAGAAGTA ACAAAAGTAT AGTTTTGACA GAGTTGGTGT TCATAATTTC    8349

AGTTCTAGTT GATTGCGAGA ATTTTCAAAT AAGGAAGAGG GGTCTTTTAT CCTTGTCGTA    8409

GGAAAACCAT GACGGAAAGG AAAAACTGAT GTTTAAAAGT CCACTTTTAA AACTATATTT    8469

ATTTATGTAG GATCTGTCAA AGAAAACTTC CAAAAAGATT TATTAATTAA ACCAGACTCT    8529
```

GTTGCAATAA GTTAATGTTT TCTTGTTTTG TAATCCACAC ATTCAATGAG TTAGGCTTTG 8589

CACTTGTAAG GAAGGAGAAG CGTTCACAAC CTCAAATAGC TAATAAACCG GTCTTGAATA 8649

TTTGAAGATT TAAAATCTGA CTCTAGGACG GGCACGGTGG CTCACGACTA TAATCCCAAC 8709

ACTTTGGGAG GCTGAGGCGG GCGGTCACAA GGTCAGGAGT TCAAGACCAG CCTGACCAAT 8769

ATGGTGAAAC CCCATCTCTA CTAAAAATAC AAAAATTAGC CAGGCGTGGT GGCAGGTGCC 8829

TGTAGGTCCC AGCTAGCCTG TGAGGTGGAG ATTGCATTGA GCCAAGATC 8878

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 4:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 491 aminokyselin
(B) TYP: aminokyselina
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: protein

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 4:

Met Lys Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys
1 5 10 15

His Leu Leu Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Gln Gly
20 25 30

Val Asn Asp Asn Glu Glu Gly Phe Phe Ser Ala Arg Gly His Arg Pro
35 40 45

Leu Asp Lys Lys Arg Glu Glu Ala Pro Ser Leu Arg Pro Ala Pro Pro
50 55 60

Pro Ile Ser Gly Gly Gly Tyr Arg Ala Arg Pro Ala Lys Ala Ala Ala
65 70 75 80

Thr Gln Lys Lys Val Glu Arg Lys Ala Pro Asp Ala Gly Gly Cys Leu
85 90 95

His Ala Asp Pro Asp Leu Gly Val Leu Cys Pro Thr Gly Cys Gln Leu
100 105 110

Gln Glu Ala Leu Leu Gln Gln Glu Arg Pro Ile Arg Asn Ser Val Asp
115 120 125

Glu Leu Asn Asn Asn Val Glu Ala Val Ser Gln Thr Ser Ser Ser Ser
130 135 140

Phe Gln Tyr Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys
145 150 155 160

Gln Val Lys Asp Asn Glu Asn Val Val Asn Glu Tyr Ser Ser Glu Leu
165 170 175

Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser Asn Ile Pro
180 185 190

Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu Arg Ser Lys
195 200 205

Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu Tyr Cys Arg
210 215 220

Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser Gly Lys Glu
225 230 235 240

Cys Glu Glu Ile Ile Arg Lys Gly Gly Glu Thr Ser Glu Met Tyr Leu
245 250 255

Ile Gln Pro Asp Ser Ser Val Lys Pro Tyr Arg Val Tyr Cys Asp Met
260 265 270

Asn Thr Glu Asn Gly Gly Trp Thr Val Ile Gln Asn Arg Gln Asp Gly
275 280 285

Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln Gly Phe Gly
290 295 300

Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly Leu Pro Gly
305 310 315 320

Glu Tyr Trp Leu Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly
325 330 335

Pro Thr Glu Leu Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val
340 345 350

Lys Ala His Tyr Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr
355 360 365

```
Gln Ile Ser Val Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met
    370                 375                 380

Asp Gly Ala Ser Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His
385                 390                 395                 400

Asn Gly Met Phe Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly Trp Leu
                405                 410                 415

Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp Gly Gly Gly Trp
            420                 425                 430

Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly Arg Tyr Tyr Trp
        435                 440                 445

Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly Thr Asp Asp Gly
    450                 455                 460

Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser Met Arg Lys Met
465                 470                 475                 480

Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln
                485                 490
```

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 5:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 10564 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: dvojité
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: DNA (genomová)

(vii) PŘÍMÝ ZDROJ:

(B) KLON: Řetězec τ lidského fibrinogenu

(ix) ZNAKY:

(A) JMÉNO/OZNAČENÍ: CDS
(B) POZICE: spoj. (1799..1876, 1973..2017, 2207..2390, 2513..2603, 4211..4341, 4645, 4778, 5758..5942, 7426..7703, 9342..9571).

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 5:

CTACACACTT CTTGAAGGCA AAGGCAATGC TGAAGTCACC TTTCATGTTC AAATCATATT 60

AAAAAGTTAG CAAGATGTAA TTATCAGTGT ACTATGTAAA TCTTTGTGAA TGATCAATAA 120

TTACATATTT TCATTATATA TATTTTAGTA GATAATATTT ATATACATTC AACATTCTAA 180

ATATAGAAAG TTTACAGAGA AAAATAAAGC CTTTTTTTCC AATCCTGTCC TCCACCTCTG 240

CATCCCATTC TTCTTCACAG AGGCAACTGA TTCAAGTCAT TACATAGTTA TTGAGTGTTA 300

ACTACAACTA TGTTAAGTAC AGCTATATAT GTTAGATGCC GTAGCCACAG AAATCAGTTT 360

ACAATCTAAT GCAGTGGATA CAGCATGTAT ACATATAATA TAAGGTTGCT ACAAATGCTA 420

TCTGAGGTAG AGCTGTTTGA AAGAATACTA ATACTTAAAT GTTTAATTCA ACTGACTTGA 480

TTGACAACTG ATTAGCTGAG TGGAAAAGAT GGATGAGAAA GATTGTGAGA CTTAATTGGC 540

TGGTGGTATG GTGATATGAT TGACAATAAC TGCTAAGTCA GAGAGGGATA TATTAAGGAG 600

GAGAAGAAAA GCAACAAATC TGGTTTTGAT GTGTTCACTT TGTTATAATT ATTGATTATT 660

TACTGAATAT GAATATTTAT CTTTGTTTTT GAGTCAATAA ATATACCTTT GTAAAGACAG 720

AATTAAAGTA TTAGTATTTC TTTCAAACTG GAGGCATTTC TCCCACTAAC ATATTTCATC 780

AAAACTTATA ATAAGCTTGG TTCCAGAGGA AGAAATGAGG GATAACCAAA AATAGAGACA 840

TTAATAATAG TGTAACGCCC AGTGATAAAT CTCAATAGGC AGTGATGACA GACATGTTTT 900

CCCAAACACA AGGATGCTGT AAGGGCCAAA CAGAAATGAT GGCCCCTCCC CAGCACCTCA 960

TTTTGCCCCT TCCTTCAGCT ATGCCTCTAC TCTCCTTTAG ATACAAGGGA GGTGGATTTT 1020

TCTCTTCTCT GAGATAGCTT GATGGAACCA CAGGAACAAT GAAGTGGGCT CCTGGCTCTT 1080

TTCTCTGTGG CAGATGGGGT GCCATGCCCA CCTTCAGACA AAGGGAAGAT TGAGCTCAAA 1140

AGCTCCCTGA GAAGTGAGAG CCTATGAACA TGGTTGACAC AGAGGGACAG GAATGTATTT 1200

CCAGGGTCAT TCATTCCTGG GAATAGTGAA CTGGGACATG GGGGAAGTCA GTCTCCTCCT 1260

GCCACAGCCA CAGATTAAAA ATAATAATGT TAACTGATCC CTAGGCTAAA ATAATAGTGT 1320

TAACTGATCC CTAAGCTAAG AAAGTTCTTT TGGTAATTCA GGTGATGGCA GCAGGACCCA 1380

```
TCTTAAGGAT AGACTAGGTT TGCTTAGTTC GAGGTCATAT CTGTTTGCTC TCAGCCATGT   1440

ACTGGAAGAA GTTGCATCAC ACAGCCTCCA GGACTGCCCT CCTCCTCACA GCAATGGATA   1500

ATGCTTCACT AGCCTTTGCA GATAATTTTG GATCAGAGAA AAAACCTTGA GCTGGGCCAA   1560

AAAGGAGGAG CTTCAACCTG TGTGCAAAAT CTGGGAACCT GACAGTATAG GTTGGGGGCC   1620

AGGATGAGGA AAAAGGAACG GGAAAGACCT GCCCACCCTT CTGGTAAGGA GGCCCCGTGA   1680

TCAGCTCCAG CCATTTGCAG TCCTGGCTAT CCCAGGAGCT TACATAAAGG GACAATTGGA   1740

GCCTGAGAGG TGACAGTGCT GACACTACAA GGCTCGGAGC TCCGGGCACT CAGACATC      1798

ATG AGT TGG TCC TTG CAC CCC CGG AAT TTA ATT CTC TAC TTC TAT GCT      1846
Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala
  1               5                  10                  15

CTT TTA TTT CTC TCT TCA ACA TGT GTA GCA GTAAGTGTGC TCTTCACAAA       1896
Leu Leu Phe Leu Ser Ser Thr Cys Val Ala
            20                  25

ACGTTGTTTA AAATGGAAAG CTGGAAAATA AAACAGATAA TAAACTAGTG AAATTTTCGT   1956

ATTTTTTCTC TTTTAG TAT GTT GCT ACC AGA GAC AAC TGC TGC ATC TTA       2005
                  Tyr Val Ala Thr Arg Asp Asn Cys Cys Ile Leu
                                  30                  35

GAT GAA AGA TTC GTAAGTAGTT TTTATGTTTC TCCCTTTGTG TGTGAACTGG        2057
Asp Glu Arg Phe
        40

AGAGGGGCAG AGGAATAGAA ATAATTCCCT CATAAATATC ATCTGGCACT TGTAACTTTT   2117

TAAAAACATA GTCTAGGTTT TACCTATTTT TCTTAATAGA TTTTAAGAGT AGCATCTGTC   2177

TACATTTTTA ATCACTGTTA TATTTTCAG GGT AGT TAT TGT CCA ACT ACC TGT     2230
                                  Gly Ser Tyr Cys Pro Thr Thr Cys
                                                  45

GGC ATT GCA GAT TTC CTG TCT ACT TAT CAA ACC AAA GTA GAC AAG GAT     2278
Gly Ile Ala Asp Phe Leu Ser Thr Tyr Gln Thr Lys Val Asp Lys Asp
 50                  55                  60                  65
```

```
CTA CAG TCT TTG GAA GAC ATC TTA CAT CAA GTT GAA AAC AAA ACA TCA      2326
Leu Gln Ser Leu Glu Asp Ile Leu His Gln Val Glu Asn Lys Thr Ser
                70                  75                  80

GAA GTC AAA CAG CTG ATA AAA GCA ATC CAA CTC ACT TAT AAT CCT GAT      2374
Glu Val Lys Gln Leu Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro Asp
            85                  90                  95

GAA TCA TCA AAA CCA  A GTGAGAAAAT AAAGACTACT GACCAAAAAA               2420
Glu Ser Ser Lys Pro
        100

TAATAATAAT AATCTGTGAA GTTCTTTTGC TGTTGTTTTA GTTGTTCTAT TTGCTTAAGG    2480

ATTTTTATGT CTCTGATCCT ATATTACAG  AT ATG ATA GAC GCT GCT ACT TTG       2532
                                Asn Met Ile Asp Ala Ala Thr Leu
                                    105                 110

