



УКРАЇНА

(19) UA (11) 41322 (13) C2

(51) 7 C12N9/20, C12N15/00, C12N15/55,
A01K67/027МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІОПИС
ДО ПАТЕНТУ НА ВИНАХІД

(54) ВИДІЛЕНА МОЛЕКУЛА ДНК, ЩО КОДУЄ БІОЛОГІЧНО ФУНКЦІОНАЛЬНІ СТИМУЛЬОВАНУ СОЛЯМИ ЖОВЧНИХ КИСЛОТ ЛІПАЗУ / КАРБОКСИЛ-СКЛАДНОЕФІРНУ ЛІПАЗУ (ССЖЛ/КСЛ), РЕПЛІКОВАНИЙ ЕКСПРЕСУЮЧИЙ ВЕКТОР, СПОСІБ ОДЕРЖАННЯ ССЖЛ/КСЛ ЛЮДИНИ, СПОСІБ ОДЕРЖАННЯ ТРАНСГЕННОГО ВІДМІННОГО ВІД ЛЮДИНИ ССАВЦЯ, ЗДАТНОГО ЕКСПРЕСУВАТИ ССЖЛ/КСЛ ЛЮДИНИ (ВАРІАНТИ), СПОСІБ ОДЕРЖАННЯ ДИТЯЧОГО ХАРЧУВАННЯ

(21) 94129155

(22) 09.06.1993

(24) 17.09.2001

(31) 9201809-2, 9201826-6, 9202088-2, 9300902-5

(32) 11.06.1992, 12.06.1992, 03.07.1992, 19.03.1993

(33) SE, SE, SE, SE

(86) PCT/SE93/00515, 09.06.1993

(46) 17.09.2001, Бюл. № 8, 2001 р.

(72) Бьюрселль Карл Гуннар, SE, Карлссон Петер
Нільс Івар, SE, Енербек Курт Свен Магнус, SE,
Ханссон Стіг Леннарт, SE, Лідберг Ульф ФредрікПонтус, SE, Нільссон Жанетт Анніка, SE, Тьорнелл
Ян Біргер Фредрік, SE

(73) АСТРА АКТИЕБОЛАГ, SE

(56) WO, 9118923 A1, 1991.

2. WO, 9115234 A1, 1991

(57) 1. Выделенная молекула ДНК, кодирующая
биологически функциональные стимулируемую
солями желчных кислот липазу/карбоксил-
сложноэфирную липазу (ССЖЛ/КСЛ), причем ука-
занная молекула ДНК включает (а) нуклеотидную
последовательность

GGATCCCTCG AACCCAGGAG TTCAAGACTG CAGTGAGCTA TGATTGTGCC ACTGCACTCT	60
AGCCTGGGTG ACAGAGACCC TGTCTCAAAA AAACAAACAA ACAAAAAACC TCTGTGGACT	120
CCGGGTGATA ATGACATGTC AATGTGGATT CATCAGGTGT TAACAGCTGT ACCCCCTGGT	180
GGGGATGTT GATAACGGGG GAGACTGGAG TGGGGCGAGG ACATACGGGA AATCTCTGTA	240
ATCTTCCTCT AATTTGCTG TGAACCTAAA GCTGCTCTAA AAATGTACAT AGATATAAAC	300
TGGGGCTTC CTTCCTCTCT GCCCTGCCCC AGCCCTCCCC CACCTCCTTC CTCCTCCTGC	360
TGCCTCCCCT CTGCCCTCCC CTTCCTCCT TAGCCACTGT AAATGACACT GCAGCAAAGG	420
TCTGAGGCAA ATGCCCTTGC CCTGGGGCGC CCCAGCCACC TGCAGGCCCC TTATTTCTCTG	480
TGGCCGAGCT CCTCCTCCCA CCTCCAGTC CTTCCTCCAG CCTCCTCTGC CCACTAGGCC	540
TCCTGAATTG CTGGCACCAG CTGTGGTCGA CAGACAGAGG GACAGACGTG GCTCTGCAGG	600
TCCACTCGGT CCTTGGCACC GGCCGAGGG GTGGCAGAAC GGGAGTGTGG TTGGTGTGGG	660
AAGCACAGGC CCCAGTGTCT CCTGGGGGAC TGTGGGTGG GAAGGCTCTG GCTGCCCTCA	720
CCCTGTTCCC ATCACTGCAG AGGGCTGTGC GGTGGCTGGA GCTGCCACTG AGTGTCTCGG	780
TGAGGGTGAC CTCACACTGG CTGAGCTTAA AGGCCCCATC TGAAGACTTT GTTCGTGGTG	840
TTCTTTCACT TCTCAGAGCC TTTCTGGCT CCAGGATTAA TACCTGTTCA CAGAAAATAC	900
GAGTCGCCTC CTCCTCCACA ACCTCACACG ACCTTCTCCC TTCCCTCCCG CTGGCCTCTT	960
TCCCTCCCCT TCTGTCACTC TGCTGGGCA TGCCCCAGGG CCTCGGCTGG GCCCTTTGTT	1020
TCCACAGGGA AACCTACATG GTTGGGCTAG ATGCCTCCGC ACCCCCCAC CCACACCCCC	1080
TGAGCCTCTA GTCCTCCCTC CCAGGACACA TCAGGCTGGA TGGTGACACT TCCACACCCT	1140

(19) UA (11) 41322 (13) C2

TGAGTGGGAC	TGCCCTGTGC	TGCTCTGGGA	TTCCGCACCCA	GCTTGGAATA	CCCCTCCAC	1200
GGGCCCCCAGG	AAAAGCTCGT	ACAGATAAAG	TCAGCCACAT	GAGTGGAGGG	CCTGCAGCAT	1250
GCTGCCCTTT	CTGTCCCAGA	AGTCACGTGC	TCGGTCCCCT	CTGAAGCCCC	TTTGGGGACC	1320
TAGGGGACAA	GCAGGGCATG	GAGACATGGA	GACAAAGTAT	GCCCTTTTCT	CTGACAGTGA	1386
CACCAAGCCC	TGTGAACAAA	CCAGAAGGCA	GGGCACTGTG	CACCCTGCCC	GGCCCCACCA	1440
TCCCCCTTAC	CACCCGCCAC	CTTGCCACCT	GCCTCTGCTC	CCAGGTAAGT	GGTAACCTGC	1500
ACAGGTGCAC	TGTGGGTTTG	GGGAAAACTG	GATCTCCCTG	CACCTGAGGG	GGTAGAGGGG	1550
AGGGAGTGCC	TGAGAGCTCA	TGAACAAGCA	TGTGACCTTG	GATCCAGCTC	CATAAATACC	1620
CGAGGCCCCAG	GGGGAGGGCC	ACCCAGAGGC	TG ATG CTC ACC ATG GGG CGC CTG Met Leu Thr Met Gly Arg Leu -23 -20			1673
CAA CTG GTT GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC Gln Leu Val Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala -15 -10 -5						1721
CGC AAG GTAAGAGCCC AGCAGAGGGG CAGGTCTCTG TGCTCTCTCG CTCAATCAGA Ala Lys 1						1777
TCTGGAAACT	TCGGGGCCAGG	CTGAGAAAGA	GCCCAGCAC	GCCCCGCAGC	AGATCCCGGG	1837
CACTCAGCCT	CATTTCTATG	GGGACAGGTG	CCAGGTAGAA	CACAGGATGC	CCAATTCCAT	1897
TTGAATTTC	GATAAACTGC	CAAGAACTGC	TGTGTAAGTA	TGTCCCATGC	AATATTTGAA	1957
ACAAATTTCT	ATGGGGCCGG	CGCAGTGGCT	CACACCTGCA	ATCCCACCAG	TTTGGGAGGC	2017
CGAGGTGGGT	GGATCACTTG	AGGTCAGGAG	TTGGAGACCA	GCCTGGCCAA	CATGGTGAAA	2077
CCCCGTCTCT	ACTAAAAATA	CAAATATTAA	TCGGGGCTGG	TGGTGGGTGC	CTGTAATCCC	2137
AGCTACTCGG	GAGGCTGAGG	CAGGAGAACC	GCTTGAAGCT	GGGAGGTGGA	GATTGCGGTG	2197
AGCTGAGATC	ACGCTACTGC	ACTCCAGCCT	GGGTGACAGG	GCGAGACTCT	GTCTCAAAAA	2257
ATAGAAAAAG	AAAAAATGA	AACATACTAA	AAAACAATTC	ACTGTTTACC	TGAAATTCAA	2317
ATGTAAGTGG	GCCTCTTGAA	TTTACATTTG	CTAATCCTGG	TGATTCCACC	TACCAACCTC	2377
TCTGTTGTTT	CCATTTTACA	GAAGGGGAAA	CGGGCCCCAGG	GGCAGGGAGT	GTGGAGAGCA	2437
GGCAGACGGG	TGGAGAGAAG	CAGGCAGGCA	GTTTGCCCTG	CATGGCACAG	CTGCTGCCTC	2497
CTATTCTCTG	GCAGGAAGCT	GAAAGCCGGG	CTACTCCACA	CCCCGGTCCG	GGTCCCTCCA	2557
GAAAGAGAGC	CGGCAGGCAG	GAGCTCTCTC	GAGGCATCCA	TAAATTCTAC	CCTCTCTGCC	2617
TGTGAAGGAG	AAGCCACAGA	AACCCCAAGC	CCCACAGGAA	GCCGGTGTGG	GTGCCCCGGC	2677
CAGTCCCTGC	CCCCAGCAGG	AGTCACACAG	GGGACCCCAG	ATCCCAACCA	CGCTGTTCTG	2737
CTGCTTGCGG	TGTCTCAGGC	CCTGGGGACT	CCTGTCTCCA	CCTCTGCTGC	CTGCTCTCCA	2797
CACTCCCTGG	CCCTGGGACC	GGGAGSTTTG	GGCAGTGGTC	TTGGGCTCCT	GACTCAAAGG	2857
AGAGGTCACC	TTCTTCTTGG	GCGAGCTCTT	CTTGGGGTGC	TGAGAGGCCT	TCGGCAGGTC	2917
ATCAGGACCC	CTCCCCATTT	CCCCACCCTG	AGGCCCTCTG	GCCAGTCTCA	ATTGCACAGG	2977
GATCAGCCCA	CTGGCACAAG	GAGACACAGA	TGCCTCGCAG	GGGATGCCCA	CGATGCCTGC	3037
ATGTGTTGCT	TCTGGTTTCT	TTCTTCCAGT	TCCAACCGCC	GCACTCTCCC	ACACCAGTGT	3097
GACAGGGGGC	CCATCACCTT	AGACTTCAGA	GGGCTGCTGG	GACCCTGGCT	GGGCTGGGG	3157

GTGTAGGGCC ACCCTGCCCT TCCCCACCTG GAACCTGGCA CAGGTGACAG CCAGCAAGCA	3217
ATGACCTGGT CCCACCATGC ACCACGGGAA GAGGGAGCTG CTGCCCCAAGA TGGACAGGAG	3277
GTGGCACTGG GGCAGACAGC TGCTTCTCAA CAGGGTGACT TCAAGCCCCA AAGCTGCCCA	3337
GCCTCAGTTC CGTCAGGGAC AGAGGGTGGA TGAGCACCAA CCTCCAGGCC CCTCCTGGGG	3397
GTGGACAGCT TGGTGCACAG AGGCCATTTT CATGGCACAG GGAAGCGTGG CGGGGGTGGG	3457
AGGTGTGGTC CCTAGGGGGT TCTTTACCAG CAGGGGGCTC AGGAACTGTG GGGACTTGGG	3517
CATGGGGCCA TCGACTTTGT GCCCAGCCAG CTAGGCCCTG TGCAGGGAGA TGGGAGGAGG	3577
GAAAAGCAGG CCCCACCCCT CAGAAAGGAG GAAGGTTGGT GTGAAACATC CCGGGTACAC	3637
TGAGCATTTG GTACACTCCT CCCGGGAGCT GGACAGGCCT CCCATGTGAT GSCAAACAGG	3697
CCGACAGGAG ACACGGCTGT TGCTCGTCTT CCACATGGGG AAATGAGGA TCGAGTCAA	3757
AGCTGGGCGG CCATAGCCAG AACCCAAACC TCCATCCAC CTCTTGGCCG GCTTCCCTAG	3817
TGGGAACACT GGTGGAACCA GTTTCCTCTA AGATTCTGGG AGCAGGACAC CCCCAGGGAT	3877
AAGGAGAGGA ACAGGAATCC TAAAGCCCTG AGCATTGCAG GGCAGGGGGT GCTGCCTGGG	3937
TCTCCTGTGC AGAGCTGTCC TGCTTTGAAG CTGTCTTTGC CTCTGGGCAC GCGGAGTCGG	3997
CTTGCCCTTG CCCCTCCGGA TTCAGGCCGA TGGGGCTTGA GCCCCCCTGA CCCTGCCCGT	4057
GTCTCCCTCG CAG CTG GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu	4105
GGC GTC AAT AAG AAG CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC Gly Val Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe	4154
AAG GGC ATC CCC TTC GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG Lys Gly Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln	4203
CCA CAT CCT GGC TGG CAA G GTGGGAGTGG GTGGTGCCGG ACTGGCCCTG Pro His Pro Gly Trp Gln	4251
CGGCGGGGCG GGTGAGGGCG GCTGCCTTCC TCATGCCAAC TCCTGCCACC TGCAG GG Gly	4308
ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC ACC ATC Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala Thr Ile	4356
ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC AAC ATT Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu Asn Ile	4404
TGG GTG CCC CAG GGC AGG AAG CAA G GTCTGCCTCC CCTCTACTCC Trp Val Pro Gln Gly Arg Lys Gln	4449
CCAAGGGACC CTCCCATGCA GCCACTGCCC CGGGTCTACT CCTGGCTTGA GTCTGGGGGC	4509
TGCAAGCTG AACTTCCATG AAATCCCACA GAGGCGGGGA GGGGAGCGCC CACTGCCGTT	4569
GCCCAGCCTG GGGCAGGGCA GCGCCTTGGA GCACCTCCCT GTCTTGCCCC CAGGCACCTG	4629
CTGCACAGGG ACAGGGGACC GGCTGGAGAC AGGGCCAGGC GGGGCGTCTG GGGTCACCAG	4689

