

PŘIHLÁŠKA VYNÁLEZU

zveřejněná podle § 31 zákona č. 527/1990 Sb.

(21) Číslo dokumentu:

966-99

(19)

ČESKÁ
REPUBLIKA



ÚŘAD
PRŮMYSLOVÉHO
VLASTNICTVÍ

(22) Přihlášeno: **29. 09. 97**

(32) Datum podání prioritní přihlášky: **01.10.96, 16.01.97**

(31) Číslo prioritní přihlášky: **96/027633, 97/97100583**

(33) Země priority: **US, EP**

(40) Datum zveřejnění přihlášky vynálezu: **11. 08. 99**
(Věstník č. 8/99)

(86) PCT číslo: **PCT/EP97/05355**

(87) PCT číslo zveřejnění: **WO 98/14568**

(13) Druh dokumentu: **A3**

(51) Int. Cl.⁶:

C 12 N 15/12
C 12 N 5/10
C 07 K 14/47
C 07 K 16/18
C 12 Q 1/68
A 61 K 38/18
A 01 K 67/027

(71) Přihlášovatel:

RAPPOLD-HOERBRAND Gudrun,
Heidelberg, DE;

(72) Původce:

Rappold-Hoerbrand Gudrun, Heidelberg, DE;
Rao Ercole, Riedstadt, DE;

(74) Zástupce:

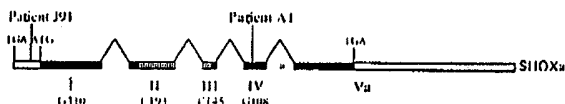
PATENTSERVIS PRAHA a.s., Jivenská 1,
Praha 4, 14000;

(54) Název přihlášky vynálezu:

**Isolovaná lidská molekula
nukleové kyseliny kodující polypeptidy,
růstový protein a jeho použití a způsob**

(57) Anotace:

Popisuje se molekula izolované lidské nukleové kyseliny, která kóduje polypeptidy, jež obsahují oblast homeoboxu tvořenou 60 aminokyselinami a mají aminokyselinovou sekvenci SEQ ID NO: 1 a vykazují regulační aktivitu lidského růstu. Identifikovaly se tři nové geny nacházející se v kritické oblasti malé výšky postavy o přibližné velikosti 500 kb na chromozomu X a Y. Alespoň jeden z těchto genů je odpovědný za fenotyp s malou výškou postavy. DNA odpovídající tomuto genu se může použít jako diagnostický nástroj při další charakterizaci molekulární podstaty fenotypu s malou výškou postavy. Identifikace genového produktu genu poskytuje nový způsob vývoje dokonalejší léčby malé výšky postavy.



CZ 966-99 A3

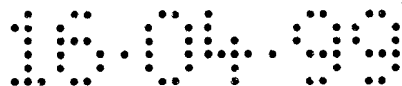
Molekula nukleové kyseliny kódující polypeptidy a způsoby jejího použití.

Oblast techniky

Vynález se týká izolace, identifikace a charakterizace nově identifikovaných lidských genů, které odpovídají za vady lidského růstu, zvláště za malou výšku postavy nebo Turnerův syndrom. Dále se vynález týká diagnózy a terapie takových vad. Izolovaná genomová DNA nebo její fragmenty se mohou použít pro farmaceutické účely nebo jako diagnostické nástroje nebo činidla pro identifikaci nebo charakterizaci genetických defektů, které se podílejí na uvedených vadách. Vynález dále popisuje lidské růstové proteiny (transkripční faktory A, B a C), které se exprimují po transkripci uvedené DNA na RNA nebo mRNA a které se mohou použít při terapii vad spojených s mutacemi v uvedených genech. Dále se popisují vhodné sekvence cDNA, které se mohou použít při přípravě rekombinantních proteinů vhodných pro léčbu takových vad. Předmětem vynálezu jsou dále plazmidové vektory vhodné pro expresi DNA těchto genů a vhodné buňky, které obsahují takovou DNA. Dále se popisují metody genetické léčby uvedených vad v oboru molekulární medicíny za použití expresivního plazmidu připraveného inkorporací DNA downstream expresivního promotoru, který umožňuje expresi v savčí hostitelské buňce.

Dosavadní stav techniky

Růst je základní aspekt vývoje organismu. Je regulovaný vysoce organizovaným a komplexním systémem. Výška postavy je multifaktorový rys, který je ovlivněn faktory prostředí a genetickými faktory. Vývojové malformace, které zahrnují výšku postavy, jsou běžným jevem u lidí všech ras. V případě lidí je četnost výskytu 3 ze 100 a retardace růstu vede k malé



velikosti postavy, za což ve většině případů odpovídají vrozené vady.

Turnerův syndrom je běžná chromozomální vada (Rosenfeld RG, Tesch L-G, Rodriguez-Rigau LJ, McCauley E, Albertsson-Wikland K, Asch R, Cara J, Conte F, Hall JG, Lippe B, Nagel TC, Neely EK, Page DC, Ranke M, Saenger P, Watkins JM, Wilson DM (1994): Recommendations for diagnosis, treatment, and management of individuals with Turner syndrom. *The Endocrinologist* 4(5): 351-358) s četností výskytu 1:2 500 u živě narozených žen. Odhaduje se, že 1 až 2 % oplodnění u lidí má karyotyp 45,X a že až 99 % takových plodů nepřichází na svět v termínu (Hall J.G., Gilchrist D.M. (1990): Turner syndrome and its variants. *Pediatr. Clin. North Am.* 37: Robinson A (1990): Demography and prevalence of Turner syndrome. In: *Turner Syndrome.*, edited by Rosenfield R.G., Grumbach M.M., pp. 93 - 100, New York, Marcel Dekker). U lidí s Turnerovým syndromem (nebo s Ullrich-Turnerovým syndromem) existuje podstatná klinická variabilita fenotypu (Ullrich, 1930; Turner, 1938). Malá výška postavy je konsistentní nález a spolu s gonádovou dysgenezí se považuje za vedoucí symptom uvedených vad. Turnerový syndrom je skutečná multifaktorová vada. Uvažuje se, že za embryonální letalitu, malou výšku postavy, gonádovou dysgenezi a charakteristické somatické rysy odpovídají monosomy genů, které jsou běžné u X a Y chromozomů. Ukazuje se, že pro normální vývoj člověka jsou nutné diploidy těchto homologních genů X-Y. Očekává se, že Turnerovy geny (nebo anti-Turnerovy geny) se u žen exprimují jak z aktivního tak z neaktivního chromozomu X nebo chromozomu Y, aby se zajistila správná dávka genového produktu. Haplo-nedostatečnost (nedostatečnost způsobená pouze jednou aktivní kopií) naznačuje mechanismus vedoucí k onemocnění.

Dosud se objasnila řada mechanismů, které vedou k malé výšce postavy. Nedostatečnost růstového hormonu a receptorů růstového hormonu, stejně jako vady skeletu se popisují jako

příklady fenotypu s malou výškou postavy (Martial et al., 1979; Phillips et al., 1981; Leung et al., 1987; Goddard et al., 1995). Identifikovaly se mutace ve třech genech, které kódují receptor lidského fibroblastového růstového faktoru (FGFR 1-3) a způsobují řadu vad skeletu. Tyto vady zahrnují většinu forem trpasličího vzrůstu, achondroplázie (Shiang et al., 1994; Rousseau et al., 1994; Muenke and Schell, 1995). Dobře známá a častá chromozomální vada (objevuje se u jedné z 2 500 žen) Turnerův syndrom (45,X) je také spojena s malou výškou postavy. Všechny tyto důvody však odpovídají pouze za malou část pacientů s malou výškou postavy. Většina případů zatím zůstává nevysvětlena.

Věří se, že chromozomy X a Y nesou geny, které ovlivňují výšku postavy (Ogata T, Tyler-Smith C, Purvis-Smith S, Turner G. (1993): Chromosomal localisation of a gene(s) for Turner stigmata on Yp. *J. Med. Genet.* 30: 918-922). To se může u pacientů s abnormalitami pohlavních chromozomů odvodit ze vztahu genotyp-fenotyp. Cytogenetické studie dokazují, že delece konců krátkých ramen buď chromozomu X nebo Y vedou u některých jedinců k malé výšce postavy (Zuffardi et al., 1982); Curry et al., 1984). Uvádí se více než 20 uspořádání chromozomu, které se spojují s delecemi konců chromozomu Xp a Yp a které nesou gen(y) odpovědné za malou výšku postavy, s pseudoautozomální oblastí (PAR1) (Ballabio A., Bardoni A., Carrozzo R., Andria G., Bick D., Campbell L., Hamel B., Ferguson-Smith M.A., Gimelli G., Fraccaro M., Maraschio P., Zuffardi O., Guilo S., Camerino G., (1989): Contiguous gene syndromes due to deletions in the distal short arm of the arm of the human X chromosome. *Proc. Natl. Acad. Sci USA* 86: 10001-10005.; Schaefer L., Ferrero G.B., Grillo A., Bassi M.T. Roth E.J. Wapenaar M.C., Van Ommen -J.B., Mohandas T.K., Rocchi M., Zoghbi H.Y. Ballabio A. (1993): A high resolution deletion map of human chromosome Xp22. *Nature genetics* 4: 272-279). Pozice těchto genů se zúžila na distální DNA o velikosti

700 kb oblasti PAR1, kterou lemuje marker DXYS15 (Ogata T., Goodfellow P., Petit C., Aya M., Matsuo N. (1992): Short stature in a girl with a terminal Xp deletion distal to DXYS15: localization of a growth gene(s) in the pseudoautosomal region. *J. Med. Genet.* 29: 455-459; Ogata T., Yoshizawa A., Muroya K., Matsuo N., Fukushima Y., Rappold G.A., Yokoya S. (1995): Short stature in a girl with partial monosomy of the pseudoautosomal region distal to DXYS15: further evidence for the assignment of the critical region for a pseudoautosomal growth gene(s). *J. Med. Genet.* 32: 831-834; Ogata T., Matsuo N. (1995): Turner syndrome and female sex chromosome aberrations: deduction of the principle factors involved in the development of clinical features. *Hum. Genet.* 95: 607-629).

Regulace růstu u savců je organizována jako komplexní systém. Je možné, že vysoce organizovaným způsobem spolu reaguje více genů (proteinů), které podporují růst. Jeden z těchto genů, který odpovídá za výšku postavy, se zmapoval až do pseudoautozomální oblasti PAR1 (Ballabio A., Bardoni A., Carrozzo R., Andria G., Bick D., Campbell L., Hamel B., Ferguson-Smith M.A., Gimelli G., Fraccaro M., Maraschio P., Zuffardi O., Guilo S., Camerino G., (1989): Contiguous gene syndromes due to deletions in the distal short arm of the arm of the human X chromosome. *Proc. Natl. Acad. Sci USA* 86: 10001-10005), což je oblast, o které se ví, že je volně zaměnitelná mezi chromozomy X a Y (Rappold G.A. (1993): The pseudoautosomal region of the human sex chromosomes. *Hum. Genet.* 92: 315-324). Celá oblast RAP1 je přibližně 2 700 kb velká.

U pacientů s delecí se definovala kritická oblast ovlivňující výšku postavy. K malé výšce postavy dochází v případě, když se deletuje celá oblast o velikosti 700 kb nebo když je specifický gen v této kritické oblasti přítomen v haploidním stavu nebo když je přerušen nebo mutován (jako je případ



idiotypické malé výšky postavy nebo Turnerův syndrom). Frekvence výskytu Turnerova syndromu ve světě je jedna z 2 500 žen; četnost výskytu tohoto druhu idiopatické malé výšky postavy se odhaduje na 1 případ ze 4 000 až 5 000 osob. Ženám s Turnerovým syndromem a některým jednotlivcům s malou výškou postavy se obvykle po několik let až po několik desítek let aplikuje růstový hormon (GH), ačkoli je známo, že jejich hladina růstového hormonu v těle je normální a nedostatečnost GH zde není problémem. Léčba takových pacientů je velmi nákladná (odhaduje se přibližně na 30 000 USD ročně). Proto je nutné získat způsob rozlišení pacientů malého vzrůstu, kteří mají genetický defekt a pacienti, kteří nevykazují genetický defekt uvedeného genu. U pacientů s genetickým defektem tohoto genu - buď s úplnou delecí genu (jako je v případě Turnerova syndromu) nebo s bodovou mutací (jako je tomu u idiopatické malé výšky postavy)- je možné aplikovat alternativní léčbu bez lidského růstového hormonu.

Korelace genotyp/fenotyp podporují existenci růstového genu v proximální části Yq a v distální části Yp. Malá výška postavy se také objevuje u jedinců s terminálními delecemi Xp. Proběhlo rozsáhlé vyhledávání pacientů s částečnými monosomiemi pseudoautozomální oblasti. Na základě korelací genotyp-fenotyp se stanovila minimální běžná delece oblasti DNA o velikosti 700 kb (Ogata T., Matsuo N. (1995): Turner syndrome and female sex chromosome aberrations: deduction of the principle factors involved in the development of clinical features. Hum. Genet. 95: 607-629; Ogata T., Goodfellow P., Petit C., Aya M., Matsuo N. (1992): Short stature in a girl with a terminal Xp deletion distal to DXYS15: localization of a growth gene(s) in the pseudoautosomal region. J. Med. Genet. 29: 455-459; Ogata T., Yoshizawa A., Muroya K., Matsuo N., Fukushima Y., Rappold G.A., Yokoya S. (1995): Short stature in a girl with partial monosomy of the pseudoautosomal region distal to DXYS15: further evidence for the assignment of the

critical region for a pseudoautosomal growth gene(s). *J. Med. Genet.* 32: 831-834). Zjistilo se, že uvedená oblast leží mezi genetickými markery DXYS20 (3cosPP) a DXYS15 (113D) a u všech genů, které aspirují na řízení růstu, a leží v oblasti PAR1 (například receptor hemopoietického růstového faktoru; CSF2RA) (Gough N.M., Gearing D.P., Nicola N.A., Baker E., Pritchard M., Callen D.F., Sutherland G.R. (1990). Localization of the human GM-CSF receptor gene to the X-Aby pseudoautosomal region. *Nature* 345: 734736), se odstranily na základě jejich fyzikální pozice (Rappold G. A., Wilson T.A., Henke A., Gough N.M. (1992): Arrangement and localization of the human GM-CSF receptor α chain gene CSF2RA within the X-Aby pseudoautosomal region. *Genomics* 14: 455-461). Geny se nacházely v deleční oblasti o velikosti 700 kb PAR1 oblasti o velikosti 2 700 kb. Nedávno se objevily u jednotlivců s malým vzrůstem delece pseudoautozomální oblasti (PAR1) pohlavních chromozomů a následně se také definovala minimální běžná oblast delece o velikosti 700 kb. Na základě analýzy DNA Southernovým přenosem u pacientů AK a SS za použití odlišných pseudoautozomálních markerů se identifikovala Xp terminální delece přibližně 700 kb od DXYS15 (113D) (Ogata T., Matsuo N. (1995): Turner syndrome and female sex chromosome aberrations: deduction of the principle factors involved in the development of clinical features. *Hum. Genet.* 95: 607-629; Ogata T., Goodfellow P., Petit C., Aya M., Matsuo N. (1992): Short stature in a girl with a terminal Xp deletion distal to DXYS15: localization of a growth gene(s) in the pseudoautosomal region. *J. Med. Genet.* 29: 455-459; Ogata T., Yoshizawa A., Muroya K., Matsuo N., Fukushima Y., Rappold G.A., Yokoya S. (1995): Short stature in a girl with partial monosomy of the pseudoautosomal region distal to DXYS15: further evidence for the assignment of the critical region for a pseudoautosomal growth gene(s). *J. Med. Genet.* 32: 831-834).

Podstata vynálezu

Oblast genu odpovídající za malý vzrůst postavy se identifikovala jako oblast o velikosti přibližně 500 kb , s výhodou o velikosti přibližně 170 kb v PAR1 oblasti chromozomů X a Y. V této oblasti se identifikovaly tři geny jako kandidáti, kteří odpovídají za malý vzrůst. Tyto geny se označily SHOX (také jako SHOX93 nebo HOX93) (SHOX znamená gen obsahující homeobox malého vzrůstu), pET92 a SHOT (gen nesoucí homeobox podobný SHOX na chromozomu 3). Zvláště významný je gen SHOX, který má dvě oddělená místa sestřihu, což vede ke vzniku dvou variací (SHOXa a b). Během předchozích výzkumů se analyzovala podstatná část nukleotidové sekvence genu malého vzrůstu (SEQ ID NO: 8). Mohou se předpovědět a identifikovat exony nebo jejich části (například exon I [G310]; exon II [ET93]; exon IV [G108]; pET92). Získané informace o sekvenci se pak mohou použít pro navržení vhodných primerů nebo nukleotidových sond, které hybridizují s částmi genu SHOX nebo jeho fragmenty. Gen SHOX se pak může izolovat běžnými metodami. Na základě další analýzy sekvence DNA genů zodpovědných za malý vzrůst se izoluje nukleotidová sekvence exonu I až V (uvedeno na obrázku č. 1 až 2). Gen SHOX obsahuje homeobox (SEQ ID NO: 1) o velikosti přibližně 180 bp (obrázek č. 2 a obrázek č. 3). Začíná nukleotidem, který kóduje aminokyselinu v pozici 117 (Q), končí nukleotidem kódujícím aminokyselinu v pozici 176 (E), to znamená od CAG(440) do GAG(619). U pacientů z Německa (A1) s malou výškou postavy a u pacientů Japonska, kde se testovalo 250 jednotlivců s idiopatickým malým vzrůstem, se identifikoval homeobox jako sekvence homeobox-pET93(SHOX) a zjistily se dvě bodové mutace. Obě bodové mutace se našly ve stejné pozici a vedou ke zkrácení proteinu v pozici aminokyseliny 195. To naznačuje, že zde může existovat tzv. horké místo. Na základě skutečnosti, že obě mutace, které vedou ke zkrácení proteinu, se vyskytují

ve stejné pozici, je pravděpodobné, že putativní horké místo rekombinace existuje v exonu 4 (G108). Primery specifické pro exon jsou například GCA CAG CCA ACC ACC TAG (forward) nebo TGG AAA GGC ATG ATC CGT AAG (reverzní). Shora uvedený nový gen obsahující homeobox, SHOX, který se nachází v intervalu 170 kb, je alternativně sestřižen a dává vznik dvěma proteinům s odlišnými funkcemi. Při demonstraci skutečnosti, že malá výška postavy je způsobena mutacemi v SHOX, se použila analýza mutace a sekvenování DNA.

Identifikace a klónování kritické oblasti malého vzrůstu podle vynálezu se uskutečnila následujícím způsobem. U 15 jednotlivců s částečnými monosomy v pseudoautozomální oblasti (PAR1) se provedla rozsáhlá fyzikální mapovací studie. Korelací výšky těchto jedinců s delečními body se definovala kritická oblast malého vzrůstu (SS) o přibližné velikosti 700 kb. Tato oblast se následně klonovala jako překrývající se kosmidový kontig za použití kvasinkových umělých chromozomů (YAC) z PAR1 (Ried K., Mertz A., Nagaraja R., Trusnich M., Riley J., Anand R., Page D., Lehrach H., Elloiso J., Rappold G.A. (1995): Characterization of a yeast artificial chromosome contig spanning the pseudoautosomal region. *Genomics* 29: 787-792) a pomocí metody „chodícího kosmidu“ (cosmid walking). Za účelem vyhledání genů odpovědných za SS v tomto intervalu se na přibližně 600 kb velkou oblast mezi distálním koncem kosmidu 56G10 a proximálním koncem 51D11 aplikovala řada metod. Za použití selekce cDNA, trapování exonu a CpG klónování se identifikovaly dva nové geny.

Pozice kritického intervalu malého vzrůstu se zúžila na menší interval DNA o velikosti 170 kb tím, že se charakterizovaly tři další jedinci (GA, AT a RY), kteří vykazují malý vzrůst postavy. Za účelem přesného stanovení pozice bodu přerušení u těchto jedinců se provedla na chromozomech v metafázi za použití kosmidů vhodných pro kontig fluorescenční hybridizace *in situ* (FISH). Pacient GA, který vykazuje terminální delecii a

normální výšku vzrůstu, definuje distální hranici kritické oblasti (bod přerušeni se nachází na kosmidu 110E3) a pacient AT vykazující inverzi X chromozomu a normální výšku vzrůstu definoval proximální hranici (bod přerušeni se nachází na kosmidu 34F5). Bod přerušeni chromozomu Y u pacienta RY, který vykazuje terminální delecii a malý vzrůst, se také nachází na kosmidu 34F5, což naznačuje, že tato oblast obsahuje sekvence s predispozicí k přestavění chromozomů.

Celá oblast ohraničená Xp/Yp se klónovala jako sada přesahujících kosmidů. Ke studiu šesti pacientů s přestavěným chromozomem X, kdy tři z nich jsou normálního vzrůstu a tři mají malou výšku vzrůstu, se použila fluorescenční hybridizace *in situ* (FISH) s kosmidy s této oblasti. Korelace genotypu-fenotypu zúžily kritický interval pro malý vzrůst na 270 kb DNA nebo na dokonce menší interval, jako je 170 kb, který obsahuje gen nebo geny důležité pro růst člověka. Tento interval pokrývá minimálně šest až osm kosmidů a jsou nyní dostupné pro hybridizaci FISH interfáze a metafáze, která je dostupný nástroj pro diagnostiku pacientů s idiopatickým malým vzrůstem postavy.

Skutečnost, že cílový gen vede k vadám růstu u lidí (např. oblast malého vzrůstu), nebyla dříve známa. Biologická a klinická asociace pacientů s uvedenou delecí vysvětluje funkci uvedeného genu. V této studii se použila fluorescenční *in situ* hybridizace (FISH) za účelem testování metafáze a interfáze jádra lymfocytu u šesti pacientů. Cílem bylo testovat u všech přesahujících kosmidů jejich použitelnost jako sondy pro FISH hybridizaci a stanovit oblasti porušení genu u všech čtyř případů, a tímto způsobem stanovit minimální kritickou oblast genu malého vzrůstu.

Duplikace nebo delece genomové DNA se může uskutečnit kvantitativní PCR nebo odhadem dávky na Southernových blotech nebo použitím RFLP.

Zvláště vhodnou metodou pro přesné rozlišení mezi jednou a zdvojenou dávkou markerů je hybridizace FISH, přičemž její klinická aplikace je v současnosti rutinní záležitostí. Zatímco při FISH hybridizaci interfáze se může zhodnotit absence nebo přítomnost molekulového markeru, hybridizace FISH metafáze chromozomů může poskytnout semi-kvantitativní stanovení delecí uvnitř kosmidu. Uvedeným způsobem se může stanovit delece o velikosti přibližně 10 kb (což odpovídá 25 % redukci signálu). Zvláště je důležité, že prakticky všechny geny onemocnění na lidském X chromosomu jsou spojeny s menšími a většími delecemi v rozmezí od několika kilobází do několika megabází DNA (Nelson, D.L., Ballabio A., Cremers F., Monaco A.P., Schlessinger D., (1995). -Report of the sixth international workshop on the X chromosome mapping. Cytogenet. Cell Genet. 71: 308-342).

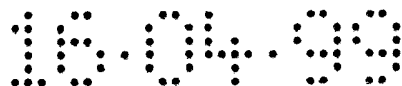
Vynález dále popisuje sekvence DNA nebo jejich fragmenty, které jsou části genů odpovědných za lidský růst (nebo v případě defektu v těchto genech odpovídají za malý vzrůst). Identifikovaly se tři geny odpovědné za lidský růst. Jsou to SHOX, pET92 a SHOT). Sekvence DNA nebo fragmenty těchto genů stejně jako sekvence DNA těchto genů v plné délce se mohou transformovat do vhodného vektoru a mohou se dále transfekovat do buněk. Když se takové vektory zavedou do buněk vhodným způsobem, jako jsou přítomny u zdravých jedinců, je možné uvedeným způsobem léčit onemocnění spojené s malým vzrůstem, to znamená například Turnerův syndrom. Onemocnění spojené s malým vzrůstem se může léčit odstraněním nebo mutací růstových genů odpovědných za malý vzrůst. Je také možné stimulovat geny, které kompenzují působení genů odpovědných za malou výšku postavy, to znamená inzercí sekvencí DNA před, za nebo do růstových genů/genů odpovídajících za malou výšku za účelem zvýšit expresi zdravých alel. Uvedenými modifikacemi genů se růstové geny/geny odpovídajících za malou výšku aktivují nebo deaktivují. To se může uskutečnit inzercí

sekvencí DNA do vhodných míst, které jsou v genu nebo blízko genu umístěny tak, že tyto inzerované sekvence DNA interferují s růstovými geny/geny odpovídajícími za malou výšku a tak aktivují nebo brání jejich transkripci. Je také možné před růstové geny začlenit regulační element (například promotorovou sekvenci), aby se geny stimulovaly a staly se aktivní. Dále je možné stimulovat promotorovou sekvenci za účelem nadměrné exprese - v případě Turnerova syndromu - správně fungujících alel a tak kompenzovat chybějící alely. Modifikace genů se může obecně dosáhnout inzercí exogenních sekvencí DNA do růstového genu/geny odpovídajícího za malý vzrůst prostřednictvím homologní rekombinace.

Sekvence DNA podle vynálezu se mohou také použít k transformaci uvedených sekvencí do zvířat, jako jsou savci, prostřednictvím vhodného vektorového systému. Tyto transgenní zvířata se pak mohou použít pro zkoumání *in vivo* při testování a identifikaci farmaceutických činidel, která jsou použitelná při léčbě onemocnění spojených s malým vzrůstem. Jestliže zvířata pozitivně reagují na aplikaci sloučeniny nebo činidla, pak uvedené činidlo nebo sloučenina nebo jejich deriváty jsou vhodná jako farmaceutická činidla. Sekvence DNA podle vynálezu se mohou také využívat v genetických experimentech za účelem nalezení způsobů, jak kompenzovat ztráty genů odpovědných za malý vzrůst postavy.

Dále vynález popisuje sekvence DNA, které se mohou využít při transformaci buněk. Tyto buňky se mohou použít k identifikaci farmaceutických činidel, jenž se mohou použít při léčbě onemocnění spojených s malou výškou postavy nebo k testování takových látek nebo knihovny látek. Ve vhodném testovacím systému se mohou identifikovat změny fenotypu nebo změny expresivního paternu těchto buněk, což umožňuje identifikaci činidla vhodného pro přípravu léků.

Sekvence DNA podle vynálezu se také mohou použít při tvorbě vhodných primerů, které za přísných podmínek hybridizují se



segmenty genů odpovídajících za malý vzrůst nebo s jejich fragmenty. Mohou se konstruovat vhodné sekvence primerů, které se používají při diagnóze lidí, které mají genetický defekt způsobující malý vzrůst. Proto je nutné poznamenat, že dvě mutace nalezené v identické pozici naznačují, že existuje mutační horké místo.

Sekvence DNA podle vynálezu zahrnují i takové sekvence, které se degenerovaly na specifické sekvence na základě degenerace genetického kódu nebo které hybridizují za přísných podmínek se specificky danými sekvencemi.

Vynález zdůrazňuje zvláště následující rysy:

- a) Molekulu izolované nukleotidové kyseliny kódující polypeptidy obsahující oblast homeoboxu šedesáti aminokyselin, které mají aminokyselinovou sekvenci SEQ ID NO: 1 a vykazují regulující aktivitu lidského růstu.
- b) Molekulu izolované DNA obsahující nukleotidovou sekvenci, jak je uvedeno na obrázku č. 2, 3 nebo 4 a zvláště, jak je uvedeno v SEQ ID NO: 10; SEQ ID NO: 12 nebo SEQ ID NO: 15.
- c) Molekuly DNA jsou schopny hybridizovat s molekulami DNA popsány v odst. b).
- d) Molekuly DNA podle odstavce c) jsou schopny hybridizovat s molekulami DNA podle odst. 2 při teplotě 60 °C až 70 °C a v přítomnosti roztoku standardního pufru.
- e) Molekuly DNA obsahují nukleotidové sekvence, které vykazují 70 % nebo vyšší homologii s nukleotidovou sekvencí SEQ ID NO: 10; SEQ ID NO: 12 nebo SEQ ID NO: 15 a kódují polypeptid, který má regulační aktivitu lidského růstu.
- f) Lidské růstové proteiny mají aminokyselinovou sekvenci SEQ ID NO: 11, 13 nebo 16 nebo jejich funkční fragment.
- g) Protilátky získané při imunizaci zvířat lidskými růstovými proteiny podle odst. f) nebo jejich antigenní varianty.
- h) Farmaceutické kompozice obsahující lidské růstové proteiny nebo jejich funkční fragmenty pro léčbu vad způsobených genetickými mutacemi lidského růstového genu.

- i) Způsob testování látky účinné při léčbě vad zmiňovaných shora v textu podle odst. h) zahrnující detekci mediátorové RNA hybridizující s libovolnými molekulami DNA popsány v odstavcích a) až e), přičemž se měří zesílení exprese molekuly DNA jako odezvy na léčbu hostitelské buňky substancí.
- j) Expresivní vektor nebo plazmid obsahující libovolné molekuly nukleové kyseliny popsané v odstavcích a) až e), které umožňují, aby molekuly DNA se exprimovaly v savčích buňkách.
- k) Způsob stanovení genu nebo genů v biologických vzorcích tkání nebo tělních tekutin odpovědných za malý vzrůst.

V případě metody popsané shora v textu v odstavci k) se pro detekci specifických nukleotidových sekvencí s výhodou používají nukleotidové amplifikační metody PCR, které jsou dobře známy v oboru a popisují se například v publikacích Mullis et al., 1986, Cold Spring Harbor Symposium Quant. Biol. 51, 263-273 a Saiki et al., Science 239, 487-491. Stanovené nukleotidové sekvence odpovídající za malý vzrůst jsou hlavně ty, které jsou reprezentovány sekvencemi SEQ ID NO: 2 až SEQ ID NO: 7.

V principu všechny oligonukleotidové primery a sondy vhodné pro amplifikaci a detekci genetického defektu, který odpovídá za poškození lidského růstu v biologických vzorcích, jsou vhodné pro amplifikaci cílové sekvence spojované s malým vzrůstem. V tabulce č. 1 jsou uvedeny vhodný pár primerů specifický pro exon podle vynálezu. Následuje vhodná detekce, přičemž se například provádí radioaktivní nebo jiné značení.

Tabulka č. 1

exon	Sense primer	Antisense primer	Produkt (bp)	Ta (°C)



5'-I (G310)	SP1	ASP1	194	58
3'-I (G310)	SP2	ASP2	295	58
II (ET93)	SP3	ASP3	262	76/72/68
III (ET45)	SP4	ASP4	120	65
IV (G108)	SP5	ASP5	154	62
Va (SHOXa)	SP6	ASP6	265	61

Vysvětlení zkratk u primerů:

SP 1: ATTTCCAATGGAAAGGCGTAAATAAC

SP 2: ACGGCTTTTGTATCCAAGTCTTTTG

SP 3: GCCCTGTGCCCTCCGCTCCC

SP 4: GGCTCTTCACATCTCTCTCTGCTTC

SP 5: CCACACATGACACCTGCTCCCTTTG

SP 6: CCCGCAGGTCCAGGCTAGCTG

ASP1: CGCCTCCGCCGTTACCGTCCTTG

ASP2: CCCTGGAGCCGGCGCGCAAAG

ASP3: CCCC GCCCCCGCCCCCGG

ASP4: CTT CAGGTCCCCCAGTCCCG

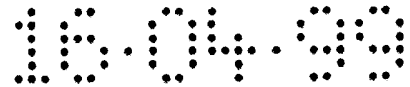
ASP5: CTAGGGATCTTCAGAGGAAGAAAAG

ASP6: GCTGCGCGGCGGGTCAGAGCCCCAG

Jako targetová DNA může také sloužit jednořetězcová RNA. Metody zpětné transkribce RNA na cDNA jsou dobře známy v oboru a popisuje je Sambrook et al., "Molecular Cloning" A Laboratory Manual, New York, Cold Spring Harbor Laboratory 1989. V jiném případě preferované metody reverzní transkripce využívají termostabilní DNA polymerázy, které vykazují RT aktivitu.

Shora popsaná metoda se může použít při selekci těch osob ze skupiny lidí, jejichž malý vzrůst postavy se spojuje s genetickým defektem, což umožňuje specifitější léčbu.

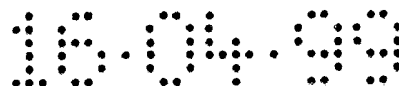
Vynález dále popisuje transkripční faktory A, B a C, které se mohou použít jako farmaceutická činidla. Tyto transkripční faktory inicijují stále neznámou kaskádu biologických účinků na molekulární úrovni, které se spojují s lidským růstem. Tyto



proteiny nebo jejich funkční fragmenty mají mitogenní účinek na různé buňky. Zvláště mají osteogenní účinek. Mohou se použít při léčbě onemocnění kostí, jako je například osteoporéza, a zvláště všech onemocnění spojených s poškozením regulace vápníku v kostech.