AAG TCC AGG ATA ATG TTA GAA GAA ATT ATG AAA TAT GAA GCA TCG ATT      2580
Lys Ser Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile
            115                 120                 125

TTA ACA CAT GAC TCA AGT ATT  CG  GTAAGGATTT TTGTTTTAAT TTGCTCTGCA    2633
Leu Thr His Asp Ser Ser Ile  Arg
        130

AGACTGATTT AGTTTTTATT TAATATTCTA TACTTGAGTG AAAGTAATTT TTAATGTGTT    2693

TTCCCCATTT ATAATATCCC AGTGACATTA TGCCTGATTA TGTTGAGCAT AGTAGAGATA    2753

GAAGTTTTTA GTGCAATATA AATTATACTG GGTTATAATT GCTTATTAAT AATCACATTG    2813

AAGAAAGATG TTCTAGATGT CTTCAAATGC TAGTTTGACC ATATTTATCA AAAATTTTTT    2873

CCCCATCCCC CATTTATCTT ACAACATAAA ATCAATCTCA TAGGAATTTG GGTGTTGAAA    2933

ATAAAATCCT CTTTATAAAA ATGCTGACAA ATTGGTGGTT AAAAAAATTA GCAAGCAGAG    2993

GCATAGTAAG GATTTTGGCT CCTAAAGTAA ATTATATTGA ATGTGGAGCA GGAAGAAACA    3053

TGTCTTGAGA GACTAAGTGT GGCAAATATT GCAAAGCTCA TATTGATCAT TGCAGAATGA    3113

ACCTGCATAG TCTCTTCCCT TCATTTGGAA GTGAATGTCT CTGTTAAAGC TTCTCAGGGA    3173

CTCATAAACT TTCTGAACAT AAGGTCTCAG ATACAGTTTT AATATTTTTC CCCAATTTTT    3233

TTTTCTGAAT TTTTCTCAAA GCAGCTTGAG AAATTGAGAT AAATAGTAGC TAGGGAGAAG    3293
```

```
TGGCCCAGGA AAGATTTCTC CTCTTTTTGC TATCAGAGGG CCCTTGTTAT TATTGTTATT   3353

ATTATTACTT GCATTATTAT TGTCCATCAT TGAAGTTGAA GGAGGTTATT GTACAGAAAT   3413

TGCCTAAGAC AAGGTAGAGG GAAAACGTGG ACAAATAGTT TGTCTACCCT TTTTTACTTC   3473

AAAGAAAGAA CGGTTTATGC ATTGTAGACA GTTTTCTATC ATTTTTGGAT ATTTGCAAGC   3533

CACCCTGTAA GTAACTACAA AAGGAGGGTT TTTACTTCCC CCAGTCCATT CCCAAAGCTA   3593

TGTAACCAGA AGCATTAAAG AAGAAAGGGG AAGTATCTGT TGTTTTATTT TACATACAAT   3653

AACGTTCCAG ATCATGTCCC TGTGTAAGTT ATATTTTAGA TTGAAGCTTA TATGTATAGC   3713

CTCAGTAGAT CCACAAGTGA AAGGTATACT CCTTCAGCAC ATGTGAATTA CTGAACTGAG   3773

CTTTTCCTGC TTCTAAAGCA TCAGGGGGTG TTCCTATTAA CCAGTCTCGC CACTCTTGCA   3833

GGTTGCTATC TGCTGTCCCT TATGCATAAA GTAAAAAGCA AAATGTCAAT GACATTTGCT   3893

TATTGACAAG GACTTTGTTA TTTGTGTTGG GAGTTGAGAC AATATGCCCC ATTCTAAGTA   3953

AAAAGATTCA GGTCCACATT GTATTCCTGT TTTAATTGAT TTTTTGATTT GTTTTTCTTT   4013

TTCAAAAAGT TTATAATTTT AATTCATGTT AATTTAGTAA TATAATTTTA CATTTTCCTC   4073

AAGAATGGAA TAATTTATCA GAAAGCACTT CTTAAGAAAA TACTTAGCAG TTTCCAAAGA   4133

AAATATAAAA TTACTCTTCT GAAAGGAATA CTTATTTTTG TCTTCTTATT TTTGTTATCT   4193

TATGTTTCTG TTTGTAG A TAT TTG CAG GAA ATA TAT AAT TCA AAT AAT CAA     4244
                    Tyr Leu Gln Glu Ile Tyr Asn Ser Asn Asn Gln
                    135                 140                 145

AAG ATT GTT AAC CTG AAA GAG AAG GTA GCC CAG CTT GAA GCA CAG TGC     4292
Lys Ile Val Asn Leu Lys Glu Lys Val Ala Gln Leu Glu Ala Gln Cys
                150                 155                 160

CAG GAA CCT TGC AAA GAC ACG GTG CAA ATC CAT GAT ATC ACT GGG AAA G   4341
Gln Glu Pro Cys Lys Asp Thr Val Gln Ile His Asp Ile Thr Gly Lys
            165                 170                 175

GTAACTGATG AAGGTTATAT TGGGATTAGG TTCATCAAAG TAAGTAATGT AAAGGAGAAA   4401

GTATGTACTG GAAAGTATAG GAATAGTTTA GAAAGTGGCT ACCCATTAAG TCTAAGAATT   4461
```

```
TCAGTTGTCT AGACCTTTCT TGAATAGCTA AAAAAAACAG TTTAAAAGGA ATGCTGATGT    4521

GAAAAGTAAG AAAATTATTC TTGGAAAATG AATAGTTTAC TACATGTTAA AAGCTATTTT    4581

TCAAGGCTGG CACAGTCTTA CCTGCATTTC AAACCACAGT AAAAGTCGAT TCTCCTTCTC    4641

TAG  AT TGT CAA GAC ATT GCC AAT AAG GGA GCT AAA CAG AGC GGG CTT       4688
    Asp Cys Gln Asp Ile Ala Asn Lys Gly Ala Lys Gln Ser Gly Leu
            180                 185                 190

TAC TTT ATT AAA CCT CTG AAA GCT AAC CAG CAA TTC TTA GTC TAC TGT      4736
Tyr Phe Ile Lys Pro Leu Lys Ala Asn Gln Gln Phe Leu Val Tyr Cys
        195                 200                 205

GAA ATC GAT GGG TCT GGA AAT GGA TGG ACT GTG TTT CAG AAG               4778
Glu Ile Asp Gly Ser Gly Asn Gly Trp Thr Val Phe Gln Lys
    210                 215                 220

GTAATTTTTT CCCCACCATG TGTATTTAAT AAATTCCTAC ATTGTTTCTG CCATATGGCA    4838

GATACTTTTC TAAGCACCTT GTGAACCGTA GCTCATTTAA TCCTTGCAAT AGCCCTAAGA    4898

GGAAGGTACT TCTGTTACTC CTATTTACAG AAAAGGAAAC TGAGGCACAC AAGGTTAAAT    4958

AACTTGCCCA AGACCACATA ACTAATAAGC AACAGAGTCA GCATTTGAAC CTAGGCAGTA    5018

TAGTTTCAGA GTTTGTGACT TGACTCTATA TTGTACTGGC ACTGACTTTG TAGATTCATG    5078

GTGGCACATA ATCATAGTAC CACAGTGACA AATAAAAAGA AGGAAACTCT TTTGTCAGGT    5138

AGGTCAAGAC CTGAGGTTTC CCATCACAAG ATGAGGAAGC CCAACACCAC CCCCCACCAC    5198

CCCACCACCA TCACCACCCT TTCACACACC AGAGGATACA CTTGGGCTGC TCCAAGACAA    5258

GGAACCTGTG TTGCATCTGC CACTTGCTGA TACCCACTAG GAATCTTGGC TCCTTTACTT    5318

TCTGTTTACC TCCCACCACT GTTATAACTG TTTCTACAGG GGGCGCTCAG AGGGAATGAA    5378

TGGTGGAAGC ATTAGTTGCC AGACACCGAT TGAGCAATGG GTTCCATCAT AAGTGTAAGA    5438

ATCAGTAATA TCCAGCTAGA GTTCTGAAGT CGTCTAGGTG TCTTTTTAAT ATTACCACTC    5498

ATTTAGAATT TATGATGTGC CAGAAACCCT CTTAAGTATT TCTCTTATAT TCTCTCTCAT    5558

GATCCTTGCA GCAACCCTAA GAAGTAACCA TCATTTTTCC TATTTGATAC ATGAGGAAAC    5618

TGAGGTAGCT TGGCCAAGAT CACTTAGTTG GGAGTTGATA GAACCAGTGC TCTGTATTTT    5678
```

```
TGACAAAATG TTGACAGCAT TCTCTTTACA TGCATTGATA GTCTATTTTC TCCTTTTGCT   5738

CTTGCAAATG TGTAATTAG AGA CTT GAT GGC AGT GTA GAT TTC AAG AAA AAC   5790
                      Arg Leu Asp Gly Ser Val Asp Phe Lys Lys Asn
                                  225                 230

TGG ATT CAA TAT AAA GAA GGA TTT GGA CAT CTG TCT CCT ACT GGC ACA   5838
Trp Ile Gln Tyr Lys Glu Gly Phe Gly His Leu Ser Pro Thr Gly Thr
    235                 240                 245

ACA GAA TTT TGG CTG GGA AAT GAG AAG ATT CAT TTG ATA AGC ACA CAG   5886
Thr Glu Phe Trp Leu Gly Asn Glu Lys Ile His Leu Ile Ser Thr Gln
250                 255                 260                 265

TCT GCC ATC CCA TAT GCA TTA AGA GTG GAA CTG GAA GAC TGG AAT GGC   5934
Ser Ala Ile Pro Tyr Ala Leu Arg Val Glu Leu Glu Asp Trp Asn Gly
                270                 275                 280