CCGCTCCCC ATCTCAG TC TCC CGG GAC CTG CCC GTT ATG ATC TGG ATC	4738
Val Ser Arg Asp Leu Pro Val Met Ile Trp Ile	
95 100	
TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT GGG GCC AAC TTC CTC	4786
Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His Gly Ala Asn Phe Leu	
105 110 115 120	
AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC ACA CGC GGA AAC GTC	4834
Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn Val	
125 130 135	
ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC CTT GGG TTC CTC AGC	4882
Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser	
140 145 150	
ACT GGG GAC GCC AAT CTG CCA G GTGCGTGGGT GCCTTCGGCC CTGAGGTGGG	4934
Thr Gly Asp Ala Asn Leu Pro	
155	
GGCACCAGCA TGCTGAGCCC AGCAGGGAGA TTTTCCTCAG CACCCCTCAC CCCAAACAAC	4994
CAGTGGCGGT TCACAGAAAG ACCCGGAAGC TGGAGTAGAA TCATGAGATG CAGGAGGCCC	5054
TTGGTAGCTG TAGTAAATA AAAGATGCTG CAGAGGCCGG GAGAGATGGC TCACGCCTGT	5114
AATCCCAGCA CTTTAGGAGG CCCACACAGG TGGGTCACCT GAGCGCAGAA GTTCAAGACC	5174
AGCCTGAAAA TCACTGGGAG ACCCCCATCT CTACACAAAA ATTAATAATT AGCTGGGGAC	5234
TGGGCGCGGC GGCTCAGCTC TGTAATCCCA GCACGTTGGG AGCCCAAGGT GGGTAGATCA	5294
CCTGAGGTCA GGAGTTTGAG ACCAGCCTGA CTAATAATGGA GAAACCTCTT CTCTACTAAA	5354
AATACAAAAAT TAGCCAGGCG TGGTGGCGCT TGCCCTGTAAT CCCAGCTACT CGGGAGGCTG	5414
AGGCAGGAGA ATCGCTTGAA CTCAGGAGGC GGAGGTTGCG GTGAGCCGAG ATCATGCCAC	5474
TGCACTCCAG CCTGGAGAAC AAGAGTAAAA CTCTGTCTCA AAAAAAAAAA AAAAAAAAAA	5534
ATAGCCAGGC GTGGTATCTC ATGCCTCTGT CCTCAGCTAC CTGGGAGGCA GAGGTGGAAG	5594
GATCGCTTGA GCCCAGGGGT TCAAAGCTGC AGTGAGCCGT GGTCCGTGCCA CTGCACTCCA	5654
GCCTGGGCGA CAGAGTGAGG CCCCATCTCA AAAATAAGAG GCTGTGGGAC AGACAGACAG	5714
GCAGACAGGC TGAGGCTCAG AGAGAAACCA GGAGAGCAGA GCTGAGTGAG AGACAGAGAA	5774
CAATACCTTG AGGCAGAGAC AGCTGTGGAC ACAGAACTGG CAGGACACAG ACAGGAGGGA	5834
CTGGGGCAGG GGCAGGAGAG GTGCATGGGC CTGACCATCC TGCCCCCGAC AAACACCACC	5894
CCCTCCAGCA CCACACCAAC CCAACCTCCT GGGGACCCAC CCCATACAGC ACCGCACCCG	5954
ACTCAGCCTC CTGGGACCCA CCCACTCCAG CAACCAACGT GACCTAGTCT CCTGGGACCC	6014
ACCCCTCCA GCACCCTACC CGACCCAGCT TCTTAGGGAC CCACCATTTG CCAACTGGGC	6074
TCTGCCATGG CCCCACCTCT GTTGAGGGCA TTTCCACCCC ACCTATGCTG ATCTCCCCTC	6134
CTGGAGGCCA GGCCTGGGCC ACTGGTCTCT AGCACCCCTT CCCCTGCCCT GCCCCAG GT	6194
Gly 160	
AAC TAT GGC CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG	6242
Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg	
165 170 175	

AAT ATC GCG GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly 180 185 190	6290
GAG TCT GCT GGA GGT GCC AGC GTC TCT CTG CAG GTCTCGGGAT CCCTGTGGGG Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln 195 200	6343
ACGGCCTGCC CCACAGGTTG AGAGGAAGCT CAAACGGGAA GGGGAGGGTG GGAGGAGGAG	6403
CGTGAGAGCTG GGGCTGTGGT GCTGGGGTGT CCTTGTCCCA GCGTGCGGTG GGCAGAGTGG	6463
GGAGCGGCCT TGGTGACGGG ATTTCTGGGT CCCGTAG ACC CTC TCC CCC TAC AAC Thr Leu Ser Pro Tyr Asn 205	6518
AAG GGC CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser 210 215 220 225	6566
CCC TGG GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys 230 235	6608
GTAAACGGAG GAGGGCAGGG CTGGGCGGGG TGGGGGCTGT CCACATTTC GTTCTTTATC	6668
CTGGACCCCA TCCTTGCCCTT CAAATGGTTC TGAGCCCTGA GCTCCGSCCT CACCTACCTG	6728
CTGGCCTTGG TTCTGCCCCC AG GTG GCT GAG AAG GTG GGT TGC CCT GTG GGT Val Ala Glu Lys Val Gly Cys Pro Val Gly 240 245	6780
GAT GCC GCC AGG ATG GCC CAG TGT CTG AAG GTT ACT GAT CCC CGA GCC Asp Ala Ala Arg Met Ala Gln Cys Leu Lys Val Thr Asp Pro Arg Ala 250 255 260 265	6828
CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA GGC CTG GAG T GTGAGTAGCT Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala Gly Leu Glu 270 275	6878
GCTCGGGTTG GCCCATGGGG TCTCGAGGTG GGGGTTGAGG GGGGTACTGC CAGGGAGTAC	6938
TCCGGAGGAG AGAGGAAGGT GCCAGAGCTG CGGTCTTGTC CTGTCACCAA CTAGCTGGTG	6998
TCTCCCCTCG AAGGCCCCAG CTGTAAGGGA GAGGGGGTGC CGTTTCTTCT TTTTTTTTGA	7058
GATGGAGTCT CACTGTTGCC CAGGCTGGAG TGCAGTGTC AATCTCAGC TCACTGCAAC	7118
CTCCACCTCC TGGGTTCAAG TGATTCTCTG ACTCAACCTC CCATGTAGCT GGGACTACAG	7178
GCACATGCCA CCATGCCCAG ATAATTTTTT TGTGTGTTTA GTAGGGATGG AGTTTCATCG	7238
TGTTAGCTAG GATGATCTCG GTCTTGGGAC CTCATGATCT GCCACCTCG GCCTCCCAAA	7298
GTGCTGGAAT TACAGGCGTG AGCCACTGTG CCCGGCCCCCT TCTTTATTCT TATCTCCCAT	7358
GAGTTACAGA CTCCCCTTTG AGAAGCTGAT GAACATTTGG GGCCCCCTCC CCCACCTCAT	7418
GCATTCATAT GCAGTCATTT GCATATAATT TTAGGGAGAC TCATAGACCT CAGACCAAGA	7478
GCCTTTGTGC TAGATGACCG TTCATTCAAT CGTTCATTCA TTCAGCAAAC ATTTACTGAA	7538
CCGTAGCACT GGGGCCCAGC CTCCAGCTCC ACTATTCTGT ACCCCGGGAA GGCCTGGGGA	7598
CCCATTCCAC AAACACCTCT GCATGTCAGC CTTACCAGCT TGCTACGCTA AGGCTGTCCC	7658

TCACTCATTC TTCTATGGCA ACATGCCATG AAGCCAAGTC ATCTGCACGT TTACCTGACA	7718
TGAGCTCAAC TGCACGGGCT GGACAAGCCC AAACAAAGCA ACCCCCACGG CCCCCTAGTA	7778
AGCAAAACCT GCTGTGCTGG GCCCAGTGAC AGCCAGGCCC CGCCTGCCTC AGCAGCCACT	7838
GGGTCTCTTA GGGGCCCCGC CAGGGGTCTG GAGTACAATG CAGACCTCCC ACCATTTTTC	7898
GCTGATGGAC TGAACCCAG CCCTGAGAGA GGGAGCTCCT TCTCCATCAG TTCCCTCAGT	7958
GGCTTCTAAG TTTCTCTCTT CCTGCTTCAG GCCCAGCAAA GAGAGAGAGG AGAGGGAGGG	8018
GCTGCCGCTG AAGAGGACAG ATCTGGCCCT AGACAGTGAC TCTCAGCCTG GGGACGTGTG	8078
GCAGGGCCTG GAGACATCTG TGATTGTCAC AGCTGGGGAG GGGGTGCTCC TGGCACCTCG	8138
TGGGTGAGG CCGGGGATGC TCTAAACATC CTACAGGGCA CAGGATGCCC CTGATGGTGC	8198
AGAATCAACC CTGCCCCAAG TGTCCATAGA TCAGAGAAGG GAGGACATAG CCAATTCCAG	8258
CCCTGAGAGG CAAGGGGCGG CTCAGGGGAA ACTGGGAGGT ACAAGAACCT GCTAACCTGC	8318
TGGCTCTCCC ACCCAG AC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT	8366
Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro	
280 285	
GTC ATT GAT GGA GAC TTC ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC	8414
Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala	
290 295 300 305	
AAC GCC GCC GAC ATC GAC TAT ATA GCA GGC ACC AAC AAC ATG GAC GGC	8463
Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp Gly	
310 315 320	
CAC ATC TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC AAC AAG	8513
His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn Lys	
325 330 335	
AAA GTC ACG GA GTAAGCAGGG GGCACAGGAC TCAGGGGCGA CCGTGCGGG	8551
Lys Val Thr Glu	
340	
AGGGCCGCGG GGAAGCACT GGCAGGGGG CCAGCCTGGA GGAGGAAGGC ATTGAGTGGA	8621
GGACTGGGAG TGAGGAAGTT AGCACCGGTC GGGGTGAGTA TGCACACACC TTCCTGTTGG	8681
CACAGGCTGA GTGTCAGTGC CTACTTGATT CCCCCAG G GAG GAC TTC TAC AAG	8734
Glu Asp Phe Tyr Lys	
345	
CTG GTC AGT GAG TTC ACA ATC ACC AAG GGG CTC AGA GGC GCC AAG ACG	8782
Leu Val Ser Glu Phe Thr Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr	
350 355 360	
ACC TTT GAT GTC TAC ACC GAG TCC TGG GCC CAG GAC CCA TCC CAG GAG	8830
Thr Phe Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu	
365 370 375	
AAT AAG AAG AAG ACT GTG GTG GAC TTT GAG ACC GAT GTC CTC TTC CTG	8878
Asn Lys Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe Leu	
380 385 390	
GTG CCC ACC GAG ATT GCC CTA GCC CAG CAC AGA GCC AAT GCC AA	8922
Val Pro Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys	
395 400 405	
GTGAGGATCT GGGCAGCGGG TGGCTCCTGG GGGCCTTCCT GGGGTGCTGC ACCTTCCAGC	8982
CGAGGCCTCG CTGTGGGTGG CTCTCAGGTG TCTGGGTTGT CTGGGAAAGT GGTGCTTGAG	9042

TCCCCACCTG TGCTGCCTG ATCCACTTTG CTGAGGCCCTG GCAAGACTTG AGGGCCTCTT	9102
TTTACCTCCC AGCCTACAGG GCTTTACAAA CCCTATGATC CTCTGCCCTG CTCAGCCCTG	9162
CACCCCATGG TCCTTCCCAC TGGAGAGTTC TTGAGCTACC TTCCATCCCC CATGCTGTGT	9222
GCACTGAGAG AACACTGGAC AATAGTTTCT ATCCACTGAC TCTTATGGGC CTCAACTTTG	9282
CCCATAATTT CAGCCCACCA CCACATTAAA AATCTTCATG TAATAATAGC CAATTATAAT	9342
AAAAAATAAG GCCAGACACA GTAGCTCATG CCTGTAATCC CAGCACATTG GGAGGTCAAG	9402
GTGGGAGGAT CACTTGAGGT CAGGAGTCTG AGACTAGTCT GGCCAACATG GCAAAACCCC	9462
ATCTCTACTA AAAATACAAA AATTATCCAG GCATGGTGGT GCATGCCTAT AATCCTAGCT	9522
ACTCAGGAGG CTGAGGTAGC AGAATTGATT GACCCAGGGA GGTGGAGGTT GCAGTGAGCC	9582
GAGATTACGC CACTGCACTC CAGCAGGGGC AACACAGTGA GACTGTGTCT CGAATAAATA	9642
AGTAAATAAA TAATAAAAAT AAAAAATAAG TTAGGAATAC GAAAAGATA GGAAGATAAA	9702
AGTATACCTA GAAGTCTAGG ATGAAAGCTT TGCAGCAACT AAGCAGTACA TTTAGCTGTG	9762
AGCCTCCTTT CAGTCAAGGC AAAAAGGGAA ACAGTTGAGG GCCTATACCT TGTCCAATCT	9822
AATTGAAGAA TGCACATTCA CTTGGAGAGC AAAATATTTT TTGATACTGA ATTCTAGAAG	9882
GAAGGTGCCT CACAATGTTT TGTGGAGGTG AAGTATAAAT TCAGCTGAAA TTGTGGAACC	9942
CATGAATCCA TGAATTTGGT TCTCAGCTTT CCCTTCCCTG GGTGTAAGAA GCCCCATCTC	10002
TTCATGTGAA TTCCCCAGAC ACTTCCCTGC CCCTGCCCCG GGACCTCCCT CCAAGTCCGG	10062
TCTCTGGGCT GATCGGTCCC CAGTGAGCAC CCTGCCTACT TGGGTGGTCT CTCCCTCCA	10122
G G AGT GCC AAG ACC TAC GCC TAC CTG TTT TCC CAT CCC TCT CGG ATG Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser His Pro Ser Arg Met 410 415 420	10159
CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT GCA GAT GAC ATT CAG Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His Ala Asp Asp Ile Gln 425 430 435 440	10217
TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG GGC TAC CGG CCC CAA Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro Gln 445 450 455	10265
GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC TGG ACC AAC TTT GCC Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala 460 465 470	10313
AAA ACA GG GTAAGACGTG GGTGAGTGC AGGGCGGAGG GCCACAGCCG Lys Thr Gly 475	10361
AGAAGGGCCT CCCACCACGA GGCCTTGTTT CCTCATTTGC CAGTGGAGGG ACTTTGGGCA	10421
AGTCACTTAA CCTCCCCCTG CATCGGAATC CATGTGTGTT TGAGGATGAG AGTTACTGGC	10481
AGAGCCCCAA GCCCATGCAC GTGCACAGCC AGTGCCAGT ATGCAGTGAG GGGCATGGTG	10541
CCCAGGGCCA GCTCAGAGGG CGGGGATGGC TCAGGCGTGC AGGTGGAGAG CAGGGCTTCA	10601
GCCCCCTGGG AGTCCCCAGC CCCTGCACAG CCTCTTCTCA CTCTGCAG G GAC CCC Asp Pro	10656

AAC ATG GGC GAC TCG GCT GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr 480 485 490	10704
GAA AAC AGC GGC TAC CTG GAG ATC ACC AAG AAG ATG GGC AGC AGC TEC Glu Asn Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser 495 500 505	10752
ATG AAG CGG AGC CTG AGA ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC Met Lys Arg Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr 510 515 520 525	10800
TAT CTG GCG CTG CCC ACA GTG ACC GAC CAG GAG GCC ACC CCT GTG CCC Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln Glu Ala Thr Pro Val Pro 530 535 540	10848
CCC ACA GGG GAC TCC GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser 545 550 555	10895
GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val 560 565 570	10944
CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp 575 580 585	10992
TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro 590 595 600 605	11040
GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly 610 615 620	11088
GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro 625 630 635	11136
CCC GTG CCG CCC ACG GGT GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG Pro Val Pro Pro Thr Gly Asp Ala Gly Pro Pro Pro Val Pro Pro Thr 640 645 650	11184
GGT GAC TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala 655 660 665	11232
CCC CCC GTG ACC CCC ACG GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC Pro Pro Val Thr Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro 670 675 680 685	11280
ACG GGT GAC TCC GGG GCC CCC CCT GTG CCC CCC ACG GGT GAC TCT GAG Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Glu 690 695 700	11328
GCT GCC CCT GTG CCC CCC ACA GAT GAC TCC AAG GAA GCT CAG ATG CCT Ala Ala Pro Val Pro Pro Thr Asp Asp Ser Lys Glu Ala Gln Met Pro 705 710 715	11376
GCA GTC ATT AGG TTT TAGCGTCCCA TGAGCCTTGG TATCAAGAGG CCACAAGAGT Ala Val Ile Arg Phe 720	11431
GGGACCCCAG GGGCTCCCCT CCCATCTTGA GCTCTTCCTG AATAAAGCCT CATACCCCTG	11491
TCGGTGTCTT TCTTTGCTCC CAAGGCTAAG CTGCAGGATC	11531,

обозначенную как SEQ ID NO: 1, или (б) нуклеотидную последовательность, которая гибридизуется в жестких условиях гибридизации с комплементом последовательности ДНК из (а), или (в) нуклеотидную последовательность, которая является генетически вырожденной нуклеотидной последовательностью из (а).

2. Молекула ДНК по п. 1, **отличающаяся** тем, что представляет собой последовательность, обозначенную как SEQ ID NO: 1.

3. Реплицируемый экспрессирующий вектор, несущий и способный опосредовать экспрессию молекулы ДНК, охарактеризованной в любом из пп. 1, 2, где вектор представляет собой pS452 (DSM 7499).