Termín „izolovaný“ znamená původní derivaci molekuly DNA klónováním. Vynález se týká jak přirozeně se vyskytujících sekvencí tak i synteticky připravených.

Molekuly DNA podle vynálezu se mohou použít ve formách genové terapie, která zahrnuje použití expresivních plazmidů připravených inkorporací vhodné sekvence DNA podle vynálezu downstream od expresivního promotoru, který ovlivňuje expresi v savčí hostitelské buňce. Vhodné hostitelské buňky jsou prokaryontní a eukaryontní buňky. Prokaryontní hostitelské buňky jsou například *E. coli*, *Bacillus subtilis* a podobně. Prostřednictvím transfekce hostitelských buněk replikony, které pocházejí z druhů adaptovatelných na hostitele, to jsou plazmidové vektory obsahující počáteční bod replikace a regulační sekvence, se mohou uvedené hostitelské buňky transfekovat požadovaným genem nebo cDNA. Takové vektory jsou s výhodou ty, které mají sekvenci, jenž vybaví transfekované buňky vlastnostmi (fenotypem), na jejichž základě se mohou selektovat. V typickém případě se používají jako hostitelé bakterie *E. coli* kmen K12a, z vektorů se používá buď pBR322 nebo plazmidy pUC. Příklady vhodných promotorů v případě hostitelů *E. coli* jsou promotor *trp*, *lac* promotor nebo *lpp* promotor. Jestliže je to nutné, sekrece expresivního produktu skrz buněčnou membránu se může ovlivnit spojením sekvence DNA kódující sekvenci signálního peptidu s 5'upstream částí genu. Eukaryontní hostitelské buňky zahrnují buňky získané z obratlovců nebo kvasinek, atd. Jako hostitelské buňky obratlovců se mohou použít buňky COS (Cell, 1981, 23: 175-182) nebo buňky CHO. Aby došlo k expresi mohou se s výhodou použít promotory, které jsou umístěny 5'upstream genu a mají místa



sestřihu RNA, polyadenylační a transkripční terminační sekvence.

Při léčbě vad způsobených mutacemi lidských růstových genů se mohou použít jako činidla podporující růst transkripční faktory A, B a C podle vynálezu. Na základě polymorfizmu, který je znám u eukaryontních genů, se může substituovat jedna nebo více aminokyselin. Jedna nebo více aminokyselin v polypeptidech může být deletována nebo začleněna do jednoho nebo více míst aminokyselinové sekvence polypeptidů SEQ ID NO: 11, 13 nebo 16. Takové polypeptidy jsou nazývány ekvivalentní polypeptidy, pokud vykazují biologickou aktivitu nemodifikovaného polypeptidu, jenž zůstává v podstatě beze změn.

Přehled obrázků na výkrese

Na **obrázku č. 1** je zobrazena genetická mapa genu SHOX, který zahrnuje pět exonů, které se označily následujícím způsobem: exon I: G310; exon 2: ET93; exon III: ET45; exon IV: G108; a exony Va a Vb, přičemž exony Va a Vb vedou ke vzniku dvou různých míst sestřihu genu SHOX. Exon II a III obsahuje homeobox o velikosti 180 nukleotidů.

Na **obrázcích č. 2 a 3** jsou nukleotidy a předpovězené aminokyselinové sekvence SHOXa a SHOXb.

SHOXa: Předpovězený začátek translace začíná nukleotidem 92, první v rámci je stop kodon (TGA) - nukleotidy 968 - 970, přičemž vzniká otevřený čtecí rámeček o velikosti 876 bp, který kóduje předpovězený protein tvořený 292 aminokyselinami (označený jako transkripční faktor A nebo protein SHOX). V rámci je 5' stop kodon v pozici nukleotidu 4, počáteční kodon a předpovězený terminační stop kodon znázorněný tučným písmem. Homeobox je ohraničen boxem (začínající aminokyselinou v pozici 117 (Q) až 176 (E), to v nukleotidové sekvenci odpovídá CAG až GAG). Pozice intronů je označena šipkami. Dva

putativní polyadenylační signály v 3' nepřekládané oblasti jsou podtrženy.

SHOXb: Otevřený čtecí rámec o velikosti 876 bp se nachází od A v prvním metioninu nukleotidu 92 do terminačního kodonu v rámci v pozici 767 - 769, což tvoří otevřený čtecí rámec o velikosti 675 bp, který kóduje předpovězený protein o velikosti 225 aminokyselin (transkripční faktor B nebo protein SHOXb). Pozice intronů jsou označeny šipkami. Exony I až IV jsou identické s SHOXa, exon V je specifický pro SHOXb. Putativní polyadenylační signál v 3' nepřekládané oblasti je podtržen.

Na obrázku č. 4 je nukleotidová sekvence a předpovězená aminokyselinová sekvence SHOT. Předpovězený počátek translace začíná v pozici nukleotidu 43, první v rámci je terminační kodon (TGA) - pozice nukleotidů 613 - 615, přičemž vzniká otevřený čtecí rámec o velikosti 573 bp, který kóduje předpovězený protein tvořený 190 aminokyselinami (označený jako transkripční faktor C nebo protein SHOT). Homeobox je ohraničen boxem (začínající aminokyselinou v pozici 11(Q) až 70 (E), to v nukleotidové sekvenci odpovídá CAG až GAG). Pozice intronů je označena šipkami. Dva putativní polyadenylační signály v 3' nepřekládané oblasti jsou podtrženy.

Na obrázku č. 5 je uvedena organizace exon/intron lidského genu SHOX a pozice v nukleotidové sekvenci.

Popis sekvencí SEQ ID:

SEQ ID NO. 1: překládaná aminokyselinová sekvence oblasti homeoboxu (180 bp)

SEQ ID NO. 2: exon II (ET93) genu SHOX

SEQ ID NO. 3: exon I (G310) genu SHOX

SEQ ID NO. 4: exon III (ET45) genu SHOX

SEQ ID NO. 5: exon IV (G108) genu SHOX

SEQ ID NO. 6: exon Va genu SHOX

- SEQ ID NO. 7: exon Vb genu SHOX
SEQ ID NO. 8: předchozí nukleotidová sekvence genu SHOX
SEQ ID NO. 9: gen ET92
SEQ ID NO. 10: sekvence SHOXa (také uvedeno na obrázku č. 2)
SEQ ID NO. 11: transkripční faktor A (také uvedeno na obrázku č. 2)
SEQ ID NO. 12: sekvence SHOXb (také uvedeno na obrázku č. 3)
SEQ ID NO. 13: transkripční faktor B (také uvedeno na obrázku č. 3)
SEQ ID NO. 14: gen SHOX
SEQ ID NO. 15: sekvence SHOX (také uvedeno na obrázku č. 4)
SEQ ID NO. 16: transkripční faktor C (také uvedeno na obrázku č. 4).

Příklady provedení vynálezu

Příklad 1: Pacienti

Všech šest studovaných pacientů vykazuje de novo odchylky na pohlavním chromozomu.

Pacient s označením CC je dívka s karyotypem 45,X/46,X psu dic (X) (Xqter →Xp22.3::Xp22.3→Xqter). Při posledním testu jí bylo 6 1/2 roku, její výška byla 114 cm (25 až 50 % percentil). Výška postavy její matky je 155 cm, otec nebyl zařazen do analýzy. Detailněji je případ popsán v publikaci Henke A., Wapenaar M., van Ommen G-J., Maraschio P., Camerino O., Rappold G.A. (1991): Deletions within the pseudoautosomal region help map three new markers and indicate role of this region in linear growth. Am. J. Hum. Genet. 49: 811-819.

Pacient s označením GA je dívka s karyotypem 46,X der X (3pter→3p23::Xp22.3→Xqter). Při posledním testu jí bylo 17 let a zjistila se normální výška (159 cm). Výška její matky je 160 cm, její otec měří 182 cm. Detailněji se pacientka popisuje v publikaci Kulharya A.S., Roop H., Kukolich M.K., Nachtman R.G., Belmont J.W., Garcia-Heras J., (1995): Mild

phenotypic effects of a de novo deletion Xpter→Xp22.3 a duplication 3pter→3p23. Am. J. Med. Genet. 56: 16-21.

Pacient s označením SS je dívka s karyotypem 46,X rea (X) (Xqter→Xq26::Xp22.3→Xq26:). V jedenácti letech její výška byla pod 3 percentil křivky růstu japonských dívek.; její předpovězená výška v dospělosti (148,5 cm) leží pod její cílovou výškou (163 cm) a cílové rozmezí je (155 až 191 cm). Detailněji je případ popsán v publikaci Ogata T., Goodfellow P., Petit C., Aya M., Matsuo N. (1992): Short stature in a girl with a terminal Xp deletion distal to DXYS15: localization of a growth gene(s) in the pseudoautosomal region. J. Med. Genet. 29: 455-459.

Pacient z označením AK je dívka s karyotypem 46,X rea (X) (Xqter→Xp22.3::Xp22.3→Xp21.3:). Ve 13-ti letech její výška je pod 2 percentil křivky růstu japonských děvčat; její předpovídaná výška v dospělosti (142,8 cm) je pod její cílovou výškou (155,5 cm) a cílové rozmezí je (147,5 až 163,5). Detailněji se případ popisuje v publikaci Ogata T., Yoshizawa A., Muroya K., Matsuo N., Fukushima Y., Rappold G.A., Yokoya S. (1995): Short stature in a girl with partial monosomy of the pseudoautosomal region distal to DXYS15: further evidence for the assignment of the critical region for a pseudoautosomal growth gene(s). J. Med. Genet. 32: 831-834.

Pacient označený RY má karyotyp 46,X,r(Y)/46,Xdic r(Y)/45,X[95:3:2], jak se zjistilo testem 100 lymfocytů; V 16-ti letech je její výška 148 cm; jeho tři bratři dosáhly normální výšky 170 cm (bratr 1, 16 let), 164 cm (bratr 2, 9 let) a 128 cm (bratr 3, 9 let). Růstová retardace tohoto pacienta je tak silná, že by měla být kompatibilní s další delecí lokusu GCY na Yq.

Pacient s označením AT je chlapec s ataxií a inv(X); v sedmi letech dosahuje normální výšky 116 cm, výška rodiču je 156 cm a 190 cm.

Pacienti pro analýzu mutací

U 250 jednotlivců s idiopatickým malým vzrůstem postavy se vyšetřovaly mutace genu SHOXa. Pacienti se vybraly podle následujících kritérií: výška chronologického věku byla pod 3 centil národního výškového standardu minus 2 standardní odchylky (SDS); nebylo známé žádné kauzativní onemocnění zvláště : normální váha (délka) v prenatálním stavu, normální tělesné proporce, žádné chronické onemocnění, normální příjem potravy, žádné psychiatrické obtíže, žádná dysplázie skeletu, žádná nedostatečnost růstového hormonu nebo hormonů štítné žlázy.

Rodina A:

Případy 1 a 2 jsou děti malého vzrůstu německé nepokrevné rodiny. Chlapec (případ 1) se narodil v 38 týdnu těhotenství císařským řezem. Porodní váha byla 2 660 g, porodní délka byla 47 cm. Vyvíjel se normálně s výjimkou subnormálního růstu. Při schůzce ve věku 6,4 roku byl proporcionálně malý (106,8 cm, -2,6 SDS) a obézní (22,7 kg) ale jinak normální. Stáří jeho kostí nebylo retardované (odpovídalo 6 letům) a dysplázie kostí se vyloučila na základě vyšetření X-paprsky. Nepravděpodobná je nedostatečnost v hladinách IGF-I a IGFBP-3 stejně jako parametry štítné žlázy, které se v séru prokazují GH nebo nedostatečnost hormonů štítné žlázy. Dívka (případ 2) se narodila v termínu císařským řezem. Porodní váha byla 2 920 g, porodní délka byla 47 cm. Její vývoj byl normální, ale ve 12 měsících se projevil její malý růst (délka 67 cm, -3,0 SDS). Ve čtyřech letech měřila 89,6 cm (-3,6 SDS). Neprojevily se žádné dysmorfie ani dysproporcionální rysy. Nebyla obezní (13 kg). Stáří jejich kostí odpovídalo 3,5 letům a vyloučila se dysplázie kostí. Hormonální parametry byly normální. Je zajímavé poznamenat, že růst děvčete i chlapce odpovídá 50 percentil růstové křivky žen s Turnerovým syndromem. Matka je nejmenší z rodiny a má střední rhizomelickou disproporci

(142,3 cm, -3,8 SDS). Jedna ze sester (150 cm, -2,5 SDS) a babička z matčiny strany (153 cm, -2,0 SDS) jsou malé postavy, aniž mají nějakou disproporci. Druhá sestra má normální výšku postavy (167 cm, +0,4 SDS). Výška otce je 166 cm (-1,8 cm) a výška dědečka z matčiny strany je 165 cm (-1,9 SDS). Další pacient je japonského původu a vykazuje identické mutace.

Příklad 2: Identifikace genu odpovídajícího za malý vzrůst

A. Hybridizace *in situ*

a) Fluorescenční hybridizace *in situ*

Provedla se fluorescenční hybridizace *in situ* (FISH) za použití kosmidů, které leží v Xp/Yp pseudoautozomální oblasti (PAR1). Uskutečnily se FISH studie za použití kosmidů 64/75cos (LLNLc110H032), E22cos(2e2), F1/14cos (110A7), M1/70cos (110E3), P99F2cos(43C11), P99cos (LLNLc110P2410), B6cosb (1CRFc104H0425), F20cos (34F5), F21cos (ICRFc104G0411), F3cos2 (9E3), F3cos1 (11E6), P117cos (29B11), P6cos1(ICRFc104P0117), P6cos2 (LLNLc110E0625) a E4cos (15G7) podle metody popsané v publikacích (Lichter P., Cremer T., Human Cytogenetics: A practical Approach, IRL Press 1992, Oxford, New York, Tokyo). Jeden mikrogram kosmidového klónu se značil biotinem a hybridizoval se s metafází lidských chromozomů za podmínek, které redukují signál z repetitivních sekvencí DNA. Detekce hybridizačního signálu probíhá prostřednictvím avidinu spojeného s FITC. Zviditelnění FITC proběhlo za použití chlazeného zdvojeného kamerového systému (Photometrics, Tucson, AZ).

b) Fyzikální mapování

Kosmidy se získaly z Lawrence Livermore National Laboratory X- a Y-chromozomové knihovny a z X-chromozomové knihovny instituce Imperial Cancer Research Fund London (nyní Max Planck Institute for Molecular Genetics Berlin). Za použití

kosmidů distálních k DXYS15, jmenovitě E4cos, P6cos2, P6cos1, P117cos a F3cos1 lze stanovit, že jsou stále přítomny dvě kopie E4cos, P6cos2, P6cos1 a jedna kopie P117cos a F3cos1. Body porušení na mapě kosmidu P6cos1 obou pacientů AK a SS jsou od sebe v maximální vzdálenosti 10kb. Zjistilo se, že abnormální X-chromozomy AK a SS měly deletováno přibližně 630 kb DNA.

Další kosmidy se získaly z ICRF X-chromozomové specifické kosmidové knihovny (ICRFc104), z Lawrence Livermore X-chromozomové specifické kosmidové knihovny (LLNLc110) a Y-chromozomové specifické knihovny (LLC03M1), stejně jako ze samostatně připravené kosmidové knihovny zahrnující celý genom. Kosmidy se identifikovaly hybridizací se všemi známými sondami, které mapují uvedenou oblast a použitím celé YAC jako sondy. Za účelem ověření překryvů v případech, kde nebylo možno použít známých sond, se použily terminační sondy z několika kosmidů.

c) Hybridizace Southernovým přenosem

Analýza Southernovým přenosem za použití pseudoautozomálních markerů poskytuje důkaz, že bod porušení X-chromozomu u pacienta CC leží mezi DXYS20(3cosPP) a DXYS60(U7A (Henke A., Wapenaar M., van Ommen G-J., Maraschio P., Camerino O., Rappold G.A. (1991): Deletions within the pseudoautosomal region help map three new markers and indicate role of this region in linear growth. Am. J. Hum. Genet. 49: 811-819). Za účelem potvrdit tento nálezn a definovat pozici bodu porušení, se použily kosmidy 64/75cos, E22cos, F1/14cos, M1/70cos, F2cos, P99F2cos a P99cos jako sondy při hybridizaci FISH. Pozice bodu porušení na abnormálním chromozomu X u pacienta CC je mezi kosmidy 64/75cos (jedna kopie) a F1/14cos (dvě kopie) na E22PAC. Pacient CC s normální výškou následně ztratil přibližně 269 až 290 kb DNA.

Hybridizace Southernovým přenosem se provedla za přísných podmínek v Churchově puftru (0,5 M NaPi pH7,2, 7 % SDS, 1 mM EDTA) při teplotě 65 °C a promývání se provedlo 40 mM NaPi, 1 % SDS při teplotě 65 °C.

d) Analýza hybridizací FISH

Biotinem značená kosmidová DNA (velikost inzertu 32 až 45 kb) nebo kosmidové fragmenty (10 až 16 kb) se hybridizovaly s metafází chromozomů, které pochází ze stimulovaných lymfocytů pacientů, za podmínek, jež se popisují dříve v textu (Lichter P., Cremer T., Human Cytogenetics: A practical Approach, IRL Press 1992, Oxford, New York, Tokyo). Hybridizované sondy se detekovaly prostřednictvím avidinu spojeného s FITC.

e) Amplifikace PCR

Všechny reakce PCR se provedly v objemech 50 ul. V tomto objemu je obsaženo 100 pg až 200 pg templátu, 20 pmol každého primeru, 200 uM dNTP (Pharmacia), 1,5 mM MgCl₂, 75 mM Tris/HCl pH9, 20 mM (NH₄)₂SO₄ 0,01 % (hmot./objem) Tween 20 a 2 U Goldstar DNA polymerázy (Eurogentec). Termocykly se provedly na zařízení GeneE (Techne).

f) Amplifikace exonu

Za účelem amplifikace exonu se použily čtyři kosmidové sebrané roztoky, kdy každý zahrnuje čtyři až pět klónů z kosmidových kontigů. Kosmidy v každém kosmidovém roztoku se částečně štěpily restričním enzymem Sau3A. Frakce čištěné na gelu v rozmezí velikostí 4 až 10 kb se klonovaly do vektoru pSPL3B štěpeného restričním enzymem BamHI (Burn et al., 1995) a použily se při experimentech amplifikace exonu, jak se popisuje dříve v textu (Church et al., 1994).

g) Sekvenování genů

Sonifikované fragmenty dvou kosmidů LLOYNCO3'M'15D10 a LLOYNCO3'M'34F5 se odděleně sub-klónovaly do vektorů M13mp18. Z každé kosmidové knihovny se vybralo alespoň 1 000 plaků, připravila se DNA M13 a sekvenovala se za použití barvivem značených terminátorů, termo-sekvenázy (Amersham) a univerzálního M13-primeru (MWG-BioTech). Gely se provedly na sekvenátoru ABI-377 a data se uspořádala a editovala pomocí programu GAP4 (Staden).

Ze všech šesti pacientů pacient GA měl nejméně chrakterizovaný bod porušení chromozomu. Nejvzdálenější markery, jejichž přítomnost na chromozomu X se testovala dříve, byly DXS1060 a DXS996, mapují přibližně 6 Mb od telomeru (Nelson, D.L., Ballabio A., CremersF., Monaco A.P., Schlessinger D., (1995). -Report of the sixth international workshop on the X chromosome mapping. Cytogenet. Cell Genet. 71: 308-342). Testovalo se několik kosmidů, které obsahují odlišné sekvence genů z PAR1 (MIC2, ANT3, CSF2RA a XE7). Všechny byly přítomny na translokačním chromozomu. Kosmidy pokrývající kritickou oblast odpovídající za malý vzrůst postavy, například z chromozomu, přičemž translokační bod porušení je na kosmidu M1/70cos. Kvantitaivní srovnání intenzity signálu M1/70cos mezi normálním a znovu uspořádaným chromozomem X ukazuje, že je deletováno přibližně 70 % kosmidu.

Tabulka č. 2

	CC	GA	AK	SS
64/75cos	-	-		
E22cos	-	-		
F1/14cos	+	-		
M1/70cos	+	(+)		
F2cos	+	+		
P99F2cos	+	+		
P99cos	+	+		
B6cos		+		

F20cos				
F21cos				
F3cos2				
F3cos1			-	-
P117cos			-	-
P6cos1			+	+
P6cos2			+	+
E4cos			+	+

Tabulka č. 2: Tabulka udává souhrn dat z hybridizace FISH pro 16 kosmidů testovaných u čtyř pacientů

(-) jedna kopie; indikuje, že kosmid je deletován na znovu uspořádaném chromozomu X, ale je přítomen na normálním chromozomu X

(+) dvě kopie; indikuje, že kosmid je přítomen na znovu uspořádaném i na normálním chromozomu X

((+)) oblast bodu porušení; indikuje, že bod porušení se vyskytuje na kosmidu, jak ukazuje hybridizace FISH.

Molekulová analýza u šesti pacientů se znovu uspořádaným chromozomem X za použití kosmidových fluorescenčně značených kosmidových sond a hybridizace *in situ* indikuje, že oblast odpovídající za malý vzrůst je zúžena na interval o velikosti 270 kb, který je ohraničený bodem porušení pacienta GA z jeho centromerové distální strany a u pacientů AK a SS na jeho centromerové proximální straně.

Korelace genotyp-fenotyp může být informativní a může posloužit k zobrazení kritického intervalu, který odpovídá za malý vzrůst postavy, na lidském chromozomu X a Y. V této studii se použila analýza hybridizací FISH ke studiu rozšíření metafáze a interfáze jádra lymfocytů u pacientů, kteří nesou delece a translokace na chromozomu X a body porušení v Xp22.3. Tyto body porušení se jeví být přítomny u dvou ze čtyř pacientů (AK a SS), což je způsobeno přítomností sekvencí

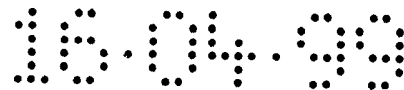


s predispozicí k znovu uspořádání chromozomu. Zjistil se další pacient s kruhovým Y s porušením nacházejícím se v kritické oblasti o velikosti 270 kb, což redukuje kritický interval na 170 kb.

Korelací výšky šesti jedinců s jejich místem porušení se mapoval interval o velikosti 170 kb v pseudoautozomální oblasti, jehož přítomnost nebo absence má podstatný účinek na výšku vzrůstu. Tento interval je ohraničen místem porušení na chromozomu X pacienta GA v místě vzdáleném 340 kb distálně od telemoru (Xptel) a místem porušení u pacientů AT a RY proximálně 510/520 kb Xptel. Toto uspořádání zahrnuje podstatnou redukci kritického intervalu na téměř jednu čtvrtinu jeho původní velikosti (Ogata T., Goodfellow P., Petit C., Aya M., Matsuo N. (1992): Short stature in a girl with a terminal Xp deletion distal to DXYS15: localization of a growth gene(s) in the pseudoautosomal region. *J. Med. Genet.* 29: 455-459; Ogata T., Yoshizawa A., Muroya K., Matsuo N., Fukushima Y., Rappold G.A., Yokoya S. (1995): Short stature in a girl with partial monosomy of the pseudoautosomal region distal to DXYS15: further evidence for the assignment of the critical region for a pseudoautosomal growth gene(s). *J. Med. Genet.* 32: 831-834; Ogata T., Matsuo N. (1995): Turner syndrome and female sex chromosome aberrations: deduction of the principle factors involved in the development of clinical features. *Hum. Genet.* 95: 607-629). Pro experimenty FISH za účelem testování prevalence a signifikance uvedeného genomového lokusu na velké sérii pacientů s idiopatickým malým vzrůstem je nyní dostupná malá sada šesti až osmi kosmidů.

B. Identifikace kandidáta genu odpovídajícího za malý vzrůst

Za účelem hledání transkripčních jednotek v nejmenší kritické oblasti o velikosti 170 kb se provedlo trapování exonu a selekce cDNA u šesti kosmidů (110E3, F2cos, 43C11, 02410, 15D10, 34F5). Trapováním exonu se izolovaly tři různé



pozitivní klóny (ET93, ET45 a G108). Všechny z nich se mapovaly zpět ke kosmidu 34F5. Předchozí studie za použití cDNA selekčních protokolů a přístupu k 25 různým knihovnám cDNA vedly k neúspěchu, což naznačuje, že geny v tomto intervalu se exprimují s velmi malým nadbytkem.

Za účelem zjištění, zda nějaký gen v uvedeném intervalu chybí, se stanovila za použití náhodné M13 metody a terminátorové chemie nukleotidová sekvence vzdálená přibližně 140 kb od uvedené oblasti PAR1. Za účelem sekvenční analýzy se vybraly kosmidy se vzájemným minimálním překryvem, aby spolu pokryly kritický interval. Provedla se sekvenční analýza a následná předpověď proteinu pomocí programu „X Grail“, verze 1.3c, stejně jako trapování exonu pomocí programu FEXHB, což potvrdilo všechny 3 dříve klónované exony. Nedetekoval se žádný jiný gen kódující protein, než co už byl dříve izolován.

C. Izolace genu SHOX odpovídajícího za malý vzrůst

Za předpokladu, že všechny tři exonové klóny ET93, ET45 a G108 jsou součástí stejného genu, použily se spolu jako sondy k testování 14 různých knihoven cDNA vytvořených z 12 různých fetálních (plíce, játra, mozek 1 a 2) a dospělých tkání (vaječníky, placenta 1 a 2, fibroblasty, skeletové svalstvo, kostní dřeň, mozek, mozkový kmen, hypotalamus, podvěsek mozkový). Mezi přibližně 14 milióny klóny nanesenými na plotny se nedetekoval ani jeden klón. Za účelem izolace transkriptu v plné délce se uskutečnil 3' a 5'RACE. Za účelem 3'RACE se použily primery z exonu G108 na RNA z placenty, skeletového svalstva a fibroblastů z kostní dřene a z tkáně, kde se exprimuje G108. Ze všech tří tkání se získaly dva různé 3'RACE klóny velikosti 1173 a 652 bp, což naznačuje, že existují dva různé 3'exony a a b. Tyto dvě odlišné formy se nazývají SHOXa a SHOXb.

Za účelem zvýšení šancí pro izolaci celé 5'části genu, o kterém se ví, že je exprimován s malým přebytkem, se buněčná

linie Hela ošetřila retinoovou kyselinou a forbolesterem PMA. RNA z takové indukované buněčné linie a RNA z placenty a skeletového svalstva se použila ke konstrukci „Mararhonovy knihovny cDNA“. Ze všech tří tkání se izolovaly identické klóny 5'RACE cDNA.

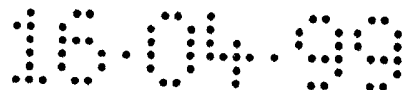
Experimentální postup:

RT-PCR a konstrukce knihovny CDNA.

Od firmy Clontech se získala lidská polyA⁺RNA ze srdce, pankreasu, skeletového svalstva, fetálních ledvin a jater. Celková RNA se izolovala z buněčné linie fibroblastů z kostní dřeně činidlem TRIZOL (Gibco-BRL), způsobem, který popisuje výrobce. První řetězec syntézy cDNA se získal pomocí kitu „Superscript first strand cDNA synthesis kit (Gibco-BRL), kdy počátečním materiálem je 100 ng polyA⁺RNA nebo 10 ug celkové RNA za použití oligo(dt)-ada⁺RNAptorového primeru (GGCCACGCGTCGACTAGTAC[dT]₂₀N. Po syntéze prvního řetězce cDNA se reakce naředila 1/10. Při dalších experimentech PCR se použilo 5 ul tohoto ředění. „Marathonova knihovna cDNA“ se zkonstruovala z polyA⁺RNA skeletového svalstva a placenty amplifikačním kitem pro cDNA (Clontech), jak popisuje výrobce. Knihovny cDNA z fetálního mozku (katalogové číslo # HL5015b), fetální plíce (HL3022a), vaječníků (HL1098a), podvěsku mozkového (HL1097v) a hypotalamu (HL1172b) se získaly od firmy Clontech. Knihovny cDNA mozku, ledvin, jater a plic se staly součástí rychlého testu panelu lidské cDNA knihovny (Clontech). cDNA knihovna fetálního svalstva se získala z instituce UK Human Genome Mapping Project Resources Center.

D. Sekvenační analýza a struktura genu SHOX

Na základě analýzy sekvencí získaných klónů pomocí 5' a 3'RACE se uspořádala sekvence SHOXa a SHOXb (1349 a 1870 bp). Identifikoval se jedne otevřený čtecí rámeček o velikosti 1870 bp (SHOXa) a 1349 bp (SHOXb), který dává vzniku dvou proteinů



o 292 aminokyselinách (SHOXa) a 225 aminokyselinách (SHOXb). Oba transkripty a i b se dělí o běžný 5'konec, ale mají odlišný poslední 3'exon . Toto zjištění naznačuje, že se používá alternativní signály sestřihu. Dosáhlo se úplného uspořádání mezi dvěma cDNA a sekvenovanou genomovou DNA z kosmidů LLOYNCO3"M"15D10a LLOYNCO3"M"34F5, což umožňuje ustanovení struktury exon-intron (obrázek č. 4). Gen se skládá ze 6 exonů, jejichž velikost se pohybuje v rozmezí od 58 bp (exon III) do 1146 bp (exon Va). Exon I obsahuje CpG, startovací kodon a 5'oblast. Stop kodon stejně jako 3'-nekódující oblast se nachází v každém alternativně sestřiženém exonu Va a Vb.

Příklad 3:

Identifikovaly se dvě cDNA, které tvoří 160 kb velkou oblast, která je kritická pro malý vzrůst. Tyto cDNA korespondují s geny SHOX a pET92. cDNA se se identifikovaly hybridizací subklónů kosmidů s knihovnamí cDNA.

Sada kosmidových klónů, které úplně pokrývají kritickou oblast, je genetickým materiálem pro identifikaci kauzativního genu. Projekty pozičního klónování, které vedou k izolaci genů z uvedené oblasti, se uskutečnily trapováním exonu a selekcí cDNA. Zjištěním jejich polohy v pseudoatozomální oblasti se tyto geny mohou vyhnout X-deaktivaci a může se uplatnit účinek dávky.

Klónování genu, který když není přítomen (haploid) nebo je deficitní odpovídá za malý vzrůst, je dalším krokem přesnosti diagnózy a poskytuje základy pro analýzu mutací v genu pomocí například jednořetězcového konformačního polymorfizmu (SSCP). Navíc klónování uvedeného genu a jeho následná biochemická charakterizace otevřela cestu k hlubšímu porozumění biologických procesů, které jsou zahrnuty v řízení růstu.

Sekvence DNA podle vynálezu poskytují první molekulový test pro identifikaci jednotlivců se specifickou vadou genu

v komplexní heterogenní skupině pacientů, kteří vykazují idiopatický malý vzrůst.

Příklad 4: Expresní patern SHOXa a SHOXb

Analýza northernovým přenosem za použití jednotlivých exonů jako hybridizačních sond ukázala odlišný expresivní profil každého exonu, což silně naznačuje, že pruhy o různé velikosti a intenzitě reprezentují produkty křížové hybridizace s jinými genovými sekvencemi bohatými na G,C. Aby se dosáhlo více realistického expresivního profilu genů SHOXa a b, provedly se s RNA z různých tkání RT-PCR experimenty. Zatímco exprese SHOXa se pozorovala ve fibroblastech skeletového svalstva, placenty, pankreasu, srdce a kostní dřeně, exprese SHOXb se omezila na fibroblasty fetální ledviny, skeletového svalstva a kostní dřeně, přičemž zdaleka nejsilnější expresi vykazují fibroblasty kostní dřeně.

Expresní SHOXa v několika cDNA knihovnách připravených z fetálního mozku a ze svalstva, z mozku dospělých, z plic a z podvěsku mozkového a skutečnost, že v žádné uvedené knihovně není exprimován SHOXb, podává další důkaz, že exprese jedné formy sestřihu (SHOXa) je více rozšířena než druhá (SHOXb), která se exprimuje převážně tkáňově specifickým způsobem.

K odhadu transkripční aktivity SHOXa a SHOXb na X a Y chromozomu se použila RT-PCR RNA, která se extrahovala z různých buněčných linií, jenž vykazují aktivní X, neaktivní X nebo Y chromozom. Všechny buněčné linie vykazují produkt amplifikace očekávané délky 119 bp (SHOXa) a 541 bp (SHOXb), což je důkaz, že jak SHOXa i b unikly deaktivaci chromozomu X. Geny SHOXa a b kódují nové proteiny homeodomén. Gen SHOX je vysoce konzervativní mezi jednotlivými živočišnými druhy od savců po ryby a ptáky. Na samém 5'konci a 3'konci vedle homeodomény jsou pravděpodobně konzervativní oblasti mezi člověkem a myší, což znamená funkční signifikanci. Během

evoluce mezi člověkem a myší se v těchto oblastech aminokyselin neakumulovaly rozdíly.