AGA ACC  AG  GTACTGTTTT GAAATGACTT CCAACTTTTT ATTGTAAAGA          5982
Arg Thr  Ser

TTGCCTGGAA TGTGCACTTT CCAACTATCA ATAGACAATG GCAAATGCAG CCTGACAAAT   6042

GCAAACAGCA CATCCAGCCA CCATTTTCTC CAGGAGTCTG TTTGGTTCTT GGGCAATCCA   6102

AAAAGGTAAA TTCTATTCAG GATGAATCTA AGTGTATTGG TACAATCTAA TTACCCTGGA   6162

ACCATTCAGA GTAATAGCTA ATTACTGAAC TTTTAATCAG TCCCAGGAAT TGAGCATAAA   6222

ATTATAATTT TATCTAGTCT AAATTACTAT TTCATGAAGC AGGTATTATT ATTAATCCCA   6282

TTTTATAGAT TAACTTGCTC AAAGTCACAT TGCTGATAAG TGGTAGAGGT AGAATTCAGA   6342

CTCAAGTAGT TTAACTTTAG AGCCTGTCCT CTTAACAACT ATCCTGGTTG AAAAGCAAAT   6402

ACAGCCTCTT CAGACTTCTC AGTGCCTTGA TGGCCATTTA TTCTGTCAAA TCATGAGCTA   6462

CCCTAAAAGT AAACCAGCTA GCTCTTTTGA TGATCTAGAG GCTTCTTTTT GCTTGAGATA   6522

TTTGAAGGTT TTAAGCATTG TTACCTAATT AAAATGCAGA AAAATATCCA ACCCTCTTGT   6582

TATGTTTAAG GAATAGTGAA ATATATTGTC TTCAAACACA TGGACTTTTT TTTATTGCTT   6642

GGTTGGTTTT TAATCCAGAA AGTGCTATAG TCAGTAGACC TTCTTCTAGG AAAGGACCTT   6702
```

```
CCATTTCCCA GCCACTGGAG ATTAGAAAAT AAGCTAAATA TTTTCTGGAA ATTTCTGTTC    6762

ATTCATTAAG GCCCATCCTT TCCCCCACTC TATAGAAGTG TTGTCCACTT GCACAATTTT    6822

TTCCAGGAAA GAATCTCTCT AACTCCTTCA GCTCACATGC TTTGGACCAC ACAGGGAAGA    6882

CTTTGATTGT GTAATGCCCT CAGAAGCTCT CCTTCTTGCC ACTACCACAC TGATTTGAGG    6942

AAGAAAATCC CTTTAGCACC TAACCCTTCA GGTGCTATGA GTGGCTAATG GAACTGTACC    7002

TCCTTCAAGT TTTGTGCAAT AATTAAGGGT CACTCACTGT CAGATACTTT CTGTGATCTA    7062

TGATAATGTG TGTGCAACAC ATAACATTTC AATAAAAGTA GAAAATATGA AATTAGAGTC    7122

ATCTACACAT CTGGATTTGA TCTTAGAATG AAACAAGCAA AAAAGCATCC AAGTGAGTGC    7182

AATTATTAGT TTTCAGAGAT GCTTCAAAGG CTTCTAGGCC CATCCCGGGA AGTGTTAATG    7242

AGCTGTGGAC TGGTTCACAT ATCTATTGCC TCTTGCCAGA TTTGCAAAAA ACTTCACTCA    7302

ATGAGCAAAT TTCAGCCTTA AGAAACAAAG TCAAAAATTC CAAGGAAGCA TCCTACGAAA    7362

GAGGGAACTT CTGAGATCCC TGAGGAGGGT CAGCATGTGA TGGTTGTATT TCCTTCTTCT    7422

CAG T ACT GCA GAC TAT GCC ATG TTC AAG GTG GGA CCT GAA GCT GAC         7468
      Thr Ala Asp Tyr Ala Met Phe Lys Val Gly Pro Glu Ala Asp
      285                 290                 295

AAG TAC CGC CTA ACA TAT GCC TAC TTC GCT GGT GGG GAT GCT GGA GAT      7516
Lys Tyr Arg Leu Thr Tyr Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp
    300                 305                 310

GCC TTT GAT GGC TTT GAT TTT GGC GAT GAT CCT AGT GAC AAG TTT TTC      7564
Ala Phe Asp Gly Phe Asp Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe
315                 320                 325                 330

ACA TCC CAT AAT GGC ATG CAG TTC AGT ACC TGG GAC AAT GAC AAT GAT      7612
Thr Ser His Asn Gly Met Gln Phe Ser Thr Trp Asp Asn Asp Asn Asp
                335                 340                 345

AAG TTT GAA GGC AAC TGT GCT GAA CAG GAT GGA TCT GGT TGG TGG ATG      7660
Lys Phe Glu Gly Asn Cys Ala Glu Gln Asp Gly Ser Gly Trp Trp Met
          350                 355                 360