4. Вектор по п. 3, **отличающийся** тем, что включает молекулу ДНК, охарактеризованную в любом из пп. 1, 2, способную кодировать биологически функциональные ССЖЛ/КСЛ, и регуляторные элементы гена молочного белка, способные управлять экспрессией в молочной железе отличного от человека млекопитающего.

5. Способ получения ССЖЛ/КСЛ человека, **отличающийся** тем, что включает

а) вставку молекулы ДНК, охарактеризованной в пп. 1-2, в вектор, способный реплицироваться в конкретной клетке-хозяине;

б) введение полученного рекомбинантного вектора в клетку-хозяина;

в) культивирование полученной клетки в или на культуральной среде с целью экспрессии полипептида; и

г) выделение полипептида.

6. Способ получения трансгенного отличного от человека млекопитающего, способного экспрессировать ССЖЛ/КСЛ человека, **отличающийся** тем, что включает

а) введение вектора, охарактеризованного в п. 4, в оплодотворенную яйцеклетку или клетку эмбриона отличного от человека млекопитающего с тем, чтобы ввести указанный вектор в зародышевую линию млекопитающего, и

б) развитие полученной введенной оплодотворенной яйцеклетки или клетки эмбриона во взрослую самку отличного от человека млекопитающего.

7. Способ получения трансгенного отличного от человека млекопитающего, способного экспрессировать ССЖЛ/КСЛ человека и, по существу, неспособного экспрессировать ССЖЛ/КСЛ самого млекопитающего, **отличающийся** тем, что включает

а) устранение способности млекопитающего экспрессировать ССЖЛ/КСЛ с тем, чтобы, по существу, не происходила экспрессия ССЖЛ/КСЛ млекопитающего, и введение вектора, охарактеризованного в п. 4, в зародышевую линию млекопитающего таким образом, чтобы в организме млекопитающего происходила экспрессия ССЖЛ/КСЛ человека; и/или

б) замену гена ССЖЛ/КСЛ млекопитающего или его части вектором, охарактеризованным в п. 4.

8. Способ получения детского питания путем введения в состав детского питания полипептида, кодируемого молекулой ДНК, охарактеризованной в одном из пп. 1 и 2.

9. Молекула ДНК по п. 1, **отличающаяся** тем, что указанная молекула ДНК применима для получения ССЖЛ/КСЛ человека.

10. Молекула ДНК по п. 9, **отличающаяся** тем, что указанная молекула ДНК применима для получения трансгенного отличного от человека млекопитающего, экспрессирующего ССЖЛ/КСЛ человека.

11. Молекула ДНК по п. 9, **отличающаяся** тем, что указанная молекула ДНК применима для получения молока, содержащего ССЖЛ/КСЛ человека, полученного от трансгенного отличного от человека млекопитающего.

12. Молекула ДНК по п. 9, **отличающаяся** тем, что указанная молекула ДНК применима для получения детского питания, содержащего молоко, содержащее ССЖЛ/КСЛ человека, полученное от трансгенного отличного от человека млекопитающего.

13. Молекула ДНК по п. 1 или 2, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения патологического состояния, связанного с эндокринной недостаточностью поджелудочной железы.

14. Молекула ДНК по п. 13, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения муковисцидоза.

15. Молекула ДНК по п. 13, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения хронического панкреатита.

16. Молекула ДНК по п. 13, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения нарушения всасывания жиров.

17. Молекула ДНК по п. 13, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения нарушения всасывания жирорастворимых витаминов.

18. Молекула ДНК по п. 13, **отличающаяся** тем, что указанная молекула ДНК применима в производстве лекарственного препарата для лечения нарушения всасывания жиров, развивающегося в силу физиологических факторов.

Настоящее изобретение относится к молекуле ДНК, содержащей интронные последовательности и кодирующей человеческий белок, который в зависимости от места действия называется стимулируемой солями желчных кислот липазой (ССЖЛ) или карбоксил-сложноэфирной липазой (КСЛ). ДНК молекула с успехом применима в получении рекомбинантных ССЖЛ/КСЛ человека, предпочти-

тельно посредством продукции в трансгенных млекопитающих (но не человеке). Рекомбинантные ССЖЛ/КСЛ человека могут быть использованы в качестве компонента детского питания в виде заменителя материнского молока, предназначенного для кормления младенцев, или для изготовления лекарственных средств против, например: малаб-

сорбции жиров, муковисцидоза и хронического панкреатита.

Гидролиз пищевых липидов

Пищевые липиды являются важным источником энергии. Богатые энергией триацилглицерины составляют более 95% таких липидов. Некоторые липиды, например: определенные жирные кислоты и растворимые в жирах витамины являются существенными компонентами пищи. Перед абсорбцией желудочно-кишечным трактом триацилглицерины, а также присутствующие в меньших количествах компоненты, например этерифицированные растворимые в жирах витамины и холестерин, и диацилфосфатидилглицерины должны подвергнуться гидролизу по сложноэфирной связи с образованием менее гидрофобных абсорбируемых продуктов. Реакция гидролиза катализируется особой группой ферментов, называемых липазами.

В организме взрослого человека к важнейшим липамам относятся желудочная липаза, панкреатическая колипаза - зависимая липаза (гидролиз три- и диацилглицеринов), панкреатическая фосфолипаза A2 (гидролиз диацилфосфатидилглицеринов) и карбоксил-сложноэфирная липаза (КСЛ) (гидролиз сложных эфиров холестерина и растворимых в жирах витаминов). У вскармливаемых грудью новорожденных стимулируемая солями желчных кислот липаза (ССЖЛ) играет существенную роль в гидролизе некоторых из вышеупомянутых липидов. Совместно с солями желчных кислот продукты гидролиза липидов образуют смешанные мицеллы, из которых и происходит абсорбция.

Стимулируемая солями желчных кислот липаза

В лактирующей молочной железе человека происходит синтез и секреция с молоком стимулируемой солями желчных кислот липазы (ССЖЛ) (Bläckberg и др., 1987), которая после особой активации первичными солями желчных кислот придает вскармливаемому грудью младенцу эндогенную способность усваивать в кишечнике жиры. Указанный фермент, составляющий примерно 1% всего молочного белка (Bläckberg & Hernell, 1981), не разрушается при прохождении с молоком через желудок, а от содержащихся в двенадцатиперстной кишке панкреатических протеаз, например: трипсина и химотрипсина липаза защищена от дезактивации солями желчных кислот. И тем не менее липаза дезактивируется при пастеризации молока, например, при нагревании при 62,5°C, 30 мин (Björkstén и др., 1980).

Опыты на моделях *in vitro* предполагают, что конечные продукты гидролиза триацилглицерина отличаются в присутствии ССЖЛ (Bernbäck и др., 1990; Hernell & Bläckberg, 1982). Вследствие низкой внутрипросветной концентрации солей желчных кислот в течение неонатального периода, это может иметь важное значение для абсорбции продукта.

Карбоксил-сложноэфирная липаза

Карбоксил-сложноэфирная липаза (КСЛ) человеческого панкреатического сока (Lombardo и др.) функционально, видимо, идентична или по меньшей мере очень схожа, с ССЖЛ (Bläckberg и др., 1981). Кроме того, обе липазы характеризуются

общими эпитопами, имеют идентичные N-концевые последовательности аминокислот (Abouaki и др., 1988) и ингибируются ингибиторами серин-эстераз, например эсерином и диизопропилфторфосфатом. В недавних исследованиях, проведенных в нескольких лабораториях, выявлено строение кДНК как молочной липазы, так и панкреатической липазы (Baba и др. 1991; Hui и др., 1991; Nilsson и др., 1990; Reue и др., 1991), и исследователи пришли к выводу, что и молочный фермент, и панкреатический фермент являются продуктами одного и того же гена (в данном описании называется CEL-геном, ЕС 3.1.1.1). Последовательность кДНК и определенная на ее основе аминокислотная последовательность описаны в WO 91/15234 (Оклахома Медикал Рисерч Фаундейшн) и в WO 91/18923 (Актизболагет Астра).

Таким образом, предполагается, что КСЛ идентичен ССЖЛ, и полипептид, кодируемый CEL-геном, в контексте настоящего описания называется ССЖЛ/КСЛ.

Малабсорбция липидов

Общие причины малабсорбции липидов, а следовательно нарушенного питания заключается в пониженном внутрипросветного содержания панкреатической колипаза-зависимой липазы и/или солей желчных кислот. Типичные примеры случаев такого дефицита липазы включают больных, страдающих муковисцидозом - обычным генетическим нарушением, сопровождающих 80% больных с таким дефицитом на протяжении всей их жизни, и хроническим панкреатитом, часто связанным с хроническим алкоголизмом.

Лечение больных, страдающих дефицитом панкреатической липазы, в настоящее время заключается во введении очень больших доз сырого препарата панкреатических ферментов свиньи. Однако, колипаза-зависимая панкреатическая липаза дезактивируется при низких значениях pH, преобладающих в желудке. Этот феномен не может быть полностью преодолен применением больших доз фермента. Таким образом, вводимые в больших дозах препараты не приносят облегчение большинству больных, а сами препараты загрязнены и неприятны на вкус.

Приготовлены таблетки, способные пройти через кислотные области желудка с выделением фермента только в сравнительно щелочном окружении тощей кишки. Однако, многие больные, страдающие панкреатическими нарушениями, отличаются необычайно высоким содержанием кислоты в тощей кишке, и в этих случаях фермент из таблеток может и не выделиться.

Более того, поскольку выпускаемые в настоящее время промышленностью препараты животного происхождения, имеется риск иммунных реакций, могущих нанести вред больным или уменьшить эффективность лечения. Еще один недостаток существующих препаратов заключается в том, что для них не установлено проявлений других форм липолитической активности, отличной от активности колипаза-зависимой липазы. В самом деле, большинство таких препаратов в очень незначительной степени проявляют ССЖЛ/КСЛ активность. В этом может быть одна из причин того, что страдающие муковисцидозом больные, несмотря на поддерживающую терапию,

испытывают недостаток в растворимых в жирах витаминах и в необходимых жирных кислотах.

Таким образом, существует большая необходимость в продуктах со свойствами и строением, аналогичными свойствам и строению человеческих липаз, и широкой специфичностью субстрата, при этом такие продукты могут быть введены перорально больным, страдающим дефицитом одного или нескольких панкреатических липолитических ферментов. Продукты, которые могут быть приготовлены использованием настоящего изобретения, отвечают такой необходимости, как сами по себе, так и в комбинации с препаратами, содержащими другие липазы.

Детское питание

Хорошо известно, что вскармливание младенца материнским молоком считается более полезным, чем искусственное кормление. И не только потому, что молоко человека является хорошо сбалансированным источником питательных веществ, но еще и потому, что молоко матери легко усваивается младенцем. Так, некоторые биологически активные вещества, которые, как известно, выполняют физиологические функции в организме младенца, являются либо компонентами человеческого молока, либо образуются в ходе его усвоения, включая компоненты, участвующие в защите от инфекции, и компоненты, облегчающие усвоение питательных веществ человеческого молока.

Несмотря на серьезные попытки в направлении получения детского питания, так и не удалось создать состав, хоть в какой-то заметной степени отвечающий преимуществам материнского молока. Так, детское питание, часто приготовляемое на основе коровьего молока, как правило, не полностью усваивается младенцем, и в нем отсутствуют вещества, которые, как известно, оказывают влияние на физиологические функции организма младенца. С целью получения детского питания, аналогичного по своей питательной ценности материнскому, в состав питания предлагались разнообразные добавки, в том числе: белковые фрагменты, витамины, минеральные вещества и т.д. с последующим риском повышенной нагрузки на важнейшие органы, например печень и почки с их возможным длительным разрушением. Другой недостаток, связанный с применением составов на основе коровьего молока, заключается в повышенном риске возникновения у младенца аллергии к коровьим белкам.

В качестве альтернативы детскому питанию на основе коровьего молока используется человеческое молоко, которое можно получить в так называемых молочных банках. Однако, кормление новорожденных младенцев молоком из молочных банков в последние годы в значительной степени стараются избежать из боязни присутствия в молоке инфекционных агентов, например: ВИЧ или ВОВ (вирус огуречной мозаики). Для разрушения инфекционных агентов в человеческом молоке перед употреблением необходимо пастеризовать. Однако, при пастеризации питательная ценность и биологическое действие компонентов молока уменьшается, например, как указано выше, происходит дезактивация ССЖЛ.

Введение в детское питание липаз

Функционирование поджелудочной железы и печени при рождении проявляется не в полной мере, что особенно заметно в случае недоношенного ребенка. По физиологическим причинам малабсорбция жиров является наиболее общим нарушением, которое, как полагают, возникает в результате низкой внутрипросветной концентрации панкреатической колипаза-зависимой липазы и солей желчных кислот. Однако, вследствие присутствия ССЖЛ такая малабсорбция отмечается гораздо реже в случаях вскармливания грудью младенцев по сравнению с младенцами, вскармливаемыми пастеризованным человеческим молоком или детским питанием (Bernbäck и др., 1990).

Чтобы обойти вышеуказанные недостатки, связанные с применением пастеризованного молока и детского питания на основе коровьего молока, желательно, таким образом, создать такое детское питание, состав которого был бы особенно близок составу материнского молока, то есть состав, содержащий белки материнского молока.

ССЖЛ/КСЛ обладают несколькими уникальными свойствами, делающими их идеально применимыми в качестве добавки к детскому питанию:

Самой природой эти липазы предназначены для перорального введения. Так липазы устойчивы к прохождению через желудок, и активируются под действием содержимого тонкой кишки.

Вследствие своей широкой специфичности к субстрату липазы со своей стороны обладают потенциалом, способствующим полному усвоению большинства пищевых липидов, в том числе и растворимых в жирах сложных эфиров витаминов.

ССЖЛ/КСЛ могут превосходить панкреатическую колипаза-зависимую липазу по способности гидролизовать сложноэфирные связи в эфирах, содержащих длинноцепочечные полиненасыщенные жирные кислоты.

В присутствии желудочной липазы и в отсутствие или при низком содержании колипаза-зависимой липазы ССЖЛ/КСД способны обеспечить полный гидролиз триацилглицерина *in vitro* даже при низкой концентрации солей желчных кислот, как, например, в случае новорожденных младенцев. В присутствии ССЖЛ/КСЛ конечными продуктами гидролиза триацилглицерина являются свободные жирные кислоты и свободный глицерин, а не моноацилглицерин, образующийся в присутствии двух других липаз (Bernbäck и др., 1990). Это может благоприятствовать абсорбции продукта, особенно при низком внутрипросветном содержании солей желчных кислот.

Для применения ССЖЛ/КСЛ в качестве добавки к детскому питанию необходимо, однако, иметь доступ к большим количествам этого продукта. Хотя белки человеческого молока могут быть выделены непосредственно из молока, такой подход не является реалистичным и достаточно экономичным для получения больших количеств продуктов, необходимых для широкомасштабного производства детского питания. В связи с этим необходимо создать другие методы для того, чтобы иметь возможность получать детское питание, содержащее белки человеческого молока. Настоящим изо-

бретением даются такие способы получения в больших количествах ССЖЛ/КСЛ.

Продуцирование белков с молоком трансгенных животных

Выделение генов, кодирующих фармакологически активные белки, позволяет получать такие белки более дешевым путем в гетерологичных системах. Приемлемая экспрессионная система для молочных белков может быть представлена трансгенным животным (обзор см. Hennighausen и др., 1990). Пищевые составы, содержащие активированные солями желчных кислот липазы, полученные, например, технологией с использованием трансгенных животных, описаны в EP 317,355 (Оклахома Медикал Рисерч Фаундейшн).

В случае трансгенного животного кодирующая белок последовательность может быть введена в виде кДНК или в виде геномной последовательности. Поскольку для регулируемой экспрессии гена могут потребоваться интроны (Brinster и др., 1988; Whitelaw и др., 1991), во многих случаях рекомендуется применять геномную форму, а не кДНК форму структурного гена. В патенте WO 90/05188 (ФармасьютикалПротеинс Лимитид) описано применение в трансгенном животном кодирующей белок ДНК, включающей хотя бы один, но не все интроны, встречающиеся в природном кодирующем белок гене.