Experimentální postupy:

a) 5' a 3' RACE

Za účelem klónování transkriptů 5' konce SHOXa a b se provedla metoda 5' RACE za použití zkonstruované „Marathonových cDNA knihoven“. Použily se následující oligonukleotidové primery: SHOX B rev. GAAAGGCATCCGTAAGGCTCCC (pozice 697-718, reverzní řetězec [r]) a adaptorový primer AP1. Proběhla PCR za těchto podmínek: 94 °C po dobu 2 minut, 94 °C po dobu 30 vteřin, 70 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 5 cyklů. 94 °C po dobu 30 vteřin, 66 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 5 cyklů. 94 °C po dobu 30 vteřin, 62 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 25 cyklů. Druhé kolo amplifikace se uskutečnilo za použití 1/100 produktu PCR a následujících oligonukleotidových primerů: SHOX A rev, GACGCCTTTATGCATCTGATTCTC (pozice 617-640 r) a adaptorového primeru AP2. PCR se provedlo během 35 cyklů s teplotou renaturace 60 °C.

Za účelem klónování transkriptů 3' konce SHOXa a b se provedla metoda 3' RACE, jak popisuje v publikaci Frohman et al., 1988, za použití prvního řetězce cDNA s primerem oligo(dT)adaptoru. Použily se následující oligonukleotidové primery: SHOX A for GAATCAGATGCATAAAGGCGTC (pozice 619-640) a oligo(dT)adaptor. Proběhla PCR za těchto podmínek: 94 °C po dobu 2 minut, 94 °C po dobu 30 vteřin, 62 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 35 cyklů. Druhé kolo amplifikace se uskutečnilo za použití 1/100 produktu PCR a následujících oligonukleotidových primerů: SHOX B for, GGGAGCCTTACGGATGCCTTTC (pozice 697-718) a oligo(dT)adaptoru. PCR se provedlo během 35 cyklů s teplotou renaturace 62 °C.

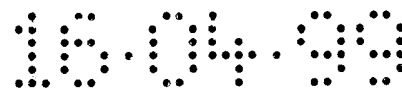
Za účelem potvrdit sekvence transkriptů SHOXa a SHOXb se uskutečnila PCR 5'oligonukleotidovým primerem a 3'oligonukleotidovým primerem. V případě SHOXa se použily následující primery: G310 for AGCCCCGGCTGCTCGCCAGC (pozice 59-78) a SHOX D rev CTGCGCGGCGGGTCAGAGCCCCAG (pozice 959-982 r). V případě SHOXb se použily následující primery: G310 for AGCCCCGGCTGCTCGCCAGC a SHOX2A rev GCCTCAGCAGCAAAGCAAGATCCC (pozice 1215-1238 r). Obě PCR se provedly za podmínek: 94 °C po dobu 2 minut, 94 °C po dobu 30 vteřin, 70 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 5 cyklů. 94 °C po dobu 30 vteřin, 70 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 5 cyklů. 94 °C po dobu 30 vteřin, 65 °C po dobu 30 vteřin, 72 °C po dobu 2 minut po 35 cyklů. Produkty se čistily na gelu a klónovaly se za účelem sekvenační analýzy.

b) Analýza SSCP

Amplifikovaná genomová DNA pocházející z pacientů se analyzovala způsobem SSCP, který se popisuje v publikaci Orita M., Suzuki Y., Sekiya T., and Hayashi K., (1989): Rapid and sensitive detection of point mutations and polymorphisms using the polymerase chain reaction. *Genomics* 5:874-879). 1 až 5 ul produktů PCR se smíchalo s 5 ul denaturačního roztoku, který obsahuje 95 % formamid a 10 mM EDTA pH8, a denaturovaly se při teplotě 95 °C po dobu 10 minut. Vzorky se ihned ochladily na ledu a nanuly se na 10 % polyakrylamidový gel (akrylamid : bisakrylamid = 37,5 : 1 a 29 : 1; gel s více sloty, báze TGGE, Qiagen), který obsahuje 2 % glycerol a 1 x TBE. Elektroforéza proběhla při teplotě 15 °C při napětí 500 V po dobu 3 až 5 hodin a obarvila se stříbrem, jak se popisuje TGGE manuálu (Qiagen, 1993).

c) Klónování a sekvenování produktů PCR

Produkty PCR se klónovaly do vektoru pMOSBlue za použití pMOSBlueT-sady vektorů od firmy Amersham. Kultura z jedné



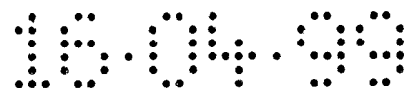
kolonie kultivovaná přes noc se lyzovala povařením ve 100 ul vody po dobu 10 minut. Lyzáty se použily jako templáty pro PCR se specifickými primery. SSCP produktů PCR umožnila identifikaci klónů, které obsahují různé alely. Klóny se sekvenovaly za použití vektorových primerů Uni a T7 značených CY5 způsobem sekvenování v cyklech, jak popisuje výrobce (ThermoSequenase Kit (Amersham)) na automatickém sekvenátoru ALF (Pharmacia).

d) Testování cDNA knihoven pomocí PCR

Aby se detekovala exprese SHOXa a b provedlo se PCR testování několika knihoven cDNA a prvního řetězce cDNA s SHOXa a b specifickými primery. V případě knihoven cDNA se použil ekvivalent DNA 5×10^8 pfu. V případě SHOXa se použily primery SHOX E rev GCTGAGCCTGGACCTGTTGGAAAGG (pozice 713-737 r) a SHOX a for. V případě SHOXb se použily následující primery SHOX B for a SHOX2A rev. Obě PCR se provedly za podmínek: 94 °C po dobu 2 minut, 94 °C po dobu 30 vteřin, 68 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 5 cyklů. 94 °C po dobu 30 vteřin, 65 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 5 cyklů. 94 °C po dobu 30 vteřin, 62 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 35 cyklů

e) PCR testování knihoven cDNA

Aby se detekovala exprese SHOXa a b provedlo se PCR testování několika knihoven cDNA s SHOXa a b specifickými primery. V případě knihoven cDNA se použil ekvivalent DNA 5×10^8 pfu. V případě SHOXa se použily primery SHOX E rev GCTGAGCCTGGACCTGTTGGAAAGG (pozice 713-737 r) a SHOX a for. V případě SHOXb se použily následující primery SHOX B for a SHOX2A rev. Obě PCR se provedly za podmínek: 94 °C po dobu 2 minut, 94 °C po dobu 30 vteřin, 68 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 5 cyklů. 94 °C po dobu 30 vteřin, 65 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 5 cyklů. 94 °C



po dobu 30 vteřin, 62 °C po dobu 30 vteřin, 72 °C po dobu 40 vteřin po 35 cyklů.

Příklad 5: Expresivní patern OG12, což je putativní homolog SHOX a SHOT

Za účelem stanovení expresivního paternu se provedla hybridizace *in situ* u embryí myši z 5. dne p.c. a 18,5 dne p.c. , u fetálních a u nově narozených zvířat. Expresi bylo možno ukázat u vyvíjejících se zárodků noh, v mesodermu nasálních postupů, které se podílejí na tvorbě nosu a patra, u očních víček, u aorty, ve vyvíjejících se žláz samic, ve vyvíjející se páteři (omezených na diferenciaci motorických neuronů) a v mozku. Na základě tohoto expresivního paternu a na mapování pozice jeho lidského homologu SHOT, SHOT reprezentuje pravděpodobně kandidáta Cornelia de Lange syndromu, který zahrnuje malý vzrůst.

Příklad 6: Izolace nového genu homeoboxu podobného SHOX SHOT na chromozomu 3, který souvisí s lidským růstem/s malým vzrůstem postavy.

U člověka se izoloval nový gen nazývaný SHOT (SHOX homolog na chromozomu 3), který vykazuje nejvíce homologie s myším genem OG12 a lidským SHOX. Lidský gen SHOT a myší OG12 jsou vysoce homologní. Vykazují 99 % identity na úrovni proteinu. Ačkoli se to ještě neprokázalo, díky homologii mezi SHOT a SHOX (jde o identitu pouze v homeooblasti) je pravděpodobné, že gen SHOT také odpovídá za malý vzrůst nebo lidský růst.

Gen SHOT se izoloval za použití primerů ze dvou nových lidských EST (HS 1224703 a HS 126759) z databáze EMBL, přičemž se amplifikuje reverzně přepsaná RNA z linie fibroblastů kostní dřeně (Rao E., Weiss B., Mertz A., et al., (1995): Construction of a cosmid contig spanning the short stature candidate region in the pseudosomal region PAR 1. In: Turner

syndrome in a life span respective: Research and clinical aspects. Proceeding of the 4th International Symposium on Turner Syndrome, Gothenburg, Sweden, 18-21 May, 1995., edited by Albertsson-Wikland K., Ranke M.B., pp. 19-24, Elsevier.

3' konec genu SHOT se generoval RACE-PCR z fibroblastové knihovny kostní dřeně, která se zkonstruovala podle publikace Rao E., Weiss B., Fukami M., Rump A., Niesler B., Metz A., Muroya K., Binder G., Kirsch S., Winkelmann M., Nordsiek G., Heinrich U., Breuning M. H., Ranke M.B., Rosenthal A., Ogata T., Rappold G.A. (1997): Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nature Genet* 15:54-62. Gen SHOT se mapoval analýzou FISH chromozomu 3q25/q26 a dále se mapoval myší homolog syntetické oblasti na myším chromozomu 3. Na základě expresivního paternu OG12, což je myší homolog genu SHOT, gen SHOT reprezentuje kandidáta odpovídajícího za Cornelia Lange syndrom (který vykazuje malou výšku vzrůstu a další rysy, kam patří kraniofaciální abnormality), který se mapoval v tomto intervalu na chromozomu 3q25/26.

Příklad 7: Vyhledávání mutací u pacientů s idiopatickým malým vzrůstem

Sekvence DNA podle vynálezu se používají při PCR, LCR a při jiných známých technologiích, za účelem stanovení, zda takoví jednotlivci s malou výškou vzrůstu vykazují malé delece nebo bodové mutace v genu odpovídajícím za malý vzrůst.

U 91 (z celkového počtu 250 jednotlivců) nezávislých pacientů mužského a ženského pohlaví s idiopatickou malou výškou vzrůstu (odhad četnosti idiopatické malé výšky vzrůstu v obecné populaci se odhaduje na 2 až 2,5 %) se testovala přítomnost malých změn v uspořádání nebo bodové mutace v genu SHOXa. Navrhlo se šest sad primerů PCR, ne pouze za účelem amplifikace jednotlivých exonů, ale také sekvencí, které

lemují exon a malé části 5'UTR. V případě největšího exonu, exon I, se generovaly dva další primery uvnitř exonu. Primery, které se používaly při PCR, se uvádějí v tabulce č. 2.

Provedl se jednořetězcový konformační polymorfismus (SSCP) všech amplifikovaných exonů, jehož velikost se pohybuje v rozmezí od 120 do 295 bp. Pouze v případě dvou jednotlivců s malou výškou vzrůstu (Y91 a A1) se identifikoval posun pohyblivosti pruhů. Klónovaly se a sekvenovaly fragmenty, které vykazují pozměněné paterny SSCP (unikátní konformery SSCP). Aby se zabránilo vzniku PCR a sekvenačních artefaktů, sekvenování se provedlo na dvou řetězcích za použití dvou nezávislých reakcí PCR. Mutace u pacienta Y91 se nachází v pozici 28 bp 5'konce startovacího kodonu v 5'UTR a zahrnuje substituci, kdy guanin se nahradí citidinem. Aby se zjistilo, jestli tato mutace reprezentuje řídký polymorfismus nebo je odpovědná za fenotyp regulací exprese genu, například slabším navázáním translačních iniciačních faktorů, testovaly se pacientovi rodiče a sestra. Přestože sestra a otec pacienta dosahují normální CP (data nejsou uvedeny). Uvedená substituce bází reprezentuje řídký polymorfismus, který neovlivňuje fenotyp.

Klónování a sekvenování unikátních konformerů SSCP pacienta A1 vykazuje transici bází, kdy citidin nahradí thymidin (nukleotid 674), která zavede terminační kodon do pozice aminokyseliny 195 předpovězených aminokyselinových sekvencí 225 a 292. Aby se stanovilo, zda tato nesmyslná mutace je geneticky asociována s malou výškou vzrůstu rodiny, provedla se rodová analýza. Zjistilo se, že všech šest jednotlivců s malou výškou vzrůstu (definovanou jako výška nižší než 2 standardní odchylky) vykazuje nenormální posun SSCP a transici cytidinu za thymidin. Ani otec, ani jedna z tet a dědeček z matčiny strany, kteří dosáhly normální výšky nevykazují mutaci, což naznačuje, že babička přenesla mutovanou alelu na její dvě dcery a na jejich dvě vnoučata. U této rodiny

existuje shoda mezi přítomností mutovaných alel a fenotypu vykazujícího malou výšku vzrůstu.

Stejná situace, která se popisuje shora v textu, se zjistila u jiného pacienta s malou výškou vzrůstu, který je japonského původu.

Příklad 8:

Sekvence DNA podle vynálezu se používají pro charakterizaci funkce genu nebo genů. Sekvence DNA se mohou použít jako otázky při vyhledávání bází v databázích aminokyselin a nukleových kyselin za účelem identifikace příbuzných genů nebo genových produktů. Jako otázka pro vyhledávání v databázi aminokyselin se použila částečná aminokyselinová sekvence SHOX93. Test vykazuje velmi vysokou homologii s mnoha známými proteiny homeoboxů. Sekvence cDNA podle vynálezu se mohou použít při produkci peptidu rekombinantní metodou. Při produkci proteinu rekombinantním způsobem se mohou použít různé expresivní systémy, které jsou dobře známy v oboru.

Běžnou syntézou peptidu (syntéza proteinů podle Merrifieldovy metody) se syntetizoval peptid se sekvencí CSKSFQKSKDGNGG a podle standardního protokolu se získaly polyklonální protilátky z králíků a kuřat.

Seznam literatury

Ashworth A., Rastan S., Lovell-Badge R., Kay G., (1991): X-chromosome inactivation may explain the difference in viability of XO humans and mice. *Nature* 351: 406-408.

Blagowidow N., Page D.C., Huff D., Mennuti MT (1989): Ullrich-Turner syndrome in an XY female fetus with deletion of the sex-determining portion of the Y chromosome. *Am. J. med. Genet.* 34: 159-162.

Cantrell M.A., Bicknell J.N., Pagon RA et al. (1989): Molecular analysis 46,XY females and regional assignment of a new Y-chromosome-specific probe. *Human. Genet.* 83. 88-92.

Connor J.M., Loughlin S.A.R. (1989): Molecular genetics of Turner's syndrome. *Acta Paediatr. Scand. (Suppl.)* 356: 77-80.

Disteche C.M., Casanova M., Saal H., Friedman C., Sybert V., Graham J., Thline H., Page D.C., Fellous M., (1986): Small deletions of the short arm of the Y-chromosome in 46,XY females. *Proc. Natl. Acad. Sci USA* 83: 7841-7844.

Ferguson-Smith M.A. (1965): Karyotype-phenotype correlations in gonadal dysgenesis and their bearing on the pathogenesis of malformations. *J. med. Genet.* 2: 142-155.

Ferrari D., Koshier R.A., Dealy C.N. (1994): Limb mesenchymal cells inhibited from undergoing cartilage differentiation by a tumor promoting phorbol ester maintain expression of the homeobox-containing gene *MSX1* and fail to exhibit gap junctional communication. *Biochemical and Biophysical Research Communications.* 205(1):429-434.

Fischer M., Bur-Romero P., Brown L.G. et al., (1990): Homologous ribosomal protein genes in the human X- and Y-chromosomes escape from X-inactivation and possible implementation for Turner syndrome. *Cell* 63: 1205-1218.

Freund C., Horsford D.J., McInnes R.R. (1996): Transcription factor genes and the developing eye: a genetic perspective: *Hum. Mol. Genet.* 5: 1471-1488.

Gehring W. J., Qian Y.Q., Billeter, Furukubo-Tokunaga K., Schier A. F., Resendez-Perez D., Affolter M., Otting G., Wuthrich K., (1994): Homeodomain-DNA recognition . *Cell* 78: 211-223.

Grumbach M.M., Conte F.A. (1992): Disorders of sexual differentiation. In: Williams textbook of endocrinology, 8th edn.m edited by Wilson J.D., Foster D.W., pp. 853-952, Philadelphia, WB Saunders.

Hernandez D., Fisher E.M.C. (1996): Down syndrome genetics: unravelling a multifactorial disorder. *Hum. Mol. Genet.* 5: 1411-1416.

Kenyon C., (1994): If birds can fly, why can't we? Homeotic genes and evolution . *Cell* 78: 175-180.

Krumlauf R., (1994): *Hox* genes in vertebrate development. *Cell* 78: 191-201.

Lawrence P.A., Morata G., (1994): Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell* 78: 181-189.

Lehrach H., Drmnac R., Hoheisel J.D., Larin Z., Lemon G., Monaco A.P., Nizetic D., et al., Hybridization finger printing in genome mapping and sequencing. In Davies K.E., Tilghman S., Eds. Genome Analysis 1990: 39-81 Cold Spring Harbor, NY.

Levilliers J., Quack B., Weissenbach J., Petit C., (1989): Exchange of terminal portions of X- and Y-chromosomal short arms in human XY females. Proc. Natl. Acad. Sci. USA 86: 2296-2300.

Lippe B.M. (1991): Turner Syndrome. Endocrinol. Metab. Clin. North Am. 20: 121-152. Magenis R.E., Tochen M.L. Holahan K.P., Carey T., Allen L., Brown M.G. (1984): Turner syndrome resulting from partial deletion of Y-chromosome short arm: localization of male determinantns. J. Pediatr 105: 916-919.

Pohlschmidt M., RappoldG.A., Krause M., Ahlert D., Hosenfeld D., Weissenbach J., Gal A. (1991): Ring Y chromosome: Molecular characterization by DNA probes. Cytogenet. Cell Genet. 56: 65-68.

Qiagen (1993) TGGE Handbook, Diagen GmbH, TGMA 41123/93.

Osenfeld R.G. (1992): Turner syndrome: a guide for physicians. Second edition. The Turner's Syndrome Society.

Rovescalli A.C., Asoh S., Nirenberg M. (1996): Cloning and characterization four murine homeobox genes. Proc. Natl. Acad. Sci. USA 93: 10691-10696.

Shalet S.M. (1993): Leukemia in children treated with growth hormone. Journal of Pediatric Endocrinology 6: 109-111

15.04.99

41

Vimpani G.V., Vimpani A.F., Lidgard G.P., Cameron E.H.D.,
Farquhar J.W. (1977) Prevalence of severe growth hormone
deficiency. Br. Med. J. 2: 427-430.

Zinn A.R., Page D.C., Fisher E.M.C. (1993): Turner syndrome:
the case of the missing sex chromosome. TIG 9(3): 90-93.

Sekvenční protokol

- (1) Obecné informace:
- (i) Navrhovatel:
 - (A) JMÉNO: Rappold-Hoerbrand, Gudrun, Dr.
 - (B) ULICE: Hausackerweg 14
 - (C) MĚSTO: Heidelberg
 - (D) ZEMĚ: Německo
 - (E) ZIP: 69118
 - (ii) Název vynálezu: Molekula nukleové kyseliny kódující polypeptidy a způsoby jejího použití.
 - (iii) Počet sekvencí: 16
 - (iv) Počítačem čitelná forma:
 - (A) TYP MÉDIA: disketa
 - (B) POČÍTAČ: kompatibilní s IBM PC
 - (C) OPERAČNÍ SYSTÉM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, verze #1.30
EPO
 - (vi) Data předchozích přihlášek
 - (A) PŘIHLÁŠKA č. US 60/027,633
 - (B) DATUM PODÁNÍ: 01-10-1996
 - (vi) Data předchozích přihlášek
 - (A) PŘIHLÁŠKA č. US 60/027,633
 - (B) DATUM PODÁNÍ: 01-10-1996
- (2) Informace o sekvenci SEQ ID NO: 1:
- (i) CHARAKTERISTIKA SEKVENCE:
 - (A) DÉLKA: 60 aminokyselin
 - (B) TYP: aminokyselinová
 - (C) DRUH ŘETĚZCE: jednořetězcová
 - (D) TOPOLOGIE: lineární



(ii) TYP MOLEKULY: peptid

(xi) POPIS SEKVENCE: SEQ ID NO: 1

Gln	Arg	Arg	Ser	Arg	Thr	Asn	Phe	Thr	Leu	Glu	Gln	Leu	Asn	Glu	Leu
1			5						10					15	
Glu	Arg	Leu	Phe	Asp	Glu	Thr	His	Tyr	Pro	Asp	Ala	Phe	Met	Arg	Glu
		20						25					30		
Glu	Leu	Ser	Gln	Arg	Leu	Gly	Leu	Ser	Glu	Ala	Arg	Val	Gln	Val	Trp
		35				40						45			
Phe	Gln	Asn	Arg	Arg	Ala	Lys	Cys	Arg	Lys	Gln	Glu				
	50					55					60				

(2) Informace o sekvenci SEQ ID NO: 2:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 209 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

(A) POPIS: /popis = „exon II: ET93“

(v) TYP FRAGMENTU: lineární

(xi) POPIS SEKVENCE: SEQ ID NO: 2

GGATTTATGA ATGCAAAGAG AAGCGCGAGG ACGTGAAGTC GGAGGACGAG GACGGGCAGA	60
CCAAGCTGAA ACAGAGGGCG AGCCGCACCA ACTTCACGCT GGAGCAGCTG AACGAGCTCG	120
AGCGACTCTT CGACGAGACC CATTACCCCG ACGCCTTCAT GCGCGAGGAG CTCAGCCAGC	180
GCCTGGGGCT CTCCGAGGCG CGCGTGCG	209

(2) Informace o sekvenci SEQ ID NO: 3:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 368 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

(A) POPIS: /popis= „exon I: G310“

(xi) POPIS SEKVENCE: SEQ ID NO: 3:

GTGATCCACC CGCGCGCACG GGCCGTCCTC TCCGCGCGGG GAGACGCGCG CATCCACCAG	60
CCCCGGCTGC TCGCCAGCCC CGGCCCCAGC CATGGAAGAG CTCACGGCTT TTGTATCCAA	120
GTCTTTTGAC CAGAAAAGCA AGGACGGTAA CGGCGGAGGC GGAGGCGGCG GAGGTAAGAA	180
GGATTCCATT ACGTACCGGG AAGTTTGGGA GAGCGGACTG GCGCGCTCCC GGGAGCTGGG	240
GACGTCGGAT TCCAGCCTCC AGGACATCAC GGAGGGCGGC GGCCACTGCC CGGTGCATT	300
GTTCAAGGAC CACGTAGACA ATGACAAGGA GAAACTGAAA GAATTCGGCA CCGCGAGAGT	360
GGCAGAAG	368

(3) Informace o sekvenci SEQ ID NO: 4:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 58 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „exon III: ET45“

(xi) POPIS SEKVENCE: SEQ ID NO: 4:

GTTTGTTCC AGAACCGGAG AGCCAAGTGC CGCAAACAAG AGAATCAGAT GCATAAAG

58

(2) Informace o sekvenci SEQ ID NO: 5:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 89 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „exon IV: G108“

(xi) POPIS SEKVENCE: SEQ ID NO: 5:

GCGTCATCTT GGGCACAGCC AACCCACTAG ACGCCTGCCG AGTGGCACCC TACGTCAACA

60

TGGGAGCCTT ACGGATGCCT TTCCAACAG

89

(2) Informace o sekvenci SEQ ID NO: 6:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 1166 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „exon : Va“

(xi) POPIS SEKVENCE: SEQ ID NO: 6:

```

GTCCAGGCTC AGCTGCAGCT GGAAGGCGTG GCCCAGCGGC ACCCGCACCT GCACCCGCAC      60
CTGGCGGGCG ACGCGCCCTA CCTGATGTTC CCCC GCCCGCCTTCCGCT GCCCATCGCG      120
TCGCTGGCCG AGTCCGCCTC GGCCGCCGCC GTGGTCGCCG CCGCCGCCAA AAGCAACAGC      180
AAGAATTCCA GCATCGCCGA CCTGCGGCTC AAGGCGCGGA AGCACGCGGA GGCCCTGGGG      240
CTCTGACCCG CCGCGCAGCC CCCC GCCCGCCTTCCGCT GCCCATCGCG      300
GCACCCGCAGC TCCTGCACTC AACCCCGCCT GGAGCTCCTT CCGCGGCCAC CGTGTCCGG      360
GCACCCCGGG AGCTCCTGCA AGAGGCCTGA GGAGGGAGGC TCCCGGGACC GTCCACGCAC      420
GACCCAGCCA GACCCTCGCG GAGATGGTGC AGAAGGCGGA GCGGGTGAGC GGCCGTGCGT      480
CCAGCCCGGG CCTCTCCAAG GCTGCCCGTG CGTCTGGGA CCCTGGAGAA GGTTAAACCC      540
CCGCTGGCT GCGTCTTCTT CTGCTATACC CTATGCATGC GGTTAACTAC ACACGTTTGG      600
AAGATCCTTA GAGTCTATTG AACTGCAAA GATCCCGGAG CTGGTCTCCG ATGAAAATGC      660
CATTCTTCG TTGCCAACGA TTTCTTTAC TACCATGCTC CTTCTTCAT CCCGAGAGGC      720
TGCGGAACGG GTGTGGATTT GAATGTGGAC TTCGGAATCC CAGGAGGCAG GGGCCGGGCT      780
CTCCTCCACC GCTCCCCCGG AGCCTCCCAG GCAGCAATAA GGAAATAGTT CTCTGGCTGA      840
GGCTGAGGAC GTGAACCGG GGCTTTGGAA AGGGAGGGGA GGGAGACCCG AACCTCCCAC      900
GTTGGGACTC CCACGTTCCG GGGACCTGAA TGAGGACCGA CTTTATAACT TTTCCAGTGT      960
TTGATTCCCA AATTGGGTCT GGTTTTGTGTT TGGATTGGTA TTTTTTTTTT TTTTTTTTTT     1020
TGCTGTGTTA CAGGATTCAG ACGCAAAGA CTTGCATAAG AGACGGACGC GTGGTTGCAA     1080
GGTGTACATC TGATATGCAG CATTAACTTT ACTGACATGG AGTGAAGTGC AATATTATAA     1140
ATATTATAGA TAAAAAAA AATAGC                                             1166

```

(2) Informace o sekvenci SEQ ID NO: 7:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 625 párů bází
- (E) TYP: nukleová kyselina
- (F) DRUH ŘETĚZCE: jednořetězcová

(G) TOPOLOGIE: lineární

(iii) TYP MOLEKULY: jiná nukleová kyselina

(A) POPIS: /popis = „exon Vb“

(xi) POPIS SEKVENCE: SEQ ID NO: 7:

ATGGAGTTTT GCTCTGTCTG CCCAGGCTGG AGTATAATGG CATGATCTCG ACTCACTGCA	60
ACCTCCGCCT CCCGAGTTCA AGCGATTCTC CTGCCTCAGC CTCCCAGTA GCTGGGATTA	120
CAGGTGCCCA CCACCATGTC AAGATAATGT TTGTATTTTC AGTAGAGATG GGGTTTGACC	180
ATGTTGGCCA GGCTGGTCTC GAACTCCTGA CCTCAGGTGA TCCACCCGCC TTAGCCTCCC	240
AAAGTGCTGG GATGACAGGC GTGAGCCCCT GCGCCCGGCC TTTGTAACCT TATTTTAAAT	300
TTTTTTTTTT TTTAAGAAA GACAGAGTCT TGCTCTGTCA CCCAGGCTGG AGCACACTGG	360
TGGATCATA GCTCACTGCA GCCTCAAACCT CCTGGGCTCA AGCAATCCTC CCACCTCAGC	420
CTCCTGAGTA GCTGGGACTA CAGGCACCCA CCACCACACC CAGCTAATTT TTTTGATTTT	480
TACTAGAGAC GGGATCTTGC TTTGCTGCTG AGGCTGGTCT TGAGCTCCTG AGCTCCAAAG	540
ATCCTCTCAC CTCCACCTCC CAAAGTGTTA GAATTACAAG CATGAACCAC TGCCCGTGGT	600
CTCCAAAAAA AGGACTGTTA CGTGG	625

(2) Informace o sekvenci SEQ ID NO: 8:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 15577 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „HOX93“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: exon
- (B) POZICE: 1498..1807
- (D) JINÉ INFORMACE: /funkce = „část exonu I (G310)“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: misc_
- (B) POZICE: 3844..4068
- (D) JINÉ INFORMACE: /funkce = „oblast pET92 (první část)“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: misc_
- (B) POZICE: 4326..4437
- (D) JINÉ INFORMACE: /funkce = „oblast pET92 (druhá část)“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: misc_
- (B) POZICE: 4545..4619
- (D) JINÉ INFORMACE: /funkce = „oblast pET92 (třetí část)“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: misc_
- (B) POZICE: 5305..5512
- (D) JINÉ INFORMACE: /funkce = „část exonu II (ET93)“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: misc_

(B) POZICE: 11620..11729

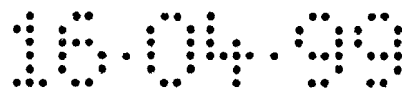
(D) JINÉ INFORMACE: /funkce = „část exonu IV
(G108)“

(xi) POPIS SEKVENCE: SEQ ID NO: 8:

CTCTCCCTGT TGTGTCTCTC TTTCTCTCTC TCCATCTCTC TCCGTCTTTC CCCCTCTGTC	60
TCTTTCTCTG TCTCCATCCC TCTGTCTCTC CCTTCTCTC TGTCTTTCCT TGTCTCTCTC	120
TTTCTCTCTC TCTCTCCATC TCTCTCTCTC CCGGTCTCTC TCTCTCCATC TCCCCGTCTC	180
TCCGTTTCTC TCTCTGCCTC TCCCTGTCG TCTCTCTCTT TGTGTGTGTT ACACACACCC	240
CAACCCACCG TCACTCATGT CCCCCACTG CTGTGCCATC TCACACAAGT TCACAGCTCA	300
GCTGTCATCC TGGGTCCCCA GGCCCCGCCG GGGAGGAAGA TGCGCCGTGG GGTACGGGA	360
GGAAGGGGAC TCCGGGCCTC CTGGTGCCCC ACTTTATTG CAGAAGGTCC TTGGCAGGAA	420
CCGTGACGCG TTTGGTTTCC AGGACTTGA AAACGAATTT CAGGTCGCGA TGGCGAGCAC	480
CGGCTTCCCC TGAAGCACAT TCAATAGCGA GAGGCGGGAG GGAGCGAGCA GGAGCATCCC	540
ACCATGAAAA CCAAAAACAC AAGTATTTTT TTCACCCGGT AAATACCCCA GACGCCAGGG	600
TGACAGCGCG GCGCTAAGGG AGGAGGCCTC GCGCCGGGGT CCGCCGGGAT CTGGCGCGGG	660
CGGAAAGAAT ATAGATCTT ACGAACCGGA TCTCCCGGGG ACCTGGGCTT CTTTCTGCGG	720
GCGCTGGAAA CCCGGGAGGC GGCCCCGGGG ATCCTCGGCC TCCGCCGCCG CCGCCTCCCA	780
AGCGCCCGCG TCCCGGTTG GGGACACCCG GCCCCTCTT CTCACTTTCG GGGATTCTCC	840
AGCCGCGTTC CATCTACCA ACTCTCCATC CAAGGGCGCG CCGCCACCAA CTGGAGCTC	900
ATCTTCTCCC AAAATCGTGC GTCCCCGGGG CGCCCGGGTC CCCCCCTCG CCATCTCAAC	960
CCCGGCGCGA CCCGGGCGCT TCCTGGAAAG ATCCAGGCGC CGGGCTCTGC GCTCCTCCCG	1020
GGAGCGAGGG CGGCCGACA ACTGGGACCC TCCTCTCTCC AGCCGTGAAC TCCTGTCTC	1080