AAC AAG TGT CAC GCT GGC CAT CTC AAT GGA GTT TAT TAC CAA  G           7703
Asn Lys Cys His Ala Gly His Leu Asn Gly Val Tyr Tyr Gln
       365                 370                 375
```

GTATGTTTTC CTTTCTTAGA TTCCAAGTTA ATGTATAGTG TATACTATTT TCATAAAAAA 7763

TAATAAATAG ATATGAAGAA ATGAAGAATA ATTTATAAAG ATAGTAGGGA TTTTATCATG 7823

TTCTTTATTT CAACTAAGTT CTTTGAAACT GGAAGTGGAT AATACCAAGT TCATGCCTAA 7883

AATTAGCCCT TCTAAAGAAA TCCACCTGCT GCAAAATATC CAGTAGTTTG GCATTATATG 7943

TGAAACTATC ACCATCATAG CTGGCACTGT GGGTTGTGGG ATCTCCTTTA GACATACAAC 8003

ATAAATGATC TGGATGGATT AACATTACTA CATGGATGCT TGTTGACACA TTAACCTGGC 8063

TTCCCATGAG CTTTGTGTCA GATACACGCA GTGAACAGGT GTTTGGAGGA ACAGAATAAA 8123

GAGAAGGCAA GCACTGGTAA GGGCAGGGGT TTGTGAAAGC TTGAGAGAAG AGACCAGTCT 8183

GAGGACAGTA GACACTTATT TTAGGATGGG GGTTGGATGA GGAGGCTATA GTTTGCTATA 8243

AGCTTGGAAT GGTTTGGAAC ACTGGTTTCA CTCACCTACC CAGCAGTTAT GTGTGGGGAA 8303

GCCTTACCGA TGCTAAAGGA TCCATGTTAC AATAATGGCA TTATTTGGAA ATCCCAGTGG 8363

TATTCCATGA ATAAAACCAC TATGAAGATA ATCCCACTCA ACAGACTCTC CGTTGGAGAA 8423

GGACAGCAAC ACCACCCTGG GAAAGCCAAA CAGTCAGACC AGACCTGTTT AGCATCAGTA 8483

GGACTTCCCT ACCATATCTG CTGGGTAGAT GAGTGAAACC AGTGTTCCAA ACCACTCCGG 8543

GCTTGTAGCA AACCATAGTC TCCTCATCTA CCAAGATGAG CAACCTTACC TCCTGATGTC 8603

CTAGCCAATC ACCAACTAGG AAACTTTGCA CAGTTTATTT AAAGTAACAG TTTGATTTTC 8663

ACAATATTTT TAAATTGGAG AAACATAACT TATCTTTGCA CTCACAAACC ACATAATGAG 8723

AAGAAACTCT AAGGGAAAAT GCTTGATCTG TGTGACCCGG GGCGCCATGC CAGAGCTGTA 8783

GTTCATGCCA GTGTTGTGCT CTGACAAGCC TTTTACAGAA TTACATGAGA TCTGCTTCCC 8843

TAGGACAAGG AGAAGGCAAA TCAACAGAGG CTGCACTTTA AAATGGAGAC ATAAAATAAC 8903

ATGCCAGAAC CATTTCCTAA AGCTCCTCAA TCAACCAACA AAATTGTGCT TTCAAATAAC 8963

CTGAGTTGAC CTCATCAGGA ATTTTGTGGC TCCTTCTCTT CTAACCTGCC TGAAGAAAGA 9023

TGGTCCACAG CAGCTGAGTC CGGGATGGAT AAGCTTAGGG ACAGAGGCCA ATTAGGGAAC 9083

TTTGGGTTTC TAGCCCTACT AGTAGTGAAT AAATTTAAAG TGTGGATGTG ACTATGAGTC 9143

ACAGCACAGA TGTTGTTTAA TAATATGTTT ATTTTATAAA TTGATATTTT AGGAATCTTT 9203

GGAGATATTT TCAGTTAGCA GATAATACTA TAAATTTTAT GTAACTGGCA ATGCACTTCG 9263

TAATAGACAG CTCTTCATAG ACTTGCAGAG GTAAAAAGAT TCCAGAATAA TGATATGTAC 9323

ATCTACGACT TGTTTTAG GT GGC ACT TAC TCA AAA GCA TCT ACT CCT AAT 9373
Gly Gly Thr Tyr Ser Lys Ala Ser Thr Pro Asn
380 385

GGT TAT GAT AAT GGC ATT ATT TGG GCC ACT TGG AAA ACC CGG TGG TAT 9421
Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr Arg Trp Tyr
390 395 400

TCC ATG AAG AAA ACC ACT ATG AAG ATA ATC CCA TTC AAC AGA CTC ACA 9469
Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn Arg Leu Thr
405 410 415

ATT GGA GAA GGA CAG CAA CAC CAC CTG GGG GGA GCC AAA CAG GTC AGA 9517
Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys Gln Val Arg
420 425 430 435

CCA GAG CAC CCT GCG GAA ACA GAA TAT GAC TCA CTT TAC CCT GAG GAT 9565
Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr Pro Glu Asp
440 445 450

GAT TTG TAGAAAATTA ACTGCTAACT TCTATTGACC CACAAAGTTT CAGAAATTCT 9621
Asp Leu

CTGAAAGTTT CTTCCTTTTT TCTCTTACTA TATTTATTGA TTTCAAGTCT TCTATTAAGG 9681

ACATTTAGCC TTCAATGGAA ATTAAAACTC ATTTAGGACT GTATTTCCAA ATTACTGATA 9741

TCAGAGTTAT TTAAAAATTG TTTATTTGAG GAGATAACAT TTCAACTTTG TTCCTAAATA 9801

TATAATAATA AAATGATTGA CTTTATTTGC ATTTTTATGA CCACTTGTCA TTTATTTTGT 9861

CTTCGTAAAT TATTTTCATT ATATCAAATA TTTTAGTATG TACTTAATAA AATAGGAGAA 9921

CATTTTAGAG TTTCAAATTC CCAGGTATTT TCCTTGTTTA TTACCCCTAA ATCATTCCTA 9981

TTTAATTCTT CTTTTTAAAT GGAGAAAATT ATGTCTTTTT AATATGGTTT TTGTTTTGTT 10041

ATATATTCAC AGGCTGGAGA CGTTTAAAAG ACCGTTTCAA AAGAGATTTA CTTTTTTAAA 10101

GGACTTTATC TGAACAGAGA GATATAATAT TTTTCCTATT GGACAATGGA CTTGCAAAGC 10161

TTCACTTCAT TTTAAGAGCA AAAGACCCCA TGTTGAAAAC TCCATAACAG TTTTATGCTG 10221

ATGATAATTT ATCTACATGC ATTTCAATAA ACCTTTTGTT TCCTAAGACT AGATACATGG 10281

TACCTTTATT GACCATTAAA AAACCACCAC TTTTTGCCAA TTTACCAATT ACAATTGGGC 10341

AACCATCAGT AGTAATTGAG TCCTCATTTT ATGCTAAATG TTATGCCTAA CTCTTTGGGA 10401

GTTACAAAGG AAATAGCAAT TATGGCTTTT GCCCTCTAGG AGATACAGGA CAAATACAGG 10461

AAAATACAGC AACCCAAACT GACAATACTC TATACAAGAA CATAATCACT AAGCAGGAGT 10521

CACAGCCACA CAACCAAGAT GCATAGTATC CAAAGTGCAG CTG 10564

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 6:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 453 aminokyselin
(B) TYP: aminokyselina
(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: protein

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 6:

Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala
1 5 10 15

Leu Leu Phe Leu Ser Ser Thr Cys Val Ala Tyr Val Ala Thr Arg Asp
20 25 30

Asn Cys Cys Ile Leu Asp Glu Arg Phe Gly Ser Tyr Cys Pro Thr Thr
35 40 45

Cys Gly Ile Ala Asp Phe Leu Ser Thr Tyr Gln Thr Lys Val Asp Lys
50 55 60

Asp Leu Gln Ser Leu Glu Asp Ile Leu His Gln Val Glu Asn Lys Thr
65 70 75 80

Ser Glu Val Lys Gln Leu Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro
85 90 95

Asp Glu Ser Ser Lys Pro Asn Met Ile Asp Ala Ala Thr Leu Lys Ser
100 105 110

Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile Leu Thr
115 120 125

His Asp Ser Ser Ile Arg Tyr Leu Gln Glu Ile Tyr Asn Ser Asn Asn
130 135 140

Gln Lys Ile Val Asn Leu Lys Glu Lys Val Ala Gln Leu Glu Ala Gln
145 150 155 160

Cys Gln Glu Pro Cys Lys Asp Thr Val Gln Ile His Asp Ile Thr Gly
165 170 175

Lys Asp Cys Gln Asp Ile Ala Asn Lys Gly Ala Lys Gln Ser Gly Leu
180 185 190

Tyr Phe Ile Lys Pro Leu Lys Ala Asn Gln Gln Phe Leu Val Tyr Cys
195 200 205

Glu Ile Asp Gly Ser Gly Asn Gly Trp Thr Val Phe Gln Lys Arg Leu
210 215 220

Asp Gly Ser Val Asp Phe Lys Lys Asn Trp Ile Gln Tyr Lys Glu Gly
225 230 235 240

Phe Gly His Leu Ser Pro Thr Gly Thr Thr Glu Phe Trp Leu Gly Asn
245 250 255

Glu Lys Ile His Leu Ile Ser Thr Gln Ser Ala Ile Pro Tyr Ala Leu
260 265 270

Arg Val Glu Leu Glu Asp Trp Asn Gly Arg Thr Ser Thr Ala Asp Tyr
275 280 285

Ala Met Phe Lys Val Gly Pro Glu Ala Asp Lys Tyr Arg Leu Thr Tyr
290 295 300

Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp Ala Phe Asp Gly Phe Asp
305 310 315 320

Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe Thr Ser His Asn Gly Met
325 330 335

Gln Phe Ser Thr Trp Asp Asn Asp Asn Asp Lys Phe Glu Gly Asn Cys
340 345 350

Ala Glu Gln Asp Gly Ser Gly Trp Trp Met Asn Lys Cys His Ala Gly
355 360 365

His Leu Asn Gly Val Tyr Tyr Gln Gly Gly Thr Tyr Ser Lys Ala Ser
370 375 380

Thr Pro Asn Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr
385 390 395 400

Arg Trp Tyr Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn
405 410 415

Arg Leu Thr Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys
420 425 430

Gln Val Arg Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr
435 440 445

Pro Glu Asp Asp Leu
450

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 7:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 10807 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: dvojité
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ovčí beta–laktoglobulin

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 7:

ACGCGTGTCG ACCTGCAGGT CAACGGATCT CTGTGTCTGT TTTCATGTTA GTACCACACT 60

GTTTTGGTGG CTGTAGCTTT CAGCTACAGT CTGAAGTCAT AAAGCCTGGT ACCTCCAGCT 120

CTGTTCTCTC TCAAGATTGT GTTCTGCTGT TTGGGTCTTT AGTGTCTCCA CACAATTTTT 180

AGAATTGTTT GTTCTAGTTC TGTGAAAAAT GATGCTGGTA TTTTGATAAG GATTGCATTG 240

AATCTGTAAA GCTACAGATA TAGTCATTGG GTAGTACAGT CACTTTAACA ATATTAACTC 300

TTCACATCTG TGAGCATGAT ATATTTTCCC CCTCTATATC ATCTTCAATT CCTCCTATCA 360

GTTTCTTTCA TTGCAGTTTT CTGAGTACAG GTCTTACACC TCCTTGGTTA GAGTCATTCC 420

TCAGTATTTT ATTCCTTTGA TACAATTGTG AATGAGGTAA TTTTCTTAGT TTCTCTTTCT 480

GATAGCTCAT TGTTAGTGTA TATATAGAAA AGCAACAGAT TTCTATGTAT TAATTTTGTA 540

TCCTGCAACA GATTTCTATG TATTAATTTT GTATCCTGCT ACTTTACGGA ATTCACTTAT 600

TAGCTTTTTG GTGACATCTT GAGGATTTTC TGAAGAAAAT GGCATGGTAT GGTAGGACAA 660

GGTGTCATGT CATCTGCAAA CAGTGGCAGT TTTCCTTCTT CCCTTCCAAC CTGGATTTCT 720

TTGATTTCTT TCTGTCTGAG TACGACTAGG ATTCCCAATA CTATACCGAA TAAAAGTGGC 780

AAGAGTGGAC ATCCTTGTCT TATTTTTCTG ACCTTAGAGG AAATGCTTTC AGTTTTTCAC 840

CATTAATTAT AATGTTTACT GTGGGCTTGT CATATGTGGC CTTCATTATA TGGAGGTCTA 900

TTCCCTCTAT ACCCACCTTG TTGAGAGTTT TTATCATAAA AGTATGTTGA ATTTTGTCAA 960

AAGTTTTTCC TGCATCTATT GAGATGATTT TTACTCTTCA ATTCATTAAT GATTTTTATT 1020

CTTCATTTTG TTAATGATTT CCATTCTTCA ATTTGTTAAC GTGGTATATC ACATTGATTG 1080

ATTTGTGGAT ACCTTTGTAT CCCTGGGATA AACCTCACTT GATCATGAGC TTTCAATGTA 1140

TTTTTGAATT CACTTTGCTA ATATTCTGTT GGGTATTTTT GCATCTCTAT TCATCAATGA 1200

TATTGGCCTA AGAAAGGTTT TGTCTGGTTT TAGTATCAGG GTGATGCTGG CCTCATAGAG 1260

AGAGTTTAGA AGCATTTCCT CCTCTTTGAT TTTTCGGAAT AGTTTGAGTA GGATAGGTAT 1320

TAACTCTTCT TTAAATGTTT GGGGACTTCC CTGGTGAGCC GGTGGTTGAG AATCCGCCTC 1380

AGGGATGTGG GTTTGATCCC TGGTCAGGGA ACCATTAATA AGATCCCACA TGCTGCAGGC 1440

AACAAGCCCC CAAGCTGCAA CCACTGAGCT GCAACCGCTG CAGTGCCCAC AGGCCACGAC 1500

CAGAGAAAGC CCACATACAG CAGGGAAGAC CCAGCACAAC CGGAAAAAGG AGTTTGGTGG 1560

AATACAGCTG TGAAGCCGTC TGGTCCTGGA CTCCTGCTTG AGGGAATTTT TTAAAAATTA 1620

TTGATTCAAT TTCATTACTG GTAACTGGTC TGTTCATATT TTCTATTTCT TCCGGGTTCA 1680

GTCTTGGGAG ATTGTACATG CCTAGGAATG TGTCCGTTTC TTCTAGGTTG TCCATTTTAT 1740

TGGACATGCA TGGGAGCACA CAGCACCGAC CAGCGAGACT CATGCTGGCT TCCTGGGGCC 1800

AGGCTGGGGC CCCAAGCAGC ATGGCATCCT AGAGTGTGTG AAAGCCCACT GACCCTGCCC 1860

AGCCCCACAA TTTCATTCTG AGAAGTGATT CCTTGCTTCT GCACTTACAG GCCCAGGATC 1920

TGACCTGCTT CTGAGGAGCA GGGGTTTTGG CAGGACGGGG AGATGCTGAG AGCCGACGGG 1980

GGTCCAGGTC CCCTCCCAGG CCCCCCTGTC TGGGGCAGCC CTTGGGAAAG ATTGCCCCAG 2040

TCTCCCTCCT ACAGTGGTCA GTCCCAGCTG CCCCAGGCCA GAGCTGCTTT ATTTCCGTCT 2100

CTCTCTCTGG ATGGTATTCT CTGGAAGCTG AAGGTTCCTG AAGTTATGAA TAGCTTTGCC 2160

CTGAAGGGCA TGGTTTGTGG TCACGGTTCA CAGGAACTTG GGAGACCCTG CAGCTCAGAC 2220

GTCCCGAGAT TGGTGGCACC CAGATTTCCT AAGCTCGCTG GGGAACAGGG CGCTTGTTTC 2280

TCCCTGGCTG ACCTCCCTCC TCCCTGCATC ACCCAGTTCT GAAAGCAGAG CGGTGCTGGG 2340

GTCACAGCCT CTCGCATCTA ACGCCGGTGT CCAAACCACC CGTGCTGGTG TTCGGGGGGC 2400

TACCTATGGG GAAGGGCTTC TCACTGCAGT GGTGCCCCCC GTCCCCTCTG AGATCAGAAG 2460

TCCCAGTCCG GACGTCAAAC AGGCCGAGCT CCCTCCAGAG GCTCCAGGGA GGGATCCTTG 2520

CCCCCCCGCT GCTGCCTCCA GCTCCTGGTG CCGCACCCTT GAGCCTGATC TTGTAGACGC 2580

CTCAGTCTAG TCTCTGCCTC CGTGTTCACA CGCCTTCTCC CCATGTCCCC TCCGTGTCCC 2640

CGTTTTCTCT CACAAGGACA CCGGACATTA GATTAGCCCC TGTTCCAGCC TCACCTGAAC 2700

AGCTCACATC TGTAAAGACC TAGATTCCAA ACAAGATTCC AACCTGAAGT TCCCGGTGGA 2760

TGTGAGTTCT GGGGCGACAT CCTTCAACCC CATCACAGCT TGCAGTTCAT CGCAAAACAT 2820

GGAACCTGGG GTTTATCGTA AAACCCAGGT TCTTCATGAA ACACTGAGCT TCGAGGCTTG 2880

TTGCAAGAAT TAAAGGTGCT AATACAGATC AGGGCAAGGA CTGAAGCTGG CTAAGCCTCC 2940

TCTTTCCATC ACAGGAAAGG GGGGCCTGGG GGCGGCTGGA GGTCTGCTCC CGTGAGTGAG 3000

CTCTTTCCTG CTACAGTCAC CAACAGTCTC TCTGGGAAGG AAACCAGAGG CCAGAGAGCA 3060

AGCCGGAGCT AGTTTAGGAG ACCCCTGAAC CTCCACCCAA GATGCTGACC AGCCAGCGGG 3120

CCCCCTGGAA AGACCCTACA GTTCAGGGGG GAAGAGGGGC TGACCCGCCA GGTCCCTGCT 3180

ATCAGGAGAC ATCCCCGCTA TCAGGAGATT CCCCCACCTT GCTCCCGTTC CCCTATCCCA 3240

ATACGCCCAC CCCACCCCTG TGATGAGCAG TTTAGTCACT TAGAATGTCA ACTGAAGGCT 3300

TTTGCATCCC CTTTGCCAGA GGCACAAGGC ACCCACAGCC TGCTGGGTAC CGACGCCCAT 3360

GTGGATTCAG CCAGGAGGCC TGTCCTGCAC CCTCCCTGCT CGGGCCCCCT CTGTGCTCAG 3420

CAACACACCC AGCACCAGCA TTCCCGCTGC TCCTGAGGTC TGCAGGCAGC TCGCTGTAGC 3480

CTGAGCGGTG TGGAGGGAAG TGTCCTGGGA GATTTAAAAT GTGAGAGGCG GGAGGTGGGA 3540

GGTTGGGCCC TGTGGGCCTG CCCATCCCAC GTGCCTGCAT TAGCCCCAGT GCTGCTCAGC 3600

CGTGCCCCCG CCGCAGGGGT CAGGTCACTT TCCCGTCCTG GGGTTATTAT GACTCTTGTC 3660

ATTGCCATTG CCATTTTTGC TACCCTAACT GGGCAGCAGG TGCTTGCAGA GCCCTCGATA 3720

CCGACCAGGT CCTCCCTCGG AGCTCGACCT GAACCCCATG TCACCCTTGC CCCAGCCTGC 3780

AGAGGGTGGG TGACTGCAGA GATCCCTTCA CCCAAGGCCA CGGTCACATG GTTTGGAGGA 3840

GCTGGTGCCC AAGGCAGAGG CCACCCTCCA GGACACACCT GTCCCCAGTG CTGGCTCTGA 3900

CCTGTCCTTG TCTAAGAGGC TGACCCCGGA AGTGTTCCTG GCACTGGCAG CCAGCCTGGA 3960

CCCAGAGTCC AGACACCCAC CTGTGCCCCC GCTTCTGGGG TCTACCAGGA ACCGTCTAGG 4020

CCCAGAGGGG ACTTCCTGCT TGGCCTTGGA TGGAAGAAGG CCTCCTATTG TCCTCGTAGA 4080

GGAAGCCACC CCGGGGCCTG AGGATGAGCC AAGTGGGATT CCGGGAACCG CGTGGCTGGG 4140

GGCCCAGCCC GGGCTGGCTG GCCTGCATGC CTCCTGTATA AGGCCCCAAG CCTGCTGTCT 4200

CAGCCCTCCA CTCCCTGCAG AGCTCAGAAG CACGACCCCA GGGATATCCC TGCAGCCATG 4260

AAGTGCCTCC TGCTTGCCCT GGGCCTGGCC CTCGCCTGTG GCGTCCAGGC CATCATCGTC 4320

ACCCAGACCA TGAAAGGCCT GGACATCCAG AAGGTTCGAG GGTTGGCCGG GTGGGTGAGT 4380

TGCAGGGCGG GCAGGGGAGC TGGGCCTCAG AGAGCCAAGA GAGGCTGTGA CGTTGGGTTC 4440

CCATCAGTCA GCTAGGGCCA CCTGACAAAT CCCCGCTGGG GCAGCTTCAA CCAGGCGTTC 4500

ACTGTCTTGC ATTCTGGAGG CTGGAAGCCC AAGATCCAGG TGTTGGCAGG GCTGGCTTCT 4560

CCTGCGGCCG CTCTCTGGGG AGCAGACGGC CGTCTTCTCC AGTCCTCTGC GCGCCCTGAT 4620

TTCCTCTTCC TGTGAGGCCA CCAGGCCTGC TGGAAACACG CCTGCCTGCG CAGCTTCACA 4680

CGACCTTTGT CATCTCTTTA AAGGCCATGT CTCCAGAGTC ATGTGTTGAA GTTCTGGGGG 4740

TTAGTGGGAC ACAGTTCAGC CCCTAAAAGA GTCTCTCTGC CCCTCAAATT TTCCCCACCT 4800

CCAGCCATGT CTCCCCAAGA TCCAAATGTT GCTACATGTG GGGGGGCTCA TCTGGGTCCC 4860

TCTTTGGGTT CAGTGTGAGT CTGGGGAGAG CATTCCCCAG GGTGCAGAGT TGGGGGGAGT 4920

ATCTCAGGGC TGCCCAGGCC GGGGTGGGAC AGAGAGCCCA CTGTGGGGCT GGGGGCCCCT 4980

TCCCACCCCC AGAGTGCAAC TCAAGGTCCC TCTCCAGGTG GCGGGGACTT GGCACTCCTT 5040

GGCTATGGCG GCCAGCGACA TCTCCCTGCT GGATGCCCAG AGTGCCCCCC TGAGAGTGTA 5100

CGTGGAGGAG CTGAAGCCCA CCCCCGAGGG CAACCTGGAG ATCCTGCTGC AGAAATGGTG 5160