Целью настоящего изобретения является создание средств для получения с высоким выходом и по приемлемой цене человеческих рекомбинантных ССЖЛ/КСЛ, предназначенных для применения в детском питании с тем, чтобы избежать недостатков, характерных для пастеризованного молока и составов на основе коровьих белков.

Цель изобретения достигнута клонированием и секвенированием человеческого гена CEL. Для повышения выхода ССЖЛ/КСЛ при продуцировании человеческих ССЖЛ/КСЛ в трансгенном млекопитающем (но не человеке) вместо известной последовательности кДНК применена полученная молекула ДНК, содержащая интронные последовательности.

Соответственно, в одном из своих аспектов настоящее изобретение относится к молекуле ДНК, приведенной в списке последовательностей как SEQ ID NO: 1, или аналог указанной молекулы ДНК, который гибридизуется с комплементом молекулы ДНК, приведенной в списке последовательностей как SEQ ID NO: 1, или ее конкретной частью, в жестких условиях гибридизации.

Методика, применяемая для выделения молекулы ДНК человеческих ССЖЛ/КСЛ, приведена в нижеследующих примерах.

Жесткие условия гибридизации, упомянутые выше, следует понимать в их обычном значении, то есть как гибридизацию, проводимую согласно стандартному лабораторному руководству, например, Sambrook и др. (1989).

Другим аспектом настоящего изобретения является экспрессирующая система млекопитающего, включающая последовательность ДНК, кодирующую человеческие ССЖЛ/КСЛ, встроенную в ген, кодирующий молочный белок отличного от человека млекопитающего, с образованием, в результате, гибридного гена, способного экспресси-

роваться в молочной железе взрослой самки млекопитающего, принявшей гибридный ген, с последующим продуцированием человеческих ССЖЛ/КСЛ при экспрессировании гибридного гена.

И в еще одном своем аспекте настоящее изобретение относится к способу создания трансгенного отличного от человека млекопитающего, способного экспрессировать ССЖЛ/КСЛ человека. Способ состоит с инъектировании вышеохарактеризованной экспрессионной системы млекопитающего в оплодотворенное яйцо или клетку эмбриона млекопитающего с введением в результате экспрессирующей системы в зародышевую линию млекопитающего с последующим развитием инъектированного оплодотворенного яйца или эмбриона во взрослую самку млекопитающего.

Молекула ДНК, показанная в перечне последовательностей, как ПОСЛЕДОВАТЕЛЬНОСТЬ № 1, общей длины в 11531 п. о. имеет следующие признаки:

Признак	От основания	До основания
Фланкирующая 5'-область	1	1640
Рамка Хогнесса	1611	1617
Экзон 1	1641	1727
Начало трансляции	1653	1653
Экзон 2	4071	4221
Экзон 3	4307	4429
Экзон 4	4707	4904
Экзон 5	6193	6323
Экзон 6	6501	6608
Экзон 7	6751	6868
Экзон 8	8335	8521
Экзон 9	8719	8922
Экзон 10	10124	10321
Экзон 11	10650	11490
Фланкирующая 3'-область	11491	11531

В контексте настоящего изобретения термин "ген" применяется для обозначения ДНК последовательности, участвующей в продуцировании полипептидной цепи и включающей области до и после области кодирования (5'-восходящая и 3'-нисходящая последовательности), а также промежуточные последовательности (так называемые интроны), расположенные между отдельными кодирующими сегментами (так называемые экзоны) или в 5'-восходящей или 3'-нисходящей области. Восходящая 5'-область включает регуляторную последовательность, контролирующую экспрессию гена, как правило, промотор. Нисходящая 3'-область включает последовательности, участвующие в прекращении транскрипции гена, и возможно последовательности, ответственные за полиаденилирование транскрипта и нетранслируемой 3'-области.

Молекулы ДНК изобретения, о которых идет здесь речь, могут представлять собой естественные или синтетические ДНК последовательности, причем источником естественной последовательности обычно является непосредственно геномная ДНК, как правило, млекопитающего, например, описанная ниже. Синтетическая последовательность может быть получена обычными способами синтеза молекулы ДНК. Кроме того, ДНК по-

следовательность может быть смешанного геномного и синтетического происхождения.

В еще одном своем аспекте настоящее изобретение относится к реплицируемому вектору экспрессии, который осуществляет и который способен поддерживать экспрессию ДНК последовательности, кодирующей человека ССЖЛ/КСЛ.

В контексте настоящего изобретения термин "реплицируемый" означает, что вектор способен реплицироваться в клетке-хозяине данного типа, в которую был введен. Непосредственно в восходящем направлении от ДНК последовательности человеческих ССЖЛ/КСЛ может быть предусмотрена последовательность, кодирующая сигнальный пептид, присутствие которого гарантирует секрецию человеческих ССЖЛ/КСЛ, экспрессируемых клетками-хозяевами, принявших вектор. Сигнальная последовательность может быть последовательностью, связанной в природе с ДНК последовательностью человеческих ССЖЛ/КСЛ, или иметь иной источник происхождения.

Вектор может быть представлен любым вектором, который может быть использован в методах биотехнологии, и выбор вектора часто зависит от клетки-хозяина, в которую вектор будет вводиться. Так вектор может быть представлен автономно реплицирующимся вектором, то есть вектором, существующим в виде внехромосомной частицы, репликация которой не зависит от хромосомной репликации. Примеры подобных векторов включают: плазмиды, фаги, космиды, минихромосомы или вирусы. Или же вектор может быть таким, который при введении в клетку-хозяина интегрируется в геном клетки-хозяина и реплицируется вместе с хромосомой (-ами), с которой был интегрирован. Примеры приемлемых векторов включают бактериальные вектора экспрессии и дрожжевые вектора экспрессии. Вектор изобретения может нести любую молекулу ДНК изобретения, охарактеризованную выше.

Настоящее изобретение, кроме того, относится к клетке, принявшей вышеохарактеризованный реплицируемый вектор экспрессии. В принципе, такая клетка может быть клеткой любого типа, то есть прокариотной клеткой, одноклеточным эукариотным микроорганизмом или клеткой, происходящей из многоклеточного организма, например млекопитающего. Клетки млекопитающего особенно пригодны для целей изобретения, и более подробно обсуждаются ниже.

В другом своем важном аспекте изобретение относится к способу продуцирования рекомбинантных ССЖЛ/КСЛ человека, в которой ДНК последовательность, кодирующую ССЖЛ/КСЛ человека, вводят в вектор, способный реплицироваться в специфичной клетке-хозяине, причем полученный рекомбинантный вектор вводят в клетку-хозяина, выращиваемую в соответствующей культурной среде в приемлемых для экспрессии ССЖЛ/КСЛ человека условиях с последующим выделением ССЖЛ/КСЛ человека.

Применяемой для выращивания клеток средой может служить любая пригодная для таких целей среда. Приемлемый вектор может быть представлен любым вышеописанным вектором и клетка-хозяин может относиться к любому из вышеперечисленных клеточных типов. Для конст-

руирования вектора и осуществления его введения в клетку-хозяина могут быть использованы любые известные способы, применяемые для этих целей в биотехнологии. Экспрессируемые клетками рекомбинантные ССЖЛ/КСЛ человека могут секретироваться, например, переноситься через клеточную мембрану по механизму, зависящему от типа клетки и состава вектора.

Если ССЖЛ/КСЛ человека продуцируется рекомбинантным хозяином внутриклеточно, то есть не секретизируется клеткой, в этом, случае их выделяют стандартными методами, включающими разрушение клетки механическими средствами, например: обработкой ультразвуком или гомогенизацией или ферментативными или химическими средствами с последующей очисткой.

Для возможности секретирования ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, ей должна предшествовать последовательность, кодирующая сигнальный пептид, присутствие которой гарантирует секретирование из клетки ССЖЛ/КСЛ человека таким образом, что по меньшей мере значительная часть экспрессируемых ССЖЛ/КСЛ человека секретизируется в культурную среду с последующим их выделением из среды.

Рекомендуемый способ продуцирования рекомбинантных ССЖЛ/КСЛ человека настоящего изобретения состоит в использовании трансгенных, отличных от человека млекопитающих, способных выделять ССЖЛ/КСЛ человека в свое молоко. Использование трансгенных, отличных от человека млекопитающих имеет то преимущество, что рекомбинантные ССЖЛ/КСЛ человека могут быть получены с высокими выходами и при разумных затратах, особенно, если отличным от человека млекопитающим является корова. В этом случае рекомбинантные ССЖЛ/КСЛ человека продуцируются с молоком, являющимся обычным компонентом, например, детского питания, в результате чего нет необходимости в интенсивной очистке при использовании рекомбинантных ССЖЛ/КСЛ человека в качестве питательной добавки к продуктам на основе молока. Более того, продуцирование в высшем организме, например, отличном от человека млекопитающем, обычно ведет к точному процессингу белка млекопитающего, например, с точки зрения посттрансляционного процессинга, обсуждаемого выше, и надлежащей укладки белка. Кроме того, по существу чистые ССЖЛ/КСЛ человека могут быть получены в больших количествах.

Соответственно, еще одним своим важным аспектом настоящее изобретение относится к экспрессирующей системе млекопитающего, включающей кодирующую ССЖЛ/КСЛ человека ДНК последовательность, встроенную в ген, кодирующий молочный белок отличного от человека млекопитающего, с образованием в результате гибридного гена, экспрессируемого в молочной железе взрослой самки млекопитающего, принявшей гибридный ген.

В качестве ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, рекомендуется ДНК последовательность, показанная в перечне последовательностей, как ПОСЛЕДОВАТЕЛЬНОСТЬ № 1, или геномный СЕЛ-ген человека, или его аналог.

Молочная железа в качестве экспрессирующей ткани и гены, кодирующие молочные белки, как правило, считаются наиболее пригодными для применения с целью продуцирования гетерологичных белков в трансгенных, отличных от человека млекопитающих, поскольку молочные белки в естественных условиях продуцируются молочной железой на высоком уровне. Кроме того, молоко легко собирается и доступно в больших количествах. В этой связи, применение генов молочного белка в продуцировании рекомбинантных ССЖЛ/КСЛ человека имеет еще и то преимущество, что ССЖЛ/КСЛ продуцируются в условиях, аналогичных условиям их естественного продуцирования, с точки зрения регулирования экспрессии и места продуцирования (молочная железа).

В контексте настоящего изобретения термин "гибридный ген" относится к ДНК последовательности, содержащей, с одной стороны, ДНК последовательности, кодирующую вышеохарактеризованные ССЖЛ/КСЛ человека, а с другой стороны, ДНК последовательности гена молочного белка, способного поддерживать экспрессию продукта гибридного гена. Термин "ген, кодирующий молочный белок", относится ко всему гену, а также к его последовательности, способных поддерживать и направлять экспрессию гибридного гена в представляющей интерес ткани, то есть в молочной железе. Обычно такая последовательность - это последовательность, включающая по меньшей мере одну или несколько областей промотора, участок начал транскрипции, некодирующие 3'- и 5'-области и структуральные последовательности. ДНК последовательности, кодирующая ССЖЛ/КСЛ человека предпочтительно по существу не содержит прокариотных последовательностей, например, векторных последовательностей, которые могут быть связаны с ДНК последовательностью после, например, ее клонирования.

Гибридный ген рекомендуют создавать вставкой *in vitro* известными специалистам методами в ген молочного белка ДНК последовательности, кодирующей ССЖЛ/КСЛ человека. Или же ДНК последовательности, кодирующая ССЖЛ/КСЛ человека, может быть вставлена *in vivo* гомологичной рекомбинацией.

Как правило, ДНК последовательности, кодирующая ССЖЛ/КСЛ человека, вставляется в один из первых экзонов выбранного гена молочного белка или в его эффективную последовательность, содержащую первые экзоны и предпочтительно значительную часть фланкирующей 5'-области, которые, как полагают, играют важную регулируемую роль.

Рекомендуется, чтобы гибридный ген включал последовательности, кодирующие сигнальный пептид, необходимый для секретирования надлежащим образом продукта гибридного гена в молочную железу. Сигнальный пептид - это, как правило, пептид, находящийся обычно в рассматриваемом гене молочного белка, или пептид, связанный с ДНК последовательности, кодирующей ССЖЛ/КСЛ человека. Однако, также возможны и другие сигнальные последовательности, способные поддерживать секрецию продукта гибридного гена в молочную железу. Разумеется, различные элементы гибридного гена должны быть слиты пу-

тем, позволяющим правильную обработку и экспрессию генного продукта. Так, ДНК последовательности, кодирующая выбранный сигнальный пептид, как правило, должна быть точно слита с N-концевой частью ДНК последовательности, кодирующей ССЖЛ/КСЛ человека. В гибридном гене ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, обычно включает их терминирующий кодон, но без их участка сигнала отщепления и полиаденилирования. В нисходящем направлении от ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, обычно сохраняются процессинговые последовательности мРНК гена молочного белка.

Считается, что целый ряд факторов отвечает за реальный уровень экспрессии конкретного гибридного гена. Эффективность промотора, а также других вышеупомянутых регуляторных последовательностей, участок интеграции экспрессионной системы в геноме млекопитающего, участок интеграции ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, в кодирующем молочный белок гене, элементы, создающие послетранскрипционную регуляцию, и другие аналогичные факторы могут иметь жизненно важное значение для достигаемого уровня экспрессии. Основываясь на знании разнообразных факторов, влияющих на уровень экспрессии гибридного гена, специалист способен создать экспрессионную систему, применимую для целей настоящего изобретения.

Молочной железой секретируется целый ряд различных молочных белков. Существует основная группа молочных белков, а именно: казеины и сывороточные белки. Состав молока от животных различного вида меняется качественно, а также количественно по содержанию указанных белков. Большинство отличных от человека млекопитающих продуцируют казеин 3 различных типов, а именно: α -казеин, β -казеин и χ -казеин. К наиболее обычным коровьим сывороточным белкам относятся α -лактальбумин и β -лактальбумин. Состав молока различного происхождения подробно раскрывается в работе Clark и др. (1987).

Применяемый ген молочного белка может происходить от того же вида, что и вид, в который вводится экспрессионная система, или может происходить от других видов животных. В этой связи показано, что регуляторные элементы, направляющие экспрессию гена на молочную железу, функциональны в границах вида, что может быть связано с возможным общим предшественником (Hennighausen и др., 1990).

Примеры приемлемых генов, кодирующих молочный белок, или его эффективных субпоследовательностей, предназначенных для конструирования экспрессирующей системы изобретения, как правило, могут быть обнаружены среди генов сывороточных белков разнообразных млекопитающих, например: гена сывороточного кислотного белка (WAP-ген), предпочтительно мышинного происхождения и гена β -лактоглобулина, предпочтительно овечьего происхождения. Кроме того, для трансгенного продуцирования ССЖЛ/КСЛ человека, могут оказаться пригодны гены казеина различного происхождения например: коровий α S1-казеин и кроличий β -казеин. Рекомендуемым для

применения в настоящем изобретении геном является мышинный WAP-ген, поскольку обнаружено, что этот ген способен обеспечить на высоком уровне экспрессию ряда чужеродных белков человека в молоке различных трансгенных животных (Hennighausen и др. 1990).

Другой рекомендуемой последовательностью, связанной с экспрессирующей системой изобретения, является так называемая стабилизирующая экспрессирующая последовательность, способная поддерживать высокий уровень экспрессии. Существуют сильные доказательства того, что такая стабилизирующая последовательность находится вблизи и в восходящем направлении от генов молочных белков.

ДНК последовательность, кодирующая ССЖЛ/КСЛ человека и подлежащая вставке в экспрессирующую систему изобретения, может быть геномной или искусственного происхождения, или их комбинацией. Некоторые экспрессионные системы для достижения удовлетворительной экспрессии, как обнаружено, требуют присутствия интронов и иных регуляторных областей (Hennighausen и др., 1990). В некоторых случаях может оказаться желательным введение в векторные конструкции в качестве кодирующего полипептид элемента геномных структур, а не элементов кДНК (Brinster и др.). Применение интронных и экзонных структур может привести к более высокому уровню мРНК в устойчивом состоянии, чем в случае применения векторов на основе кДНК.