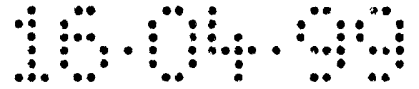
TCTGTCTCTC TCTGCAGGAA AACTGGAGTT TGCTTTTCCT CCGGCCACGG AAAGAACGCG	1140
GGTAACCTGT GTGGGGGGCT CGGGCGCCTG CGCCCCCTC CTGCGCGCGC GCTCTCCCTT	1200
CCAAAAATGG GATCTTTCCC CCTTCGCACC AAGGTGTACG GACGCCAAC AGTGATGAAA	1260
TGAGAAGAAA GCCAATTGCC GGCCTGGGGG GTGGGGGAGA CACAGCGTCT CTGCGTGCCT	1320
CCGCCGCGGA GCCCGGAGAC CAGTAATTGC ACCAGACAGG CAGCGCATGG GGGGCTGGGC	1380
GAGGTCGCCG CGTATAAATA GTGAGATTTC CAATGGAAAG GCGTAAATAA CAGCGCTGGT	1440
GATCCACCCG CGCGCACGGG CCGTCCTC TC CGCGGGGA GACGCGCGCA TCCACCAGCC	1500
CCGGCTGCTC GCCAGCCCCG GCCCCAGCCA TGGAAGAGCT CACGGCTTTT GTATCCAAGT	1560
CTTTTGACCA GAAAAGCAAG GACGGTAACG GCGGAGGCGG AGGCGGCGGA GGTAAGAAGG	1620
ATTCCATTAC GTACCGGAA GTTTTGGAGA GCGGACTGGC GCGCTCCCGG GAGCTGGGGA	1680
CGTCGGATTC CAGCCTCCAG GACATCACGG AGGGCGGCGG CCACTGCCCG GTGCATTTGT	1740
TCAAGGACCA CGTAGACAAT GACAAGGAGA AACTGAAAGA ATTCGGCACC GCGAGAGTGG	1800
CAGAAGGTAA GTTCCTTTCG GCGCCGCTC CAGGGGGGCC TCCTGGGGT TCGGCGCCTC	1860
CTGCCACGG AGTCGGCCCC GCGCGCCCT CGCTGTGCAC ATTTGCAGCT CCCGTCTCGC	1920
CAGGTAAGG CCCGGCCCGT CAGGCTTTCG CTAAGAAAGG AAGGAAGGCA GGAGTGGACC	1980
CGACCGGAGA CGCGGGTGGT GGGTAGCGGG GTGCGGGGGG ACCCAGGGAG GGTGCGAGCG	2040
GGGGCCGCGC GCGTGGGCAC CGACACGGGA AGGTCCCGGG CTGGGGTGGGA TCCGGGTGGC	2100
TGTGCTGAA GCCGTAGGGC CTGAGATGTC TTTTTCATTT TCTTTTCTT TCCTTTCCTT	2160
TTTTTGTGG TTTGTTTGTG TGTTGAGAC AGAGTCTCGC TCTGTCCCC AGGCTGGAGT	2220
GCAGTGGTGC GATCTCGGCT CACTGCAACC TCCGCTCCT GGGTTCAAGC GATTCTCCTG	2280
CCTCAGCCTC CCCAGTAGCT GGGATTACAG GCATGCACCA CCACGCCTGG CTAATTTTTG	2340
TGCTTTTAGT AAAGACGGGG ATTCACCATG TTGGCCAGGC TGGTCTCGAA CTCCTGACCT	2400
CAGGTGATCC ACCCGCCTCG GCCTCCCAA GTGCTGGGAT GACAGGCGTG AGGCACCGG	2460
CCCGGCCTGG GTCCTGACGG CTTAGGATGT GTGTTTCTGT CTCTGCCTGT CTGCCTTGTA	2520
TTTACGGTCA CCCAGACGCA CAGAGGAGCC GTCTCCACGC GCCTTCCCAG CGCTCAGCGC	2580
CTGCCGGGCC CCCGGAGATC ACGGGAAGAC TCGAGGCTGC GTGGTAGGAG ACGGGAAGGC	2640
CCCGGGTCAG CTCGGTTCTG TTTNCTTTA AGGAACCCTT CATTATTATT TCATTGTTTT	2700
CCTTTGAACG TCGAGGCTTG ATCTTGGCGA AAGCTGTTGG GTCATAAAA ACCACTCCCG	2760
TGAGCGGAGG TGGCCGGGAT CTGGATGGGG CGCGAGGGC CCCGGGGAAG CTGGCGGCTT	2820
CGCGGGCGCG TCCTAAGTCA AGGTTGTCAG AGCGCAGCCG GTTGTGCGCG GCCCGGGGNG	2880
AGCTCCCCTC TGGCCCTTCC TCCTGAGACC TCAGTGGTGG GTCGTCCCGT GGTGGAAATC	2940
GGGGAGTAAG AGGCTCAGAG AGAGGGGCTG GCCCGGGGA TCTCTGTGCA CACACGACAA	3000
CTGGGCGGCA TACATCTTAA GAATAAAATG GGCTGGCTGT GTCGGGGCAC AGCTGGAGAC	3060



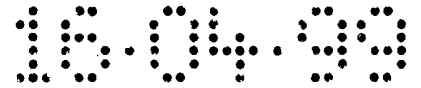
GGCTATGGAC GCCTGTTATG TTTTCATTAC AAAGACGCAG AGAATCTAGC CTCGGCTTTT	3120
GCTGATTTCG AAAGTTGAGG TGCGAGGGTG AATGCCCAA AGGTAATTCT TCCTAAGACT	3180
CTGGGGCTAC CTGCTCTCCG GGGCCCTGCA TTTGGGGTGT GGAGTGGCCC CGGGAAATAG	3240
CCCTTGATT CGTAGGAGGC ACCAGGCAGC TTCCAAGGC CCTGACTTTG TCGAAGCAGA	3300
AAGCTGTGGC TACGGTTTAC AAAGCAGTCC CCGTTTCTG ACCGTCTAAG AGGCAGGAGC	3360
CCAGCCTGCC TTTGACAGTG AGAGGAGTTC CTCCTACAC ACTGCTGGG GCACCCGGCA	3420
CTGTAATTCA TACACAGAGA GTTGGCCTTC CTGGACGCAA GGCTGGGAGC CGCTTGAGGG	3480
CCTGCGTGTA ATTTAAGAGG GTTCGCANGC GCCCGCGGC CGCTTCTGNT GGGTTGCTT	3540
TTTGGTTGTC CTTNGCAA CACCGTTTTG CTCCTCTNGN AACTCTCTCT TNCTCCCCN	3600
TGGCCNGTNG GACCCGGGNA NGAGCAAAGT GTCTCCAGA CCNTTTTGAA ANGTGAGAGG	3660
AAAATAAAGA CCAGGCCAAA NNGACCCAGG GCCACAGGAG AGGAGACAGA GAGTCCCCGT	3720
TACATTTTNC CCCTTGGCTG GGTGCAGAAA GACCCCGGG CCAGGACTGC CACCCAGGCT	3780
ACTATTTTAT CATCAGATCC AAGTTAAATC GAGGTTGGAG GGCAGGGGAG AGTCTGAGGT	3840
TACCGTGGAA GCCTGGAGTT TTTGGGNAAC AGCGTGTCCC CGCCGAGCCT GGGAGCCCGT	3900
GGGTTCTGCA AAGCCTGCGG GTGTTTGAGG ACTTTGAAGA CCAGTTTGTC AGTTGGGCTC	3960
AATTNCCTGG GGTTCAGACT TAGAGAAATG AAGGAGGGAG AGCTGGGGTC GTCTCCAGGA	4020
AACGATTCAC TTGGGGGGAA GGAATGGAGT GTTCTTGCA GACATGTCT GTTAGGAGGT	4080
GAAACAGAAAT GTGAAATCCA CGTTGGAGTA AGCGTCCAGC GCTGAATGTA GCTCGGGGTG	4140
GGGTGGGAGG GCCCTGGTGT GGATCGTGA AGGNAAGAAA GACAGAACAG GGTGCTAGTA	4200
TTTACCCCGT TNCCCTGTAG ACACCCTGGA TTTGTCAGCT TTGCAAGCTT CTGGTTGCA	4260
GCGGCCTTGC CTGTGCCCT TTAGACTGT TTCCAGACTA AACTTCCAAA TGTCAGCCCC	4320
TTACCCTTGA CAGCAAGGGA CATCTCATT GGGCATCGCG TGCTTCTCAT CTGTGNCTCA	4380
GCAGGCCNG AGATAGGAAN CANGAGGGC NGTTGNAGA TGCNCACTC CACCAGCCCT	4440
GGGNTTGAAG GGGANGCGAN GGGANGACNA CCTTTTANCT TAAACCCCTN GAGCTTGGTN	4500
CAGAGAGGNC TGAATGTCTA AAATGAGGAA GAAAAGGTTT TTCACCTGGA AACGCTGAG	4560
GGCTGAGTCT TCTGCCNTT CTGACNTCCC CCAGCAAATA CAGACAGGTC ACCAANTAC	4620
TGGAGATGAG AAAGTGCCAT TTTTGGCACA CTCTGGTGGG GTAGGTGCC GACCGCGTGT	4680
GAAAAANGTG GGAANNNGAG AGATTTCTGN CGCACCGGT TCAGCCCCCA GGCGGGNTG	4740
GCNGCATTCN AGGNTACTCA GACGCGGTTG TGCTGTCTG CTGAGAAACA GGCTTCGGGT	4800
AGGGGCTCCT AGCTCCGCCA GATCGCGGAG GGACCCCG CCCTCCTGCG CTGCAGCGGT	4860
GGGGATAGCG TCTCTCCGTA GGCCTAGAAT CTGCAACCCG CCCCAGGTCC TCCCCGTGC	4920
CTTCCCGGGC GTCCCGCCGG GGATCCACA GTTGGCAGCT CTTCCCTCAA TTCTTTCCCT	4980
TAAAAATAGG ATTTGACACC CCACTCTCCT TAAAAAATA AAATAAGAAA AAAAGGTTAG	5040
GTTATGTCAA CAGAGGTGAA GTGGATAATT GAGGAAACGA TTCTGAGATG AGGCCAAGAA	5100



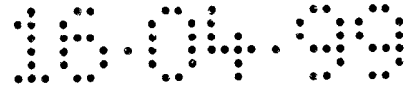
AACAACGCTC GTGCAAAGCC CAGGTTTTTG GGAAAGCAGC GAGTATCCTC CTCGGCTTTT	5160
GCGTTATGGA CCCCACGCAG TTTTTCGCTC AAAGCGCATT GGTTCGAG GGCCCCCTTT	5220
CCACCGCGGG ATGCACGAAG GGGTTCGCCA CGTTGCGCAA AACCTCCCCG GCCTCAGCCC	5280
TGTGCCCTCC GCTCCCCACG CAGGGATTTA TGAATGCAAA GAGAAGCGCG AGGACGTGAA	5340
GTGGGAGGAC GAGGACGGGC AGACCAAGCT GAAACAGAGG CGCAGCCGCA CCAACTTCAC	5400
GCTGGAGCAG CTGAACGAGC TCGAGCGACT TTTTGACGAG ACCCATTACC CCGACGCCTT	5460
CATGCGCGAG GAGCTCAGCC AGCGCCTGGG GCTTCCGAG GCGCGCGTGC AGGTAGGAAC	5520
CCGGGGGCGG GGGCGGGGGG CCCGGAGCCA TCGCCTGGTC CTCGGGAGCG CACAGCACGC	5580
GTACAGCCAC CTGCGCCCGG GCCGCCCGCC TCCCCTCCC GGAGCGCGGG GAGGTTGGGT	5640
GAGGGACGGG CTGGGGTTCC TGGACTTTTG GAGACGCCTG AGGCCTGTAG GATGGGTTC	5700
TTGCGTTTGT TTTTACCAA CAGCAAACAA ATATATATAC ATATATATTA TACAATAAC	5760
AAATAAATAT ATATGTTATA CAGATGGGTA TATTGTATAT ATTATAGATA TTTGTTCGTC	5820
CTGGTGCAA AGACACCCGG TGAACCCATA TATTGGCTCC TGACTGCCTT CGGTTCCCCT	5880
GGGATTGGTT ATAGGGGCAA CACATGCAAA CAAAACCTTC CCTGGATTAT ACTTAGGAGA	5940
CGAAGCTACA GATGCGTTTG ATCCAGAGTG TTTTACAAGA TTTTTCATT AAAAATAAT	6000
GTGTCTTTTG GCCCTGATT CCCCTCCGTC TTCCCGTGTG GCTGCATTGA AAAGGTTTCC	6060
TTAGGATGAA AGGAGAGGGG TGTCCTCTGT CCCTAGGTGG AGAGAAACAG GGTCTTCTCT	6120
TTCTCCGTT TTTTACCTA CCGTTTCTAT CTCCCTCCTC CCCTCTCCAG CCCTGTCCTC	6180
TGCTACAAAC CACCCCTCC TCCCTCCGGC TGTGGGGAGC GCAGGAGCAC GTTGGGCATC	6240
TGGATGAGCG GNAGACTATT AGCGGGGCAC GGGGGCTCCC CGAGGAGCGC GCGAATTCAC	6300
GCTGCCCCAT GAGACCAGGC ACCGGGGGGC GGAGGGGCCT TGGGTGTCCG CAGAGGGAGC	6360
GGCGGGCAGA GCCTTCTTCC GCATTCTAAA CATTCACTTA AAGGTATGAG TTTANTTTCA	6420
GGGGTGCTGC TGGGAGAGCC TCCAAATGGC TTCTTCCAGC CCCTGCCTGA CAGTTCAGCT	6480
CCCCTGGAAG GTCAACTCCT CTAGTCCTTT CTCCTGGTTC TGGGCAGGAC AGAAGTGGGG	6540
GGAGGGAGAG AGAGAGAGAG AGAGAGAGAG ACGGTCAGGA TCCCCGACC CTGGGGAACC	6600
CGTCAAAAAT AAATGAAATT AAGATTGCCG ACCAGAGAGA GAACCGTGAC AAAGCAAACG	6660
GCGTTCAAAG CAAAGAGACG AACTGAAAGC CCGTTCCCGT AGGACTGGTT ATGAGGTCAA	6720
CACATTCAAA CACAGCTTGC TCTGGATTTT GCTGAGCAGA GGAAGATACA GATGCATTTG	6780
ATCCAAAGTG TGTTACATCT TTCATTATAT GTGTGTCTAT ATATATAAAC ATATATAAAT	6840
ATATAAACAT ACATAAATGT ATGTAAATAT ATATAATCTA TATACATATA TAAATATATA	6900
AACACATATA TAATATATAA ATCTATAAAC ATATATAATA TATAAACATA AATATATAAA	6960
CATATAAAT ATATAAATAT ATTAACATAT ATAAAATATG TATAAATATA TATAAACATA	7020
TAAACATATA TAAATATATA AACATATAAA TATATAAACA TATATAAATA TATACAAACA	7080



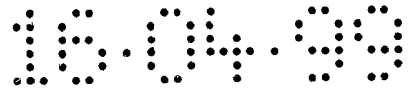
TATTGTATAT ATATAAATAT ATATAAAAAC ATATATATAC ATATAAAAAT ATATATAAAC	7140
ATATATACAT ATAAAGAAAT ATATATAAAC ATATATACAT ATAAATATAC ATATATAAAC	7200
ATATATATAC ATAAAATATA TATAAACATA TATACATATA AAAATATATA TATATTAACA	7260
TATATATACA TATAAAAATA TATATATTAA CATATATATA CATATAAAAA TATATATATA	7320
TTTTTGCCC CTGATTCCCT TCGGTTCCCTG TGGGATGGGT GATTGAGTCA ACACATTCAA	7380
ACACAACTTT TCCATCGATG TTGCTTAGGA GATGAGGATA CAGATGCGTT TGATGGAGAG	7440
GGTTTTACAA GCTCTTTCAT TTAAATATAT ATATATATAT ATATATATTT TTTGGCTCCT	7500
GATTCTCTTC CGTCTTCCCA TGTGGCTGCA TTTTAAAAGG CTTCCCTAAG ATCGTTACGA	7560
TTAAATCAAC CCTCCCCAGG CATCTTTACC GAGGGCTGTG GTCCCAAAG CGATACAGCC	7620
CAGGAGGGAG AGAGGCTTTG GTGACTTGGG GGAAGGACTG TGTCCCTCCT TAGGGCGTCT	7680
GTGGCCTCAG TGAGGGAAGG AAGCTGCATC AGACAGGGGT TTCCTCGCTG TCCACCCCTC	7740
TGGCAGAAGA TGGATTGGGC TGCCCCGNTA TAAATTAATG AAAAGATTAA AGTTTCGCTA	7800
AAGGGGACAT CGAGTTTATG TGTCATCTCC TGGTGNTCTG TGTGCCNTGG GATNCTGCAA	7860
TATATCCCAN NGCCCTTGAT GNNNTACTGT TTNCTATAAA AANNTAAATN TACTTGTNNA	7920
ATTTAANTTC CNNNACACTA TTNCTTTCC NNGTNAGTCT NATTANCCGA NCGAGAGCAN	7980
CGNTTAGTIN CAGCTNGCGG AAAATTGGTT GTGGGGTGTG TGCGGACCCC NGAGNAACGC	8040
CCNNTAAAT NAAAGACAAA NTCNGGGGAC AAGNCTNGGG GGTATCGNN ATTGCNNAGG	8100
GGTCGNATG AAAANTTTAA CGACGGTAAA TAATAATAAA AANNCAAACA TGGGAATGNC	8160
AATAAAGAC ATAATTCTCC NNATCSCCGC GGGGGGAAAG GATCCTATAG TAAAGGCGAG	8220
TGCGCTTTGA GGGGTCAATA AAATCAATTA GTTCCAACAC CCACGTCCCG CGTTGAGGGG	8280
ACGGGGACGA GCAGGGACAG AAAAAGAAAC CATATTTGAA TCCCATCTCT CTGTGAATC	8340
TTGGGTCAACA TGCGTCTCAG TACAGCCCGT CCCGTGCTGT GACCGGATAG AGTTTCAATT	8400
TACTGTGGAA ATTTGCTGTA AATAAATTGA GCATCCGATA GAAGCTGTG CTGATTAACC	8460
TTTTATTTTT AGCGTGCCCC TGCAAAGTCG TATCACCCAG CTGTCAGGCT TCTAATCGAA	8520
AGTTATGAGA CCACGGTGAG GGCAGGCGG TAATTTAATT ACAACAAATA TCTTTGGGTT	8580
TATGGCGCAG AGCTAAATTA AATGTCATTA TTCACTGTCT GTNAATGGNA AATCAAAAANN	8640
GGAAATCGCA NTTACGGNCA TTTGGGNNA ANGAAAGCGG GGNAGTGCTC TTTAATNGAA	8700
NNGAAATAAC TGCTTAAGC AGTGTCACAC ACTTCACTTA CCATATTCGN GGCCTNAATT	8760
GGAANNTGGA TCGTNNGAAT CACTCCNAAG ACTNGATTTA TTANGCGCTT CACGNCAGCN	8820
NGGCNTAATT CATCNACTN NGTATTCTTC ATCNNNNATT TTTTTTTTTC CTCTCNNGCC	8880
GTGTTNNGAA GGGAGAGTGA ATGAGGCTTT CCACGTTTCA GGAGGATTTT CTTTTTTGAA	8940
AAATGCCCTT CCAGAGGCTT TTGGGTGGCT GGCTTGCTTT CTGGGCCCTG GAGGANGACA	9000
GGCGGANGAG TCCAGGTGGG CATGGAGAGG CACAGTGGCA GGTCACCTGG ATGGTCAGTG	9060
GAGGTGGAGG TCTGAAGGCG CCAGCTTTGG AAATTATTGG TGAATTTCCA TGTCAGCACC	9120



AGGNCAGGGG CCTTTTGGC GGGGGTGTGA GGGANGGATG ANCTTGCTG GGAAANNCAG	9180
GATCAGGTTT TCCAGGCGCA CTGCAGCCCG GTAGGACCCA CTTTGGAAAT GAAAAGCCAG	9240
TTNCCGAAAG CTGGGCTGGA AGCTTCCGTG TTGGGTTCAA GAGCAAGTTC ACCTTGCGCT	9300
GTGTAGACTC CTGGCTGCTC CCAAACCTG AGGTTTTTCT GAGGTTCCCT TCATAGGGGC	9360
ACCGGCCCTG GGCCATGCAC AGTGCGTAAG GGTGGCTGTG GGCCGAGGGA CCCAGCACGT	9420
GTTTTGCCCA CAACAGCCGG AGTACTGGT TCACTCACCG CCTTGGCGGA GGACGCCTGT	9480
TCTCTGGACG AATCATTCT CTGGGTGGT GACTGCCTTG TGGGTCAAGG TGCAGGTTTT	9540
CTGCCACAGA AAACCTGTTA GGAGGAATTA AGCGACTAAG ACTGTCAGGG AGGTGGTGGT	9600
GGGGGANGAG GNAGGGGGTG GTGTCCAGAT TACCAGGCAT AGGCTAAACT GCCTGCACTC	9660
TCCAGCTGGT CTGTCTGTGG AGGAGGGGAT TGCAATACT GGGAGAGCAG AGGAGGCTCG	9720
TAGGAGGTGA GAGGGGGTGG AATTGTCATG CAAATCTTCA CATGAGGCCT GTGTGAATTT	9780
CTCCAGCCTC CTGAGGGTCC CCTGCGCTAT TGCCTCAAC TTCTTGATAG TTTACCCCAA	9840
GACTCAGAAG TCCTTAGAGG GGCAGAATGC CCCACCACA AAGCCTGCTA TCCTTGGGCG	9900
TCCTCAGGAC CCTTGGTCAT GAATGGGACC CTTTCATGTA TGGGGACCCT TGTAATATG	9960
AATGGGACGC CTCAGCTCC CCAGGGCTTC CGAGGAGGCC GAGAAGGGCA AAGACACTTC	10020
CGAGGAGGCC GAGAAGGGCA AAGACATTTT CTGGGCTTGG TGTGTCAAGA GCTAGATTGG	10080
AGAAGGGGCT GGATTTGGAA CTCTTTAGCC ATCAGCTCAC CCTCTCCGTT TGTGGCTAAA	10140
GTCTGAAGGT GGAAACTTCG GTTCTCCTAC AGGGTCTACA GGAGTTGGGG GGCGGGGCGC	10200
CCACACAGAA CGCTGGAAAG TTCGACAGTC CACTTCCACT GGCTCGGAAC TCCTTTTTC	10260
ACCTTAAGTT CATCAGCGGT AACGCATAGG TCTCACTTAG GCAGGGCAGG GATGATTTAA	10320
CAATTTCTAC TTCTAGGTCA GGTGCGGTGG CTCACACCTC TAATCCAGC ACTTTGGGAG	10380
GCCCAGGAGG GTGGATCGCT TGAGGTCAGG AGTTTGAGAC CAGCCTGGCC AACATGGTGA	10440
AACCCCGTCT CTAATAAAT ACGAAAATTA GCCAGGCATG GTGGTGAGCA CCTGTAATTC	10500
CAGCTACTCG GGAGGCTGAG GCAGGAGAAT CGCTTGAACC TGGGAGGTGG ACCTTGCACT	10560
GAGGTGAGAT CACACCACTG CACTCCAGCC TGGATGAGAG AGCAAGACTC TGTCTCAAAA	10620
ACAAAATAAA AAAAAACAA AACAAAATC AAAAAAGAAA ACCCAATTTT CAGTTCTAGG	10680
CCAGGTGCAG TGGCTACGC CTGTCATCCC AGCACTTTGG GAGGCCCAGG AGGGTGGATC	10740
GCTTGAGGTC AGGAGTTCGA GACCAGCCTG GCCAACATGG TGAAACCCCA TCTTTACTAA	10800
AAATACAAAC GTTAGCTGGG TGTGGTGGTG TGCCTGTGTA ATCCCAGCTA CTCGGGAAGC	10860
TGAGGCTGGA GAATGCTTG AATCTGGGAG GTGGAGGTTG CAGGGAGGCG AGATAGTGCC	10920
ACTGCAGTCC AGCCTGGACC AGAGAGCAAG ACTCCGTCTC AAAAAACAAA GAAAGCAAAA	10980
ACAAAAACA AGAGACCAGC CTGGCCAACA TGGTGAACC GCGTCTTTAC TAAAATACAA	11040
AATTAGCCGG GCATGGTGGT GGGCACCTGT AGTCCAGCT ACTCGGGAGG CTGAGGCAGG	11100



AGAATGGCTT GAACCTGGGA GGTGGAGCTT GCAGTGAGCC GAGATAGTGC CACTGCACTC 11160
 CAGCCTGGGC GACAGAGCGA GACTTGATT CAGAACCACC ACCACCACAA CAAAACAAA 11220
 CAAAAATCC AAAAAACCC CAATTTCCAG TACTAGGTAG TCAGTGATGC AGGGCTGGAG 11280
 ACAGAGGGGC GGTAAGTGC TGGGCGCCCA CCATCAGTCA CCTCCCAGCT CCCANGAGGT 11340
 GCAAAGTGCT TGGTTCAGCC TCATGGGAAG GATGCTCCCT GGGGAGGCTG GGCTGGGTTC 11400
 ACAGGGCTCT TCACATCTCT CTCTGCTTCT NCCCAAGGT TTGGTTNCCA GAACCGGAGA 11460
 GCCAAGTGCC GNCAACAAG AGAATCAGAT GCATAAAGGT GGGTGTCTGG ACTGGGGGA 11520
 CCTGAAGCTG GGGATCCTG CTCCAGGAGG GATGGGGTCG ACAAGGTGCT GGCTACACC 11580
 AGGACCACCA CACTGACACC TGCTCCCTTT GGACACAGGC GTCATCTTGG GCACAGCCAA 11640
 CCACCTAGAC GCCTGCCNGA GTGGCACCT ACGTCAACAT GGGAGCCTTA CGGATGCCTT 11700
 TCCAACAGGT AGCTCACTTT TTCTTCTCT GNAAGATCCC TAGGGACCTG CTGCTCCCTT 11760
 CCCCTTTCCC CTATTTGCTG CCGCATCTG AACTCCTAG TCCCTCCCTG CCCCTGCAGA 11820
 CTTCTCAGCT GGCCCTTGA AAAAAAGCCT CTTTTCCGAG GAGGCATTA CAGGCACCTT 11880
 GGCACCTATG AAATCAGGCT GGGCCAGGCG GGGTGGCTCA CACCTGTCAT CCCAGCACTT 11940
 TGGGAGGCCA AGGTTAGGAG TTTGAGACCA GCCTGGACAA CATAGCAAAA GCCTGTCTCT 12000
 ACTAAAAATA CAAAAAATA TTAACAGGA GTGGTGGTGG GCACCTGTAA TCCAGCTAC 12060
 TTGGGAGGCT GAGGCAGGAG AATCACTGA ACCCGGAGG CCGAGGTTGC GGTGAGCCGA 12120
 GATCGTGCCA TTGCACTCCA GGCTGGGCGA CAGAGTGAGA CTCTGTCTCA AAAAATAAAT 12180
 AAATAAATA ATGTAATAA ATAAAAATAG GTCGGGCACG GTGGCTCACG TCTGTAATCC 12240
 CAGCACTTG GAAGGCCGAG GTGGGTGGAT GACAGGGTCA AGAGATTGAG ACCATCTGG 12300
 CCAACATGGC AAAATGCCGT CTCTACTAAA AAATACAAA ATTAGCGGG CGTGGTGGCG 12360
 GGTGCCTGTA ATCCCAGCTA CTCGGGAGGC TGAGGCAGGA GAATCGGTTG AACCCGGGAT 12420
 GCGGAGGTTG CAGTGAGCGG AGATCACATC ACTGCACTCC AGGCTGGGCA ACAAGAGCGA 12480
 AACTGCGTCT TACAATAAAT AAATAGATAA ATAAATAAAC AAATAAECTT TACTTTAGAA 12540
 ACAAATCCCT GTCCGTGTTT GTCTTTTTCAC CTGTCTGCA GGGAAAACAA AACATAAAT 12600
 GTCAAGGCAA ATAGTAGTGA TTTTATTCCG GGAAAAAGAA AGTGGATGTT TGCCTTACC 12660
 CTTTCTCGTC CTTCTCTGG TGCTCCTCAN GGCCCANGG NAGAGGGTGG AAAGTNCAGA 12720
 GGAAGAAAGA CGGGGCTGGG GGGGGGGTCC CGTGGGGACC CAGGCAGGCA TGTTCCCNAT 12780
 TTCCNTGTCT TCACNTCAA AGNAGGGGCC CCTCGNCTCT GGAATGAGGC CTACGGTTTC 12840
 CTTCCNGA AGAGTNNCC CTTTGTGAGC TTACGGCTTC GGAGTGAACC TCGGTGCAAC 12900
 CTGTTATTAA AACACACAGA GGCTAATGCC AGCAAAAACA CGCCCCCGC TCCTGGTTTC 12960
 AGAGGGAAGA AAAAAATCA TAAGCACGGC CATGCTTTTC TAATAAAAT TCATTAAATA 13020
 ATCGTTATAA GGGATGAAGC CGGGAGGGGA GAGGAGAGGA ACACAATCAA GAGACTTTCT 13080
 TTGAACTTTT TCTCCCTGCT TCAAATACAA AGCAATCTTC TGTGGGCTG GGCTGGGG 13140



GTTTCCCCCT TTCTCTGCAG CCCATTGGGA GGAAGAAAAT GCTTCCCTGA ANGTGCTGC 13200
AAAATGTTT CTGTTTTTCT TTTCTTTTC TTTTTTTTTT TTTTTTGAGA CGGAGTCTCG 13260
CTCTGTCACC AGGCTGGAGT GCAATGGTAT GATCTCAGCT CACTGCAACC TCCAGTTCC 13320
TGTTTCAAGT CATTCTCCTG CCTCAGCCTC CTGAGTAGCT GGGACTACAG GCGCCCGCCA 13380
CCACGCCCGG CTAGTGTGTTG TATTTTTAGA AAAGACAGGG TTTCCCATG TTGGCCAGGC 13440
TGGTCTTGAA CTCCTGTCCT CAAGTGATCT GCCTGCCTCG GCCTCCCAA GTGCTGTGTT 13500
TCTGTTTTTC TTTCCCGCT TTCTTAGGAG GCCATCGGGA AGAATAAAAT GCTTTCCTTG 13560
AAGTTGATGC AAAATGTTT CTGTTTTTCT TTTCTTTTTT CTTTCTTTTT GAGATGGAGT 13620
CTCGCTCTTT CACCCAGGCT GGAGGGCAGT GGC GCGACCT CGGCTCACTG CAACCTCCGC 13680
CTCCCGGGTT CAAGCGATTC TCCTGCCTCA GCCTCCGGAG TAGCTGGGAT TACAGGCACC 13740
TGCCACTATG CCTGGCTAAT TTTATTATTT TTAGTAGAGA CGGGGTTTCA CCATGTTGGC 13800
CAGGCTGGTC TCAAACCTCT GACCTCAGGT GATCCGCCCC CCTCGCCTCC CAAAGTGATG 13860
GGATGANCAG GNCATNGAGC NCACCGTCC CGGCCCTCTA ACTCTTTACC AGACATAAAG 13920
TCTCCNNTTC CCCTTCTAA ATGTATATAT TGTGTTTTTA AAAGTTAACA GCAGGGATCC 13980
CACCTCATTN CCCCCTNCT CTCCCAAGA CCTGTCTGC ACGTTGCACA CAGCAGGTGT 14040
GCCCTGGACA TATCCCAAAC CCACGCTGAA AGAAAGAGGG TCTCACTACA CGTATGATAT 14100
CTGTGNATCC TTTAAACATC TCCGTGGCTT CCAGGCAACA CAGCCATAAA TAGGAATCTC 14160
ATGTCTGACA TGATACCGGG ACCATGTATG GNAAAATCT GGGTGTGAAG TTCCAGCTAC 14220
CCCCGCAGAG GCANCCATTG CATACCCTCC AGAACTCCC CTGCCGTTNC AAGCCAAAGA 14280
CACAACACAA ACAGCNTCCG AGAGAGGGTG TCATTGAAAA TCAATACCAT CATAAGAGCA 14340
CACAGCACCG TCTTCTCTT CTGCCCGTTG ATACACAATT ATGAGCAATT TGCTAACACT 14400
GACAACTCGT GGCAAGAACA GGTCGTGTTG ATACGGTTGC CTCGTGAGGA CCCATCTGTC 14460
TTCTGGGGTC TTGCCTGGAA CGGAGATCGG AGTTCAGGGT GGCTAATAGA ATCATTACTC 14520
ACCTAGGGAC ACAGAATNAT GAGGGTTACC CCCAGTTAAG TGCATACAGT CAAACGGACG 14580
GCTGCTCTGG AAGGTACAGT GACGTGAACA GCTTTTATGA AATGCCTAGA TCTGGACCTT 14640
CCATACCTGA GCCACCGTTC CAAAGCACTG GCGTTTTTTC AGATACTTTC ATGAGAAATG 14700
TTGTCAACAC CGCAAGTTG CAGTACACAG TCTGAAAGAT ATTCTTGAT ATGTAGATGT 14760
CTGTAGATGC CCTGAAGGTG TGTAGACTTT AGACACCCAG AAGGTGTGTA GATGTCTGTA 14820
GACACCTTCT ATGTGTGTAG ATGTCTGTAG ACGCCCTGCA GGTGTGTAGA TATATCTAGA 14880
TGGTCTGCCT GTGTATGATA CAGGCTAAAA AGACATTTGT GGTGGACACT AGTTGATTAT 14940
TTAGGACTAT GAGATGGGAA AGGAAGNAGC AACCAGCAGT GAAAGGCATG TGGTGGGTGG 15000
GGGGTTGGCA TTGCAGTGGG GTCCTCNTGA NGCAGGTGAC ACCCACTATA GGGCTGCCCT 15060
TGGNATGGAC GCTTTGTNGA AGCTGTTGA TTTCACCACA CCAAGCCTGG AGGCACGGAC 15120