GGCGTCTCTC CCCAACATGG AACCCCCACT CCCCAGGGCT GTGGACCCCC CGGGGGGTGG 5220

GGTGCAGGAG GGACCAGGGC CCCAGGGCTG GGGAAGAGGG CTCAGAGTTT ACTGGTACCC 5280

GGCGCTCCAC CCAAGGCTGC CCACCCAGGG CTTTTTTTTT TTTTAAACTT TTATTAATTT 5340

GATGCTTCAG AACATCATCA AACAAATGAA CATAAAACAT TCATTTTTGT TTACTTGGAA 5400

GGGGAGATAA AATCCTCTGA AGTGGAAATG CATAGCAAAG ATACATACAA TGAGGCAGGT 5460

ATTCTGAATT CCCTGTTAGT CTGAGGATTA CAAGTGTATT TGAGCAACAG AGAGACATTT 5520

TCATCATTTC TAGTCTGAAC ACCTCAGTAT CTAAAATGAA CAAGAAGTCC TGGAAACGAA 5580

GCAGTGTGGG GATAGGCCCG TGTGAAGGCT GCTGGGAGGC AGCAGACCTG GGTCTTCGGG 5640

CTCAAGCAGT TCCCGCTACC AGCCCTGTCC ACCTCAGACG GGGGTCAGGG TGCAGGAGAG 5700

AGCTGGATGG GTGTGGGGGC AGAGATGGGG ACCTGAACCC CAGGGCTGCC TTTTGGGGGT 5760

GCCTGTGGTC AAGGCTCTCC CTGACCTTTT CTCTCTGGCT TCATCTGACT TCTCCTGGCC 5820

CATCCACCCG GTCCCCTGTG GCCTGAGGTG ACAGTGAGTG CGCCGAGGCT AGTTGGCCAG 5880

CTGGCTCCTA TGCCCATGCC ACCCCCCTCC AGCCCTCCTG GGCCAGCTTC TGCCCCTGGC 5940

CCTCAGTTCA TCCTGATGAA AATGGTCCAT GCCAATGGCT CAGAAAGCAG CTGTCTTTCA 6000

GGGAGAACGG CGAGTGTGCT CAGAAGAAGA TTATTGCAGA AAAAACCAAG ATCCCTGCGG 6060

TGTTCAAGAT CGATGGTGAG TCCGGGTCCC TGGGGGACAC CCACCACCCC CGCCCCCGGG 6120

GACTGTGGAC AGGTTCAGGG GGCTGGCGTC GGGCCCTGGG ATGCTAAGGG ACTGGTGGTG 6180

ATGAAGACAC TGCCTTGACA CCTGCTTCAC TTGCCTCCCC TGCCACCTGC CCGGGGCCTT 6240

GGGGCGGTGG CCATGGGCAG GTCCCGGCTG GCGGGCTAAC CCACCAGGGT GACACCCGAG 6300

CTCTCTTTGC TGGGGGGCGG GCGGTGCTCT GGGCCCTCAG GCTGAGCTCA GGAGGTACCT 6360

GTGCCCTCCC AGGGGTAACC GAGAGCCGTT GCCCACTCCA GGGGCCCAGG TGCCCCACGA 6420

CCCCAGCCCG CTCCACAGCT CCTTCATCTC CTGGAGACAA ACTCTGTCCG CCCTCGCTCA 6480

TTCACTTGTT CGTCCTAAAT CCGAGATGAT AAAGCTTCGA GGGGGGGTTG GGGTTCCATC 6540

AGGGCTGCCC TTCCGCCGGG CAGCCTGGGC CACATCTGCC CTTGGCCCCC TCAGGACTCA 6600

CTCTGACTGG AGGCCCTGCA CTGACTGACG CCAGGGTGCC CAGCCCAGGG TCTCTGGCGC 6660

CATCCAGCTG CACTGGGTTT GGGTGCTGGT CCTGCCCCCA AGCTGCCCGG ACACCACAGG 6720

CAGCCGGGGC TGCCCACTGG CCTCGGTCAG GGTGAGCCCC AGCTGCCCCC GCTCAGGGCT 6780

TGCCCCGACA ATGACCCCAT CCTCAGGACG CACCCCCCTT CCCTTGCTGG GCAGTGTCCA 6840

GCCCCACCCG AGATCGGGGG AAGCCCTATT TCTTGACAAC TCCAGTCCCT GGGGGAGGGG 6900

GCCTCAGACT GAGTGGTGAG TGTTCCCAAG TCCAGGAGGT GGTGGAGGGT CCTGGCGGAT 6960

CCAGAGTTGA CAGTGAGGGC TTCCTGGGCC CCATGCGCCT GGCAGTGGCA GCAGGGAAGA 7020

GGAAGCACCA TTTCAGGGGT GGGGGATGCC AGAGGCGCTC CCCACCCCGT CTTCGCCGGG 7080

TGGTGACCCC GGGGGAGCCC CGCTGGTCGT GGAGGGTGCT GGGGGCTGAC TAGCAACCCC 7140

TCCCCCCCCG TTGGAACTCA CTTTTCTCCC GTCTTGACCG CGTCCAGCCT TGAATGAGAA 7200

CAAAGTCCTT GTGCTGGACA CCGACTACAA AAAGTACCTG CTCTTCTGCA TGGAAAACAG 7260

TGCTGAGCCC GAGCAAAGCC TGGCCTGCCA GTGCCTGGGT GGGTGCCAAC CCTGGCTGCC 7320

CAGGGAGACC AGCTGCGTGG TCCTTGCTGC AACAGGGGGT GGGGGGTGGG AGCTTGATCC 7380

CCAGGAGGAG GAGGGGTGGG GGGTCCCTGA GTCCCGCCAG GAGAGAGTGG TCGCATACCG 7440

GGAGCCAGTC TGCTGTGGGC CTGTGGGTGG CTGGGGACGG GGGCCAGACA CACAGGCCGG 7500

GAGACGGGTG GGCTGCAGAA CTGTGACTGG TGTGACCGTC GCGATGGGGC CGGTGGTCAC 7560

TGAATCTAAC AGCCTTTGTT ACCGGGGAGT TTCAATTATT TCCCAAAATA AGAACTCAGG 7620

TACAAAGCCA TCTTTCAACT ATCACATCCT GAAAACAAAT GGCAGGTGAC ATTTTCTGTG 7680

CCGTAGCAGT CCCACTGGGC ATTTTCAGGG CCCCTGTGCC AGGGGGGCGC GGGCATCGGC 7740

GAGTGGAGGC TCCTGGCTGT GTCAGCCGGC CCAGGGGGAG GAAGGGACCC GGACAGCCAG 7800

AGGTGGGGGG CAGGCTTTCC CCCTGTGACC TGCAGACCCA CTGCACTGCC CTGGGAGGAA 7860

GGGAGGGGAA CTAGGCCAAG GGGGAAGGGC AGGTGCTCTG GAGGGCAAGG GCAGACCTGC 7920

AGACCACCCT GGGGAGCAGG GACTGACCCC CGTCCCTGCC CCATAGTCAG GACCCCGGAG 7980

GTGGACAACG AGGCCCTGGA GAAATTCGAC AAAGCCCTCA AGGCCCTGCC CATGCACATC 8040

CGGCTTGCCT TCAACCCGAC CCAGCTGGAG GGTGAGCACC CAGGCCCCGC CCTTCCCCAG 8100

GGCAGGAGCC ACCCGGCCCC GGGACGACCT CCTCCCATGG TGACCCCCAG CTCCCCAGGC 8160

CTCCCAGGAG GAAGGGGTGG GGTGCAGCAC CCCGTGGGGG CCCCCTCCCC ACCCCCTGCC 8220

AGGCCTCTCT TCCCGAGGTG TCCAGTCCCA TCCTGACCCC CCCATGACTC TCCCTCCCCC 8280

ACAGGGCAGT GCCACGTCTA GGTGAGCCCC TGCCGGTGCC TCTGGGGTAA GCTGCCTGCC 8340

CTGCCCCACG TCCTGGGCAC ACACATGGGG TAGGGGGTCT TGGTGGGGCC TGGGACCCCA 8400

CATCAGGCCC TGGGGTCCCC CCTGTGAGAA TGGCTGGAAG CTGGGGTCCC TCCTGGCGAC 8460

TGCAGAGCTG GCTGGCCGCG TGCCACTCTT GTGGGTGACC TGTGTCCTGG CCTCACACAC 8520

TGACCTCCTC CAGCTCCTTC CAGCAGAGCT AAGGCTAAGT GAGCCAGAAT GGTACCTAAG 8580

GGGAGGCTAG CGGTCCTTCT CCCGAGGAGG GGCTGTCCTG GAACCACCAG CCATGGAGAG 8640

GCTGGCAAGG GTCTGGCAGG TGCCCCAGGA ATCACAGGGG GGCCCCATGT CCATTTCAGG 8700

GCCCGGGAGC CTTGGACTCC TCTGGGGACA GACGACGTCA CCACCGCCCC CCCCCCATCA 8760

GGGGGACTAG AAGGGACCAG GACTGCAGTC ACCCTTCCTG GGACCCAGGC CCCTCCAGGC 8820

CCCTCCTGGG GCTCCTGCTC TGGGCAGCTT CTCCTTCACC AATAAAGGCA TAAACCTGTG 8880

CTCTCCCTTC TGAGTCTTTG CTGGACGACG GGCAGGGGGT GGAGAAGTGG TGGGGAGGGA 8940

GTCTGGCTCA GAGGATGACA GCGGGGCTGG GATCCAGGGC GTCTGCATCA CAGTCTTGTG 9000

ACAACTGGGG GCCCACACAC ATCACTGCGG CTCTTTGAAA CTTTCAGGAA CCAGGGAGGG 9060

ACTCGGCAGA GACATCTGCC AGTTCACTTG GAGTGTTCAG TCAACACCCA AACTCGACAA 9120

AGGACAGAAA GTGGAAAATG GCTGTCTCTT AGTCTAATAA ATATTGATAT GAAACTCAAG 9180

TTGCTCATGG ATCAATATGC CTTTATGATC CAGCCAGCCA CTACTGTCGT ATCAACTCAT 9240

GTACCCAAAC GCACTGATCT GTCTGGCTAA TGATGAGAGA TTCCCAGTAG AGAGCTGGCA 9300

AGAGGTCACA GTGAGAACTG TCTGCACACA CAGCAGAGTC CACCAGTCAT CCTAAGGAGA 9360

TCAGTCCTGG TGTTCATTGG AGGACTGATG TTGAAGCTGA AACTCCAATG CTTTGGCCAC 9420

CTGATGTGAA GAGCTGACTC ATTTGAAAAG ACCCTGATGC TGGGAAAGAT TGAGGGCAGG 9480

AGGAGAAGGG GACGACAGAG GATGAGATGG TTGGATGGCA TCACCAACAC AATGGACATG 9540

GGTTTGGGTG GACTCCAGGA GTTGGTGATG GACAGGGAGG CCTGGCGTGC TACGGAAGCG 9600

GTTTATGGGG TCACAAAGAC TGAGTGACTG AACTGAGCTG AACTGAATGG AAATGAGGTA 9660

TACAGCAAAG TGGGGATTTT TTAGATAATA AGAATATACA CATAACATAG TGTATACTCA 9720

TATTTTTATG CATACCTGAA TGCTCAGTCA CTCAGTCGTA TCTGACTCTG TGACCTATGG 9780

ACCGTAGCCT TCCAGGTTTC TTCTGTCCAC AGAATTCTCC AAGGCAAGAA TACTGGAGTG 9840

GGTAGCCATT TCCTCCTCCA GGGGATCCTC CCGACCCAGG GATTGAACCG GCATCTCCTG 9900

TATTGGCAGG TGGATTCTTT ACCACTGTGC CACCAGGGAA GCCCGTGTTA CTCTCTATGT 9960

CCCACTTAAT TACCAAAGCT GCTCCAAGAA AAAGCCCCTG TGCCCTCTGA GCTTCCCGGC 10020

CTGCAGAGGG TGGTGGGGGT AGACTGTGAC CTGGGAACAC CCTCCCGCTT CAGGACTCCC 10080

GGGCCACGTG ACCCACAGTC CTGCAGACAG CCGGGTAGCT CTGCTCTTCA AGGCTCATTA 10140

TCTTTAAAAA AAACTGAGGT CTATTTTGTG ACTTCGCTGC CGTAACTTCT GAACATCCAG 10200

TGCGATGGAC AGGACCTCCT CCCCAGGCCT CAGGGGCTTC AGGGAGCCAG CCTTCACCTA 10260

TGAGTCACCA GACACTCGGG GGTGGCCCCG CCTTCAGGGT GCTCACAGTC TTCCCATCGT 10320

CCTGATCAAA GAGCAAGACC AATGACTTCT TAGGAGCAAG CAGACACCCA CAGGACACTG 10380

AGGTTCACCA GAGCTGAGCT GTCCTTTTGA ACCTAAAGAC ACACAGCTCT CGAAGGTTTT 10440

CTCTTTAATC TGGATTTAAG GCCTACTTGC CCCTCAAGAG GGAAGACAGT CCTGCATGTC 10500

CCCAGGACAG CCACTCGGTG GCATCCGAGG CCACTTAGTA TTATCTGACC GCACCCTGGA 10560