В еще одном своем аспекте настоящее изобретение относится к гибриднему гену, включающему ДНК последовательность, кодирующую ССЖЛ/КСЛ человека и вставленную в ген, кодирующий молочный белок отличного от человека млекопитающего, при этом ДНК последовательность вставляется в ген молочного белка таким образом, что способна экспрессироваться в молочной железе млекопитающего, принявшего гибридный ген. Гибридный ген и его составляющие подробно обсуждены выше. Гибридный ген является важным промежуточным элементом в конструировании экспрессирующей системы изобретения, раскрытой выше.

В другом своем аспекте настоящее изобретение относится к клетке отличного от человека млекопитающего, принявшей вышеохарактеризованную экспрессирующую систему. Рекомендуемой клеткой млекопитающего является клетка эмбриона, или пронуклеус. Экспрессирующую систему вводят соответствующим образом в клетку млекопитающего применением способа, разъясняемого далее и иллюстрируемого нижеследующими специальными примерами.

В еще одном своем важном аспекте настоящее изобретение относится к способу создания трансгенного, отличного от человека млекопитающего, способного экспрессировать ССЖЛ/КСЛ человека. Способ состоит в инъектировании вышеопределенной экспрессионной системы изобретения в оплодотворенную яйцеклетку или клетку эмбриона млекопитающего с введением в результате экспрессионной системы в зародышевую линию млекопитающего с последующим развитием полученного инъектированного оплодотворенного яйца

или эмбриона во взрослую самку млекопитающего.

Введение экспрессионной системы в зародышевую линию млекопитающего может быть осуществлено применением любой приемлемой методики, приведенной, например, в издании "Манипуляции с эмбрионом мыши", Лабораторное руководство, Коулд Спринг Харбор Лаборатории Пресс, 1986. К примеру, несколько сот молекул экспрессионной системы могут быть непосредственно инъектированы в оплодотворенную яйцеклетку, например, в одну ее оплодотворенную клетку или ее пронуклеус, или в эмбрион выбранного млекопитающего, после чего микроинъектированные яйцеклетки могут быть перенесены в яйцеводы псевдобеременных приемных матерей с последующим развитием эмбрионов. Как правило, не все инъектированные яйцеклетки развиваются во взрослые самки, экспрессирующие ССЖЛ/КСЛ человека. Так, около половины животных со статистической точки зрения будут самцами, от которых, однако, в следующих поколениях могут быть получены и самки.

После интегрирования в зародышевую линию ДНК последовательность, кодирующая ССЖЛ/КСЛ человека, может быть экспрессирована на высоком уровне с продуцированием правильно обработанных и функциональных ССЖЛ/КСЛ человека в стабильных линиях рассматриваемого млекопитающего.

Особый интерес представляет способ создания трансгенных, отличных от человека млекопитающих, способных экспрессировать ССЖЛ/КСЛ человека и по существу неспособных экспрессировать собственные ССЖЛ/КСЛ. Способ состоит в:

(а) уничтожении способности млекопитающим экспрессировать ССЖЛ/КСЛ с тем, чтобы по существу никакой экспрессии ССЖЛ/КСЛ не происходило, и введении вышеохарактеризованной экспрессионной системы изобретения или ДНК последовательности, кодирующей ССЖЛ/КСЛ человека, в зародышевую линию млекопитающего таким образом, что в млекопитающем экспрессируются ССЖЛ/КСЛ человека, и/или

(б) замене CEL-гена млекопитающего или части гена на вышеохарактеризованную экспрессионную систему изобретения или на ДНК последовательность, кодирующую ССЖЛ/КСЛ человека.

Способность млекопитающего экспрессировать ССЖЛ/КСЛ обычно уничтожают созданием мутаций в ДНК последовательности, отвечающей за экспрессию ССЖЛ/КСЛ. Такие мутации включают мутации, при которых ДНК последовательность выводится из рамки, вводится терминирующий кодон или происходит делеция одного или нескольких нуклеотидов ДНК последовательности.

Замена CEL-гена млекопитающего или его части на вышеохарактеризованную экспрессионную систему или на ДНК последовательность, кодирующую ССЖЛ/КСЛ человека, может быть осуществлена применением хорошо известных принципов гомологичной рекомбинации.

В еще одном своем аспекте настоящее изобретение относится к трансгенному, отличному от человека млекопитающему, созданному вышеприведенным способом.

Хотя трансгенное, отличное от человека млекопитающее изобретения в самом широком его аспекте не ограничено каким-либо конкретным типом млекопитающего, тем не менее млекопитающее обычно выбирают из группы, включающей: мышь, крысу, кролика, овцу, свинью козу и крупный рогатый скот. Для широкомасштабного продуцирования ССЖЛ/КСЛ человека обычно рекомендуются более крупные животные, например: овцы, козы, свиньи и крупный рогатый скот, что связано с их способностью давать в больших количествах молоко. Однако, интерес могут также представлять мыши, кролики и крысы вследствие большей простоты манипуляций с этими животными и более быстрое получение результатов для трансгенных животных, чем в случае, например, крупного рогатого скота.

Объемом изобретения охватывается также потомство вышеохарактеризованных трансгенных животных, способное продуцировать ССЖЛ/КСЛ человека.

В еще одном своем аспекте настоящее изобретение включает молоко от млекопитающего (но не человека), содержащее рекомбинантные ССЖЛ/КСЛ человека.

И еще в одном своем аспекте настоящее изобретение относится к детскому питанию, содержащему рекомбинантные ССЖЛ/КСЛ человека, в частности, вышеохарактеризованный полипептид изобретения. Детское питание может быть приготовлено добавлением к обычным компонентам детского питания рекомбинантных ССЖЛ/КСЛ человека или полипептида в очищенной или частично очищенной форме. Однако, как правило, рекомендуется готовить детское питание из вышеохарактеризованного молока изобретения, особенно, если это коровье молоко. Детское питание может быть изготовлено обычными методами и может содержать любые необходимые добавки, например: минеральные вещества, витамины и т.д.

Примеры

Пример 1. Геномная организация, анализ последовательности и локализация CEL-гена в хромосоме

Если нет особых указаний, применялись стандартные методики молекулярной биологии (Maniatis и др., 1982; Ausubel и др. 1987; Sambrook и др., 1989).

Выделение геномных рекомбинантов

Гибридизацией в бляшках отобраны две различные геномные фаговые библиотеки человека: λ DASH (Клонтех Лабораториз Инк., Паоло Альто, Ка., США) и λ EMBL-3 SP6/T7 (Стратаген, Ла Джолла Ка., США). В качестве зондов при гибридизации применялись различные субклонированные фрагменты рестрикции кДНК (Nilsson др., 1990). Зонды метились методом введения метки в олигонуклеотиды с применением [α - 32 P] дЦТФ (Feinberg и др., 1983).

Картирование, субклонирование и секвенирование геномных клонов

Положительные клоны гидролизуют в присутствии различных ферментов рестрикции, подвергают электрофорезу на 1% агарозных гелях и затем переносят в вакууме (Фармация ЛКБ БТГ, Упсала, Швеция) на нейлоновую мембрану. Мем-

брану подвергают гибридизации с разнообразными кДНК зондами. Фрагменты рестрикции, гибридизованные с зондами, выделяют применением метода изахофореза (Öfverstedt и др., 1984). Более мелкие фрагменты (<800 п.о непосредственно вставляют в вектора M13mp18, M13mp19, M13BM20 или M13BM21 и секвенируют применением в качестве бактерии-хозяина E. coli TG1, а более крепкие фрагменты субклонируют в вектора pTZ18R или pTZ19R применением в качестве бактерии-хозяина E. coli DH5 α и дополнительно гидролизуют. (Плазмиды pS309, pS310 и pS451, применяемые в нижеследующем примере 2, получают соответственно.) Некоторые из выделенных фрагментов применяются также в качестве зондов при гибридизации. Все последовательности нуклеотидов определены методом термации дидезоксицепи (Sanger и др., 1977) применением фермента Кленова и M13 универсального секвенирующего праймера специфичных олигонуклеотидов. Информация о последовательности получена на основе автордиограмм применением программного обеспечения MS-EdSeg по методике Sjöberg и др. (1989). Последовательности анализируют с помощью программ, входящих в пакет UWGCG программного обеспечения (Devereux и др., 1984).

Удлинение праймера

Полную РНК выделяют из человеческой поджелудочной железы и жировой ткани по методике с применением гуанидинийизотиоцианата-CsCl (Chirgwin и др., 1979). Удлинение праймера осуществляют по методике (Ausubel и др., 1987) применением полной РНК и десенсибилизированного 26-тимерного олигонуклеотида (5'-AGGTGAGGCCCAACACAACCAGTTGC-3'), положения 33-58. Гибридизацию праймера с 20 мкг полной РНК проводят при 30°C в течение ночи в 30 мкл 0,9 М NaCl, 0,15 М Хепеса (рН 7,5) и 0,3 М ЭДТН. После реакции удлинения в присутствии обратной транскриптазы продукты удлинения анализируют электрофорезом через 6% денатурированный полиакриламидный гель.

Соматические клеточные гибриды

Для хромосомного анализа CEL-гена использована ДНК из 16 соматических клеточных гибридных линий типа человек-грызун, полученных от NIGMS Хьюмен Генетик Мутант Селл Репозитори (Кориэлл Институт фор Медикал Рисерч, Камдэм, НДж). Человек-мышинные соматические клеточные гибриды GM09925-GM09950 происходят от слияния фибробластов мужского плода человека (IMR-91) с мышинной клеточной линией с дефицитом тимидин-киназы B-82 (Taggart и др., 1985; Mohandas и др., 1986). Гибриды GM10324 и GM102860 включают мышиную клеточную линию A9 с дефицитом HPRT и APRT (Callen и др., 1986), в то время как гибрид GM10611 получен микроклеточным слиянием клеточной линии GM07890 лимфобластов человека, инфицированных ретровирусным вектором SP-1, с клеточной линией UV-135 яичников китайского хомячка (Warburton и др., 1990). Гибрид GM10095 происходит от слияния лимфоцитов самки с балансированным 46, X, t (X;9)(q13;34) кариотипом с клеточной линией CHW1102 яичников китайского хомячка (Mohandas и др., 1979). Содержание хромосом человека в гибрид-

ных линиях, определенное цитогенетическим анализом, а также саузерн-блот анализом и анализом гибридизации *in vitro*, приведено в табл. 1. Высокомолекулярные ДНК, выделенные из исходных клеточных линий мыши, китайского хомячка и человека и 16 гибридных клеточных линий, гидролизуют в присутствии *EcoRI*, фракционируют на 0,8 агарозных гелях и переносят на нейлоновый фильтр. Методом мечения олигонуклеотидов (Feinberg и Vogelstein, 1983) получены [α - 32 P] дЦТФ меченный CEL кДНК зонд (кДНК полной длины), который гибридизован с содержимым фильтров. Фильтры промывают 60 мин. каждый раз при 65°C в 6xSSC/0,5%НДС и в 2xSSC/0,5%НДС.

Полимеразная цепная реакция

Полную геномную ДНК человека, выделенную из лейкоцитов, ДНК из соматических клеточных гибридов и из некоторых положительных рекомбинантов, и полную РНК из лактирующей молочной железы человека и поджелудочной железы человека амплифицируют по экзону 10 и экзону 11. Применяют 2 мкг ДНК. Применяемые праймеры перечислены в табл. 2. Осуществляют тридцать циклов ПЦР в объеме 100 мкл [10 mM Трис-HCl (pH 8,3), 50 mM KCl, 1,5 mM MgCl₂, 200 мкМ каждого дНТФ, 100 мкг/мл желатина, 100 пмоль каждого праймера, 1,5 Е Tag ДНК полимеразы (Перкин-Эльмер Цетус, Норуолк, СТ, США)] и температуре отжига 55°C для всех пар праймеров РНК последовательность амплифицируют применением комбинированной комплементарной ДНК (кДНК) и методики ПЦР. кДНК синтезируют из 10 мкг полной РНК в 40 мл раствора, содержащего 50 mM Трис-HCl (pH 8,3), 50 mM KCl, 10 mM MgCl₂, 10 мкг/мл БСА, 1 mM каждого дНТФ, 500 нг олиго(dt)₁₂₋₁₈, 40 Е ингибитора рибонуклеазы и 200 Е обратной транскриптазы (MoMuLV) (БРЛ, Бетседа Рисерч Лабореториз, Н.-И., США) выдерживанием 30 мин при 42°C. кДНК осаждают и вновь суспендируют в 25 мкл H₂O, 2 мкл суспензии амплифицируют вышеприведенным способом. Амплифицированные фрагменты анализируют на 2% агарозном геле. Некоторые фрагменты дополнительно субклонировать и секвенируют.

Генная структура человеческого GEL-гена

В каждой геномной библиотеке было скринировано 10⁶ рекомбинантов и скрининги давали несколько положительных клонов, которые были изолированы и картированы. Два клона, обозначенные λ BSSL1 и λ BSSL5A, подвергнуты дополнительному анализу. Гидролизом в присутствии нескольких ферментов рестрикции, саузерн-блоттингом с последующей гибридизацией с кДНК зондами показано, что λ BSSL5A клон охватывает GEL-ген целиком и что λ BSSL1 клон охватывает 5'-половину и примерно 10 т.п.о. 5'-фланкирующей области (фиг. 1). Совместно эти два клона охватывают 25 т.п.о. генома человека.

После субклонирования и гидролиза ферментами рестрикции получены приемлемые для секвенирования фрагменты, позволяющие определить последовательность CEL-гена полностью, включая 1640 п.о. 5'-фланкирующей области и 41 п.о. 3'-фланкирующей области. На основании этих данных установлено, что CEL-ген человека (Последовательность № 1) охватывает область в 9850 п.о., содержащую 11 экзонов, разделяемых

10 интронами (фиг. 1). Это означает, что экзоны и особенно интроны сравнительно малы. В самом деле, размер экзонов 1-10 колеблется в интервале 97-204 п.о., а длина экзона 11 - 841 п.о. Размер интронов колеблется в пределах 85-2343 п.о. Как видно из табл. 3, все границы экзон/интрон подчиняются AG/GT правилу и хорошо согласуются с консенсусной последовательностью, предложенной Mount и др. (1982). При сравнении кодирующей части GEL-гена с кДНК (Nilsson и др., 1990) обнаружено единственное отличие в нуклеотидной последовательности: вторым нуклеотидом в экзоне 1 является С, которому в кДНК последовательности соответствует Т. Поскольку это положение расположено на расстоянии в 10 нуклеотидов в восходящем направлении от кодона АТС начала трансляции, указанное отличие не влияет на аминокислотную последовательность.

В секвенированной области присутствуют семь членов Alu класса повторяющихся ДНК элементов, обозначенные Alu1-Alu7 (5'-3') (фиг. 1), один из которых находится в 5'-фланкирующей области, а шесть остальных - в пределах CEL-гена.

Сайты иницирования транскрипции и 5'-фланкирующая область

Для картирования сайта(ов) иницирования транскрипции CEL-гена человека осуществлен анализ с удлинением праймера, в котором использована полная РНК из поджелудочной железы, лактирующей молочной железы и жировой ткани человека. Результаты анализа указывают на основной сайт начала транскрипции, расположенный в 12 п.о., и меньший сайт начала, расположенный на расстоянии в 8 п.о. в восходящем направлении от инициатора метионина. Сайты начала транскрипции одинаковы как для поджелудочной железы так и для лактирующей молочной железы, в то время как для жировой ткани никакого сигнала не было обнаружено (фиг. 2). Секвенированная область включает 1640 нуклеотидов 5'-фланкирующей ДНК. На основе схожести последовательностей на расстоянии в 30 нуклеотидов в восходящем направлении от сайта инициации транскрипции обнаружена подобная рамке Хогнесса последовательность CATAAAT (фиг. 4). Никаких свидетельств присутствия в этой области структур типа рамки Хогнесса или СС блоков не обнаружено.