ATTCCAGGAT GGTGAGGAGT CTGCAAAGGA GGAGATTGGA GGAGGTGCAA TATCCCTAGA 15180
 GTACGAGAGA TGAGATAGGA GAGCTGTATA AATAGCACTA CCAGCCGGAT GCGGTGGCTC 15240
 ACGCCTGTCA TCCCAGCACT TTAGGAGGCT GAGGCAGGCG GATCACCTGA GGTCAGGAGT 15300
 TCCAGAACAG CCTGGCCAAC ACAATGAAAC CCCATCTTTA CTAAAAATAC AAGATTAGCT 15360
 GGGCACGGTG TCTCACGCCT GTCATCCCTG CACTTTGGGA GGTCGAGGTG CGCAGATCAT 15420
 GAGGTCAGTT TGGCCAACGC GCGCAAACCC CGTCTCTACT AAAAATACAA AAAAGTAGCC 15480
 GGGCGTGGTG GTGGGCACCT GTAGTCCCAG CTACTAGGGA GGCTGAGGCA GGAGAATCGC 15540
 TTGAACCCGG ATGCGGACAT TGCAGTGAGC CGAGATC 15577

(2) Informace o sekvenci SEQ ID NO: 9:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 753 párů bází
 (B) TYP: nukleová kyselina
 (C) DRUH ŘETĚZCE: jednořetězcová
 (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „segment genu ET92“

(xi) POPIS SEKVENCE: SEQ ID NO: 9:

CGTGGAAGCC TGGAGTTTTT GGAACAGCG TGTCCCGCC GAGCCTGGGA GCCCGTGGGT 60
 TCTGCAAAGC CTGCGGTGT TTGAGGACTT TGAAGACCAG TTTGTCAGTT GGGCTCAATT 120
 CCTGGGGTTC AGACTTAGAG AAATGAAGGA GGGAGAGCTG GGGTCGTCTC CAGGAAACGA 180
 TTCACTTGGG GGAAGGAAT GGAGTGTCTT TGCAGGCACA TGTCTGTTAG GAGGTGAAAC 240
 AGAATGTGAA ATCCACGTTG GAGTAAGCGT CCAGCGCTGA ATGTAGCTCG GGGTGGGGTG 300
 GGAGGGCCCT GGTGTGGATC GTGGAAGGAA GAAAGACAGA ACAGGGTGCT AGTATTTACC 360
 CCGTTCCTG TAGACACCCT GGATTTGTCA GCTTTGCAAG CTTCTTGGTT GCAGCGGCCT 420
 TGCCTGTGCC CCTTTGAGAC TGTTTCCAGA CTAACCTCC AAATGTCAGC CCCTTACCCT 480
 TGACAGCAAG GGACATCTCA TTAGGGCATC GCGTGCTTCT CATCTGTGCT CAGCAGGCC 540
 GAGATAGGAA CAGAGGGGCG TTGGAGATGC CACTTCCACC AGCCCTGGGT TGAAGGGGAG 600
 CGAGGGAGAC ACCTTTTACT TAAACCCCTG AGCTTGGTCA GAGAGGCTGA ATGTCTAAA 660
 TGAGGAAGAA AAGGTTTTTC ACCTGGAAC GCTTGAGGGC TGAGTCTTCT GCCCTTCTGA 720
 CTCCCCAGC AAATACAGAC AGGTCACCAA CTA 753



(2) Informace o sekvenci SEQ ID NO: 10:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 1890párů bazí
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „SHOXa“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: cds
- (B) POZICE: 91..968

(xi) POPIS SEKVENCE: SEQ ID NO: 10:

GTGATCCACC CGCCGCACGG GCCGTCCTCT CCGCGCGGGG AGACGCGCGC ATCCACCAGC	60
CCCGGCTGCT CGCCAGCCCC GGCCCCAGCC ATG GAA GAG CTC ACG GCT TTT GTA	114
Met Glu Glu Leu Thr Ala Phe Val	
1 5	
TCC AAG TCT TTT GAC CAG AAA AGC AAG GAC GGT AAC GGC GGA GGC GGA	162
Ser Lys Ser Phe Asp Gln Lys Ser Lys Asp Gly Asn Gly Gly Gly Gly	
10 15 20	
GGC GGC GGA GGT AAG AAG GAT TCC ATT ACG TAC CGG GAA GTT TTG GAG	210
Gly Gly Gly Gly Lys Lys Asp Ser Ile Thr Tyr Arg Glu Val Leu Glu	
25 30 35 40	
AGC GGA CTG GCG CGC TCC CGG GAG CTG GGG ACG TCG GAT TCC AGC CTC	258
Ser Gly Leu Ala Arg Ser Arg Glu Leu Gly Thr Ser Asp Ser Ser Leu	
45 50 55	
CAG GAC ATC ACG GAG GGC GGC GGC CAC TGC CCG GTG CAT TTG TTC AAG	306
Gln Asp Ile Thr Glu Gly Gly Gly His Cys Pro Val His Leu Phe Lys	
60 65 70	
GAC CAC GTA GAC AAT GAC AAG GAG AAA CTG AAA GAA TTC GGC ACC GCG	354
Asp His Val Asp Asn Asp Lys Glu Lys Leu Lys Glu Phe Gly Thr Ala	
75 80 85	
AGA GTG GCA GAA GGG ATT TAT GAA TGC AAA GAG AAG CGC GAG GAC GTG	402
Arg Val Ala Glu Gly Ile Tyr Glu Cys Lys Glu Lys Arg Glu Asp Val	
90 95 100	
AAG TCG GAG GAC GAG GAC GGG CAG ACC AAG CTG AAA CAG AGG CGC AGC	450
Lys Ser Glu Asp Glu Asp Gly Gln Thr Lys Leu Lys Gln Arg Arg Ser	
105 110 115 120	
CGC ACC AAC TTC ACG CTG GAG CAG CTG AAC GAG CTC GAG CGA CTC TTC	498
Arg Thr Asn Phe Thr Leu Glu Gln Leu Asn Glu Leu Glu Arg Leu Phe	
125 130 135	
GAC GAG ACC CAT TAC CCC GAC GCC TTC ATG CGC GAG GAG CTC AGC CAG	546
Asp Glu Thr His Tyr Pro Asp Ala Phe Met Arg Glu Glu Leu Ser Gln	
140 145 150 ?	
CGC CTG GGG CTC TCC GAG GCG CGC GTG CAG GTT TGG TTC CAG AAC CGG	594
Arg Leu Gly Leu Ser Glu Ala Arg Val Gln Val Trp Phe Gln Asn Arg	
155 160 165	
AGA GCC AAG TGC CGC AAA CAA GAG AAT CAG ATG CAT AAA GGC GTC ATC	642
Arg Ala Lys Cys Arg Lys Gln Glu Asn Gln Met His Lys Gly Val Ile	
170 175 180	



TTG GGC ACA GCC AAC CAC CTA GAC GCC TGC CGA GTG GCA CCC TAC GTC Leu Gly Thr Ala Asn His Leu Asp Ala Cys Arg Val Ala Pro Tyr Val 185 190 195 200	690
AAC ATG GGA GCC TTA CGG ATG CCT TTC CAA CAG GTC CAG GCT CAG CTG Asn Met Gly Ala Leu Arg Met Pro Phe Gln Gln Val Gln Ala Gln Leu 205 210 215	738
CAG CTG GAA GGC GTG GCC CAC GCG CAC CCG CAC CTG CAC CCG CAC CTG Gln Leu Glu Gly Val Ala His Ala His Pro His Leu His Pro His Leu 220 225 230	786
GCG GCG CAC GCG CCC TAC CTG ATG TTC CCC CCG CCG CCC TTC GGG CTG Ala Ala His Ala Pro Tyr Leu Met Phe Pro Pro Pro Pro Phe Gly Leu 235 240 245	834
CCC ATC GCG TCG CTG GCC GAG TCC GCC TCG GCC GCC GCC GTG GTC GCC Pro Ile Ala Ser Leu Ala Glu Ser Ala Ser Ala Ala Ala Val Val Ala 250 255 260	882
GCC GCC GCC AAA AGC AAC AGC AAG AAT TCC AGC ATC GCC GAC CTG CGG Ala Ala Ala Lys Ser Asn Ser Lys Asn Ser Ser Ile Ala Asp Leu Arg 265 270 275 280	930
CTC AAG GCG CGG AAG CAC GCG GAG GCC CTG GGG CTC TG ACCCGCCGCG Leu Lys Ala Arg Lys His Ala Glu Ala Leu Gly Leu 285 290	978
CAGCCCCCG CGCGCCCGGA CTCCCGGGCT CCGCGCACCC CGCCTGCACC GCGCGTCCTG	1038
CACTCAACCC CGCCTGGAGC TCCTTCCGCG GCCACCGTGC TCCGGGCACC CCGGGAGCTC	1098
CTGCAAGAGG CCTGAGGAGG GAGGCTCCCG GGACCGTCCA CGCACGACCC AGCCAGACCC	1158
TGCGGGAGAT GGTGCAGAAG GCGGAGCGGG TGAGCGGCCG TCGTCCAGC CCGGGCTCT	1218
CCAAGGCTGC CCGTGCCTCC TGGGACCCTG GAGAAGGGTA AACCCCGCC TGGCTGCGTC	1278
TTCCTCTGCT ATACCCTATG CATGCGGTTA ACTACACACG TTTGGAAGAT CCTTAGAGTC	1338
TATTGAAACT GCAAAGATCC CGGAGCTGGT CTCCGATGAA AATGCCATTT CTTCGTTGCC	1398
AACGATTTTC TTTACTACCA TGCTCCTTC TTCATCCCGA GAGGCTGCGG AACGGGTGTG	1458
GATTTGAATG TGGACTTCGG AATCCCAGGA GGCAGGGGCC GGGCTCTCCT CCACCGCTCC	1518
CCCGGAGCCT CCCAGGCAGC AATAAGGAAA TAGTCTCTG GCTGAGGCTG AGGACGTGAA	1578
CCGCGGGCTT TGGAAAGGGA GGGGAGGGAG ACCCGAACCT CCCACGTTGG GACTCCCACG	1638
TTCCGGGGAC CTGAATGAGG ACCGACTTAA TAACCTTTCC AGTGTTTGAT TCCCAAATG	1698
GGTCTGGTTT TGTTTTGGAT TGGTATTTTT TTTTTTTTTT TTTTTTGCTG TGTTACAGGA	1758
TTCAGACGCA AAAGACTTGC ATAAGAGACG GACGCGTGGT TGCAAGGTGT CATACTGATA	1818
TGCAGCATTAA ACTTTACTGA CATGGAGTGA AGTGCAATAT TATAAATATT ATAGATTAAA	1878
AAAAAATAG CA	1890

(2) Informace o sekvenci SEQ ID NO: 11:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 292 aminokyselin

(B) TYP: aminokyselinová

(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: protein

(xi) POPIS SEKVENCE: SEQ ID NO: 11:

```

Met Glu Glu Leu Thr Ala Phe Val Ser Lys Ser Phe Asp Gln Lys Ser
 1          5          10          15
Lys Asp Gly Asn Gly Gly Gly Gly Gly Gly Gly Lys Lys Asp Ser
 20          25          30
Ile Thr Tyr Arg Glu Val Leu Glu Ser Gly Leu Ala Arg Ser Arg Glu
 35          40          45
Leu Gly Thr Ser Asp Ser Ser Leu Gln Asp Ile Thr Glu Gly Gly Gly
 50          55          60
His Cys Pro Val His Leu Phe Lys Asp His Val Asp Asn Asp Lys Glu
 65          70          75          80
Lys Leu Lys Glu Phe Gly Thr Ala Arg Val Ala Glu Gly Ile Tyr Glu
 85          90          95
Cys Lys Glu Lys Arg Glu Asp Val Lys Ser Glu Asp Glu Asp Gly Gln
 100         105         110
Thr Lys Leu Lys Gln Arg Arg Ser Arg Thr Asn Phe Thr Leu Glu Gln
 115         120         125
Leu Asn Glu Leu Glu Arg Leu Phe Asp Glu Thr His Tyr Pro Asp Ala
 130         135         140
Phe Met Arg Glu Glu Leu Ser Gln Arg Leu Gly Leu Ser Glu Ala Arg
 145         150         155         160
Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Cys Arg Lys Gln Glu
 165         170         175
Asn Gln Met His Lys Gly Val Ile Leu Gly Thr Ala Asn His Leu Asp
 180         185         190
Ala Cys Arg Val Ala Pro Tyr Val Asn Met Gly Ala Leu Arg Met Pro
 195         200         205
Phe Gln Gln Val Gln Ala Gln Leu Gln Leu Glu Gly Val Ala His Ala
 210         215         220
His Pro His Leu His Pro His Leu Ala Ala His Ala Pro Tyr Leu Met
 225         230         235         240
Phe Pro Pro Pro Pro Phe Gly Leu Pro Ile Ala Ser Leu Ala Glu Ser
 245         250         255
Ala Ser Ala Ala Ala Val Val Ala Ala Ala Ala Lys Ser Asn Ser Lys
 260         265         270
Asn Ser Ser Ile Ala Asp Leu Arg Leu Lys Ala Arg Lys His Ala Glu
 275         280         285
Ala Leu Gly Leu
 290

```

(2) Informace o sekvenci SEQ ID NO: 12:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 1354 párů bazí
 (B) TYP: nukleová kyselina
 (C) DRUH ŘETĚZCE: jednořetězcová
 (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „SHOXb“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: cds
 (B) POZICE: 91..768

(xi) POPIS SEKVENCE: SEQ ID NO: 12:

GTGATCCACC CGCCGCACGG GCCGTCCTCT CCGCGCGGGG AGACGCGCGC ATCCACCAGC	60
CCCGGCTGCT CGCCAGCCCC GGCCCCAGCC ATG GAA GAG CTC ACG GCT TTT GTA	114
Met Glu Glu Leu Thr Ala Phe Val	295 300
TCC AAG TCT TTT GAC CAG AAA AGC AAG GAC GGT AAC GGC GGA GGC GGA	162
Ser Lys Ser Phe Asp Gln Lys Ser Lys Asp Gly Asn Gly Gly Gly Gly	305 310 315
GGC GGC GGA GGT AAG AAG GAT TCC ATT ACG TAC CGG GAA GTT TTG GAG	210
Gly Gly Gly Gly Lys Lys Asp Ser Ile Thr Tyr Arg Glu Val Leu Glu	320 325 330
AGC GGA CTG GCG CGC TCC CGG GAG CTG GGG ACG TCG GAT TCC AGC CTC	258
Ser Gly Leu Ala Arg Ser Arg Glu Leu Gly Thr Ser Asp Ser Ser Leu	335 340 345
CAG GAC ATC ACG GAG GGC GGC GGC CAC TGC CCG GTG CAT TTG TTC AAG	306
Gln Asp Ile Thr Glu Gly Gly Gly His Cys Pro Val His Leu Phe Lys	350 355 360
GAC CAC GTA GAC AAT GAC AAG GAG AAA CTG AAA GAA TTC GGC ACC GCG	354
Asp His Val Asp Asn Asp Lys Glu Lys Leu Lys Glu Phe Gly Thr Ala	365 370 375 380
AGA GTG GCA GAA GGG ATT TAT GAA TGC AAA GAG AAG CGC GAG GAC GTG	402
Arg Val Ala Glu Gly Ile Tyr Glu Cys Lys Glu Lys Arg Glu Asp Val	385 390 395
AAG TCG GAG GAC GAG GAC GGG CAG ACC AAG CTG AAA CAG AGG CGC AGC	450
Lys Ser Glu Asp Glu Asp Gly Gln Thr Lys Leu Lys Gln Arg Arg Ser	400 405 410
CGC ACC AAC TTC ACG CTG GAG CAG CTG AAC GAG CTC GAG CGA CTC TTC	498
Arg Thr Asn Phe Thr Leu Glu Gln Leu Asn Glu Leu Arg Leu Phe	415 420 425
GAC GAG ACC CAT TAC CCC GAC GCC TTC ATG CGC GAG GAG CTC AGC CAG	546
Asp Glu Thr His Tyr Pro Asp Ala Phe Met Arg Glu Glu Leu Ser Gln	430 435 440
CGC CTG GGG CTC TCC GAG GCG CGC GTG CAG GTT TGG TTC CAG AAC CGG	594
Arg Leu Gly Leu Ser Glu Ala Arg Val Gln Val Trp Phe Gln Asn Arg	445 450 455 460 ?



AGA GCC AAG TGC CGC AAA CAA GAG AAT CAG ATG CAT AAA GGC GTC ATC	642
Arg Ala Lys Cys Arg Lys Gln Glu Asn Gln Met His Lys Gly Val Ile	
465 470 475	
TTG GGC ACA GCC AAC CAC CTA GAC GCC TGC CGA GTG GCA CCC TAC GTC	690
Leu Gly Thr Ala Asn His Leu Asp Ala Cys Arg Val Ala Pro Tyr Val	
480 485 490	
AAC ATG GGA GCC TTA CGG ATG CCT TTC CAA CAG ATG GAG TTT TGC TCT	738
Asn Met Gly Ala Leu Arg Met Pro Phe Gln Gln Met Glu Phe Cys Ser	
495 500 505	
TGT CGC CCA GGC TGG AGT ATA ATG GCA TGA TCTCGACTCA CTGCAACCTC	788
Cys Arg Pro Gly Trp Ser Ile Met Ala *	
510 515	
CGCCTCCCGA GTTCAAGCGA TTCTCCTGCC TCAGCCTCCC GAGTAGCTGG GATTACAGGT	848
GCCCACCACC ATGTCAAGAT AATGTTTGTA TTTTCAGTAG AGATGGGGTT TGACCATGTT	908
GGCCAGGCTG GTCTCGAACT CCTGACCTCA GGTGATCCAC CCGCCTTAGC CTCCCAAAGT	968
GCTGGGATGA CAGGCGTGAG CCCCTGCGCC CGGCCTTTGT AACTTTATTT TTAATTTTTT	1028
TTTTTTTTTA AGAAAGACAG AGTCTTGCTC TGTCACCCAG GCTGGAGCAC ACTGGTGCGA	1088
TCATAGCTCA CTGCAGCCTC AAACTCCTGG GCTCAAGCAA TCCTCCCACC TCAGCCTCCT	1148
GAGTAGCTGG GACTACAGGC ACCCACCACC ACACCCAGCT AATTTTTTTG ATTTTTACTA	1208
GAGACGGGAT CTTGCTTTGC TGCTGAGGCT GGTCTTGAGC TCCTGAGCTC CAAAGATCCT	1268
CTCACCTCCA CCTCCCAAAG TGTTAGAATT ACAAGCATGA ACCACTGCCC GTGGTCTCCA	1328
AAAAAAGGAC TGTTACGTGG AAAAAA	1354

(2) Informace o sekvenci SEQ ID NO: 13:

(i) CHARAKTERISTIKA SEKVENCE:

(A) DÉLKA: 226 aminokyselin

(B) TYP: aminokyselinová

(D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: protein

(xi) POPIS SEKVENCE: SEQ ID NO: 13:

15.04.99

Met Glu Glu Leu Thr Ala Phe Val Ser Lys Ser Phe Asp Gln Lys Ser
 1 5 10 15
 Lys Asp Gly Asn Gly Gly Gly Gly Gly Gly Gly Lys Lys Asp Ser
 20 25 30
 Ile Thr Tyr Arg Glu Val Leu Glu Ser Gly Leu Ala Arg Ser Arg Glu
 35 40 45
 Leu Gly Thr Ser Asp Ser Ser Leu Gln Asp Ile Thr Glu Gly Gly Gly
 50 55 60
 His Cys Pro Val His Leu Phe Lys Asp His Val Asp Asn Asp Lys Glu
 65 70 75 80
 Lys Leu Lys Glu Phe Gly Thr Ala Arg Val Ala Glu Gly Ile Tyr Glu
 85 90 95
 Cys Lys Glu Lys Arg Glu Asp Val Lys Ser Glu Asp Glu Asp Gly Gln
 100 105 110
 Thr Lys Leu Lys Gln Arg Arg Ser Arg Thr Asn Phe Thr Leu Glu Gln
 115 120 125
 Leu Asn Glu Leu Glu Arg Leu Phe Asp Glu Thr His Tyr Pro Asp Ala
 130 135 140
 Phe Met Arg Glu Glu Leu Ser Gln Arg Leu Gly Leu Ser Glu Ala Arg
 145 150 155 160
 Val Gln Val Trp Phe Gln Asn Arg Arg Ala Lys Cys Arg Lys Gln Glu
 165 170 175
 Asn Gln Met His Lys Gly Val Ile Leu Gly Thr Ala Asn His Leu Asp
 180 185 190
 Ala Cys Arg Val Ala Pro Tyr Val Asn Met Gly Ala Leu Arg Met Pro
 195 200 205
 Phe Gln Gln Met Glu Phe Cys Ser Cys Arg Pro Gly Trp Ser Ile Met
 210 215 220
 Ala
 225

(2) Informace o sekvenci SEQ ID NO: 14:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 32367 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „kosmid: LLNOYCO3`M`34F5“

(xi) POPIS SEKVENCE: SEQ ID NO: 14:

```

TTTCTCTGTC TCCATCCCTC TGTCTCTCCC TTTCTCTCTG TCTTTCCTTG TCTCTCTCTT      60
TCTCTCTCTC TCTCCATCTC TCTCTCTCCC TGTCTCTCTC TCTCCATCTC CCCGTCTCTC      120
CGTTTCTCTC TCTGCCTCTC CCTGTCTGTC TCTCTCTTTC TGTGTCTTAC ACACACCCCA      180
ACCCACCGTC ACTCATGTCC CCCCACTGCT GTGCCATCTC ACACAAGTTC ACAGCTCAGC      240
TGTCATCCTG GGTCCCCAGG CCCCGCCGGG GAGGAAGATG CGCCGTGGGG TTACGGGAGG      300
AAGGGGACTC CGGGCCTCCT GGTGCCCCAC TTTATTTGCA GAAGGTCTTT GGCAGGAACC      360
GTGACGCGTT TGGTTTCCAG GACTTGAAA ACGAATTTC A GGTGCGGATG GCGAGCACCG      420
GCTTCCCTG AAGCACATTC AATAGCGAGA GCGGGGAGGG AGCGAGCAGG AGCATCCCAC      480
CATGAAAACC AAAAAACAAA GTATTTTTTT CACCCGGTAA ATACCCCA GAGGAGGGTG      540
ACAGCGCGGC GCTAAGGGAG GAGGCCTCGC GCCGGGGTCC GCCGGGATCT GCGCGGGGGC      600
GAAAGAATAT AGATCTTTAC GAACCGGATC TCCCGGGGAC CTGGGCTTCT TTCTGCGGGC      660
GCTGGAGACC CGGGAGGCGG CCCCGGGGAT CCTCGGCCTC CGCCGCGGCC GCCTCCCAAG      720
CGCCCGCGTC CCGGTTTGGG GACACCCGGC CCCTTCTTCT CACTTTTCGGG GATTCTCAG      780
CCGCGTTCCA TCTACCAAC TCTCCATCCA AGGGCGCGCC GCCACCAACT TGGAGCTCAT      840
CTTCTCCCAA GATCGTGCCT CCCCGGGGCG CCCGGGTCCC CCCCTCGCC ATCTCAACCC      900
CGGCGCGACC CGGGCGCTTC CTGGAAAGAT CCAGGCGCCG GGCTCTGCGC TCCTCCCGGG      960
AGCGAGGGCG GCCGGACGAC TGGGACCCTC CTCTCTCCAG CCGTGAATC CTTGTCTCTC     1020
TGTCTCTCTC TGCAGAAAA CTGGAGTTTG CTTTCTCTCC GGCCACGGAG AGAACGCGGG     1080
TAACCTGTGT GGGGGGCTCG GCGCCTGCG CCCCCCTCCT GCGCGCGCGC TCTCCCTTCC     1140
AAAAATGGGA TCTTTCCCCC TTCGCACCAA GGTGTACGGA CGCCAAACAG TGATGAAATG     1200

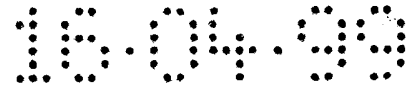
```



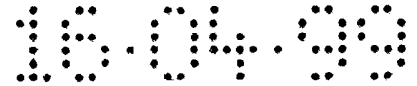
AGAAGAAAGC CAATTGCCGG CCTGGGGGGT GGGGGAGACA CAGCGTCTCT GCGTGCGTCC	1260
GCCGCGGAGC CCGGAGACCA GTAATTGCAC CAGACAGGCA GCGCATGGGG GGCTGGGCGA	1320
GGTCGCCGCG TATAAATAGT GAGATTTCCA ATGGAAAGGC GTAAATAACA GCGCTGGTGA	1380
TCCACCCGCG CGCACGGGCC GTCCTCTCCG CGCGGGGAGA CGCGCGCATC CACCAGCCCC	1440
GGCTGCTCGC CAGCCCCGGC CCCAGCCATG GAAGAGCTCA CGGCTTTTGT ATCCAAGTCT	1500
TTTGACCAGA AAAGCAAGGA CGGTAACGGC GGAGGCGGAG GCGGCGGAGG TAAGAAGGAT	1560
TCCATTACGT ACCGGGAAGT TTTGGAGAGC GGACTGGCGC GCTCCCAGGA GCTGGGGACG	1620
TCGGATTCCA GCCTCCAGGA CATCACGGAG GGCGGCGGCC ACTGCCCGGT GCATTTGTTC	1680
AAGGACCACG TAGACAATGA CAAGGAGAAA CTGAAAGAAT TCGGCACCGC GAGAGTGCCA	1740
GAAGGTAAGT TCCTTTGCGC GCCGGCTCCA GGGGGGCCCT CCTGGGGTTC GCGCCCTCCT	1800
CGCCACGGAG TCGGCCCCGC GCGCCCTCG CTGTGCACAT TTGCAGCTCC CGTCTCGCCA	1860
GGGTAAGGCC CGGGCCGTCA GGCTTTGCCT AAGAAAGGAA GGAAGGCAGG AGTGGACCCG	1920
ACCGGAGACG CGGGTGGTGG GTAGCGGGGT GCGGGGGGAC CCAGGGAGGG TCGCAGCGGG	1980
GGCCGCGCGC GTGGGCACCG ACACGGGAAG GTCCCGGCT GGGGTGGATC CGGGTGGCTG	2040
TGCCTGAAGC CGTAGGGCCT GAGATGTCTT TTTCATTTTC TTTTCTTTC CTTTCTTTT	2100
TTTGTTTGTT TGTTTGTTG TTTGAGACAG AGTCTCGCTC TGTCCCCAG GCTGGAGTGC	2160
AGTGGTGGCA TCTCGGCTCA CTGCAACCTC CGCCTCCTGG GTTCAAGCGA TTCTCCTGCC	2220
TCAGCCTCCC CAGTAGCTGG GATTACAGGC ATGCACCACC ACGCCTGGCT AATTTTGTG	2280
CTTTTAGTAA AGACGGGGAT TCACCATGTT GGCCAGGCTG GTCTGAACT CCTGACCTCA	2340
GGTGATCCAC CCGCCTCGGC CTCCCAAAGT GCTGGGATGA CAGGCGTGAG GCACCGCGCC	2400
CGGCCTGGGT CCTGACGGCT TAGGATGTGT GTTCTGTCT CTGCCTGTCT GCCTGTIATT	2460
TACGGTCACC CAGACGCACA GAGGAGCCGT CTCCACGCGC CTCCCAGCG CTCAGCGCCT	2520
GCCGGGCCCC CGGAGATCAC GGAAGACTC GAGGTGCGT GGTAGGAGAC GGAAGGCC	2580
CGGGTCAGCT CGGTTCTGTT TCCTTAAAGG AACCTTCAT TATTATTCA TTGTTTCTCT	2640
TTGAACGTCG AGGCTTGATC TTGGCGAAAG CTGTTGGGTC CATAAAAACC ACTCCCGTGA	2700
GCGGAGGTGG CCGGATCTG GATGGGGCGC GAGGGGCCCC GGGGAAGCTG GCGGCTCGC	2760
GGGCGCGTCC TAAGTCAAGG TTGTCAGAGC GCAGCCGGTT GTGCGCGGCC CGGGGGAGCT	2820
CCCCTCTGGC CCTTCTCCT GAGACCTCAG TGGTGGGTCG TCCCGTGGTG GAAATCGGGG	2880
AGTAAGAGGC TCAGAGAGAG GGGCTGGCCC CGGGATCTC TGTGCACACA CGACAACCTGG	2940
GCGGCATACA TCTTAAGAAT AAAATGGGCT GGCTGTGTCG GGGCACAGCT GGAGACGGCT	3000
ATGGACGCCT GTTATGTTT CATTACAAAG ACGCAGAGAA TCTAGCCTCG GCTTTTGCTG	3060
ATTGCGCAGAG TTGAGGTGCG AGGGTGAATG CCCCAAAGGT AATCTTCTCT AAGACTCTGG	3120
GGCTACCTGC TCTCCGGGGC CCTGCATTTG GGGTGTGGAG TGGCCCCGGG AAATAGCCCT	3180
TGTATTCGTA GGAGGCACCA GGCAGCTTCC CAAGGCCCTG ACTTTGTGCA AGCAGAAAGC	3240



TGTGGCTACG GTTTACAAAG CAGTCCCCGG TTTCTGACCG TCTAAGAGGC AGGAGCCCAG 3300
 CCTGCCTTTG ACAGTGAGAG GAGTTCCTCC CTACACACTG CTGCGGGCAC CCGGCACTGT 3360
 AATTCATACA CAGAGAGTTG GCCTTCCTGG ACGCAAGGCT GGGAGCCGCT TGAGGGCCTG 3420
 CGTGTAATTT AAGAGGGTTC GCAGCGCCCG GCGGCCGCTT CTGTGGGGTT GCTTTTTGGT 3480
 TGTCCTTCGC AGACACCGTT TTGCTCCTCT GAACCTCTCTC TTCTCCCCCT GGCCGTGGAC 3540
 CCGGGAGAGC AAAGTGTCCT CCAGACCTTT TGAAAGTGAG AGGAAAATAA AGACCAGGCC 3600
 AAAGACCCAG GGCCACAGGA GAGGAGACAG AGAGTCCCCG TTACATTTTC CCCTTGGCTG 3660
 GGTGCAGAAA GACCCCGGG CCAGGACTGC CACCCAGGCT ACTATTTATT CATCAGATCC 3720
 AAGTTAAATC GAGGTTGGAG GGCAGGGGAG AGTCTGAGGT TACCGTGGAA GCCTGGAGTT 3780
 TTTGGGAACA GCGTGTCCCC GCCGAGCCTG GGAGCCCGTG GGTCTGCAA AGCCTGCGGG 3840
 TGTTTGAGGA CTTGAAGAC CAGTTTGTC A GTTGGGCTCA ATTCTGGGG TTCAGACTTA 3900
 GAGAAATGAA GGAGGGAGAG CTGGGGTCGT CTCCAGGAAA CGATTCACCT GGGGGGAAGG 3960
 AATGGAGTGT TCTTGACGGC ACATGTCTGT TAGGAGGTGA AACAGAATGT GAAATCCACG 4020
 TTGGAGTAAG CGTCCAGCGC TGAATGTAGC TCGGGGTGGG GTGGGAGGGC CCTGGTGTGG 4080
 ATCGTGGAAG GAAGAAAGAC AGAACAGGGT GCTAGTATTT ACCCCGTTCC CTGTAGACAC 4140
 CCTGGATTTG TCAGCTTTGC AAGCTTCTTG GTTGACGCGG CCTTGCCTGT GCCCCTTTGA 4200
 GACTGTTTCC AGACTAAACT TCCAAATGTC AGCCCCTTAC CCTTGACAGC AAGGGACATC 4260
 TCATTAGGGC ATCGCGTGCT TCTCATCTGT GCTCAGCAGG CCCGAGATAG GAACAGAGGG 4320
 GCGTTGGAGA TGCCACTTCC ACCAGCCCTG GGTGGAAGGG GAGCGAGGGA GACACCTTTT 4380
 ACTTAAACCC CTGAGCTTGG TCAGAGAGGC TGAATGTCTA AAATGAGGAA GAAAAGGTTT 4440
 TTCACCTGGA AACGCTTGAG GGCTGAGTCT TCTGCCCTTC TGACTCCCC AGCAAATACA 4500
 GACAGGTCAC CAACTACTGG AGATGAGAAA GTGCCATTTT TGGCACACTC TGGTGGGGTA 4560
 GGTGCCCCGAC CGCGTGTGAA AAAGTGGGAA GGAGAGATTT CTGCGCACGC GGTTTCAGCCC 4620
 CCAGGCGCGG TGGCGCATT C AGGTA CTAG ACGCGGTTCT GCTGTTCTGC TGAGAAACAG 4680
 GCTTCGGGTA GGGGCTCCTA GCTCCGCCAG ATCGCGGAGG GACCCCCAGC CCTCCTGCGC 4740
 TGCAGCGGTG GGGATAGCGT CTCTCCGTAG GCCTAGAATC TGCAACCCGC CCCGGGTCCT 4800
 CCCCCTGTCC TTCCC GGCG TCCCGCCGGG GATCCCACAG TTGGCAGCTC TTCTCAAAT 4860
 TCTTTCCCTT AAAAATAGGA TTTGACACCC CACTCTCCTT AAAAAAAAAA AATAAGAAAA 4920
 AAAGGTTAGG TTATGTCAAC AGAGGTGAAG TGGATAATTG AGGAAACGAT TCTGAGATGA 4980
 GGCCAAGAAA ACAACGCTCG TGCAAAGCCC AGGTTTTTGG GAAAGCAGCG AGTATCTCC 5040
 TCGGCTTTTG CGTTATGGAC CCCACGCAGT TTTTGCCTCA AAGCGCAT TG GTTTTCGAGG 5100
 GCCCCCTTTC CACCGCGGGA TGCACGAAGG GGTTCCGCAC GTTGCGCAA ACCTCCCCGG 5160
 CCTCAGCCCT GTGCCCTCCG CTCCCCACGC AGGGATTTAT GAATGCAAAG AGAAGCGCGA 5220