ATTAATCGGT CCAAACTGGA CAAAAACCTT GGTGGGAAGT TTCATCCCAG AGGCCTCAAC 10620

CATCCTGCTT TGACCACCCT GCATCTTTTT TTCTTTTATG TGTATGCATG TATATATATA 10680

TATATATTTT TTTTTTTTTC ATTTTTTGGC TGTGCTGGCT GTTCGTTGCA GTTCGGTGCG 10740

CAGGCTTCTC TCTAGTTTCT CTCTAGTCTT CTCTTATCAC AGAGCAGTCT CTAGACGATC 10800

GACGCGT 10807

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 8:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 47 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 8:

AATTCCGATC GACGCGTCGA CGATATACTC TAGACGATCG ACGCGTA 47

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 9:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 24 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ

(B) KLON: BLGAMP3

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 9:

TGGATCCCCT GCCGGTGCCT CTGG 24

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 10:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 24 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: BLGAMP4

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 10:

AACGCGTCAT CCTCTGTGAG CCAG 24

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 11:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 10 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6839

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 11:

ACTACGTAGT 10

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 12:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 42 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6632

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 12:

CGACGCGGAT CCTACGTACC TGCAGCCATG TTTTCCATGA GG 42

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 13:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 21 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6627

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 13:

AGGGCTTCGG CAAGCTTCAG G 21

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 14:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 24 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6521

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 14:

GCCAAAGACT TACTTCCCTC TAGA 24

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 15:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 30 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6520

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 15:

GCATGAACGT CGCGTGGTGG TTGTGCTACC 30

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 16:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 30 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6519

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 16:

ACCACGCGAC GTTCATGCTC TAAAACCGTT 30

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 17:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 36 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6518

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 17:

GCTGCGGGAT CCTACGTACT AGGGGGACAG GGAAGG 36

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 18:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 45 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6629

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 18:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAAAGGATGG TTTCT 45

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 19:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 45 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6630

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 19:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAACATCTAT TATTG 45

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 20:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 21 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6625

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 20:

GTGAGATTTT CAGATCTTGT C 21

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 21:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 21 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6626

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 21:

AAGAATTACT GTGGCCTACC A 21

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 22:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 33 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6624

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 22:

GCTGCGGAAT TCTACGTACT ATTGCTGTGG GAA 33

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 23:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 45 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6514

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 23:

CGACGCGGAT CCTACGTACC TGCAGCCATG AGTTGGTCCT TGCAC 45

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 24:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 21 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6517

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 24:

GTCTCTGGTA GCAACATACT A 21

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 25:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 22 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6516

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 25:

GGGTTTCTAG CCCTACTAGT AG 22

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 26:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 22 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6515

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 26:

GGGTTTCTAG CCCTACTAGT AG 22

(2) INFORMACE PRO SEKVENCI IDENTIFIKAČNÍHO ČÍSLA 27:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 47 párů bází
(B) TYP: nukleová kyselina
(C) TYP VLÁKNA: jednoduché
(D) TOPOLOGIE: lineární

(vii) PŘÍMÝ ZDROJ:

(B) KLON: ZC6515

(xi) POPIS SEKVENCE: SEKVENCE IDENTIFIKAČNÍHO ČÍSLA 27:

AAGCTACGCG TCGATCGTCT AGAGTATATC GTCGACGCGT CGATCGG

PATENTOVÉ NÁROKY

1. Způsob produkce fibrinogenu, **vyznačující se tím**, že se:

poskytne první segment DNA kódující sekreční signál funkčně spojený s řetězcem Aα fibrinogenu, druhý segment DNA kódující sekreční signál funkčně spojený s řetězcem Bβ fibrinogenu a třetí segment DNA kódující sekreční signál funkčně spojený s řetězcem γ fibrinogenu, kde každý segment, tj. první, druhý i třetí, je funkčně spojený s dalšími segmenty DNA, které jsou nutné pro jeho expresi v mléčné žláze hostitelské samice savce,

zavedou segmenty DNA do oplodněného vejce savčích druhů kromě člověka, kde molární poměr segmentů DNA, prvního, druhého, a třetího, je v rozmezí 0,5 až 1 : 0,5, až 1 : 0,5 až 1,

vloží vejce do vejcovodu savčích druhů kromě člověka, nebo dělohy samice druhu, aby se získalo potomstvo nesoucí konstrukty DNA,