5'-фланкирующая последовательность подвергнута в обоих нитях компьютерному скринингу на нуклеотидные последовательности, известные, как последовательности транскрипционного фактора связывания в других специфических для молочной железы и поджелудочной железы генах. Обнаружено несколько предполагаемых последовательностей (см. фиг. 4).

Хромосомная локализация CEL-гена

В контрольной ДНК человека применением CEL кДНК зонда обнаружены четыре *EcoRI* фрагмента примерно в 13 т.п.о., 10 т.п.о., 2,2 т.п.о. и 2 т.п.о., в то время как в контрольных ДНК мыши и хомячка обнаружены единственные фрагменты соответственно в 25 т.п.о. и 8,6 т.п.о. Присутствие в гибридных клонках последовательностей CEL-гена человека коррелирует только с присутствием хромосомы 9 человека (табл. 1). Всего лишь один

из 16 проанализированных гибридов оказался положительным на присутствие CEL-гена, и данный гибрид содержит хромосому 9 в качестве единственной хромосомы человека. Никакого несоответствия в локализации этой хромосомы не обнаружено, в то время как в локализации любой другой хромосомы выявлено по меньшей мере два несоответствия (табл. 1). Для дополнительной сублокализации CEL-гена нами использован гибрид человек-китайский хомячок (GM 10095), сохранивший (9) транслокационную хромосому (9pter→9q34:Xq13→Xqter) в качестве единственной ДНК человека. Саузерн-блоттингом нам не удалось обнаружить в этом гибриде какие-либо последовательности CEL-гена, что указывает на присутствие остатков CEL-гена в пределах 9q34-qter области.

Пример 2. Конструирование векторов экспрессии

Для конструирования вектора экспрессии для продуцирования рекомбинантной человеческой КСЛ в молоке от трансгенных животных применена следующая стратегия (фиг. 5).

По вышеприведенным методикам получены три основанные на pTZ плазмиды (Фармация, Упсала, Швеция), содержащие различные части CEL-гена человека, а именно: pS309, pS310 и pS311. Плазмида pS309 содержит SphI фрагмент, охватывающий CEL-ген от 5'-нетранскрибируемой области до части четвертого интрона. Плазмида pS310 содержит SacI фрагмент, охватывающий последовательность CEL-гена от части первого интрона до части шестого интрона. Третья плазмида pS311 содержит BamHI фрагмент, охватывающий вариант CEL-гена от основной части пятого интрона до остальной части структуры интрон/экзон. В этой плазмиде повторяющаяся последовательность экзона 11, обычно кодирующая 16 повторов, мутирована с кодированием усеченного варианта, имеющего 9 повторов.

Для слияния геномных последовательностей использована другая плазмида (pS283), содержащая часть CEL кДНК человека, клонированную в плазмиду pUC19 по HindIII и SacI сайтам. Плазмида pS283 использована также для доступа к удобному участку рестрикции ферментом (KpnI), расположенному в 5'-нетранслируемой лидерной последовательности CEL-гена. Затем плазмиду pS283 гидролизуют в присутствии NcoI и SacI с последующим выделением фрагмента в 2,7 т.п.о. Гидролизом плазмиды pS309 в присутствии NcoI и BspEI выделен фрагмент в 2,3 т.п.о., содержащий 5'-часть CEL-гена. Гидролизом плазмиды pS310 в присутствии BspEI и SacI выделен фрагмент в 2,7 т.п.о., содержащий часть средней области CEL-гена. Эти три фрагмента лигируют и трансформируют в компетентном TC2 штамме *E. coli* и трансформанты выделяют отбором по ампициллину. Из ряда трансформантов получают плазмиды и одну из плазмид названную pS312 и содержащую целевую конструкцию (фиг. 6), применяют в дальнейших экспериментах.

Для получения модификации pS311, в которой для облегчения последующего клонирования расположенный в нисходящем направлении от терминирующего кодона BamHI сайт превращен в Sall сайт, используют следующий способ. Частич-

ным гидролизом в присутствии BamHI pS311 переводят в линейную форму. Выделяют фрагмент в линейной форме и в него вставляют синтетический ДНК линкер (5'-GATCGTTCGAC-3'), превращающий BamHI сайт в Sall сайт с разрушением в результате BamHI сайта. Поскольку существуют два потенциальных положения для интеграции синтетического линкера полученные плазмиды анализируют расщеплением в присутствии фермента рестрикции. Выделенную плазмиду со вставкой линкера в целевом положении, расположенном в нисходящем направлении от экзона 11, обозначают, как pS313.

Для получения вектора экспрессии, принявшего CEL геномные последовательности и кодирующего усеченный CEL вариант, применяют плазмиду pS314, предназначенную для поддержания постатдийно- и тканево-специфичной экспрессии в клетках молочной железы в период лактации. Плазмида pS314 содержит геномный фрагмент из мышинового гена сывороточного молочного белка (WAP-гена) (Campbell и др., 1984), клонированного в виде NotI фрагмента. Геномный фрагмент содержит примерно 4,5 т.п.о. восходящих регуляторных последовательностей (ВРП), полную транскрибируемую экзон-интронную область и 3 т.п.о. последовательности последнего экзона в нисходящем направлении. Единственный KpnI сайт расположен в первом экзоне на расстоянии в 24 п.о. в восходящем направлении от природного WAP кодона инициирования трансляции. Другими уникальным сайтом рестрикции ферментом является Sall сайт, расположенный в экзоне 3. В плазмиде pS314 этот Sall сайт разрушают гидролизом, наполняя применением фрагмента Кленова и повторным лигированием. Взамен вводят новый Sall сайт непосредственно в нисходящем направлении от KpnI сайта в экзоне 1. Вставку осуществляют гидролизом в присутствии KpnI и введением в указанное положение ренатурированных синтетических олигомеров SYM 2401 (5'-CGTTCGACGTAC-3') и SYM 2402 (5'-GTCTGACGGTAC-3') (фиг. 8). CEL геномную последовательность человека вставляют между указанными сайтами (KpnI и Sall) применением следующей стратегии. Во-первых, pS314 гидролизуют в присутствии KpnI и Sall и фрагмент, представляющий отщепленную плазмиду, выделяют электрофорезом. Во-вторых, плазмиду pS312 гидролизуют в присутствии KpnI и BamHI и выделяют фрагмент в примерно 4,7 т.п.о., представляющий 5'-часть CEL-гена человека. В-третьих, плазмиду pS313 гидролизуют в присутствии BamHI и Sall и выделяют 3'-часть CEL-гена человека. Три полученных фрагмента лигируют, трансформируют в компетентном штамме *E. coli* и трансформанты выделяют после отбора по ампициллину. Из нескольких трансформантов получают плазмиды, которые тщательно анализируют картированием ферментами рестрикции и секвенс-анализом. Выявлена одна плазмида, представляющая целевой вектор экспрессии и названная pS317.

Для конструирования вектора экспрессии геномного CEL, кодирующего CEL полной длины, плазмиду pS317 модифицируют следующим образом (фиг. 5). Во-первых, плазмиду pTZ18R (Фармация), содержащую BamHI фрагмент в 5,2 т.п.о.

CEL-гена человека, простирающийся от пятого интрона в нисходящем направлении одиннадцатого экзона (pS451), гидролизуют в присутствии HindIII и SacI. В результате гидролиза получают фрагмент в 1,7 т.п.о., простирающийся от HindIII сайта, расположенной в интроне 9, до SacI сайта, расположенного в экзоне 11. Во-вторых, плазмиду pS313 гидролизуют в присутствии SacI и Sail и выделяют фрагмент в 71 п. о., содержащий 3'-часть экзона 11 и созданный Sall сайт. В-третьих, из плазмиды pS317 выделяют остальную часть WAP/CEL рекомбинантного гена и плазмидные последовательности в виде Sall/HindIII фрагмента в 20 т.п.о. Три полученных фрагмента лигируют, и трансформируют в бактерию. Из нескольких трансформантов получают плазмиды. Плазмиды гидролизуют в присутствии различных ферментов рестрикции и подвергают секвенс-анализу. Идентифицирована одна плазида, содержащая целевой рекомбинантный ген. Этот конечный вектор экспрессии получил название pS452 (фиг. 7).

Для удаления прокариотных плазмидных последовательностей pS452 гидролизуют в присутствии NotI. Последующим электрофорезом на агарозном геле выделяют элемент рекомбинантного вектора, состоящий из мышиной WAP последовательности, фланкирующей человеческий CEL геномный фрагмент. Перед инъектированием в эмбрионы мыши выделенный фрагмент подвергают дополнительной очистке применением электроэлюирования.

Рекомбинантный WAT/CEL-ген для экспрессии в молочной железе трансгенных животных показан на фиг. 8.

Депозиты

Следующие плазмиды депонированы согласно Будапештскому Соглашению в (Deutsche Sammlung von Mikroorganismen und Zellkulturen).

Плазида	№ депозита	Дата депонирования
pS309	DSM 7101	12 июня 1992 г.
pS310	DSM 7102	
pS451	DSM 7498	26 февраля 1993 г.
pS452	DSM 7499	

Пример 3. Создание трансгенных животных

По методике примера 2 из плазмиды pS452 выделяют NotI фрагмент. Полученный ДНК фрагмент содержит мышиный WAP промотор, связанный с геномной последовательностью, кодирующей ССЖЛ/КСЛ человека. Выделенный фрагмент в концентрации 3 нг/мкл инъектируют в пронуклеус 350 C57Bl/ 6JxCBA/2J-f₂ эмбрионов, полученных от донорной мыши, праймированной для сверховуляции 5 МЕ сывороточного гонадотропина жеребой кобылы. Мыши линии C57Bl/6JxCBA/2J-f₁ получены от Бомхольтгард Бридинг энд Рисерч Сентр ЛТД, Рай, Дания. После сбора эмбрионов из яйцевода их обработкой гиалуронидазой в среде M2 отделяют от клеток кумулуса (Ноган и др., 1986). После промывания эмбрионы переносят в среду M16 (Ноган и др., 1986) и выдерживают в инкубаторе в атмосфере 5% CO₂. Инъектирование осуществляют в микрокапле M2 под легким парафиновым маслом с помощью гидравлического микроманипулятора фирмы Наришиги и инверси-

онного микроскопа Никон, снабженного оптикой фирмы Номарски. После инъектирования выглядывшие здоровыми эмбрионы имплантируют внутрибрюшинно псевдобеременным C57Bl/6JxCBA/2J-f₁ реципиентам, получавшим 0,37 мл 2,5% Авертина. Мыши с интегрированным трансгеном выявлены ПЦР анализом ДНК из образцов хвоста, отобранных биопсией через три недели после рождения животных. Положительные результаты подтверждены саузерн-блот анализом.

Пример 4. Экспрессия ССЖЛ/КСЛ в трансгенной мыши

Трансгенную мышь идентифицируют анализом ДНК, полученной из образца отрезанного хвоста. Образцы ткани инкубируют с протеиназой К и экстрагируют смесью фенола с хлороформом. Выделенную ДНК используют в полимеразной цепной реакции с праймерами, амплифицирующими специфические фрагменты при наличии введенной гетерологичной ДНК, представляющей фрагмент вектора экспрессии. Животные также оценивают в тестах гибридизации ДНК с целью подтверждения данных ПЦР и выявления возможных перегруппировок, строения элементов интегрированного вектора и для получения информации о числе копий элементов интегрированного вектора.

В одной серии опытов анализу подвергнуто 18 мышей. Анализ осуществлен двумя способами и полученные результаты показывают, что одна из мышей несет гетерологичный ДНК векторный элемент, происходящий из pS452. Результаты ПЦР анализа и опытов с гибридизацией были идентичны (фиг. 9).

Мышь, для которой выявлено присутствие элемента векторной ДНК (животное-основатель), спаривают и F1 помёт подвергают анализу на трансген теми же методами.

Самкам лактирующих животных инъектируют внутрибрюшинно 2 МЕ окситоцина и 10 минут спустя анестезируют введением внутрибрюшинно 0,4 мл 2,5 авертина. Устройство для сбора молока присоединяют к соску силиконовой трубочкой и молоко собирают в пробирки Эппендорфа на 1,5 мл осторожным массажем молочной железы. Количество молока менялось в зависимости от дня сбора в пределах 0,1-0,5 мл на мышь и на сбор.

Анализ на присутствие рекомбинантных ССЖЛ/КСЛ человека проводят проведением НДС-ПАГЭ, переносом на нитроцеллюлозные мембраны и инкубированием с поликлональными антителами, созданными против нативных ССЖЛ/КСЛ человека. Полученные результаты указывают на экспрессию рекомбинантных ССЖЛ/КСЛ человека в молоке трансгенных мышей. Присутствие рекомбинантных ССЖЛ/КСЛ человека в молоке от трансгенных мышей показано на фиг. 10 в виде полосы при 116,5.

Созданы стабильные линии трансгенных животных.

Аналогичным путем могут быть созданы и другие трансгенные животные, например: коровы или овцы, способные экспрессировать ССЖЛ/КСЛ человека.