GGACGTGAAG TCGGAGGACG AGGACGGGCA GACCAAGCTG AAACAGAGGC GCAGCCGCAC	5280
CAACTTCACG CTGGAGCAGC TGAACGAGCT CGAGCGACTC TTCGACGAGA CCCATTACCC	5340
CGACGCCTTC ATGCGCGAGG AGCTCAGCCA GCGCCTGGGG CTCTCCGAGG CGCGCGTGCA	5400
GGTAGGAACC CGGGGGCGGG GCGGGGGGGC CCGGAGCCAT CGCCTGGTCC TCGGGAGCGC	5460
ACAGCACGCG TACAGCCACC TGCGCCCGGG CCGCCGCCGT CCCCTTCCCG GAGCGCGGGG	5520
AGGTTGGGTG AGGGACGGGC TGGGGTTCTT GGACTTTTGG AGACGCCTGA GGCCTGTAGG	5580
ATGGGTTTCAT TGGTTTGTG TTTCACCAAC AGCAAACAAA TATATATACA TATATATTAT	5640
ACAAATAACA AATAAATATA TATGTTATAC AGATGGGTAT ATTGTATATA TTATAGATAT	5700
TTGTTCTGTC TTGGTGCAA GACACCCGGT GAACCCATAT ATTGGCTCCT GACTGCCTTC	5760
GGTCCCCTG GGATTGGTTA TAGGGGCAAC ACATGCAAAC AAAACTTTCC CTGGATTATA	5820
CTTAGGAGAC GAAGTACAG ATGCGTTTGA TCCAGAGTGT TTTACAAGAT TTTTCATTTA	5880
AAAAAAAAATG TGTCTTTTGG CCCCTGATTC CCCTCCGTCT TCCCCTGTGG CTGCATTGAA	5940
AAGGTTTCCT TAGGATGAAA GGAGAGGGGT GTCCTCTGTC CCTAGGTGGA GAGAAACAGG	6000
GTCTTCTCTT TCCTCCGTTT TTTCACCTAC CGTTTCTATC TCCCTCCTCC CCTCTCCAGC	6060
CCTGTCTCTT GCTACAAACC ACCCCCTCCT CCCTCCGGCT GTGGGGAGCG CAGGAGCAGC	6120
TTGGGCATCT GGATGAGCGG AGACTATTAG CGGGGCACGG GGGCTCCCCG AGGAGCGCGC	6180
GAATTCACGC TGCCCCATGA GACCAGGCAC CGGGGGGCGG AGGGGCCTTG GGTGTCCGCA	6240
GAGGGACGGG CGGGCAGAGC CTTCTCCGC ATTCTAAACA TTCACTTAAA GGTATGAGTT	6300
TATTTACAGG GTGCTGCTGG GAGAGCCTCC AAATGGCTTC TTCCAGCCCC TGCCTGACAG	6360
TTCAGCTCCC CTGGAAGGTC AACTCCTCTA GTCCTTCTC CTGGTTCTGG GCAGGACAGA	6420
AGTGGGGGGA GGGAGAGAGA GAGAGAGAGA GAGAGAGACG GTCAGGATCC CCGGACCCTG	6480
GGGAACCCGT CAAAAATAAA TGAATTAAAG ATTGCCGACC AGAGAGAGAA CCGTGACAAA	6540
GCAAACGGCG TTCAAAGCAA AGAGACGAAC TGAAAGCCCG TTCCCGTAGG ACTGGTTATG	6600
AGGTCAACAC ATTCAAACAC AGCTTGCTCT GGATTTTGCT GAGCAGAGGA AGATACAGAT	6660
GCATTTGATC CAAAGTGTGT TACATCTTTC ATTATATGTG TGTCTATATA TATAACATA	6720
TATAAATATA TAAACATACA TAAATGTATG TAAATATATA TAATCTATAT ACATATATAA	6780
ATATATAAAC ACATATATAA TATATAAATC TATAAACATA TATAATATAT AAACATAAAT	6840
ATATAAACAT ATATAATATA TAAATATATT AACATATATA AAATATGTAT AAATATATAT	6900
AAACATATAA ACATATATAA ATATATAAAC ATATAAATAT ATAAACATAT ATAAATATAT	6960
ACAAACATAT TGTATATATA TAAATATATA TAAAAACATA TATATACATA TAAAAATATA	7020
TATAAACATA TATACATATA AAGAAATATA TATAAACATA TATACATATA AATATACATA	7080
TATAAACATA TATATACATA AAATATATAT AAACATATAT ACATATAAAA ATATATATAT	7140
ATTAACATAT ATATACATAT AAAAATATAT ATATTAACAT ATATATACAT ATAAAAATAT	7200
ATATATATTT TTGGCCCTG ATTCCCTTCG GTTCTGTGG GATGGGTGAT TGAGTCAACA	7260



CATTCAAACA	CAACTTTTCC	ATCGATGTTG	CTTAGGAGAT	GAGGATACAG	ATGCGTTTGA	7320
TGGAGAGGGT	TTTACAAGCT	CTTTCATTTA	AATATATATA	TATATATATA	TATATTTTTT	7380
GGCTCCTGAT	TCTCTCCGT	CTTCCCATGT	GGCTGCATTT	TAAAAGGCTT	CCCTAAGATC	7440
GTTACGATTA	AATCAACCCT	CCCCAGGCAT	CTTTACCGAG	GGCTGTGGTC	CCCAAAGCGA	7500
TACAGCCCAG	GAGGGAGAGA	GGCTTTGGTG	ACTTGGAGGA	AGGACTGTGT	CCCTCCTTAG	7560
GGCGTCTGTG	GCCTCAGTGA	GGGAAGGAAG	CTGCATCAGA	CAGGGGTTC	CTCGCTGTCC	7620
ACCCCTCTGG	CAGAAGATGG	ATTGGGCTGC	CCCGTATAAA	TTAATGAAAA	GATTAAAGTT	7680
TCGCTAAAGG	GGACATCGAG	TTTATGTGTC	ATCTCCTGGT	GTCTGTGTGC	CTGGGATCTG	7740
CAATATATCC	CAGCCCTTGA	TGTACTGTTT	CTATAAAAAAT	AAATTACTTG	TAATTTAATT	7800
CCACACTATT	TCTTTCCGTA	GTCTATTACC	GACGAGAGCA	CGTTAGTTCA	GCTGCGGAAA	7860
ATTGGTTGTG	GGGTGTGTGC	GGACCCCGAG	AACGCCCTAA	AATAAAGACA	AATCGGGGAC	7920
AAGCTGGGGG	TTATCGATTG	CAGGGGTCGC	ATGAAAATTT	AACGACGGTA	AATAATAATA	7980
AAAACAAACA	TGGGAATGCA	ATAAAAGACA	TAATTCCTCA	TCGCCGCGGG	GGGAAAGGAT	8040
CCTATAGTAA	AGGCGAGTGC	GCTTTGAGGG	GTCATAAAAA	TCAATTAGTT	CCAACACCCA	8100
CGTCCCGCGT	TGAGGGGACG	GGGACGAGCA	GGGACAGAAA	AAGAAACCAT	ATTTGAATCC	8160
CATCTCTCTG	TGAATTCTTG	GGTCACATGC	GTCTCAGTAC	AGCCCGTCCC	GTGCTGTGAC	8220
CGGATAGAGT	TTCAATTTAC	TGTGGAAATT	TGCTGTAAAT	AAATTGAGCA	TCCGATAGAA	8280
GCTGTTGCTG	ATTAACCTTT	TATTTTTAGC	GTGGCCCTGC	AAAGTCGTAT	CACCCAGCTG	8340
TCAGGCTTCT	AATCGAAAGT	TATGAGACCA	CGGTGAGGGG	CAGGCGGTAA	TTTAATTACA	8400
ACAAATATCT	TTGGGTTTAT	GGCGCAGAGC	TAAATTAAT	GTCATTATTC	ACTGTCTGTA	8460
ATGGAAATCA	AAAGGAAATC	GCATTACGGC	ATTTGGGAAA	GAAAGCGGGG	AGTGCTCTTT	8520
AATGAAGAAA	TAACTGTCTT	AAGCAGTGTG	ACACACTTCA	CTTACCATAT	TCGGGCCTAA	8580
TTGGAATGGA	TCGTGAATCA	CTCCAAGACT	GATTTATTAG	CGCTTCACGC	AGCGGCTAAT	8640
TCATCACTTG	TATTCTTCAT	CATTTTTTTT	TTTCTCTCG	CCGTGTTGAA	GGGAGAGTGA	8700
ATGAGGCTTT	CCACGTTTCA	GGAGGATTTT	CTTTTTTGAA	AAATGCCCTT	CCAGAGGCTT	8760
TTGGGTGGCT	GGCTTGCTTT	CTGGGCCCTG	GAGGAGACAG	GCGGAGAGTC	CAGGTGGGCA	8820
TGGAGAGGCA	CAGTGGCAGG	TCACCTGGAT	GGTCAGTGGG	GGTGGAGGTC	TGAAGGCGCC	8880
AGCTTTGGAA	ATTATTGGTG	AATTTGATG	TCAGCACCAG	GCAGGGGCCT	TTTGGCGGG	8940
GGTGTGAGGG	AGGATGACTT	TGCTGGGAAA	CAGGATCAGG	TTCTCCAGGC	GCACTGCAGC	9000
CCGGTAGGAC	CCACTTTGGA	AATGAAAAGC	CAGTCCGAA	AGCTGGGCTG	GAAGCTCCG	9060
TGTTGGGTTT	AAGAGCAAGT	TCACGTTGCG	CTGTGTAGAC	TCCTGGCTGC	TCCCAAATC	9120
TGAGGGTTTT	CTGAGGTTCC	CTTCATAGGG	GCACCGGCC	TGGGCATGC	ACAGTGCCTA	9180
AGGGTGGCTG	TGGGCCGAGG	GACCCAGCAC	GTGTTTTGCC	CACAACAGCC	GGAGTGA CTG	9240

GTTCACTCAC CGCCTTGGCG GAGGACGCCT GTTCTCTGGA CGAATCATTT CTCTTGGGTG 9300
 GTGACTGCCT TGTGGGTCAA GGTGCAGGTT TTCTGCCACA GAAAACCTGT TAGGAGGAAT 9360
 TAAGCGACTA AGACTGTCTAG GGAGGTGGTG GTGGGGGAGA GGAGGGGGTG GTGTCCAGAT 9420
 TACCAGGCAT AGGCTAAACT GCCTGCACTC TCCAGCTGGT CTGTCTGTGG AGGAGGGGAT 9480
 TGTCAACTACT GGGAGAGCAG AGGAGGCTCG TAGGAGGTGA GAGGGGGTGG AATTTGCATG 9540
 CAAATCTTCA CATGAGGCCT GTGTGAATTT CTCCAGCCTC CTGAGGGTCC CCTGCGCTAT 9600
 TGCACTCAAC TTCTTGATAG TTTACCCCAA GACTCAGAAG TCCTTAGAGG GGCAGAATGC 9660
 CCCCACCACA AAGCCTGCTA TCCTTGGGCG TCCTCAGGAC CCTTGGTCAT GAATGGGACC 9720
 CTTTCATGTA TGGGGACCCT TGGTAATATG AATGGGACGC CTTAGCTCC CCAGGGCTTC 9780
 CGAGGAGGCC GAGAAGGGCA AAGACACTTC CGAGGAGGCC GAGAAGGGCA AAGACATTTT 9840
 CTGGGCTTGG TGTGTCAAGA GCTAGATTGG AGAAGGGGCT GGATTTGGAA CTCTTTAGCC 9900
 ATCAGCTCAC CCTCTCCGTT TGTGGCTAAA GTCTGAAGGT GGAAACTTCG GTTCTCCTAC 9960
 AGGGTCTACA GGAGTTGGGG GCGGGGGCGC CCACACAGAA CGCTGGAAAG TTCGACAGTC 10020
 CACTTCCACT GGCTCGGAAC TCACTTTTTT ACCTTAAGTT CATCAGCGGT AACGCATAGG 10080
 TCTCACTTAG GCAGGGCACG GATGATTTAA CAATTTCTAC TTCTAGGTCA GGTGCGGTGG 10140
 CTCACACCTC TAATCCCAGC ACTTTGGGAG GCCCAGGAGG GTGGATCGCT TGAGGTCAGG 10200
 AGTTTGAGAC CAGCCTGGCC AACATGGTGA AACCCCGTCT CTAATAAAAT ACGAAAATTA 10260
 GCCAGGCATG GTGGTGAGCA CCTGTAATTC CAGCTACTCG GGAGGCTGAG GCAGGAGAAT 10320
 CGCTTGAACC TGGGAGGTGG ACGTTGCAGT GAGGTGAGAT CACACCACTG CACTCCAGCC 10380
 TGGATGAGAG AGCAAGACTC TGTCTCAAAA AAAAAATAAA AAAAAACAA AACAAAAATC 10440
 AAAAAAGAAA ACCCAATTC CAGTTCTAGG CCAGGTGCAG TGGCTCACGC CTGTCTATCC 10500
 AGCACTTTGG GAGGCCCAGG AGGGTGGATC GCTTGAGGTC AGGAGTTCGA GACCAGCCTG 10560
 GCCAACATGG TGAAACCCCA TCTCTACTAA AAATACAAAC GTTAGCTGGG TGTGGTGGTG 10620
 TGCGCCTGTA ATCCCAGCTA CTCGGGAAGC TGAGGCTGGA GAATTGCTTG AATCTGGGAG 10680
 GTGGAGGTTG CAGGGAGGCG AGATAGTGCC ACTGCAGTCC AGCCTGGACC AGAGAGCAAG 10740
 ACTCCGTCTC AAAAAACAAA GAAAGCAAAA AAAAAAACA AGAGACCAGC CTGGCCAACA 10800
 TGGTGAACCC GCGTCTCTAC TAAATACAA AATTAGCCGG GCATGGTGGT GGGCACCTGT 10860
 AGTCCCAGCT ACTCGGGAGG CTGAGGCAGG AGAATGGCTT GAACCTGGGA GGTGGAGCTT 10920
 GCAGTGAGCC GAGATAGTGC CACTGCACTC CAGCCTGGGC GACAGAGCGA GACTTGATTT 10980
 CAGAACCACC ACCACCACAA CAAAACAAA CAAAAATCC AAAAAACCC CAATTTCCAG 11040
 TACTAGGTAG TCAGTGATGC AGGGCTGGAG ACAGAGGGGC GGTAAGTGTC TGGGCGCCCA 11100
 CCATCAGTCA CCTCCCAGCT CCCAGAGGTG CAAAGTGCTT GGTTACGCCT CATGGGAAGG 11160
 ATGCTCCCTG GGGAGGCTGG GCTGGGTTCA CAGGGCTCTT CACATCTCTC TCTGCTTCTC 11220
 CCCAAGGTTT GGTTCAGAA CCGGAGAGCC AAGTGCCGCA AACAAAGAGAA TCAGATGCAT 11280

AAAGGTGGGT GTCGGGACTG GGGGGACCTG AAGCTGGGGG ATCCTGCTCC AGGAGGGATG 11340
GGGTCGACGA GGTGCTGGCT ACACCCAGGA CCACCACACT GACACCTGCT CCCTTTGGAC 11400
ACAGGCGTCA TCTTGGGCAC AGCCAACCAC CTAGACGCCCT GCCGAGTGGC ACCCTACGTC 11460
AACATGGGAG CCTTACGGAT GCCTTTCCAA CAGGTAGCTC ACTTTTTCTT CCTCTGAAGA 11520
TCCCTAGGGA CCTGCTGCTC CCTTCCCCTT TCCCCTATTT GCTGCCGCAT CCTGACACTC 11580
CTAGTCCCTC CCTGCCCCTG CAGACTTCTC AGCTGGCCCT TAGAAAAAAA GCCTCTTTTC 11640
CGAGGAGGCA TTTACAGGCA CCTTGGCACC TATGAAATCA GGCTGGGCCA GCGGGGGTGG 11700
CTCACACCTG TCATCCCAGC ACTTTGGGAG GCTGAGGAGG GTGCATCACC TGAGATCAGG 11760
AGTTCAAGAC CAGCCTGGCC AACTTAACGA AACCCCGTCT ATTAAAAATA CAAAATGGGT 11820
GTGGTGGCTC ACGCCTGTCA TCCCAGCACT TTGGGAGGCC GAGGCAGGTG GATCACCTGA 11880
GGTCAGGAAT TCGAGACCAG CCTGACCAAC ATGCTGAAAC CCCGTCTCTA CTGAAAACAC 11940
AAAGCTTAGC CGGGCGTGGT GGTGCACACC TGTGATCCCA GGTACTTGGG AGGGAGAATC 12000
ACTTGAACCT GGGAGGTGGA GGTGCCGTG AGCCAATATC GCGCCACTGC ACTCCACTCT 12060
GGGTGACAGA GTGAGACTCC AAGACTCCAT CTCAAAAAA AAAAAAAAAA TCAGGCTGTA 12120
AAAATCCACT TTTGGGAAGG TGAACACACA CAAGCCCAA CAGAAATCTG AAAAAACCA 12180
GAGGGGTGAA AAGTCCACAC AGTCAGGCAC CCCCACCTGG CTTGCTGCCT GGTTAAGAAG 12240
GGCGCAGATG CCTGTGCCTG GATACCAGAG ATGGGACAGA CACCCTTCC CTTTTCATCA 12300
CCACCCCGA GTGCCCGAGG GCCTGGGGCG TCTGCCTGGC CCCTGGCCCC TGGCTTGGGC 12360
TCTGCACCTC TGAACTGGAG ACACCCTACT CAGCTCCCA CTTACTTTGG AGTGAGCAGC 12420
GCTTGGGTGC CCAGCGTGA TTTGGGGCTT CCAGGGAGTC GGGGTTCGGT CGCGGAGCCC 12480
AAGCTTCCCA AGGGCGCCCC CGCCCTGCCC TGGCTTAGTG GTGGGGATGG GATGGGGGGA 12540
AACGGGGAGC TCGTGGAAG GAGGTGAAGG GTCACAGGAG GAGAGAGCGC AGCGCCCACG 12600
TGCGCCCTGC CTGAACGCGC AGCGCAGCGC CCGGCTGCGG TGCCCTTGC CCCTTCGGTC 12660
CCTAATTTGG GGATCGGGAG TGCATGCGCG GCGGSAACGG GCTTGGGGGG GGGGCTCTGG 12720
CAGGGCGGAC GCGTGGCCTC CCTTCTTAC CGTTTTATT CAAGGGGACA GGCTGGGGAT 12780
TGTATTTGGG CGCGTGTGTTG GCTGAGGGTG CAGGACTTG GGGGGTGGCG GTGGGGAGGG 12840
CGGAAGGTAT AAACGTATAA ATCATAAGTA AACAACTCAG AAATGGACCC CGAGCGCTGG 12900
TCGCCGCTAG CTCTCCAGCT CTCCCTGGCC CAGGCCCGAA GGAGAGGGGT CCGCATCCCT 12960
CCGCGGTTCT CCTCTCCTGG GTACCTGGCC TTGAGGTGGG GGAACGAGCC TACTTCTTGT 13020
ACCGTCTTTT GCCGACGGCG GGACCCAGTG AAATTAGGCC GTTGGAGCCC GCAGGCCTGC 13080
CTGGCTTTGC GCACCGGAGT CTGGGGACC TGGTGTCCCC GGGAAAACT TGGGGACCTG 13140
GTATCCCCGG GAGAGGCTTG GGGACCTGGT GTCCCGGGAG AGGCTTGGGT ACCTGGTTTC 13200
TCTGGAAGAG GCTTGGACAC CTGGTGCCT GGGAGGGCCT TTGGGACCTG GTGTCCTGGG 13260



AGAGGCTTGG AGATCTGTTG TCCTGGGAGA GGCTTGGGA CCTGGTGTCC CTGGAGAGGC 13320
 TTGGGGACCT GGTGACCTTG GAGAGGCTTG GAGACCTGGT GTTCTGGGAG AGGCTTGGGG 13380
 ACCTGGTGT CTGGGAGAGG CTTGGGGACC TGGTGTCTCT GGAAGAGGCT TGGACACCTG 13440
 GTGACCCGGG AGGGCCTTGG GGATCTGGTG TCCCCGGAGA GCCTTGGGGA CCTGGTGTCC 13500
 TGGGAGAGGC TTGGGGACCT GGTGACCTTG GAGAGGCTTG GGGACCTGGT GTCCTGAGAG 13560
 AGCCTTGGGG ATCTGGTGTG CCAGGAGAGG CTTGGGGACC TGGTGTCTCT GGAAGAGGCT 13620
 TGGACACCTG GTGTCTGGG GAGAGGCTTG GGGACCTGGT GTCCTGGGAG AGGCTTGGGG 13680
 ACCTGGTGTG CTGGGAGAGG CTTGGAGATC TGGTGTAGCCG GGAGAGGCTT GGGGACCTGG 13740
 TGTCCCGGA GAGGCTTGGG GACTTGGTGT CCCGGGAGAG GCTTGAACAC CTGGTGTCCC 13800
 AGGAGAGGCT TGGGGACCTG GTGACCTTG AGAGGCCTGG GGACCTGGTG ACCCGGAGAGA 13860
 GCCTTGGGGA CCTGGTGTCC TGGGGAGAGC CTTGGGGACC TGGTGACCTT GGAGAGGCTT 13920
 GGGGACCTGG TGTCTCGGA GTGCCTTGGG GACCTAGTGA CCCGGGAGAG GCTTGGGGAC 13980
 CTGGTGTCCC GGGAGAGGCT TGGGGACCTG GTGTCTGGG AGAGCCTTGG GGATCTGGTG 14040
 TCCTGGGGAG AGGCTGGGGG ACCTGGTGTG TCGGGAGAGA GCCTTGGGGA CCTGGTGACC 14100
 CGGGAGAGGC TTGGACACCT GGTGTCCCGG GAGAGGCTTG GGGACCTGGT GACCCGGGAG 14160
 AGCCTTGGGG ACCTGGTGTG CTGGGGAGAG GCTGGGGAC CTGGTGTCTC GGGAGAGAGC 14220
 CTTGGGGACC TGGTGACCCG GGAGAGGCTT GGACACCTGG TGTCCCGGA GAGGCTTGGG 14280
 AGCCTGGTGT CCCGGGAGAG CTTGGGGAC CAGGTGACCT TGGAGAGGCT TGGGGACCTG 14340
 GTGATCTTGG AGAGGCTTGG GGACCTGGTG TCTCGGAGA GGTACGGGG GCTGGTGGG 14400
 GGAGAGAACG TTGTGAGCCA AAGTCCCTGA ATCCCTGCGA AAAGAGCGCA TCGGGAGCTC 14460
 CCCCTGAGGG CGTTCATTG GTGGACCCCC CTCCCATGCG CTTTGCAGGG AGCTGTTCGG 14520
 ATTCCCCTGG CCCGGCTCCC GCGGATGCAT CCAGTGGCAG CGCCAATTCT GGGCCAGGGG 14580
 GAAGGAGGAA AGGCGGGTGT GGGGTGGTCT CCACGGCTGG AGAAGGGGCG ACGCTCCCTA 14640
 GGGGAGAAGA GGCACGTTGG GGGTTCCGG GGGCGCGGG CGGAGCAGGC CCCCAGTCC 14700
 CCATCCTGCG CCCTACCCC GCCGGTCCG CTCCCGCAGG TCCAGGCTCA GCTGCAGCTG 14760
 GAAGGCGTGG CCCACGCGCA CCCGCACCTG CACCCGCACC TGGCGGCGCA CGCGCCTAC 14820
 CTGATGTTCC CCCGCGCC CTTCGGGCTG CCCATCGCGT CGCTGGCCGA GTCCGCTCG 14880
 GCCGCGCCG TGGTCGCCG CGCCGCCAAA AGCAACAGCA AGAATTCCAG CATCGCCGAC 14940
 CTGCGGCTCA AGGCGCGGAA GCACGCGGAG GCCCTGGGGC TCTGACCCG CGCGCAGCCC 15000
 CCCGCGCGCC CGGACTCCCG GGCTCCGCG ACCCCGCTG CACCGCGCGT CCTGCACTCA 15060
 ACCCCGCTG GAGCTCCTC CGCGCCACC GTGCTCGGG CACCCCGGA GCTCCTGCAA 15120
 GAGGCCTGAG GAGGGAGGCT CCCGGGACCG TCCACGCACG ACCCAGCCAG ACCCTCGGG 15180
 AGATGGTGCA GAAGGCGGAG CGGGTGGCG GCCGTGCGT CAGCCCGGGC CTCTCCAAG 15240
 CTGCCCCTGC GTCCTGGGAC CCTGGAGAAG GGTAAACCC CGCCTGGCTG CGTCTTCTC 15300

TGCTATACCC TATGCATGCG GTTAACTACA CACGTTTGGG AGATCCTTAG AGTCTATTGA 15360
 AACTGCAAAG ATCCCGGAGC TGGTCTCCGA TGAAAATGCC ATTTCTTCGT TGCCAACGAT 15420
 TTTCTTTACT ACCATGCTCC TTCCTTCATC CCGAGAGGCT GCGGAACGGG TGTGGATTTG 15480
 AATGTGGACT TCGGAATCCC AGGAGGCAGG GGCCGGGCTC TCCTCCACCG CTCCCCCGGA 15540
 GCCTCCAGG CAGCAATAAG GAAATAGTTC TCTGGCTGAG GCTGAGGACG TGAACCGCGG 15600
 GCTTTGGAAA GGGAGGGGAG GGAGACCCGA ACCTCCACG TTGGGACTCC CACGTTCCGG 15660
 GGACCTGAAT GAGGACCGAC TTTATAACTT TTCCAGTGTG TGATTCCCAA ATTGGGTCTG 15720
 GTTTTGTTTT GGATGGTAT TTTTTTTTTT TTTTTTTTTT GCTGTGTTAC AGGATTGAGA 15780
 CGCAAAGAC TTGCATAAGA GACGGACGCG TGGTTGCAAG GTGTCATACT GATATGCAGC 15840
 ATTAACTTTA CTGACATGGA GTGAAGTGCA ATATTATAAA TATTATAGAT TAAAAAATAA 15900
 ATAGCCGTGC ACTCTTGACC CCGTCAACGT CCAACGTGGA AAAGGCGTTA CCTCTTCTCC 15960
 CAGCGCTGGC CGCCTGGCCA CTGAGGGCCC TTGCAAAAA TCACGGGTGT AGAGATGGCC 16020
 CTGGGCGCGC TGGGAGTGTG GTTGTGTTTC TGAAGGGGAT AAAAGAGGGC ACGGTGGTGC 16080
 CAAGATATCA GTTTGGTACC TGAGCTGTTT CTGGTTGGGA AGCGTAAAAG CCAGGGAGAG 16140
 ATCCAGAGAG TTTTCAAGTT TTTGCAGATG TAGGTGGTTC CAGCTTTTCT TTCTCCCCTA 16200
 CTCCATCTTC TCGGTTCCCC CAGTTCTTTT ATTTCTTGT TTTTTATTTT TGAGACAGAG 16260
 ACTTGCTTTG TCGCCAGGC TGGAGTGCAG TGGCGCAATG TCAGCTCACT GCCACCTCCA 16320
 CCTCCCGGGT TCAAGCGATG CTCCTGCCTC AGCCTCCGA GTAGCTGGGA CTACAGGCAC 16380
 CTGCCACCAC CCCC GGCTAA TTTTTGTAT TTATAGTAGA GACGGGGTTT CACCGTGTG 16440
 GCCAGGCTCG TCTCGAATC CTGACCTCAG GTGATCTGCC CGCCTCGGCC TCCCAACGTG 16500
 CCCCCAGTTT TATAACAGC AGATAGCAAC TTGTCGTAC AGCTGGCATG GGCTGGACAG 16560
 TTGCTTGAAG TGACCTAACC AAAAAACATC AAGGGTTCG CCCCCAGATT TCGGGAGATC 16620
 CACGTTCCAT GTTCTGATTG GTTTTCTGGG AACACAGCAA GGGGTTTGGT GACCTCCGAG 16680
 AAGATCCATC TGCATGATTG GCATTAGTTA CCACAGCCTG CCCAGAGAGA AACTATCTTC 16740
 TCCCAACATT TACTAACATC CACTGGTCAA CTCTCTTATT TCCATAACAC ATTTGCATCT 16800
 TTCTGGATTC AAGCTTGGTG GTTTTCTTTC CTAACCTCTG ATTTAGATAC TTCTCCCGA 16860
 GGTGGGGATA AAAGAAAAA AAAAAACAC TTCTTTTTT CTTCGCATA AACTTTCTA 16920
 TCTTGCACT GAGCTGAACT GTAGATCCAT TTGGACCCGT CTCATTGTA TCTTCTGATA 16980
 TTCTTTATAC AAACCAAAG TCCCCTCAA CTTTTTTAT GTCAAATGT TACAACCGCT 17040
 GTAAAATGAC GGAGAGAGAG AGAAAGAATC CCAGACATTA ACGGTATTAG AGAGTTTGGC 17100
 TCATTCATCC ATTTTCTTA AAAGCTGGAA ATTAATAAAA AAAAAGAGAG AGAGAGGCTT 17160
 TAATAGTTAA GCTGAAATTT TTATCGAAAA GAAGAATTGC ATTTTGAATC TTTGGGAAGT 17220
 AGGTTCAATC ATCAGAGTAT GTAACCTTT GGAAAAGTGG TTGGTAAGAT ATGTACAGCC 17280