chovají potomci pro produkci samičího potomstva, které exprimuje první, druhý a třetí segment DNA a produkuje mléko obsahující biologicky způsobilý fibrinogen kódovaný segmenty, sbírá mléko od samičího potomstva, a získá fibrinogen z mléka.

2. Způsob podle nároku 1, **vyznačující se tím**, že druh savce se vybere ze skupiny skládající se z ovcí, prasat, koz a hovězího dobytka.

3. Způsob podle nároku 1, **vyznačující se tím**, že každý segment DNA, tj. první, druhý a třetí, obsahuje intron.

4. Způsob podle nároku 1, **vyznačující se tím**, že každý segment DNA, tj. první, druhý a třetí, je funkčně spojen s transkripčním promotorem vybraným ze skupiny skládající se z promotorů genů kaseinu, β–laktoglobulinu, α–laktalbuminu a kyselého proteinu syrovátky.

5. Způsob podle nároku 1, **vyznačující se tím**, že exprese každého segmentu DNA, tj. prvního, druhého a třetího, se řídí promotorem B–laktoglobulinu.

6. Způsob podle nároku 1, **vyznačující se tím**, že zaváděcí krok zahrnuje injekci prvního, druhého a třetího segmentu DNA do pronuklea oplodněného vejce.

7. Způsob podle nároku 1, **vyznačující se tím**, že fibrinogen je lidský fibrinogen.

8. Způsob podle nároku 1, **vyznačující se tím**, že druhý segment DNA obsahuje sekvenci nukleotidů, jak je uvedena v sekvenci identifikačního čísla 3 od nukleotidu 470 do nukleotidu 8100.

9. Způsob podle nároku 1, **vyznačující se tím**, že druhý segment DNA obsahuje sekvenci nukleotidů, jak je uvedena v sekvenci identifikačního čísla 3 od nukleotidu 512 do nukleotidu 8100.

10. Způsob produkce fibrinogenu podle nároku 1, **vyznačující se tím**, že se:

včlení první segment DNA kódující sekreční signál funkčně spojený s řetězcem Aαfibrinogenu do β–laktoglobulinového genu za vzniku první genové fúze,

včlení druhý segment DNA kódující sekreční signál funkčně spojený s řetězcem Bβ fibrinogenu do β–laktoglobulinového genu za vzniku druhé genové fúze,

včlení třetí segment DNA kódující sekreční signál funkčně spojený s řetězcem γ fibrinogenu do β–laktoglobulinového genu za vzniku třetí genové fúze,

zavede první, druhá a třetí genová fúze do zárodečné linie savce kromě člověka, kde molární poměr segmentů DNA, prvního, druhého a třetího, je v rozmezí 0,5 až 1 : 0,5 až 1 : 0,5 až 1 tak, že segmenty DNA jsou exprimovány v mléčné žláze savce nebo jeho samičího potomstva a biologicky způsobilý fibrinogen je sekretován do mléka savce nebo jeho samičího potomstva, získávání mléka od savce nebo jeho samičího potomstva a získávání fibrinogenu z mléka.

11. Způsob podle nároku 10, **vyznačující se tím**, že savec je ovce, prase, koza nebo hovězí dobytek.

12. Způsob podle nároku 10, **vyznačující se tím**, že každá genová fúze, tj. první, druhá a třetí, obsahuje intron.

13. Způsob podle nároku 10, **vyznačující se tím**, že zaváděcí krok zahrnuje injekci první, druhé a třetí genové fúze do pronuklea oplodněného vejce a vložení vejce do vejcovodu falešně březí samice, aby se produkovalo samičí potomstvo nesoucí v zárodečné linii genové fúze.

14. Způsob produkce savčího embrya, kromě člověka, obsahujícího v jádře heterologní segmenty DNA kódující řetězce Aα, Bβ a γ fibrinogenu, **vyznačující se tím**, že se:

poskytne první segment DNA kódující sekreční signál funkčně spojený s řetězcem Aα fibrinogenu, druhý segment DNA kódující sekreční signál funkčně spojený s řetězcem Bβ fibrinogenu a třetí segment DNA kódující sekreční signál funkčně spojený s řetězcem γ fibrinogenu, kde každý segment, tj. první, druhý a třetí, je funkčně spojen s dalšími segmenty DNA nutnými pro jeho expresi v mléčné žláze hostitelské samice savce a sekreci do mléka hostitelské samice savce, kde molární poměr segmentů DNA, prvního, druhého a třetího, je v rozmezí až 1 : 0,5 až 1 : 0,5 až 1,

zavedou segmenty DNA do oplodněného vejce savčích druhů kromě člověka,

vloží vejce do vejcovodu nebo dělohy samice druhu kromě člověka, aby se získalo potomstvo nesoucí první, druhý a třetí segmenty DNA.

15. Způsob přípravy podle nároku 14, **vyznačující se tím**, že potomek je samice.

16. Způsob přípravy podle nároku 14, **vyznačující se tím**, že potomek je samec.

17. Savec kromě člověka, který je vytvořen způsobem podle nároku 14.

18. Savec kromě člověka podle nároku 17, který je samice.

19. Samice kromě člověka podle nároku 18, která produkuje mléko obsahující biologicky způsobilý fibrinogen kódovaný segmenty DNA.

20. Savec kromě člověka podle nároku 17, který je samec.

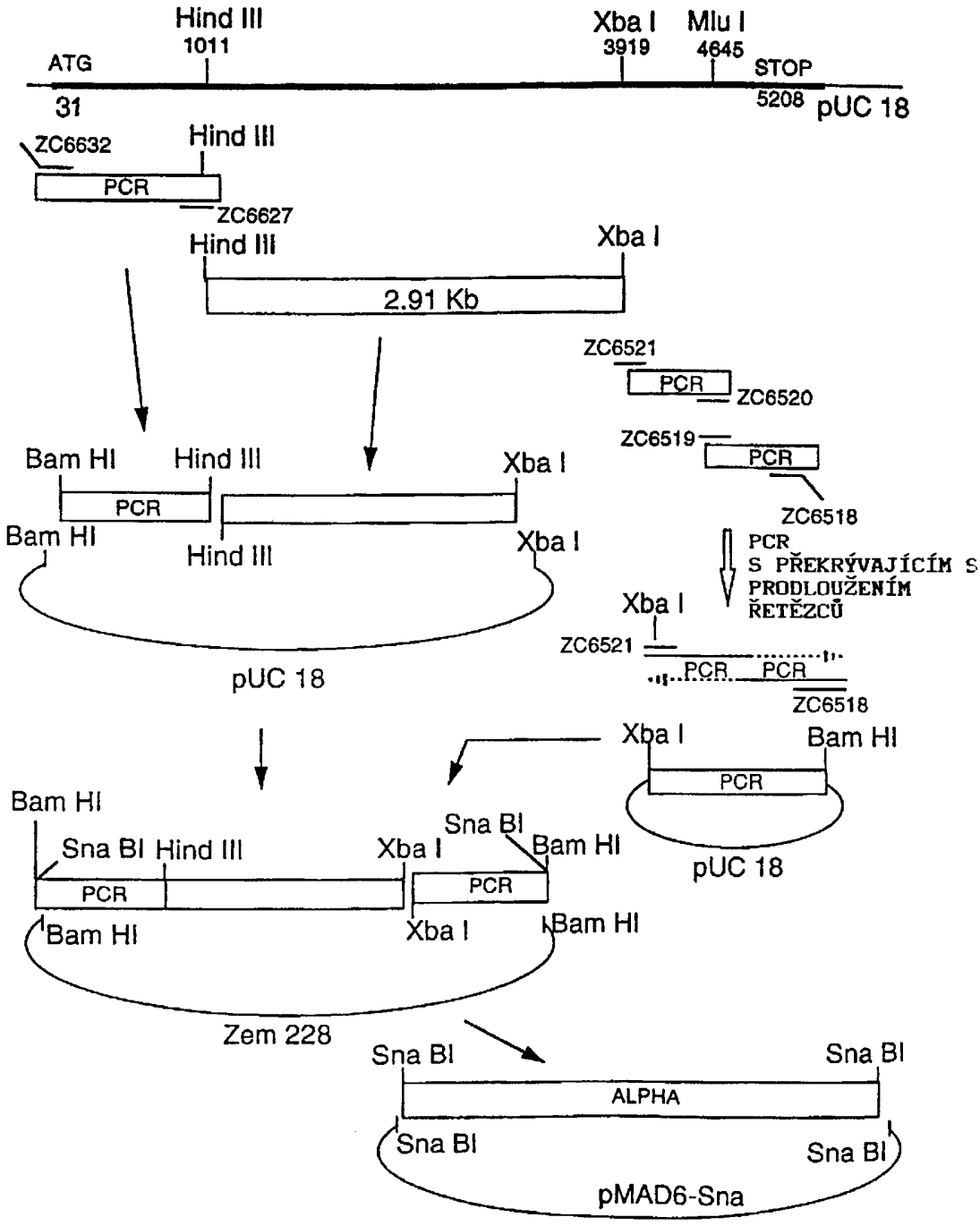
21. Savec kromě člověka, který nese ve své zárodečné linii segmenty DNA kódující řetězce Aα, Bβ a γ fibrinogenu, kdy samičí potomstvo savce exprimuje segmenty DNA v mléčné žláze za produkce biologicky způsobilého fibrinogenu.

22. Savec podle nároku 21, který je samice.

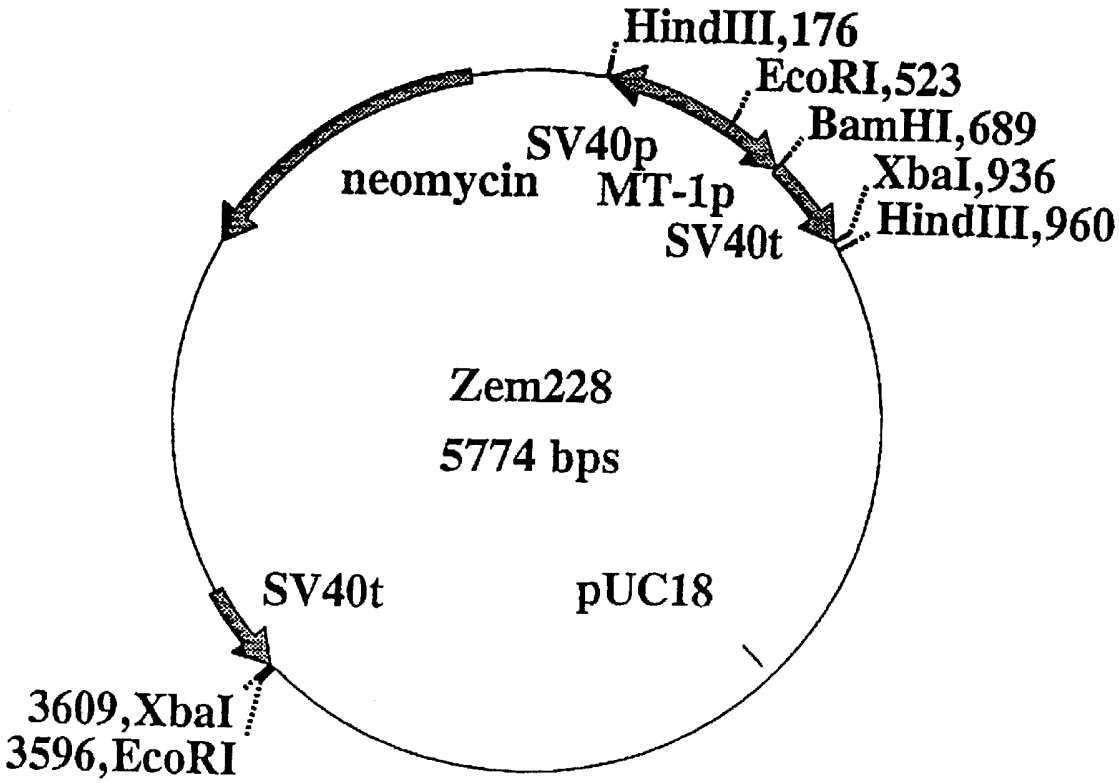
23. Savec podle nároku 21, který je samec.

5 výkresů

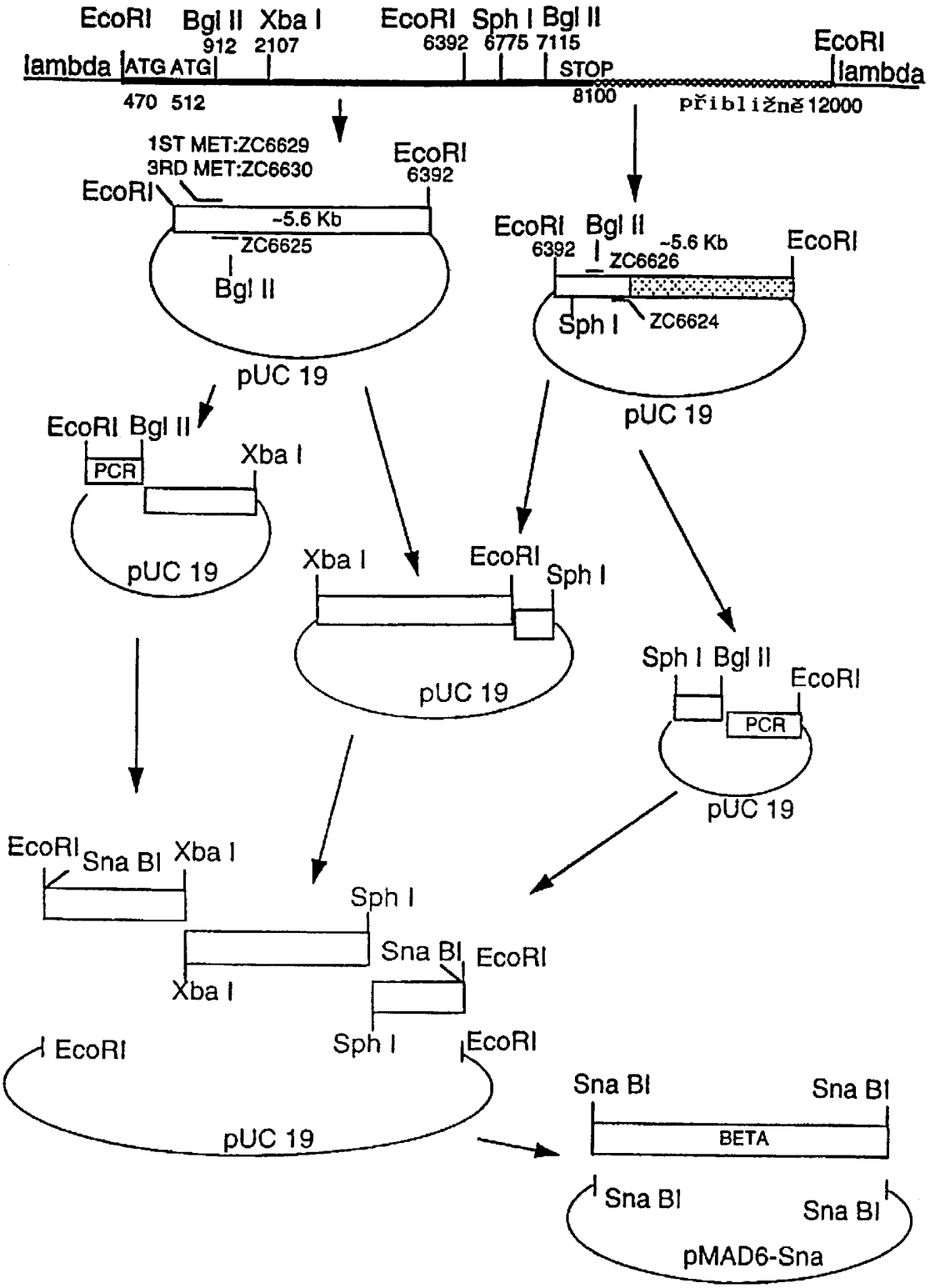
Obr. 1


Hind III
1011
Xba I
3919
Mlu I
4645
ATG
31
STOP
5208
pUC 18
ZC6632
Hind III
PCR
ZC6627
Hind III
Xba I
2.91 Kb
ZC6521
PCR
ZC6520
ZC6519
PCR
ZC6518
Bam HI
Hind III
Xba I
PCR
Bam HI
Hind III
Xba I
pUC 18
PCR
S PŘEKRÝVAJÍCÍM S
PRODLOUŽENÍM
ŘETĚZCŮ
Xba I
ZC6521
PCR
PCR
ZC6518
Xba I
Bam HI
PCR
pUC 18
Bam HI
Sna BI
Hind III
Xba I
Sna BI
Bam HI
PCR
PCR
Bam HI
Xba I
Bam HI
Zem 228
Sna BI
Sna BI
ALPHA
Sna BI
Sna BI
pMAD6-Sna

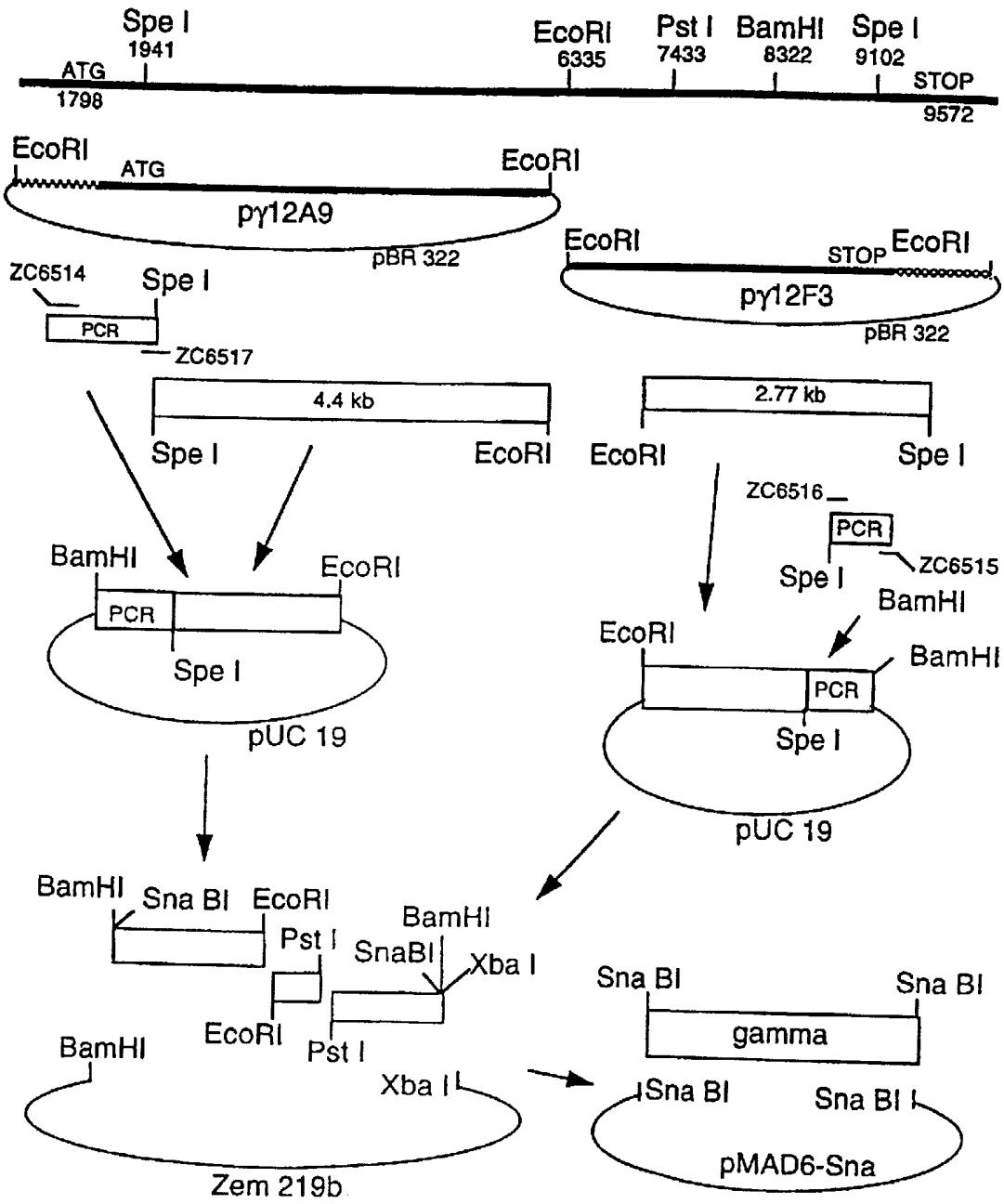
Obr. 2


HindIII,176
EcoRI,523
BamHI,689
XbaI,936
HindIII,960
SV40p
MT-1p
SV40t
neomycin
Zem228
5774 bps
SV40t
pUC18
3609,XbaI
3596,EcoRI

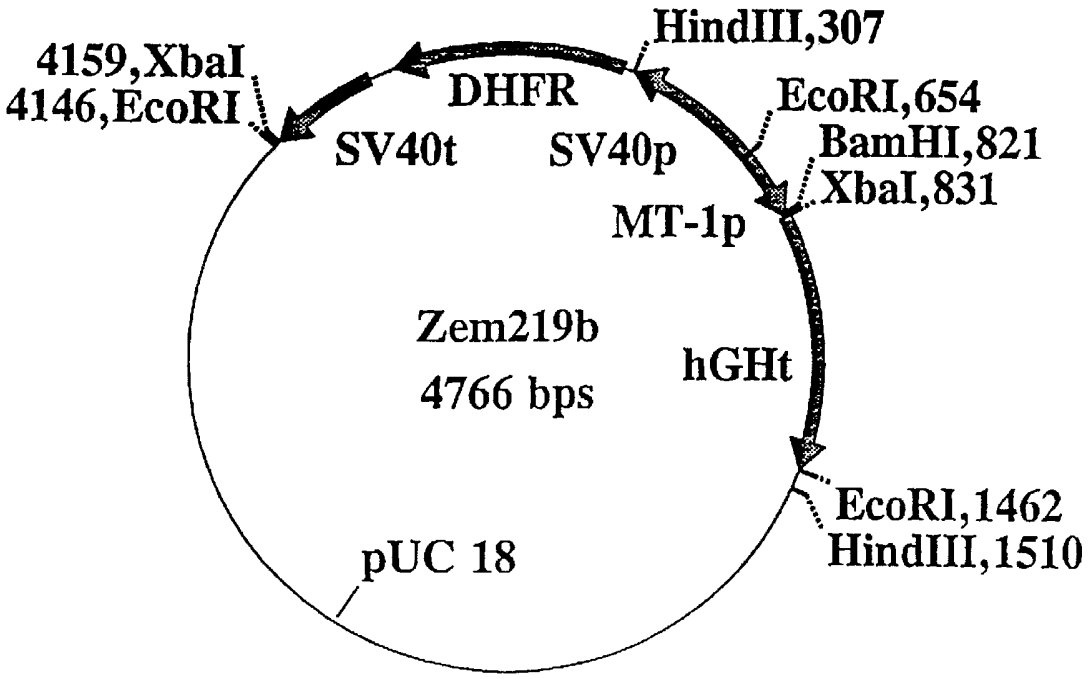
Obr. 3


EcoRI
Bgl II
Xba I
912
2107
EcoRI
Sph I
Bgl II
6392
6775
7115
lambda
ATG ATG
STOP
EcoRI
lambda
470
512
8100
přibližně 12000
1ST MET:ZC6629
3RD MET:ZC6630
EcoRI
6392
EcoRI
~5.6 Kb
ZC6625
Bgl II
pUC 19
EcoRI
Bgl II
6392
~5.6 Kb
ZC6626
EcoRI
Sph I
ZC6624
pUC 19
EcoRI
Bgl II
PCR
Xba I
pUC 19
Xba I
EcoRI
Sph I
pUC 19
Sph I
Bgl II
EcoRI
PCR
pUC 19
EcoRI
Sna BI
Xba I
Sph I
Sna BI
EcoRI
Xba I
Sph I
EcoRI
EcoRI
pUC 19
Sna BI
Sna BI
BETA
Sna BI
Sna BI
pMAD6-Sna

Obr. 4


Spe I
1941
ATG
1798
EcoRI
6335
Pst I
7433
BamHI
6322
Spe I
9102
STOP
9572
EcoRI
ATG
EcoRI
pγ12A9
pBR 322
EcoRI
STOP
EcoRI
pγ12F3
pBR 322
ZC6514
Spe I
PCR
ZC6517
4.4 kb
Spe I
EcoRI
2.77 kb
EcoRI
Spe I
ZC6516
PCR
Spe I
ZC6515
BamHI
EcoRI
PCR
Spe I
pUC 19
BamHI
EcoRI
BamHI
PCR
Spe I
pUC 19
BamHI
Sna BI
EcoRI
Pst I
BamHI
SnaBI
Xba I
EcoRI
Pst I
BamHI
Xba I
Zem 219b
Sna BI
Sna BI
gamma
Sna BI
Sna BI
pMAD6-Sna

Obr. 5



Konec dokumentu