- Источники информации
- Abouakil, N., Rogalska, E., Bonicel, J. & Lombardo, D. (1988): *Biochim. Biophys. Acta* 961, 299-308.
- Ausubel, F.M., Brent, R.E., Moore, D.D., Smiyh, J.A., Seidman, J.G. and Struhl, K.: *Current Protocols in Molecular Biology* (Wiley Interscience, New York 1987).
- Baba, T., Downs, D., Jackson, K.W., Tang, J. and Wang, C.S. (1991): *Biochemistry* 30, 500-510.
- Beato, M. (1989): *Cell* 56, 335-344.
- Bernbäck, S., Bläckberg, L. & Hernell, O. (1990): *J. Clin. Invest.* 221-226.
- Björkstén, B., Burman, L.G., deChateau, P., Fredrikzon, B., Gothefors, L. & Hernell, O. (1980): *Br. Med. J.* 201, 267-272.
- Bläckberg, L., Ängquist, K.A. & Hernell, O. (1987): *FEBS Lett.* 217, 37-41.
- Bläckberg, L. & Hernell, O. (1981): *Eur. J. Biochem* 116, 221-225.
- Bläckberg, L., Lombardo, D., Hernell, O., Guy, O. & Olivecrona, T. (1981): *FEBS Lett.* 136, 284-288.
- Boulet, A.M., Erwin, C.R. and Rutter, W.J. (1986): *Proc. Natl. Acad. Sci. U.S.A.* 83, 3599-3603.
- Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E. & Palmiter, R.D. (1988): *Proc. Natl. Acad. Sci. U.S.A.* 85, 836-840.
- Callen, D.F. (1986): *Ann. Genet.* 29, 235-239.
- Campbell, S.M., Rosen, J.M., Hennighausen, L.G., Strech-Jurk, U. and Sippel, A.E. (1984): *Nucleic Acid Res.* 12, 8685-8697.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979): *Biochemistry* 18, 5294-5299.
- Clark, A.J., Simons, P., Wilmut, I. and Lahte, R. (1987): *TIBTECH* 5, 20-24.
- Devereux, J., Haeberli, P. and Smithies (1984): *Nucleic Acids Res.* 12, 387-395.
- Feinberg, A. and Vogelstein, B. (1983): *Anal. Biochem.* 132, 6-13.
- Hennighausen, L., Ruiz, L. & Wall, R. (1990): *Current Opinion in Biotechnology* 1, 74-78.
- Hernell, O. & Bläckberg, L. (1982): *Pediatr. Res.* 16, 882-885.
- Hogan, B., Constantini, F. and Lacy, E. (1986): *Manipulating the mouse embryo. A Laboratory Manual.* Cold Spring Harbor Laboratory Press.
- Hui, D. and Kissel, J.A. (1990): *FEBS Lett.* 276, 131-134.
- Lombardo, D., Guy, O. & Figarella, C. (1978): *Biochim. Biophys. Acta* 527, 142-149.
- Maniatis, T., Fritsch, E.F. & Sambrook, J.: *Molecular Cloning. A Laboratory Manual.* (Cold Spring Harbor, NY, 1982).
- Mohandas, T., Sparkes, R.S., Sparkes, M.C., Shulkin, J.D., Toomey, K.E. and Funderburk, S.J. (1979): *Am. J. Hum. Genet.* 31, 586-600.
- Mohandas, T., Heinzmann, C., Sparkes, R.S., Wasmuth, J., Edwards, P. and Lysis, A.J. (1986): *Somatic Cell. Mol. Genet.* 12, 89-94.
- Mount, S.M. (1982): *Nucleic Acids Res.* 10, 459-472.
- Nilsson, J., Bläckberg, L., Carlsson, P., Enerbäck, S., Hernell, O. and Bjursell, G. (1990): *Eur. J. Biochem.* 192, 543-550.
- Qasba, M., and Safaya, S.K. (1984): *Nature* 308, 377-380.
- Reue, K., Zambaux, J., Wong, H., Lee, G., Leete, T.H., Ronk, M., Shively, J.E., Sternby, B., Borgström, B., Ameis, D. and Schotz, M.C. (1991): *J. Lipid. Res.* 32, 267-276.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.E.: *Molecular Cloning. A Laboratory Manual.* (Cold Spring Harbor, NY, 1989).
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977): *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Sjöberg, S., Carlsson, P., Enerbäck, S. and Bjursell, G. (1989): *Comput. Appl. Biol. Sci.* 5, 41-46.
- Taggart, R.T., Mohandas, T., Shows, T. B. and Bell, G.I. (1985): *Proc. Natl. Acad. Sci. U.S.A.* 82, 6240-6244.
- Warburton, D., Gersen, S., Yu, M.T., Jackson, C., Handelin, B. and Housman, D. (1990): *Genomics* 6, 358-366.
- Whitelaw et al. (1991): *Transgenic Research* 1, 3-13.
- Yu-Lee, L., Richter-Mann, L., Couch, C., Stewart, F., Mackinlay, G. and Rosen, J. (1986): *Nucleic. Acid. Res.* 14, 1883-1902.
- Öfverstedt, L.G., Hammarström, K., Balgobin, N., Hjerten, S., Pettersson, U. and Chattopadhyaya, J. (1984): *Biochim. Biophys. Acta* 782, 120-126.
- Пояснения к диаграммам
- Фиг. 1
- Локус CEL-гена. Показаны легализация и карта рестрикции ферментами двух частично перекрывающихся клонов: λ BSSL1 и λ BSSL5A. Экзон-интронная организация и используемые сайты рестрикции ферментами показаны ниже. Экзоны представлены блоками пронумерованными цифрами 1-11. Asp=Asp700, B=BamHI, ~~EC~~ EcoRI, S=SacI, Sa=Sall, Sp=SphI и X=XbaI. Положение и ориентация AI повторяющихся элементов показаны жирными стрелками. Буквами a-h обозначены различные субклонированные фрагменты.
- Фиг. 2
- Анализ РНК с удлинением праймера. РНК выделена из лактирующей молочной железы, поджелудочной железы и жировой ткани человека. Для праймирования РНК обратной транскрипцией использован радиомеченный с конца 26-тимерный олигонуклеотид, комплементарный нуклеотидам в положениях 33-58 CEL-гена. Вертикальная полоса А - маркер молекулярного размера (лэддер последовательности нуклеотидов), вертикальная полоса В - панкреатическая РНК, вертикальная полоса С - РНК жировой ткани и вертикальная полоса D - РНК лактирующей молочной железы.
- Фиг. 3
- Дот-плот анализ 5'-фланкирующих областей CEL-генов человека и крысы. Области гомологии отмечены буквами А-Н и прописными буквами написаны представляющие эти области последовательности: сверху - для человека и снизу - для крысы.
- Фиг. 4
- Анализ 5'-фланкирующей последовательности CEL-гена человека. Предполагаемые распознающие последовательности либо выделены подчеркиванием, либо подчеркиванием представлена комплементарная нить. Жирными буквами показано местоположение гомологичности к CEL-гену крысы (области А-Н). Рамка Хогнеса подчеркнута пунктиром.

Существуют две последовательности, каждая из которых показывает 80%-ое сходство с консенсусной последовательностью участка связывания глюкокортикоидного рецептора (GGTACANNNTGTTCT) (Beato M., 1989), первая из них находится у нуклеотида в положении -231 (IA), а вторая - у нуклеотида в положении -811 (IB). Более того, у нуклеотида в положении -861 находится последовательность, показывающая 87%-ое сходство с консенсусной последовательностью участка связывания рецептора эстрогена (AGTTCANNNTGACCT) (Beato M., 1989).

В работе Lubon и Henninghausen (1987) осуществлен анализ промотора и 5'-фланкирующих последовательностей гена сывороточного кислого белка (WAP-гена) и выявлены участки связывания белков ядра клеток лактирующей молочной железы. Один из них, а именно участок консервативной последовательности в 11 п.о (AAGAAGGAAGT) присутствует в ряде изученных генов молочных белков, например: гене крысиного α -лактальбумина (Qasba и др., 1984) и гене крысиного α -казеина (Yu-Lee и др., 1986). В 5'-фланкирующей области CEL-гена на комплементарной нити у нуклеотида в положении -1299 находится последовательность, показывающая 82%-ое сходство с указанной консервативной последовательностью.

При исследовании регуляции гена β -казеина в ядерных экстрактах беременных и лактирующих мышей обнаружен тканевоспецифичный фактор молочной железы (ФМЖ) и была идентифицирована его распознающая последовательность (ATTCTTGGNA). В 5'-фланкирующей области CEL-гена человека имеются две последовательности, одна из которых находится на комплементарной нити у нуклеотида в положении -368 (4A), а другая - у нуклеотида в положении -1095 (4B), каждая из которых показывает 82%-ое сходство с консенсусной последовательностью участка связывания ФМЖ. Кроме этих двух предполагаемых участков связывания ФМЖ в 5'-фланкирующей области на комплементарной нити у нуклеотида 275 в интроне I находится последовательность: AGTTCTTGGCA, показывающая 100%-ую идентичность консенсусной последовательности участка связывания ФМЖ.

Кроме того, существует четыре последовательности, каждая из которых показывает 65%-ое сходство с консенсусной последовательностью крысиного панкреоспецифичного усиливающего элемента (GTCACCTGTGCTTTCCCTG) (Boulet и др., 1986), одна из которых находится у нуклеотида в положении -359 (5A), вторая - у нуклеотида в положении -718 (5B), третья - у нуклеотида в положении -1140 и последняя - у нуклеотида в положении -1277 (5D).

Фиг. 5

Способ получения плазмиды pS452. Дополнительные подробности см. пример 2.

Фиг. 6

Схематическое строение плазмиды pS312.

Фиг. 7

Схематическое строение плазмиды pS452.

Фиг. 8

Физическая карта, представляющая физическое введение CEL-геномной структуры человека в первый экзон WAP-гена по методике примера 2.

Фиг. 9

А. Схематический показ локализации ПЦР-праймеров, применяемых для выявления трансгенных животных. 5'-Праймер расположен в пределах WAP последовательности, начинаясь в положении -148 п.о. в восходящем направлении от места слияния WAP и CEL. 3'-Праймер расположен в первом CEL интроне, заканчиваясь на расстоянии 398 п.о. в нисходящем направлении от места слияния.

В. Последовательности применяемых ПЦР-праймеров.

С. Агарозный гель, показывающий типичный анализ методом ПЦР потенциальных животных-основателей. М - маркеры молекулярной массы, вертикальная полоса 1 - контрольный ПЦР-продукт, полученный из плазмиды pS452, вертикальные полосы 2-13 - ПЦР, проведенные с ДНК препаратами от потенциальных животных-основателей.

Фиг. 10

Иммуноблот-анализ молока от мышинной линии, трансгенный по рекомбинантному мышинному WAP /человечьему CEL-гену плазмиды pS452. Белки разделяют методом НДС-ПАГЭ, переносят на Иммунолоновые мембраны (Миллипор) и визуально изучают продукты реакции с поликлональными кроличьими антителами, созданными с применением высокоочищенной нативной человеческой СЖЛ, с последующей реакцией со свиным антикроличьим IgG, меченным щелочной фосфатазой (Дакопаттс). Вертикальная полоса 1: низкомолекулярные маркеры 106, 80, 49,5, 32,5, 27,5 и 18,5 кДа соответственно. Вертикальная полоса 2: высокомолекулярные маркеры 205, 116,5, 80 и 49,5 кДа, соответственно. Вертикальная полоса 3: 25 нг очищенной нерекombинантной СЖЛ из человеческого молока. Вертикальная полоса 4: 2 мкл образца молока от CEL трансгенной мыши в разбавлении 1:10. Вертикальные полосы 5 и 6: образцы молока по 2 мкл от двух различных не-CEL трансгенных мышей в разбавлении 1:10 в качестве контрольных образцов.

Перечень последовательностей

(2) Информация для Последовательности № 1:

(i) Характеристики последовательности:

(A) Длина: 11531 пара оснований

(B) Тип: нуклеиновая кислота

(C) Число нитей: две

(D) Топология: линейная

(ii) Молекулярный тип: ДНК (геномная)

(vi) Первоначальный источник:

(A) Организм: Homo Sapiens

(B) Тип ткани: молочная железа

(ix) Признак:

(A) Название/ключ: CDS

(B) Местоположение: стык (1653..1727, 4071..4221, 4307..4429, 4707..4904, 6193..6323, 6501..6608, 6751..6868, 8335..8521, 8719..8922, 10124..10321, 10650..11394)

(ix) Признак:
 (A) Название/ключ: зрелый пептид
 (B) Местоположение: стык (1722..1727, 4071..4221, 4307..4429, 4707..4904, 6193..6323, 6501..6608, 6751..6868, 8335..8521, 8719..8922, 10124..10321, 10650..11394)
 (D) Прочая информация: /ЕС. номер 3.1.1.1
 /продукт = "Стимулируемая солями желчных кислот липаза"
 (ix) Признак:
 (A) Название/ключ: 5' UTR
 (B) Местоположение: 1..1640
 (ix) Признак:
 (A) Название/ключ: рамка Хогнесса
 (B) Местоположение: 1611..1617
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 1641..1727
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 4071..4221
 (ix) Признак:
 (A) Название/ключ: экзон (B)
 Местоположение: 4307..4429
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 4707..4904

(ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 6193..6323
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 6501..6608
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 6751..6868
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 8335..8521
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 8719..8922
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 10124..10321
 (ix) Признак:
 (A) Название/ключ: экзон
 (B) Местоположение: 10650..11490
 (ix) Признак:
 (A) Название/ключ: 3' UTR
 (B) Местоположение: 11491..11531
 (XI) Описание последовательности: Последовательность № 1:

GGATCCCTCG AACCCAGGAG TTCAAGACTG CAGTGAGCTA TGATTGTGCC ACTGCACTCT	60
AGCCTGGGTG ACAGAGACCC TGTCTCAAAA AAACAAACAA ACAAACAAACC TCTGTGGACT	120
CCGGGTGATA ATGACATGTC AATGTGGATT CATCAGGTGT TAACAGCTGT ACCCCCTGGT	180
GGGGGATGTT GATAACGGGG GAGACTGGAG TGGGGCGAGG ACATACGGGA AATCTCTGTA	240
ATCTTCCTCT AATTTTGCTG TGAACCTAAA GCTGCTCTAA AAATGTACAT AGATATAAAC	300
TGGGGCCTTC CTTTCCTCTT GCCCTGCCCC AGCCCTCCCC CACCTCCTTC CTCTCCCTGC	360
TGCCTCCCCT CTGCCCTCCC CTTTCCTCCT TAGCCACTGT AAATGACACT GCAGCAAAGG	420
TCTGAGGCAA ATGCCCTTTGC CCTGGGGCGC CCCAGCCACC TGCAGGCCCC TTATTTCTCTG	480
TGGCCGAGCT CCTCCTCCCA CCCTCCAGTC CTTTCCCCAG CCTCCCTCGC CCACTAGGCC	540
TCCTGAATTG CTGGCACCGG CTGTGGTCTGA CAGACAGAGG GACAGACGTG GCTCTGCAGG	600
TCCACTCGGT CCCTGGCACC GGCCGCAGGG GTGGCAGAAC GGGAGTGTGG TTGGTGTGGG	660
AAGCACAGGC CCCAGTGTCT CCTGGGGGAC TGTGTGGGTGG GAAGGCTCTG GCTGCCCTCA	720
CCCTGTTCCC ATCACTGCAG AGGGCTGTGC GGTGGCTGGA GCTGCCACTG AGTGTCTCGG	780
TGAGGGTGAC CTCACACTGG CTGAGCTTAA AGGCCCCATC TGAAGACTTT GTTCGTGGTG	840
TTCTTTCACT TCTCAGAGCC TTTCTGGCT CCAGGATTAA TACCTGTTCA CAGAAAATAC	900
GAGTCGCCTC CTCTCCACA ACCTCACACG ACCTTCTCCC TTCCCTCCCG CTGGCCTCTT	960
TCCCTCCCCT TCTGTCACTC TGCCCTGGGCA TGCCCCAGGG CCTCGGCTGG GCCCTTTGTT	1020
TCCACAGGGA AACCTACATG GTTGGGCTAG ATGCCTCCGC ACCCCCCCAC CCACACCCCC	1080
TGAGCCTCTA GTCTCCCTC CCAGGACACA TCAGGCTGGA TGGTGACACT TCCACACCCT	1140

TGAGTGGGAC	TGCCTTGTGC	TGCTCTGGGA	TTCGCACCCA	GCTTGGACTA	CCCGCTCCAC	1200
GGGCCCCAGG	AAAAGCTCGT	ACAGATAAGG	TCAGCCACAT	GAGTGGAGGG	CCTGCAGCAT	1260
GCTGCCCTTT	CTGTCCCAGA	AGTCACGTGC	TCGGTCCCCCT	CTGAAGCCCC	TTTGGGGACC	1320
TAGGGGACAA	GCAGGGCATG	GAGACATGGA	GACAAAGTAT	GCCCTTTTCT	CTGACAGTGA	1380
CACCAAGCCC	TGTGAACAAA	CCAGAAGGCA	GGGCACTGTG	CACCCTGCCC	GGCCCCACCA	1440
TCCCCCTTAC	CACCCGCCAC	CTTGCCACCT	GCCTCTGCTC	CCAGGTAAGT	GGTAACCTGC	1500
ACAGGTGCAC	TGTGGGTTTG	GGGAAAAC TG	GATCTCCCTG	CACCTGAGGG	GGTAGAGGGG	1560
AGGGAGTGCC	TGAGAGCTCA	TGAACAAGCA	TGTGACCTTG	GATCCAGCTC	CATAAATACC	1620
CGAGGCCCAG	GGGGAGGGCC	ACCCAGAGGC	TG ATG CTC ACC ATG GGG CGC CTG			1673
			Met Leu Thr Met Gly Arg Leu			
			-23 -20			
CAA CTG GTT GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC						1721
Gln Leu Val Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala						
-15 -10 -5						
GCG AAG GTAAGAGCCC AGCAGAGGGG CAGGTCTCTG TGCTCTCTCG CTCAATCAGA						1777
Ala Lys						
1						
TCTGGAAACT	TCGGGCCAGG	CTGAGAAAGA	GCCCAGCACA	GGCCCGCAGC	AGATCCCGGG	1837
CACTCACGCT	CATTTCTATG	GGGACAGGTG	CCAGGTAGAA	CACAGGATGC	CCAATTCCAT	1897
TTGAATTTCA	GATAAACTGC	CAAGAACTGC	TGTGTAAGTA	TGTCCCATGC	AATATTTGAA	1957
ACAAATTTCT	ATGGGCCCGG	CGCAGTGGCT	CACACCTGCA	ATCCCACCAG	TTTGGGAGGC	2017
CGAGGTGGGT	GGATCACTTG	AGGTCAGGAG	TTGGAGACCA	GCCTGGCCAA	CATGGTGAAA	2077
CCCCGTCTCT	ACTAAAAATA	CAAATATTAA	TCGGGCGTGG	TGGTGGGTGC	CTGTAATCCC	2137
AGCTACTCGG	GAGGCTGAGG	CAGGAGAACC	GCTTGAAGCT	GGGAGGTGGA	GATTGCGGTG	2197
AGCTGAGATC	ACGCTACTGC	ACTCCAGCCT	GGGTGACAGG	GCGAGACTCT	GTCTCAAAAA	2257
ATAGAAAAAG	AAAAAATGA	AACATACTAA	AAAACAATTC	ACTGTTTACC	TGAAATTCAA	2317
ATGTAAGTGG	GCCTCTTGAA	TTTACATTTG	CTAATCCTGG	TGATTCCACC	TACCAACCTC	2377
TCTGTTGTTT	CCATTTTACA	GAAGGGGAAA	CGGGCCCAGG	GGCAGGGAGT	GTGGAGAGCA	2437
GGCAGACGGG	TGGAGAGAAG	CAGGCAGGCA	GTTTGCCCAG	CATGGCACAG	CTGCTGCCTC	2497
CTATTCTCTG	GCAGGAAGCT	GAAAGCCGGG	CTACTCCACA	CCCGGGTCCG	GGTCCCTCCA	2557
GAAAGAGAGC	CGGCAGGCAG	GAGCTCTCTC	GAGGCATCCA	TAAATTCTAC	CCTCTCTGCC	2617
TGTGAAGGAG	AAGCCACAGA	AACCCCAAGC	CCCACAGGAA	GCCGGTGTCT	GTGCCCCGGC	2677
CAGTCCCTGC	CCCCAGCAGG	AGTCACACAG	GGGACCCCAG	ATCCCAACCA	CGCTGTTCTG	2737
CTGCCTGCGG	TGTCTCAGGC	CCTGGGGACT	CCTGTCTCCA	CCTCTGCTGC	CTGCTCTCCA	2797
CACTCCCTGG	CCCTGGGACC	GGGAGGTTTG	GGCAGTGGTC	TTGGGCTCCT	GACTCAAAGG	2857
AGAGGTCACC	TTCTTCTTGG	GCGAGCTCTT	CTTGGGGTGC	TGAGAGGCCT	TCGGCAGGTC	2917
ATCACGACCC	CTCCCCATTT	CCCCACCCTG	AGGCCCTCTG	GCCAGTCTCA	ATTGCACAGG	2977
GATCACGCCA	CTGGCACAAAG	GAGACACAGA	TGCCTCGCAG	GGGATGCCCA	CGATGCCTGC	3037