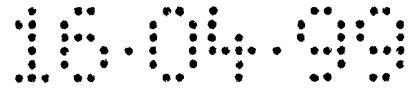


CTAGATTTTT TTTTTTTTAA CAAAAAAGGC TGAGTAATTT TGAAAAATCG AAACATAACA 17340
 GTGTGTCATC ATTTCTCCC AAGAAAAAGC TCACTCCACG TGAGTAGAAA GACATCTACC 17400
 TGGTCCCTGT AGAATCTGAA CGTTTCTCTT TAGAGACGGA ATTTCAATCT TGTGCCCAG 17460
 GCTGGAGTGC AGTGGCACAA TCTCGGCTCA CCGCAACCTC CGCCTCCCGG GTTCAAGCCA 17520
 TTCTCCTGCC TCAGTCTCCC GAGTAGCTGG GATTACAGGC ACCTGCCACC AGGCCCTGGT 17580
 AACTTTCTGG TATTTTTAGT AGAGACAGGG TTTCAGCCTC CCGAGTAGCT GGGATTACAG 17640
 GCACCTGCCA CCAGGCCTGG GTAACCTTCT GGTATTTTTA GTAGAGACAG GGTTCAGCC 17700
 TCCCAGTAG CTGGGATTAC AGGCACCTGC CACCAGGCCT GGGTAACTTT CTGGTAGTTT 17760
 TAGTAGAGAC AGGGTTTCGG CCTCCCGAGT AGCTGGGATT ACAGGCACCT GCCACCAGGC 17820
 CTGGGTAACT TTCTGGTATT TTTAGTAGAG ACAGGGTTTC GGCCTCCCGA GTAGCTGGGA 17880
 TTACAGGCAC CTGCCACCAG GCCTGGGTAA CTTTCTGGTA TTTTGTAGTAC AGACAGGGTT 17940
 TCGGCCCTCCT GAGTAGCTGG GATTACAGGC ACCTGCCACC AGGCCTGGGT AACTTTCTGG 18000
 TAGTTTTAGT AGAGACAGGG TTTCAGCCTC CCGAGTAGCT GGGATTACAG GCACCTGCCA 18060
 CCAGGCCTGG GTAATTTTTT TGCATTTTTG GTAGAGACAG GTTTTTGCCG TGTGGCCCG 18120
 GCTGGTCTCA AACTCCTGAC CTCAGGTTGA CCTGCCCGCT TTGTCCCTCG CAAAGTGCTG 18180
 GGATTACAGG CGTGAGCCAC CACACCTGGC CTGAATCTGA ACTTTTAAAA GGGAGTTACT 18240
 GACTCTCAAC TGTGCGGGGA CGGTTTCACT TTGATTTAAT ATGGAAAGAG GGCCAAGTGT 18300
 CATCCTCACA AATGGGTCCC CGAAGCAGAT CAAACGCAGA GAACTGTGAG GGTGGGACAC 18360
 GAGTGTCTGT GGACACTGGC TGCCTTTGGC TTTTCTCCTG CGAGAGAAGT TGGGTGACTT 18420
 TCTGTAGGTG GATGAGTGAT CCCTGAATGA GTGTGGGGTA CGTGTATGCT AGCTGCTTCT 18480
 TTCTCCCTGA AACTCTCGGA TGAAGGAAG TAAGAAATTC AGCTTGGGCT GTGACCAGTT 18540
 CTCACCACCA ACGCCCTCTT CTCTCTCCCT TCTCCTTCTT TCCTTCTTCT CTTCCTTTCT 18600
 TTCTTTTTCT TTCTTTCTCT CTTTCTTTCT TTTCTTTCTT TCTGTTCTT TCCTTTTAT 18660
 CTTTCTCTCT TTTTCTTTCT CTTTCCCTTT TTTGTTCTT TCTTTCTTTT TCTTCTTTC 18720
 TTTTCTTTT TTCTTTCTTT CTTCGATGAA GTCTCACTCT GTCACCCAGG CTGGAGTGCA 18780
 GTGGTGCAAT CCCAGCTCAC TGCATCTCTT ACCTCCTGGC TTCAAGAAAT TCTCCTGCCT 18840
 CAGCCTCCCA AGTAGCTGGG ATGACAGGCA CCCACCACCA TTCCCGGATA ATTTTTGTAT 18900
 TTTTGTAGTAG AGACTGGGTT TCGCCATGTT GGCCAGGCTG GTCTTGAACCT CCTGACCTCA 18960
 CATGATCCAC CCGCCTCAGC CTCCAGAGT GCTGGGATTA CGGGGTGAGG CACCGCGCCC 19020
 GGCCTCCTCT CTCTTTTCT GAGATGTTTA GGAAGGACTG GGCTGATGGG GACCCTCTGT 19080
 ATGTGATGTG CGTGGGTTTG GTTCCCGGA AGGCCCTCCA GAGACACGTT TGCGTGAACA 19140
 TTCAGCATGG AAACAACATA CGTCTCTCCA CAGGAGGTGA GAAATTGAAT TTATGGGGTG 19200
 GGTGTACGCT GGCGATTCTT GGTGCTTTTT GCTCAAACA AGGTTCTTTT GAAAGTCAGG 19260
 TTCCTGCTTT CCCTGTGGCT TCCCGGTGAG CTCGCTCGCA GAGCAAGGAA TACCACCCAG 19320

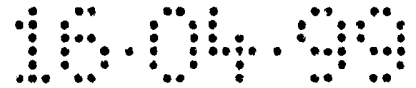
AGAGCAACGT GGGCTGTGTT CCGTTGTAAC GCCGTTGCAG AGAGAGGATT TGGTGTGTGA 19380
 GATCCGTACC AGCTCCAGCA CACTGATAGG AACACGTTGC TGGCCGAAC GAACGATGCT 19440
 GGGTTGGGTC CTGATTGATA CGTATTTTCT TCCCTCCTCT CCCCAAAAC TGGCCAAATA 19500
 GTCCGTGGAG GGTGTGTCAGT CGCCGCAGTT GAGCAAAAAA CACTTCTTCC TTTGAGTGGC 19560
 TGTTCTGGTG AAATCTGTTT CTGACATATC CACTTTTCTC TCTCTTTTCT CTCTCTCTGA 19620
 CTGCGAAGCA CCCACAGGGA GAAGGAATTG GATGTATCGG ATGTTGGTAT TAGATTTTCT 19680
 TTCTCCGTTT GAGTCTCTGA CTGGTGCATA CTTTGCAAAG GTGTGTTCCCT GGCAATTGCC 19740
 AAGAGTTAGA AAAATGCACC TTCTCTGGTG GCCGTTGGGG TGTTGTTTCA CAGGCAGTGG 19800
 TGACAGGGCC CCTGGCTGTG GGCTGTCTTC TCCAGCGCCG TGGATAAAGA GACGGGACAG 19860
 ATTCTGTGCC TCTGTACGAT TTAGAGCGTA ACTGACCGCG TCCAACACCC GTTTTTCCAC 19920
 TTACAAAGCT GGTGGTGCAG CGGGCTTGGT GTCTCCCCTA CGGGAAGGAG GCCTTTGGGC 19980
 CGCTCCAAAG ACGCCCTGTC GTAGGAATGG CCTCTCCATC CCGCCAAAGT CCAGCCAGGC 20040
 CCCCAGAAATG GTCCCATTTT CTGGAAGCC TGAGTTTCTG TTCTGGTCTT GCTGCTGTCC 20100
 TTGGCCACGT CAGCACGTGG GAGCATCTGT GGATACCGCA GAGTCTGGGG ACAGCTGGGC 20160
 GTTTAACCGA AATGAAGCCG AGACGGGTTT CAGGTTTTGG TGCCAAGCTC TGGTCAGGAT 20220
 GAAAGGGAAA TACCAGAGTC CTCTGTCCCT GCCTCTGGGT TTCATGCTGA CCTTTCTAAC 20280
 ATTTGTTTTT CCCTAAGAAC AAGCAGAAGC CTCCAGCTCC CTTTAGCTCC ACAGTTTTCC 20340
 CGGGGACATA GCGAGGATGG CACACGGCAG CCACTCCAC GACACACATT TCGGAGGCAC 20400
 TTTGCTGGAA GCCGCTTGTC TCCTCCAGCT TTGGGAGGTC TGGGGAGGAG AGAGGCTTTC 20460
 GGTGGACACG TTTGACATTA AAAAAAAAAA AAAAAAAAAA AAAAAAATG GTGCCTAATT 20520
 TATTAAGAG AATTAGCTTA GCGAGTATAT GCTGATATTC TTCGACACAC GTGGGTAAGT 20580
 TGATGCCATT TATAAATGTT TTATTGAAAT TTGATATTTA ATGAGAAGCC GGTAAAGGAA 20640
 TGTAGACAAT ATCCCGTTT AAAGCTATGA AATGTGCTAT TTATTGAAAG GGGATGTGGC 20700
 TTCACGAGTT CAGCCCATG TACGTGCAGG TCCCGTGGGA AGGAGGCAAA AGCCCTGCT 20760
 TCTTACTTTG TGATGTATGT GCATTGTTA TTTATTTTTT TTTCTTGGT CGGACGTCA 20820
 TAAATATGTA CTATTTAAT TATGTCGAGT GTAAATTTGA CATCGCGTTG CATTATTTT 20880
 TATATTTCTG AAAACTGTTG CTTTTCTTT TCCCTCCCC CATTGACGAC ATAGCGGCTCC 20940
 CCGCGTCCGG GTTACAAATA CATCTACAGA TATTTTCAGG GATTGCTTCA GATGAAAACA 21000
 AATCACACAC CGTTTCCCAA ACCAACAGTC TTCACATTC TATCCCTCTG TTATTGTCCG 21060
 CAGGCGGTGA GGGGTAGAAA AAAAAAACA AAAAAAACA AAAAAAACA CAAAAAACA 21120
 CACCCTGAGT TTCTCTGGTG ACGCCCTCAT TCTCCTAACG TTCAATAATC TCAATGTTGA 21180
 GTTGACAGCA CAGACTGTAT TTTGTGACG CCCCCTAGTA TGAATGTACA TCTTGTAATA 21240
 CTGAGATATA AATAAACTTA TAAATATTTG TATTCAAGTG TTAATAAATA AAAAATTTCT 21300



AACCTCTCCC	CTGAGGACAG	GCTTATTGGA	AAAAAAAAAA	AAAAAAAAAA	ATCCTGAGTC	21360
GGCCGTGGCT	GAACACAGAG	TGTTGTTCTG	CTCCGTGCAT	TTCCAGGGTG	GGTACCCAGT	21420
GTTGCCCCCC	AGCCTTAGAT	CGGGAGGTAC	CATTGACTTT	TGCTTGATC	CCATCCCCTT	21480
CCTTTACTGA	AACCTACCTC	CCCCTTCTC	AGCCAACGTC	CCCCAGAAG	GTGGCAAAAA	21540
AAACAGAGGA	AAAAGCCCTG	ATTTGAATCA	AGTCAGAGCT	GCTAATCTC	CACTTCTTT	21600
AATTAATTAA	TTTATTTTTT	TTTTTGAGAC	TGAGTCTCGC	TCTGTGCCC	AGGCCGGAGG	21660
AGTGCAGGGG	CGCGATCTCG	GCTCACCGCG	ACCTCCGCT	CCCGGGTCA	AGCGACTCTC	21720
CTGCCTCAGC	CTCCCAGTA	GCTGGGATGA	CAGTCACCTG	CACCACCGCG	CCCGGCTCAT	21780
TTTTGTATTT	TTAGTAGCAA	TGGGGTTTCA	CCGTGTTGGT	CAGGCTGGTC	TCGAACTCCT	21840
GACCTCGTGA	TCCACCCGCG	TCTGGGCCCG	GCCGGTGATG	TGTGTGCTTT	TAACFTTTAT	21900
TTTGTTCAG	TTTTCGACAG	TGGCACGGAT	TTCCAGCAC	GGTCTTGCAA	GGATGATTGA	21960
GTCATTTTTG	AGACAAAAAA	TATAATAATA	ATAAATGGAA	AAAGAAATCG	ACTTTTAAAA	22020
ATGACAAATT	TTTTTTTTTT	TTTTTTGCAT	AGATTTTTCT	CTCTTATGT	AAAGGAAAGT	22080
TCATGATTGG	ATTTGGCCGG	CCTGACTGCT	TCCCGGCTGT	GATAAAAAAC	ACATGTGAGC	22140
TGGGAGGGAA	GTGGGGGAGG	GACACAGCTG	CCCACACAGG	GTTCCACCG	CGGTTACAGG	22200
GTGGGCAGTG	CTGGGGGAGC	TTTCTCTGTG	GGGGGCTCAG	AGCCTGAGGA	CAGGTGAGCC	22260
TCTCCGACAC	CTCCCCAGTT	GCCTGGAGTC	TAAACCGTCC	GTTGTCTGTA	CCGTCCGTTT	22320
TTCTGCTGA	CTCCTGGTAG	TTCCTGAAAG	CTTCTCTTGG	CCAGAGAAGG	GGTTTCAGAG	22380
GCCGTGTGTC	CAGGCCATTC	TGCAAAGTGC	AACTTGACCG	TTCCTTTCCT	TTTCTGGCCT	22440
GCGTGGTCTG	AAGCTCAGAG	CCCTCTCTTC	ACCCAGCCTG	TGTGTGTCTT	GCCGGACAGA	22500
AGAAAAATGG	TGCTTTTTTG	GTGTTAGCAG	AGGTGCTTTT	CATGGCTGAC	CTCAACGCGT	22560
CCATCTCCAG	CCTTGACCAA	GCTGTTTTTT	AGGGGCAAAC	GCAGGCAAGT	TCTGAATGCA	22620
CACAGTTATT	TCATGGTTAA	ACTATTCAGC	TTTGGCCGGG	CGCAGTGTGG	CTCTCACGCC	22680
TGTCATCCCA	GCACTTTGGG	AGGCCGAGGC	GGGTGGATCA	CCTGAGGTCA	GGAGTTCGAG	22740
ACCAGCCTGG	CCAACACGGT	GAAACTCTAT	CTCTACTAAA	AATACAAAAA	TTAGCCGGGC	22800
GTGGTGGTGT	GTATCTGTAA	TCCCAGCTAC	TCAGGAAGCT	GAGGCAGGAG	AATCGCTTGG	22860
ACCCAGGAGG	CGGAGGTTGC	ACTGAGCCGA	GATCGGCCA	TTGCACTCCA	GCCTGGGCGA	22920
CAGAGCCAGA	CGCTGTCTCA	AAAAAATGAA	TAATAAAATA	AAATAACAGG	AACTAAATAA	22980
AATAAAACGT	TCAGCTTTGT	TCTGCAAATC	CACTCCTATT	GTTTTACGTG	GTTTGAGAGA	23040
CTCTGTCCCT	TAGAAATAGA	TGTTTGTGTC	CAATTGTAAT	GAATCTGTTT	CAAAAAATGAA	23100
CAGAAATATC	AAATGGTTTG	AGAGATCTTT	TCCCTTAGAA	ATAGCTTGTT	GCCAATCACA	23160
AAGAATGTTT	TTCAAAAATG	AATGGAATCT	TCCTGGATAT	CGCTCCAGA	TCTTCATTTT	23220
TTTTGCATAG	TTCAACCTGA	AAAGTAAGTG	TCTCAGCCCT	GAATTTCTTT	CTGATTTTTT	23280
CATGGGTTGT	CTTGCAGACT	TCTCTGGACT	TGACCACATT	TAAAAAAAAA	AAAATTAACT	23340



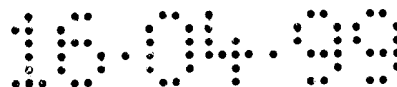
TTTTACACAG	GACACGGTTT	CAATAGGAAT	GAGATCTTTG	AGTTTTTATG	TAACAGATTC	23400
TTACCATCAG	TTCTCAGATT	CCCAAATTAC	ACACAAAAAG	CCACGGACTT	CGCCTCCTGC	23460
TAACATGTCC	TTCTGTTTCT	GAGGCTTCTG	TTGGTGTTAG	ACTTTCATGT	TTAATAGCAG	23520
ACAATGTAGG	GATTTAAAGA	AAAATGCAGA	GAAAGCAAAA	ACACTGACCA	AACACACGGA	23580
GATAAGCTTT	CTAAAGCCTT	TGTTCTTGGA	GTTGTCGTTA	AAAAAAAAAA	GTTGTTTTAA	23640
ACTTTGCAAG	CATGCCTATA	TTGAACTCAT	AAGCAAGAGA	GCCAAGAAAA	ATAGTGTCGG	23700
TCGTCTACTC	TACACGTTTT	CCCAAAACAG	ACGTATTTTA	ATTTCTTTTG	TTTGAACTCA	23760
CAGATGCTGA	GAGTTAAAAG	TTAAATTTTT	GTCATGAACA	ATAGTGGCCA	AAACCACAGT	23820
TACTTTTGCA	CTATAGCATA	ATAAGAAAAA	TACAGGCTGG	GCTCGGTGGC	TCACACCTGT	23880
AATCAAAGCA	CTTTTGGAGG	CGAAACAGCC	AGATCCCTTG	AGCCCAGGAG	ATTGAGACCA	23940
GCCTGGGCAA	CATAGCGAGA	CCCTCATCTC	TACAAAAAAG	GTTTGTTACA	TATGTAACAA	24000
ACCTGCACAT	TGTGCACATG	TACCCTAAAA	CTTAAAGTAT	AATAATAAAA	AAATTAAAAA	24060
AAAATTCACC	AATCAACTGC	CTGCTGGTGC	CTTCAAGAGA	CTCACCTAAC	ACATAAGGAC	24120
TTGCATAAAC	TTATAAAACA	ATTCAATGGA	AGAATCCTTG	AAAGTATTCT	GAGAAGACAG	24180
TATAATAAAC	TGATTTCTAA	AAAGGCTATA	AAAAATTGAA	TAAATCATTG	TTGGGCATCC	24240
TGTGCTGAAA	TATAATGCAG	CCAATAAAAA	TTACAAAATG	AATAAACATT	TTATAACAAT	24300
AAAAAAAAGT	CAAATAATTA	GGCAGGCATG	GTGGTGCTCT	CCTACGGTTG	AAGCTATTCA	24360
GCAGGCAAGA	GGATACTTTG	TTTTTGTTTT	TTAATTTTTT	TTGAGACAGA	GTCTCGCTCT	24420
GTTGCCAGGC	TGGAGTGCAG	TGGCGTGATC	TCAGCTCACT	GTAATTTCTG	CCTCCCGGGT	24480
TCAAGCGATT	TTCTGCCCC	AGCCTCCCGA	GTAGCTGGGA	TTACAGGTGC	CCGCCACCAC	24540
ACCTGGCTAA	TTTCTTTTGT	ATTTTTAGTA	GAGACGAGGT	TTCCCCATGT	TGGCCAGGCT	24600
GGTTTTGAGC	TCCCGACCTC	GGGTGATCCA	CCCGCCTCAG	CCTCCCAAAG	TGCTGGGATG	24660
ACAGGCGTGA	GCCACCGCGC	CTGGCCAGG	AGGATTATTT	GATCCCAGGA	GGTGGAGGCT	24720
GCAGGAAGCC	ATGATTGCAC	CACTGCACTC	CAGCCTGGCT	GACAGAGTGA	GACCACATCT	24780
CTAAATAAAT	GAATAAATAC	AGGCAGAAAC	TTTTTTTGT	TTGTTTTGAT	GGAGTCTTGC	24840
TCTGTACCA	GGCAGGAGTG	CAGTGGTGCC	ATCTCAGCTC	ACTGCAACCT	CCACCTCCTG	24900
GGTTCAAGCA	ATCCTCCTGC	CTCAGCCTCC	CGAGTAGCTG	GGATTACAGG	TGCCCGCCAC	24960
CACGCCCGGC	TAATTTTTTG	TATGTTTAGT	AGAGACGGGA	TTTACCCTG	TTAGCCAGGA	25020
TGGTCTTGAT	CTCTTGACTT	TGTGATCTGC	CTGCCTCAGC	CTCCCAAAGT	GCTGGGATTA	25080
CAGGCATGAG	CCCAGGAGTT	CAAGACCAGC	CTCAGCAACA	AAGTGAGACC	TTTTCTCTCC	25140
AAAAATCAA	AAATTTAGCC	AGCTGTGGTG	GCTCCTGCCC	GTGATCCAG	TACTGTGGGA	25200
GGCTGAGGCA	GAATTGCTTG	AGCCCAGGAG	TTCGAGACCA	ACCTCAGCAA	AAAGGACTCT	25260
CTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCTATAT	ATATATATAT	ATATATATAT	25320



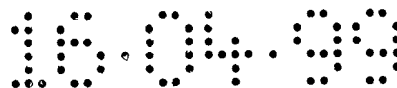
GAGTTTCAAA	AATTGCTGGG	TGACCAGCTC	ATCTACTGGT	TTCCCTTG	GGAAAGTGAA	25380	
ATTGTCATGT	ATTGAAGATT	TCCAAGGAAG	TTGTATTGAA	TGAGAAACAA	ACTCAATCTG	25440	
TTCGTGTTTA	AAGAGCTGCA	GTGCGTTTGC	TGTGTTTCCC	ATAAACTGC	ACTTCCAAAA	25500	
GACACGCTGA	GAAAGGAGAC	CAGGATTTGT	AATTCAGAAA	TTGGAAAGCA	AGTTAGGCTG	25560	
GACGTGGTAG	CTCATGCTTG	TTGTAATCTC	AGCACTCTGG	GAGGCTGAGG	CAGGAGGATC	25620	
ACTTGAGCCC	AGGAGTTCAA	GACCAGCCCG	TGCCACATGG	TGAAACCCTG	TCTCTCCAAA	25680	
AAATAAAACA	TTTAGCCAGA	TGTGGTGA	CATGCCTGTA	ATCCCGGTAT	TCTGGGAGGC	25740	
TGAGGCAGAG	TTGCTTGAGC	CCAGGAGTTC	AAGACCAGCC	TCGGCAACAA	AGTGAGACCC	25800	
TGTCTCTCCA	AAAAATAAAA	CATTTAGCCA	GCTGTGGTGA	CTCATGCTTG	TAATCTCAGT	25860	
ACTCTGGGAG	GCTGGGGCAG	AATGGCTTGA	GCCCAGGAGT	TCGAGACCAA	CCTCAGCAAC	25920	
AAAGTGAGAT	CTTGTTTCTC	CAAAAAATCA	AAAATTTAGC	CAGCTGTGCT	GGCTCATGCC	25980	
TGTAATCCCG	GTA	CTG	GGG	AGG	AGGCTGAGGC	26040	
AACCTCAGCA	ACAAAGTGAG	ATCTTGTTTC	TCCAAAAAAA	TCAAAAATTT	AGCCAGCTGT	26100	
GCTGGCTGGT	GCCTGTAATC	CCGGTACTCT	GGGAGGCTGA	GGCGGAATTG	CTTGAGCCCA	26160	
GGAGTTCAAG	ACCAGCCTCA	GCAACAAAGT	GAGATCTTGT	TTCTCCAAAA	AATAAAACAT	26220	
TTAGTCAGCT	GTGGTGGCTC	AAGCCTGTGA	TCCAGCATT	TTGGGAGGCC	GAGGCGGGCG	26280	
GATCACGAGG	TCATGAGATC	GAGACCATCC	TGGCTAACAC	GGTAAAACCC	CGTCTCTACT	26340	
AAAAATACAA	AGAAAATTAG	CCGGCGTGG	TGGCGGGCGC	CTGTAGTCCC	AGCTACTCAG	26400	
GAGGCTGAGG	CAGGAGAATG	CCGTGAGCCT	GGGAGGCGGA	CCATGCAGTG	AGTCAAGATC	26460	
GGCCACTGC	CCTCCAGCCT	GGGCCACAGA	GCAAGACTCC	GTCTCAAAAA	AAAAAAAAAAA	26520	
AAA	ACTGCTG	CCCAACCTGT	GTTTGACCA	CTGCCCTCCA	GCCTGGGCAA	CAGAGCAAGA	26580
CTCCGTCTCA	AAAAAAAAAAA	AATGCTGCC	AAGCTGTGTT	TGCACCACTG	CCCTCCAGCC	26640	
TGGGCAACAG	AGCAAGACTC	CGTCTCAAAA	AAAAAAAAAAA	AAAATGCTGC	CCAAGCTGTG	26700	
TTTGACCCAC	TGCCCTCCAG	CCTGGGCAAC	AGAGCAAGAC	TCTGTCTCAA	AAAAAAAAAAA	26760	
AATGCTGCC	AAGCTGTGTT	TGCACCACTG	CCCTCCGGCC	TGGGCAACAG	AGCAAGACTC	26820	
CGTCTCAAAA	AAAAAAAAAAA	AATGCTGCC	AAGCTGTGTT	TGCACCACTG	CCCTCCAGCC	26880	
TGGGCAACAA	AGCAAGCCTC	AGCTTTCTGC	CATCTCCACA	ACCAAGAAAG	CAATTCACAC	26940	
AGAAATCAGT	GCATCGTGCA	GTGACCTCTT	CAGAAAACCA	ATGAGTTTTC	CACCTGAGGA	27000	
ACTGTTTCTG	AGCCCCATTC	AGAAAAACAC	ATCCCTGTAA	CTGCAGGGCA	GATTACTCA	27060	
CTGTATGCCT	GTTTAAATAA	AGCTTCCAGC	CTCTGCATGG	GGTCTGTCTG	GAAGCTCCTG	27120	
TATCTGTCCC	ACATTCTTGG	AATCACAATG	CACCCTGGG	AGGAAGATAT	GTATTTAAAG	27180	
GGAGTGGATG	TTATGGTGAG	AAAATGCTGC	CCATCCTTCT	AGAAGACAAA	AGCCACACAA	27240	
AATACATCAC	AAGAACCAGT	TTTTTTCAGA	GAAGAACCCTG	CACAAAGAAC	CTGCTCCCCC	27300	
CACACCCCCA	CACACAGGTG	AATTAACAGG	ATGTATGTTT	TATCATAAAA	GCACAGGTTT	27360	



GTTTCCTATG CACTCTCTGA GGATTTGGCC ATATGCAAAG ATGTACAAAA ACCTTCTCTT 27420
 TCCCCAGGGA ACCGTAACCC GTCTGAAAAG ATGCCCTTCT CAGAAGCGAG TTGAACGATT 27480
 GTTGGAAAAG ATAAAATACG ACGTGCACAC ACACAGTAGA GAAATGTCAC CCATGCAAAAT 27540
 TATGTGTTTG AATGGAACAC ATTCAGGAAG CTAAATGGGG TATGACCACA CATTGGGTT 27600
 GATTTATTTG ACGAGTGGAA GGGGCAGATG GAAATGAATA CTGCTGTTTT CTTTGGGAAG 27660
 GCCATATATG GGAATACCAA GAGGATTACT TTGGAAGTTT AGCTTCTCCA GGTGGTCTCT 27720
 CTCTCTCTCT CTTTTTTTGA GACAGAGTCT CACTCTGTCA CCCAGGCTGC AGTGCAATGG 27780
 CGTGCTCTCG GCTCACTGCA ACCTCAGCCT CCCAGGTACA AGCGATTCTC CTGCCTCAGC 27840
 CTCCCGAGTA GCTGGGATCA CAGGTGTGCA CCACCACGCC TGGCTAATGT TTGTATTTTC 27900
 AGTAGAGATG AGGTTTTACC ATGTTGGCCA GGCTGGTCTT GAACTCCTGA CCTCAGGTGA 27960
 TCCGCCTGCC TCGGCCTCCC AAAGTGCTGG GATGACAGAC ATGAGCTAGC ACGCCCGGCC 28020
 CCAGGTGGTC TTTTAGCGG GTATTAAAGC AGCTTCTCT CTGAGCCTTA AACCATGAAG 28080
 ATAGACAGAC TCAGTGTATG GGTTTTAGAG TTGTAATTTT ATAAAAATAA GAAAAAGTCG 28140
 ACCTATCATT GATGGTTAGT ATTTTTTGTA GCAGTTGCAT GCAATATTAG GATAAGGCAT 28200
 GTTCTCAAAA AGAACTCTTT TTTTTTTTTT TTTGAGACGG AGTCTCGCTC TGTACCCAG 28260
 GCTGGAGTGC AGTGGCACGA TCTCCGCTCA CTGCAAGCTC CTCTCCCGG GTTACGCCA 28320
 TTCTCCTGCC TCAGCCTCCC CAGTAGCTGG GACTACAGGC GCCCGCCACC ACGCCCGGCT 28380
 AATTTTTTGT ATTTTTAGTA GAGACGGGGT TTCACCATGT TAGCCAGGAA GGTCTCGATC 28440
 TCCTGACCTC ATGATCCGTC CGCCTCAGCC TCCCAAAGTG CTGGGACTAC AGGCGTGAGC 28500
 CACTGCACTT GGCCTTTTTT TTTTTTTAGA TGGAGTTTTG CTCTGTCTGC CCAGGCTGGA 28560
 GTATAATGGC ATGATCTCGA CTCACTGCAA CCTCCGCCTC CCGAGTTCAA GCGATTCCTC 28620
 TGCCTCAGCC TCCCGAGTAG CTGGGATTAC AGGTGCCAC CACCATGTCA AGATAATGTT 28680
 TGTATTTTCA GTAGAGATGG GGTTTGACCA TGTGGCCAG GCTGGTCTCG AACTCCTGAC 28740
 CTCAGGTGAT CCACCCGCCT TAGCCTCCCA AAGTGCTGGG ATGACAGGCG TGAGCCCCTG 28800
 CGCCCGGCCT TTGTAACCTT ATTTTTAATT TTTTTTTTTT TTTAAGAAAG ACAGAGTCTT 28860
 GCTCTGTCAC CCAGGCTGGA GCACACTGGT GCGATCATAG CTCACTGCAG CCTCAAACCTC 28920
 CTGGGCTCAA GCAATCCTCC CACCTCAGCC TCCTGAGTAG CTGGGACTAC AGGCACCCAC 28980
 CACCACACCC AGCTAATTTT TTTGATTTTT ACTAGAGACG GGATCTTGCT TTGCTGCTGA 29040
 GGCTGGTCTT GAGCTCCTGA GCTCCAAAGA TCCTCTCACC TCCACCTCCC AAAGTGTTAG 29100
 AATTACAAGC ATGAACCACT GCCCGTGGTC TCCAAAAAAA GGACTGTTAC GTGGATGTTT 29160
 TAGCTTCCTG TTCTCGTCTT TTCTTTGTTA ATGTACAGT TTGAGGGTGT GTGTGCGTGT 29220
 GCGCACGTGT GTGTGTGCAG TCTCCTGATT TCATGTATTT AATTGTTATT ACCACCACCT 29280
 CCATCTCTCA TTCCTTCTTA CCCTCACTGT GTAAAGATAC ATGTTGTTTT TAAATTTTAT 29340



GTATTTATAT TTATTTATTT GTATTTCTGA GACAGAGTCT CACTCTGTTG CCCAGGCTAG	29400
TGGCATGATC TCAGCTCACA GCAACCTTTG CCTCCTGGGT TCAAGCGATT CTCCTGCCTC	29460
AGCCTCCCGA GTAGCTGAGA TTACAGGCAC ACACCACCAC ACCCGGCTAG TTTTGTTTTG	29520
AGACGGAGTC TCGCTCTGTT GCAGGCTGCA GTGCAGTGGC GTGATCCTGG CTCACTGCAA	29580
CCTCTGCCTC CTGGATTCAA GCGATTCTCC TGCCTCAGCC TCCCAAGTAG CTGGGATTAC	29640
AGGCGCCAC CGCCACACCT GGCTAATTTT TTATTGGTAG TAGAGACGGG GTTCTCCAT	29700
GTTGACCAGA CTGGTCTTGA ACTCCCAACC TCGGGTGATC CACCCACCTG GGCCTCCCAA	29760
AGTGCTGGGA TGACAGGCGA GGGCCACCGC GTCCAGCCTT CTCTTCTTC TTCTTTTTT	29820
TTTTTTTAAG ATGGAGTTTC ACTCTGTTGC CCAGGCTGGA GTGCAGTGGT GCAATCTCGG	29880
CTCCCTGCAA CCTCCACCTC CCAGGTTCAA GAAATCTTT TGCCTCAGCC TCCCGAGTAG	29940
CTGGGACTAC AGGTGCCCCG CACCACACCC ACCTAATGTT TGTATTTTTT TGGTAGAGAC	30000
GGGGCTTAC CACATTGGCC AGGCTGGTCT TGAACTCCTG ACTTCAGATG ATCCTCCTGC	30060
CTCAGCCTCC CAGAGTGTG GGATTACAGG CGTGAGCCAC GGTGCCCGGC CAGACGTCAT	30120
GTCTTAGGAA ATCAGAAAGT GGGTAGTTTC CGCACTCTGA GGAGAAAAG AGACGTCCGG	30180
CGAAGAGAAA GGAGAGTGAA AGGATGTCTC CTCTTGTCTG TAGCCTGTTC TCAATCGTGA	30240
GTGAGCCAAT TGCCAGAAAC TGAGGGTGCT TCATTTGGCC AGGCAAGCTT CTCAACAGAA	30300
TGTCTAAGTA CTTGTTAATG CTGAGAAGCT CTCCAAGCTA CTGCACTCCA GCCTGGGTGA	30360
CAGAGCACGA CCTTGTCTGA AAACAATTAA TTAATCAATT AATTAATATA ATGAAATCAT	30420
ACTGAACTCA GGAGACCATT GGGGTGGGCA GGGCTGGGGT TGGAAAGGAA CATAAAATAT	30480
GGTGCAATGG ACTTGTCTCC AGTCTCCCTC CCCATCTCTT CTCGCCAAGA GTCTCTGGAG	30540
GGAGCATGGG GAAGATGCTT TGGGAATCTG TAACTTCTTG TCTTGTAAC AGAATATCTA	30600
AGTAATTGTT AATGCTGAGA AGTTATAGAT TTCAAAGCC TTTCTCCAGG CTACGGACAA	30660
GGGTCAATGG TTAATCAGTG TTACAGAAAG AATGACATGG AGATGTTTGT TACATCTTAA	30720
GGAAACATGA GGGGCCAGAG TATTTTACTC TAAGTGTA GAAGTACATTG GCCACGCTG	30780
TCCCAACACC ACCAATGGTG GCACCTAACT TTTGTGTTG TGCCCCACAT TTCTTCTTCT	30840
TTTCTGACGT AAATGCAAGT GATATTCCTT GGAAACCATG CTGCAGCPAG AGGCCATCTG	30900
ACTACTAGTG ATACCCTGTA GCTCACCTAC AGCAGCTCAC TGAAGCAGC TCACCCATAG	30960
CTCAGGTATA GCTCACCTGC AGCGGCTCAC CTGTAGCTCA CGTGTAGCTC ACTGTAGCA	31020
GCTCACTGGT AGCTCACCTG CAGCAGCTCA CCTGTACCTC ACCTGTACCT CACCTGCAGC	31080
AGCTCACCTG TAGCTCACCT GTACGTGAGC CACCGTACCC GGCCAGCAAG ACCCCATTC	31140
TAAAATAAAT ACACAAAAT TAGCCGGACG CGGTGGCGCG TGCTGTAGT TGTAGTACT	31200
CAGGAGGCTG AGGTGGGAGG ATTGCTGGAG GCTGGGAGGT AGAGGCTGCA GTGAACCGTG	31260
ATCCAGCCAC TGTACTCTAG CCTGGATGAC ATAGCAAAAC CTGTCTCAA AAAACAAAAA	31320
CAAAAAACAA AACAAAGAAA CAAACAAAA ACCCACACAC ACCGGAAAAC AAAACAAAAA	31380



GCAAAAAGGA AAGAAAAGAG AGCCAGGTCC CAAATATATA TTCCTTGGGA GAACCATTTG	31440
CAAAGAGCAC ACTTAAGGCC GGGCGCGGTG GCTCACGCCT GTCATCCCGG CACTTTGGGA	31500
GGCCGAGGTG GGTGGATCAC GAGGTTGGGA GATCGAGACC ATCCTGGCCA ACATGGCGAA	31560
ACCCCATCTC TACTAAAAAT ACAAAAATC AGCCAGGTGC TGAGGCAGGT GCCTGTAGTC	31620
CCAGCCACTC AGGAGGCTGA GGCAGGAGAA TGGCATGAAC CTGGGAGGTG GAGGTTGCAG	31680
TGAGCCGAGA TCGCGCCCCT GCACTCCAGC CTGGGCGACA GAGCGAGACT CCTTCTCAA	31740
TAAATAAATA AATAAATAAC AAAGAGCAA CTTAAATTG TCTCAGAAAT CCCACGGGAT	31800
ATTGGATCTC CCTCATGCCT ATCTGATGAC ACTTTGAGTG TCTGGGGCCC CGTGCCTATT	31860
TTCTGGGGTT CCCAGAAGCT GCCGTTCTGA AAGTGTGGCT CTCGGGGACG TGGCACAGGT	31920
GTGGATGTCT GTTTTAAATG TCAGGCGTTT GGACGTTGAG GAACGTGAGG CTGAAGGTCC	31980
CCTTCGCCGA CCCCTGAGT TTAGGGTCTT GCCTTTTAAA ATCTTCCCAG CACTCTGTTG	32040
TTCACGCAAG CGTCCCATCT GTTTGGGTGG CCGTGCCGTC TGCATCTGTC TCGAACCTTC	32100
ACAGCTTTGC AGAATATCCT GTTTCTCAAT ACGGATGGAG AAACACGAGA CGCGTTTTCT	32160
GGGTTATTTT AGCCGTCACG GAGAACCCCA GACTCATGTG TGCTAATGAC CTCATTAATG	32220
ATACTCTGAG GCAGACAGCC CTGCCTGATC TTAACAACAT TTTTAAAT TCTTTTTTTG	32280
TTGTTGTGTG TACAGCATCA TTCATATAAC GTAGGAAACC GTGATCAGTA GCTTTTAGGA	32340
TATTTGCAAC AGGGTGTAA ADAAAABD	32367

(2) Informace o sekvenci SEQ ID NO: 15:

(i) CHARAKTERISTIKA SEKVENCE:

- (A) DÉLKA: 806 párů bází
- (B) TYP: nukleová kyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(iii) TYP MOLEKULY: jiná nukleová kyselina

- (A) POPIS: /popis = „SHOT“

(ix) RYSY:

- (A) NÁZEV/KLÍČ: CDS
- (B) POZICE: 43..615

(xi) POPIS SEKVENCE: SEQ ID NO: 15:

GTGTCCCCGG AGCTGAAAGA TCGCAAAGAG GATGCGAAAG GGATGGAGGA CGAAGGCCAG	60
ACCAAAATCA AGCAGAGGCG AAGTCGGACC AATTTACCC TGAACAACACT CAATGAGCTG	120
GAGAGGCTTT TTGACGAGAC CCACTATCCC GACGCCTTCA TGCGAGAGGA ACTGAGCCAG	180
CGACTGGGCC TGTCGGAGGC CCGAGTGCAG GTTTGGTTTCAAATCGAAG AGCTAAATGT	240
AGAAAACAAG AAAATCAACT CCATAAAGGT GTTTCATAG GGGCCGCCAG CCAAGTTGAA	300
GCTGTAGAG TCGCACCTTA TGCAACGTA GTTGCTTTAA GGATGCCATT TCAGCAGGTT	360
CAGGCGCAGC TGCAGCTGGA CAGCGCTGTG GCGCACGCGC ACCACCACCT GCATCCGCAC	420

```

CTGGCCGCGC ACGCGCCCTA CATGATGTTT CCAGCACCGC CCTTCGGACT GCCGCTCGCC      480
ACGCTGGCCG CGGATTCGGC TTCCGCCGCC TCGGTAGTGG CGGCCGCAGC AGCCGCCAAG      540
ACCACCAGCA AGGACTCCAG CATCGCCGAT CTCAGACTGA AAGCCAAAAA GCACGCCGCA      600
GCCCTGGGTC TGTGACVCCA ACGCCAGCAC CAATGTCGCG CCTGTCCCGC GGCACCTCAGC      660
CTGCASNCCC TNDDKANMCG TTRCTYHTCM ATTACACTTT GGGACCYCGG GDBAGVCCTT      720
TTNNAGACTT YVATKGGSCW CSCTGGBCCC TBRKGAUVAC TTGSGHYCGR GAACCGAKHT      780
GCCCABAYGA GGACCRGTTT GGAKDG                                           806

```

(2) Informace o sekvenci SEQ ID NO: 16:

(i) CHARAKTERISTIKA SEKvence:

- (A) DÉLKA: 190 aminokyselin
- (B) TYP: aminokyselina
- (C) DRUH ŘETĚZCE: jednořetězcová
- (D) TOPOLOGIE: lineární

(ii) TYP MOLEKULY: peptid

(xi) POPIS SEKvence: SEQ ID NO: 16:

```

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

```

15.04.99

PU 966-99

~~27943~~

Změněné nároky

P A T E N T O V É N Á R O K Y

1. Molekula nukleové kyseliny kódující polypeptidy, které obsahují oblast homeoboxu tvořenou 60-ti aminokyselinami a mají aminokyselinovou sekvenci SEQ ID NO: 1 a vykazují regulační aktivitu lidského růstu a dále obsahují nukleotidovou sekvenci SHOX ET93 (SEQ ID NO:2) a nukleotidovou sekvenci SHOX ET45 (SEQ ID NO: 4).

2. Molekula DNA podle nároku 1, která kóduje polypeptid, jehož délka je 150 až 350 aminokyselin.

3. Molekula DNA podle libovolného z nároků 1 nebo 2, která dále obsahuje nukleotidovou sekvenci SHOX G310 (SEQ ID NO: 3).

4. Molekula DNA podle libovolného z nároků 1 až 3, která dále obsahuje nukleotidovou sekvenci SHOX G108 (SEQ ID NO: 5).

5. Molekula DNA podle libovolného z nároků 1 až 4, která dále obsahuje nukleotidovou sekvenci SHOX Va (SEQ ID NO: 6) nebo SHOX Vb (SEQ ID NO: 7).

6. Molekula DNA podle libovolného z nároků 1 až 5, která kóduje polypeptid vybraný ze skupina, která zahrnuje:

- a) transkripční faktor A, který má v podstatě aminokyselinovou sekvenci (SEQ ID NO: 11),
- b) transkripční faktor B, který má v podstatě aminokyselinovou sekvenci (SEQ ID NO: 13).

7. Sekvence DNA podle libovolného z nároků 1 až 6, přičemž DNA je genomová nebo izolovaná DNA odpovědná za regulaci lidského růstu.

8. Sekvence DNA podle libovolného z nároků 1 až 6, přičemž DNA je cDNA.

9. cDNA podle nároku 8, která v podstatě obsahuje nukleotidovou sekvenci SHOXa (SEQ ID NO: 10) nebo SHOXb (SEQ ID NO: 12).

10. DNA podle nároku 7, která v podstatě obsahuje nukleotidovou sekvenci genu SHOX (SEQ ID NO: 14).

11. Lidský růstový protein, který má aminokyselinovou sekvenci danou v sekvenci (SEQ ID NO: 11) (transkripční faktor SHOXa), nebo jeho funkční fragment vykazující regulační aktivitu lidského růstu

12. Lidský růstový protein, který má aminokyselinovou sekvenci danou v sekvenci (SEQ ID NO: 13) (transkripční faktor SHOXb), nebo jeho funkční fragment vykazující regulační aktivitu lidského růstu.

13. Lidský růstový protein, který má aminokyselinovou sekvenci danou v sekvenci (SEQ ID NO: 16) (transkripční faktor SHOT), nebo jeho funkční fragment vykazující regulační aktivitu lidského růstu.

14. cDNA kódující protein, který má aminokyselinovou sekvenci danou v (SEQ ID NO: 11), (SEQ ID NO: 13) nebo (SEQ ID NO: 16).

15. Farmaceutická kompozice, v y z n a č u j í c í s e t í m, ž e obsahuje protein podle libovolného z nároků 11, 12 nebo 13.

15.04.99

16. Způsob léčby malé výšky postavy, v y z n a č u j í c í s e t í m, ž e zahrnuje aplikaci subjektu, je-li to nutné, terapeuticky účinné množství proteinu podle nároku 11, 12 nebo 13.

17. Použití proteinu podle nároku 11, 12 nebo 13 při přípravě farmaceutické kompozice vhodné pro léčbu malé výšky postavy.

18. Použití sekvence DNA podle libovolného z nároků 1 až 10 nebo 14 nebo jejího fragmentu, která vykazuje regulační aktivitu lidského růstu, při přípravě farmaceutické kompozice pro léčbu vad spojených s mutacemi genu odpovídajícího za malou výšku postavy.

19. Použití sekvence DNA podle libovolného z nároků 1 až 10 nebo 14 nebo jejího fragmentu, která vykazuje regulační aktivitu lidského růstu, při přípravě kitu pro identifikaci jedinců, kteří mají genetický defekt odpovědný za poškození lidského růstu.

20. Použití sekvence DNA podle libovolného z nároků 1 až 10 nebo 14 nebo jejího fragmentu, která vykazuje regulační aktivitu lidského růstu, při identifikaci genu odpovědného za malou výšku lidské postavy.

21. Způsob stanovení malé výšky postavy na základě molekul RNA nebo DNA, v y z n a č u j í c í s e t í m, ž e molekula biologického vzorku, který se zkoumá, se amplifikuje v přítomnosti dvou nukleotidových sond komplementárních s libovolnou ze sekvencí DNA uvedených v SEQ ID NO: 2 až SEQ ID NO: 7 a následně se stanoví pomocí vhodného detekčního systému.

22. Použití způsobu podle nároku 21 při identifikaci osob, které mají genetický defekt odpovědný za malou výšku postavy.

23. Transgenní zvíře nebo zvíře s deaktivovaným genem, v y z n a č u j í c í s e t í m, ž e je transformované genem odpovědným za malou výšku postavy, který obsahuje sekvenci DNA podle libovolného z nároků 1 až 10 nebo 14.

24. Buňky transformované sekvencí DNA podle libovolného z nároků 1 až 10 nebo 14.

25. Testovací systém, v y z n a č u j í c í s e t í m, ž e je vhodný pro identifikaci nebo testování činidel použitelných při léčbě malé výšky lidské postavy, který zahrnuje buňky podle nároku 24.

26. Způsob identifikace nebo testování kandidátů farmaceutických činidel použitelných při léčbě vad vztahujících se k mutacím genu odpovídajícího za malou výšku postavy, v y z n a č u j í c í s e t í m, ž e zahrnuje testovací systém podle nároku 25 a stanovení změn fenotypu uvedených buněk nebo změn produktů exprese uvedených buněk po té, co se uvedené buňky dostanou do kontaktu s uvedenými kandidáty farmaceutických činidel.

27. Expresivní vektor, v y z n a č u j í c í s e t í m, ž e obsahuje molekulu DNA podle nároků 1 až 10 nebo 14, která je schopna způsobit expresi kódovaného polypeptidu.

28. Způsob *in vivo* léčby genovou terapií vad růstu u lidí spojených alespoň s jednou mutací v genu SHOX nebo SHOT, v y z n a č u j í c í s e t í m, ž e zahrnuje zavedení expresivního plazmidu do lidských buněk, ve kterém molekula DNA podle libovolného z nároků 1 až 10 nebo 14 je začleněna ve

směru 3' konce polynukleotidu od expresivního promotoru, jenž způsobuje expresi v lidské hostitelské buňce.

29. Způsob podle nároku 28, v y z n a č u j í c í s e t í m, ž e j e vhodný pro léčbu Turnerova syndromu nebo malé výšky postavy.

30. Použití lidských růstových proteinů nebo jejich funkčních fragmentů, které mají regulační aktivitu lidského růstu, při přípravě léků pro léčbu pacientů, jenž vykazují genetickou mutaci genů lidského růstu SHOX (SEQ ID NO: 8) nebo SHOT (SEQ ID NO: 15) s tou podmínkou, že je vyloučena příprava léků pro léčbu pacientů trpících Turnerovým syndromem.

31. Použití lidských růstových proteinů nebo jejich funkčních fragmentů, které mají regulační aktivitu lidského růstu, při přípravě léků pro léčbu pacientů, jenž vykazují genetickou mutaci genů lidského růstu SHOX (SEQ ID NO: 8) nebo SHOT (SEQ ID NO: 15) a u nichž se použitím molekul DNA schopných hybridizovat se sekvencí DNA genů lidského růstu SHOX nebo SHOT zjistilo, že mají genetický defekt odpovědný za malou výšku postavy, s tou podmínkou, že je vyloučena příprava léků pro léčbu pacientů trpících Turnerovým syndromem.

32. Použití podle nároků 30 až 31, kde genetická mutace je způsobena horkým místem mutace v sekvenci DNA kódující bod zkrácení proteinu v pozici aminokyseliny 195 v lidském růstovém genu SHOX.

33. Použití podle libovolného z nároků 30 až 32, kde lidský růstový protein je lidský růstový hormon.

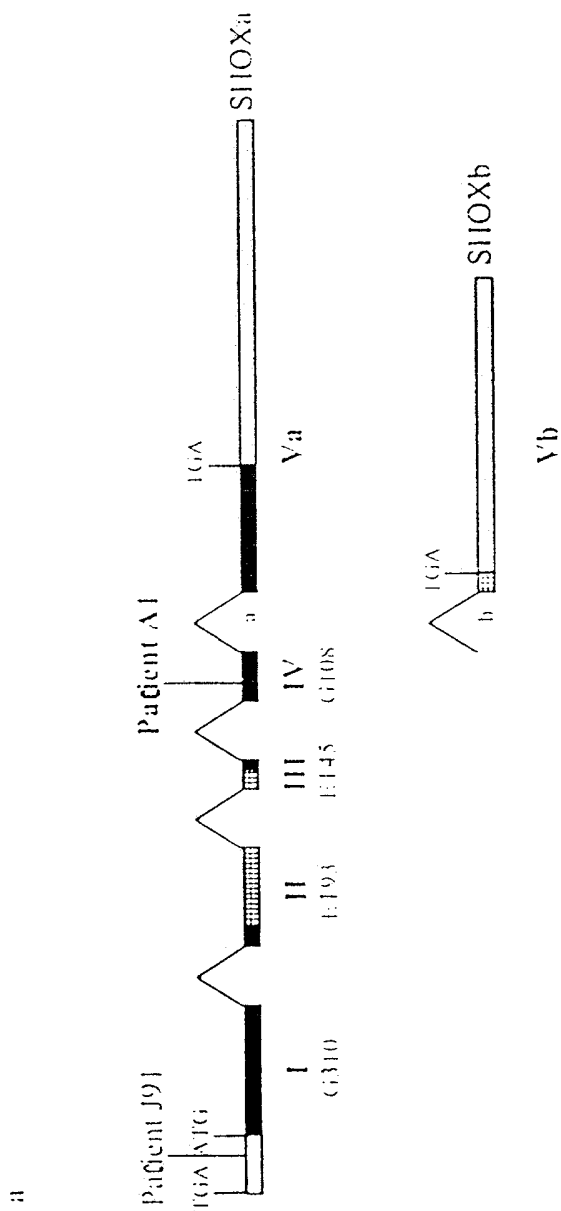
34. Způsob léčby pacientů vykazujících vady způsobené mutacemi lidských růstových genů SHOX (SEQ ID NO: 8) nebo SHOT



(SEQ ID NO: 15), v y z n a č u j í c í s e t í m, ž e zahrnuje aplikaci pacientovi, je-li to nutné, farmaceuticky aktivní množství lidského růstového proteinu s tou podmínkou, že se vyloučí léčba pacientů trpících Turnerovým syndromem.

35. Způsob léčby pacientů vykazujících vady způsobené mutacemi lidských růstových genů SHOX (SEQ ID NO: 8) nebo SHOT (SEQ ID NO: 15) a u nichž se použitím molekul DNA schopných hybridizovat se sekvencí DNA lidských růstových genů SHOX nebo SHOT zjistilo, že mají genetický defekt odpovědný za malou výšku postavy, v y z n a č u j í c í s e t í m, ž e zahrnuje aplikaci pacientovi, je-li to nutné, farmaceuticky aktivní množství lidského růstového proteinu s tou podmínkou, že je vyloučena léčba pacientů trpících Turnerovým syndromem.

Obrázek č. 1



1 GTGATCCACCCGCGCCACGGGCCGTCCTCTCCGCGCGGGGAGACGCGCGCATCCACCAG
 61 CCCCGGCTGCTCCCCAGCCCCGGCCCCAGCCATGGAAGAGCTCACGGCTTTTGTATCCAA
 : M E E L T A F V S K
 121 GTCTTTTGACCAGAAAGCAAGGACGGTAACGGCGGAGGCGGAGGCGGCGGAGGTAAGAA
 11 S F D Q K S K D G N G G G G G G G G G K K
 181 GGATTCCATTACGTACCGGGAAGTTTTGGAGAGCGGACTGGCGCGCTCCCGGGAGCTGGG
 31 D S I T Y R E V L E S G L A R S R E L G
 241 GACGTCGGATTCCAGCCTCCAGGACATCACGGAGGGCGGCGGCGGCGGACTGCCCGGTGCATTT
 51 T S D S S L Q D I T E G G G G H C P V H L
 301 GTTCAAGGACCACGTAGACAATGACAAGGAGAAACTGAAAGAATTCGGCACCGCGAGAGT
 71 F K D H V D N D K E K L K E F G T A R V
 361 GGCAGAAGGGATTATGAATGCAAAGAGAAGCGCGAGGACGTGAAGTCGGAGGACGAGGA
 91 A E G I Y E C K E K R E D V K S E D E D
 421 CGGGCAGACCAAGCTGAAACAGAGGGCGAGCCCGCACCAACTTCACGCTGGAGCAGCTGAA
 111 G Q T K L K Q R R S R T N F T L E Q L N
 481 CGAGCTCGAGCGACTTTTTGACGAGACCCATTACCCCGACGCCTTCATGCGCGAGGAGCT
 131 E L E R L F D E T H Y P D A F M R E E L
 541 CAGCCAGCGCTGGGGCTTTCCGAGGGCGCGCTGCAGGTTTGGTTCCAGAACC GGAGAGC
 151 S Q R L G L S E A R V Q V W F Q N R R A
 601 CAAGTGCCGCAAACAAGAGATCAGATGCATAAAGCGTCATCTTGGGCACAGCCAACCA
 171 K C R K Q E N Q M H K G V I L G T A N H
 661 CCTAGACGCCTGGCGAGTGGCACCCCTACGTCAACATGGGAGCCTTACGGATGCCTTTCCA
 191 L D A C R V A P Y V N M G A L R M P F Q
 721 ACAGGTCCAGGCTCAGCTGCAGCTGGAAGCGTGGCCCACGCGCACCCGACCTGCACCC
 211 Q V Q A Q L Q L E G V A H A H P H L H P
 781 GCACCTGGCGGGCAGCGGCCCTACCTGATGTTCCCCCGCCGCTTCCGGGCTGCCCAT
 231 H L A A H A P Y L M F P P P P F G L P I
 841 CGCGTCGCTGGCCGAGTCCGCCTCGGCCGCCCGCGTGGTCGCCGCCGCCGCAAAAAGCAA
 251 A L S A E S A S A A A V V A A A A K S N
 901 CAGCAAGAAATCCAGCATCGCCGACCTGCGGCTCAAGGCGCGGAAGCACGCGGAGGCCCT
 271 S K N S S I A D L R L K A R K H A E A L
 961 GGGGCTCTGACCCCGCGCGAGCCCCCGCGCGCCCGGACTCCCGGGCTCCCGCGCACCCC
 291 G L *
 1021 GCCTGCACCCGCGGCTCCTGCCTCAACCCCGCCTGGAGCTCCTTCCGCGGCCACCGTGCCT
 1081 CCGGGCACCCCGGGAGCTCCTGCAAGAGGCCCTGAGGAGGGAGGCTCCCGGGACCGTCCAC
 1141 GCACGACCCAGCCAGACCCCTCGCGGAGATGGTGCAGAAGGCGGAGCGGGTGAGCGGCCGT
 1201 GCGTCCAGCCCCGGCCTCTCCAAGGCTGCCCGTGCCTGCTGGGACCCTGGAGAAGGGTAA
 1261 ACCCCCGCCTGGCTGGCTCCTCTGCTATACCCTATGCATGCGGTTAATACACACGCT
 1321 TTGGAAGATCCTTAGAGTCTATTGAAACTGCAAAGATCCCGGAGCTGGTCTCCGATGAAA
 1381 ATGCCATTTCTTCTGTTGCCAACGATTTCTTTACTACCATGCTCCTTCTTCATCCCAG
 1441 AGGCTGCGGAACCGTGTGATTTGAATGTGGACTTCGGAATCCCAGGAGGCAGGGCCG
 1501 GGCTCTCCTCCACCGCTCCCCCGGAGCCTCCCAGGCAGCAATAAGGAAATAGTTCTCTGG
 1561 CTGAGGCTGAGGACGTGAACCGCGGGCTTTGGAAAGGGAGGGGAGGCCCGAACCTC
 1621 CCACGTTGGGACTCCCACGTTCCGGGGACCTGAATGAGGACCGACTTTATAACTTTTCCA
 1681 GTGTTTGATTCCCAAATGGGTCTGGTFTTGTFTTGGATTGGTATTTTTTTTTTTTTTTT
 1741 TTTTTGCTGTGTACAGGATTCAGACGCAAAGACTTGCATAAGAGACGGACGCGTGGTT
 1801 GCAAGGTGTCACTACTGATAAGCAGATTAACCTTACTGACATGGAGTGAAGTGAATATT
 1841 ATAAATATTATAGATTAAAAAAAAAAATAGC [A]₁

SHOXb

Obrázek č. 3

1 GTGATCCACCCGCGCGCACGGGCGCTCCTCTCCGCGCGGGGAGACGCGCGCATCCACCAG
 61 CCCCAGGCTGCTCGCCAGCCCCGGCCCCAGCCATGGAAGAGCTCACGGCTTTTGTATCCAA
 1 M E E L T A F V S K
 121 GTCTTTTGACCAGAAAAGCAAGGACGGTAACGGCGGAGGCGGAGGCGGGAGGTAAGAA
 11 S F D Q K S K D G N G G G G G G G G G G K K
 181 GGATTCATTACGTACCGGGAAGTTTTGGAGAGCGGACTGGCGCGCTCCCGGGAGCTGGG
 31 D S I T Y R E V L E S G L A R S R E L G
 241 GACGTCGGATTCCAGCCTCCAGGACATCACGGAGGGCGGGCGGCCACTGCCCGGTGCATTT
 51 T S D S S L Q D I T E G G G H C P V H L
 301 GTTCAAGGACCACGTAGACAATGACAAGGAGAACTGAAAGAATTCGGCACCGCGAGAGT
 71 F K D H V D N D K E K L K E F G T A R V
 361 GGCAGAAGGGATTTATGAATGCAAAGAGAAGCGCGAGGACGTGAAGTCGGAGGACGAGGA
 91 A E G I Y E C K E K R E D V K S E D E D
 421 CGGGCAGACCAAGCTGAAAAGAGAGGCGCAGCCGCACCAACTTCACGCTGGAGCAGCTGAA
 111 G Q T K L K Q R R S R T N F T L E Q L N
 481 CGAGCTCGAGCGACTTTTTGACGAGACCCATTACCCCGACGCCTTCATGCGCGAGGAGCT
 131 E L E R L F D E T H Y P D A F M R E E L
 541 CAGCCAGCGCCTGGGGCTTTCCGAGGCGCGCGTGCAGGTTTGGTTCCAGAACC GGAGAGC
 151 S Q R L G L S E A R V Q V W F Q N R R A
 601 CAAGTGCCGCAAACAAGAGAAATCAGATGCATAAAGGCGTCATCTTGGGCACAGCCAACCA
 171 K C R K Q E N Q M H K G V I L G T A N H
 561 CCTAGACGCCTGCCGAGTGGCACCCCTACGTCAACATGGGAGCCTTACGGATGCCTTTCCA
 191 L D A C R V A P Y V N M G A L R M P F Q
 721 ACAGATGGAGTTTTGCTCTTGCTCGCCCAGGCTGGAGTATAATGGCATGATCTCGACTCAC
 211 Q M E F C S C R P G W S I M A *
 781 TGCAACCTCCGCTCCCGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGG
 841 ATTACAGGTGCCCCACCACCATGTCAAGATAATGTTTGTATTTTTCAGTAGAGATGGGGTTT
 901 GACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCGCCTTAGCC
 961 TCCCAAAGTGCTGGGATGACAGGCGTGAGCCCCCTGCGCCCGGCCTTTGTAACTTTATTTT
 1021 TAATTTTTTTTTTTTTTAAAGAAAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGCACA
 1081 CTGGTGCGATCATACTCTGCTGCAGCCTCAAACCTCCTGGGCTCAAGCAATCCTCCACCT
 1141 CAGCCTCCTGAGTACTGCGACTACAGGCACCCACCACACCCAGCTAATTTTTTTTGA
 1201 TTTTACTAGAGACGGATCTTGCTTTGCTGCTGAGGCTGGTCTTGAGCTCCTGAGCTCC
 1261 AAAGATCCTCTCCCTCCACCTCCCAAAGTGTTAGAATTACAAGCATGAACCACTGCCCG
 1321 TGGTCTCCAAAAAAGGACTGTTACGTGG [A]

15.04.99

Obrázek č. 4

GTGTCCCCGGAGCTGAAAGATCGCAAAGAGGATGCGAAAGGGATGGAGGACGAAGGCCAG
M E D E G Q

ACCAAAATCAAGTACAGAGGCGAAGTCGGACCAATTTACCCCTGGAACAACCTCAATGAGCTG
T K I K ~~D E R H S R T N E L D E O N E R T~~

GAGAGGCTTTTTGACGAGACCCACTATCCCGACGCCTTCATGCGAGAGGAACTGAGCCAG
~~E R T D E D E T H Y P D Z E M R E E L S Q~~

CGACTGGGCCTGTCCGAGGCCCGAGTGCAGGTTTGGTTTCAAATCGAAGAGCTAAATGT
~~R N G C I S R E A H E V O V H E R O N E R A K O C~~

AGAAAACAAGAAAATCAACTCCATAAAGGTGTTCTCATAGGGGCCGCCAGCCAGTTTGAA
~~R K O E~~ N Q L H K G V L I G A A S Q F E

GCTTGTAGAGTCGCACCTTATGTCAACGTAGGTGCTTTAAGGATGCCATTTTCAGCAGGAT
A C R V A P Y V N V G A L R M P F Q Q D

AGTCATTGCAACGTGACGCCCTTGCCCTTTCAGGTTTCAGGCGCAGCTGCAGCTGGACAGC
S H C N V T P L P F Q V Q A Q L Q L D S

GCTGTGGCGCACGCGCACCACCACCTGCATCCGCACCTGGCCGCGCACGCGCCCTACATG
A V A H A H H H L H P H L A A H A P Y M

ATGTTCCACGACCCGCCCTTCGGACTGCCGCTCGCCACGCTGGCCGCGGATTTCGGCTTCC
M F P A P P F G L P L A T L A A D S A S

GCCGCTCGGTAGTGGCGGCCGAGCAGCCGCCAAGACCACCAGCAAGGACTCCAGCATC
A A S V V A A A A A A K T T S K D S S I

GCCGATCTCAGACTGAAAGCCAAAAAGCACGCCGAGCCCTGGGTCTGTGACGCCAACGC
A D L R L K A K K H A A A L G L *

CAGCACCAATGTCGCGCCTGTCCCGGGCACTCAGCCTGCACGCCCTCCGCGCCCGGCTG
CTTCTCCGTTACCCCTTGGACCTCGGGAGCCGGCCCTCTTCCCGCCTCACTGACCATC
CCTCGTCCCTATCGCATCTTGGACTCGGAAAGCCAGACTCCACGCAGGACCAGGGATCT
CACGAGGCACGCAGGCTCCGTGGCTCCTGCCGTTTTCTACTCGAGGGCCTAGAATTGG
GTTTTGTAGGAGCGGGTTTGGGGGAGTCTGGAGAGAGACTGGACAGGGTAGTGCTGGAAC
CGCGGAGTTTGGCTCACCGCAAAGCTACAACGATGGACTCTTGCATAGAAAAAAAATC
TTGTTAACAATGAAAAAATGAGCAAAACAAAAAATCGAAAGACAAACGGGAGAGAAAAAG
AGGAAGGCAACTTATTTCTTAACTGCTATTTGGCAGAAGCTGAAATTTGGAGAACCAAGGA
GCAAAAACAAATTTTAAAAATAAAGTATTTTATACATTTAAAAATATGGAAAAACAACC
AGACGATTCCTCGAGAGACTGGGGGGAGTTACCAACTTAAATGTGTGTTTTAAAAAATGCC
CTAAGAAGGCAAAGCAGAAAGAAAGAGGTATACTTATTTAAAAAACTAAGATGAAAAAAGT
GCGCAGGTGGGAAGTTCACAGGTTTGGAAACTGACCTTTTTCTGCGAAGTTCACGTTAAT
ACGAGAAATTTGATGAGAGAGGCGGGCCTCCTTTTACGTTGAATCAGATGCTTTGAGTTT
AAACCACCATGTATGGAAGAGCAAGAAAAAGAGAAAAATATTTAAACGAGGAGAGAGAAAA
ATAATGGCAAACTGTCTGGACTGCTGACAGTAAATTCGGTTTGCATGGAAAAAAA
AAAAA

Obrázek č. 5

Organizace exonu/intronu v lidském genu SHOX

Exon	cDNA-a	cDNA-b	Genom.DNA	Velikost exonu	Intron/Exon	Exon/Intron
I	UTR-368	UTR-368	1-368	368	GTGATCCACC	GTGGCAGAAAGgtaagttcct
II	369-577	369-577	3817-4025	209	ccccacgcagGGATTTATGA	GCGCGTCCAGgtaggaaccc
III	578-635	578-635	9851-9908	58	tctccccaaagTTTGGTTCC	ATGCATAAAGgtagggtgtcg
IV	636-724	636-724	10029-10117	89	ttggacacagGCGTCATCTT	TTTCCAACACAGgtagctcact
Va	725-1890	----	13364-14529	1166	gctccccgcagTCCAGGCTC	AAAAAATAGC
Vb	----	725-1349	27154-27778	625	tttttttttagATGGAGTTT	TGTTACGTGG
polyA	>1891	>1350	----	----		

Velikosti exonů se uvádějí v párech bází; sekvence exonu jsou označeny velkými písmeny; donorová a akceptorová místa sestřihu jsou podtržena. Genomové sekvence a sekvence cDNA jsou dostupné prostřednictvím GeneBank.