ATGTGTTGCT TCTGGTTCCT TTCCTCCAGT TCCAACCGCC GCACTCTCCC ACACCAGTGT	3097
GACAGGGGGC CCATCACCCT AGACTTCAGA GGGCTGCTGG GACCCTGGCT GGGCCTGGGG	3157
GTGTAGGGCC ACCCTGCCCT TCCCCACCTG GAACCTGGCA CAGGTGACAG CCAGCAAGCA	3217
ATGACCTGGT CCCACCATGC ACCACGGGAA GAGGGAGCTG CTGCCCCAAGA TGGACAGGAG	3277
GTGGCACTGG GGCAGACAGC TGCTTCTCAA CAGGGTGA CTCAAGCCCCA AAGCTGCCCA	3337
GCCTCAGTTC CGTCAGGGAC AGAGGGTGGA TGAGCACCAA CCTCCAGGCC CCTCGTGGGG	3397
GTGGACAGCT TGGTGACAG AGGCCATTTT CATGGCACAG GGAAGCGTGG CGGGGGTGGG	3457
AGGTGTGGTC CCTAGGGGGT TCTTTACCAG CAGGGGGCTC AGGAACTGTG GGGACTTGGG	3517
CATGGGGCCA TCGACTTTGT GCCCAGCCAG CTAGGCCCTG TGCAGGGAGA TGGGAGGAGG	3577
GAAAAGCAGG CCCCACCCCT CAGAAAGGAG GAAGGTTGGT GTGAAACATC CCGGGTACAC	3637
TGAGCATTGG GTACACTCCT CCCGGGAGCT GGACAGGCCT CCCATGTGAT GGCAAACAGG	3697
CCGACAGGAG ACACGGCTGT TGCTCGTCTT CCACATGGGG AACTGAGGA TCGGAGTCAA	3757
AGCTGGGCGG CCATAGCCAG AACCCAAACC TCCATCCCAC CTCTTGCCCG GCTTCCCTAG	3817
TGGGAACACT GGTGAACCA GTTTCCTCTA AGATTCTGGG AGCAGGACAC CCCCAGGGAT	3877
AAGGAGAGGA ACAGGAATCC TAAAGCCCTG AGCATTGCAG GGCAGGGGGT GCTGCCTGGG	3937
TCTCGTGTC AGAGCTGTCC TGCTTTGAAG CTGTCTTTGC CTCTGGGCAC GCGGAGTCGG	3997
CTTGCCCTGC CCCCTCCGA TTCAGGCCGA TGGGGCTTGA GCCCCCTGA CCCTGCCCCG	4057
GTCTCCCTCG CAG CTG GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA	4106
Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu	
5 10	
GGC GTC AAT AAG AAG CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC	4154
Gly Val Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe	
15 20 25 30	
AAG GGC ATC CCC TTC GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG	4202
Lys Gly Ile Pro Phe Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln	
35 40 45	
CCA CAT CCT GGC TGG CAA G GTGGGAGTGG GTGGTGCCGG ACTGGCCCTG	4251
Pro His Pro Gly Trp Gln	
50	
CGGCGGGGCG GGTGAGGGCG GCTGCCTTCC TCATGCCAAC TCCTGCCACC TGCAG GG	4308
Gly	
ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC ACC ATC	4356
Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala Thr Ile	
55 60 65	
ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC AAC ATT	4404
Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu Asn Ile	
70 75 80 85	
TGG GTG CCC CAG GGC AGG AAG CAA G GTCTGCCTCC CCTCTACTCC	4449
Trp Val Pro Gln Gly Arg Lys Gln	
90	
CCAAGGGACC CTCCCATGCA GCCACTGCCC CGGGTCTACT CCTGGCTTGA GTCTGGGGGC	4509
TGCAAAGCTG AACTTCCATG AAATCCCACA GAGGCGGGGA GGGGAGCGCC CACTGCCGTT	4569

GCCCAGCCTG GGGCAGGGCA GCGCCTTGGA GCACCTCCCT GTCTTGGCCC CAGGCACCTG	4629
CTGCAÇAGGG ACAGGGGACC GGCTGGAGAC AGGGCCAGGC GGGGCGTCTG GGGTCACCAG	4689
CCGCTCCCCC ATCTCAG TC TCC CGG GAC CTG CCC GTT ATG ATC TGG ATC Val Ser Arg Asp Leu Pro Val Met Ile Trp Ile 95 100	4738
TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT GGG GCC AAC TTC CTC Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His Gly Ala Asn Phe Leu 105 110 115 120	4786
AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC ACA CGC GGA AAC GTC Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn Val 125 130 135	4834
ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC CTT GGG TTC CTC AGC Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser 140 145 150	4882
ACT GGG GAC GCC AAT CTG CCA G GTGCGTGGGT GCCTTCGGCC CTGAGGTGGG Thr Gly Asp Ala Asn Leu Pro 155	4934
GCGACCAGCA TGCTGAGCCC AGCAGGGAGA TTTTCCTCAG CACCCCTCAC CCCAAACAAC	4994
CAGTGGCGGT TCACAGAAAG ACCCGGAAGC TGGAGTAGAA TCATGAGATG CAGGAGGCCC	5054
TTGGTAGCTG TAGTAAAATA AAAGATGCTG CAGAGGCCGG GAGAGATGGC TCACGCCTGT	5114
AATCCCAGCA CTTTAGGAGG CCCACACAGG TGGGTCACTT GAGCGCAGAA GTTCAAGACC	5174
AGCCTGAAAA TCACTGGGAG ACCCCCATCT CTACACAAAA ATTAAAAATT AGCTGGGGAC	5234
TGGGCGCGGC GGCTCACCTC TGTAAATCCA GCACGTTGGG AGCCCAAGGT GGGTAGATCA	5284
CCTGAGGTCA GGAGTTTGAG ACCAGCCTGA CTAATAAGGA GAAACCTCTT CTCTACTAAA	5354
AATACAAAAT TAGCCAGGCG TGGTGGCGCT TGCCTGTAAT CCCAGCTACT CGGGAGGCTG	5414
AGGCAGGAGA ATCGCTTGAA CTCAGGAGGC GGAGGTTGCG GTGAGCCGAG ATCATGCCAC	5474
TGCACTCCAG CCTGGAGAAC AAGAGTAAAA CTCTGTCTCA AAAAAAAAAA AAAAAAAAAA	5534
ATAGCCAGGC GTGGTATCTC ATGCCTCTGT CCTCAGCTAC CTGGGAGGCA GAGGTGGAAG	5594
GATCGCTTGA GCCCAGGGGT TCAAAGCTGC AGTGAGCCGT GGTGCTGCCA CTGCACTCCA	5654
GCCTGGGCGA CAGAGTGAGG CCCCATCTCA AAAATAAGAG GCTGTGGGAC AGACAGACAG	5714
GCAGACAGGC TGAGGCTCAG AGAGAAACCA GGAGAGCAGA GCTGAGTGAG AGACAGAGAA	5774
CAATACCTTG AGGCAGAGAC AGCTGTGGAC ACAGAAGTGG CAGGACACAG ACAGGAGGGA	5834
CTGGGGCAGG GGCAGGAGAG GTGCATGGGC CTGACCATCC TGCCCCCGAC AAACACCACC	5894
CCCTCCAGCA CCACACCAAC CCAACCTCCT GGGGACCCAC CCCATACAGC ACCGCACCCG	5954
ACTCAGCCTC CTGGGACCCA CCCACTCCAG CAACCAACGT GACCTAGTCT CCTGGGACCC	6014
ACCCCTCCA GCACCCTACC CGACCCAGCT TCTTAGGGAC CCACCATTTG CCAACTGGGC	6074
TCTGCCATGG CCCCACCTCT GTTGAGGGCA TTTCCACCCC ACCTATGCTG ATCTCCCCTC	6134
CTGGAGGCCA GGCCTGGGCC ACTGGTCTCT AGCACCCCCT CCCCTGCCCT GCCCCAG GT Gly 160	6194

AAC TAT GGC CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg 165 170 175	6242
AAT ATC GCG GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly 180 185 190	6290
GAG TCT GCT GGA GGT GCC AGC GTC TCT CTG CAG GTCTCGGGAT CCCTGTGGGG Glu Ser Ala Gly Gly Ala Ser Val Ser Leu Gln 195 200	6343
AGGGCCTGCC CCACAGGTTG AGAGGAAGCT CAAACGGGAA GGGGAGGGTG GGAGGAGGAG	6403
CGTGAGCTG GGGCTGTGGT GCTGGGGTGT CCTTGTCCCA GCGTGGGGTG GGCAGAGTGG	6463
GGAGCGGCCT TGGTGACGGG ATTCTCTGGGT CCCGTAG ACC CTC TCC CCC TAC AAC Thr Leu Ser Pro Tyr Asn 205	6518
AAG GGC CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT Lys Gly Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser 210 215 220 225	6566
CCC TGG GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG Pro Trp Val Ile Gln Lys Asn Pro Leu Phe Trp Ala Lys Lys 230 235	6608
GTAAACGGAG GAGGGCAGGG CTGGGCGGGG TGGGGGCTGT CCACATTTC GTTCTTTATC	6668
CTGGACCCCA TCCTTGCCTT CAAATGGTTC TGAGCCCTGA GCTCCGSCCT CACCTACCTG	6728
CTGGCCTTGG TTCTGCCCCC AG GTG GCT GAG AAG GTG GGT TGC CCT GTG GGT Val Ala Glu Lys Val Gly Cys Pro Val Gly 240 245	6780
GAT GCC GCC AGG ATG GCC CAG TGT CTG AAG GTT ACT GAT CCC CGA GCC Asp Ala Ala Arg Met Ala Gln Cys Leu Lys Val Thr Asp Pro Arg Ala 250 255 260 265	6828
CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA GGC CTG GAG T GTGAGTAGCT Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala Gly Leu Glu 270 275	6878
GCTCGGGTTG GCCCATGGGG TCTCGAGGTG GGGGTTGAGG GGGGTACTGC CAGGGAGTAC	6938
TCCGGAGGAG AGAGGAAGGT GCCAGAGCTG CGGTCTTGTC CTGTCACCAA CTAGCTGGTG	6998
TCTCCCTCG AAGCCCCAG CTGTAAGGA GAGGGGGTGC CGTTTCTTCT TTTTTTTTGA	7058
GATGGAGTCT CACTGTTGCC CAGGCTGGAG TGCAGTGTCA CGATCTCAGC TCACTGCAAC	7118
CTCCACCTCC TGGGTTCAAG TGATTCTCTG ACTCAACCTC CCATGTAGCT GGGACTACAG	7178
GCACATGCCA CCATGCCCAG ATAATTTTTC TGTGTGTTTA GTAGGGATGG AGTTTCATCG	7238
TGTTAGCTAG GATGATCTCG GTCTTGGGAC CTCATGATCT GCCCACCTCG GCCTCCCAAA	7298
GTGCTGGAAT TACAGGCGTG AGCCACTGTG CCCGGCCCCCT TCTTTATTCT TATCTCCCAT	7358
GAGTTACAGA CTCCCCTTTG AGAAGCTGAT GAACATTTGG GGCCCCCTCC CCCACCTCAT	7418
GCATTCATAT GCAGTCATTT GCATATAATT TTAGGGAGAC TCATAGACCT CAGACCAAGA	7478
GCCTTTGTGC TAGATGACCG TTCATTCAAT CGTTCATTCA TTCAGCAAAC ATTTACTGAA	7538
CCGTAGCACT GGGGCCCAGC CTCCAGCTCC ACTATTCTGT ACCCCGGGAA GGCTTGGGGA	7598
CCCATTCCAC AAACACCTCT GCATGTCAGC CTTACCAGCT TGCTACGCTA AGGCTGTCCC	7658

TCACTCATTC TTCTATGGCA ACATGCCATG AAGCCAAGTC ATCTGCACGT TTACCTGACA	7718
TGAGCTCAAC TGCACGGGCT GGACAAGCCC AAACAAAGCA ACCCCCACGG CCCCCTAGA	7778
AGCAAAACCT GCTGTGCTGG GCCCAGTGAC AGCCAGGCCC CGCCTGCCTC AGCAGCCACT	7838
GGGTCTCTA GGGGCCCCGTC CAGGGGTCTG GAGTACAATG CAGACCTCCC ACCATTTTGTG	7898
GCTGATGGAC TGAACCCAG CCCTGAGAGA GGGAGCTCCT TCTCCATCAG TTCCCTCAGT	7958
GGCTTCTAAG TTTCCTCCTT CCTGCTTCAG GCCCAGCAAA GAGAGAGAGG AGAGGGAGGG	8018
GCTGCCGCTG AAGAGGACAG ATCTGGCCCT AGACAGTGAC TCTCAGCCTG GGGACGTGTG	8078
GCAGGGCCTG GAGACATCTG TGATTGTAC AGCTGGGGAG GGGGTGCTCC TGGCACCTCG	8138
TGGGTCGAGG CCGGGGATGC TCTAAACATC CTACAGGGCA CAGGATGCCC CTGATGGTGC	8198
AGAATCAACC CTGCCCCAAG TGTCCATAGA TCAGAGAAGG GAGGACATAG CCAATTCCAG	8258
CCCTGAGAGG CAAGGGGCGG CTCAGGGGAA ACTGGGAGGT ACAAGAACCT GCTAACCTGC	8318
TGGCTCTCCC ACCCAG AC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT	8366
Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro	
280 285	
GTC ATT GAT GGA GAC TTC ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC	8414
Val Ile Asp Gly Asp Phe Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala	
290 295 300 305	
AAC GCC GCC GAC ATC GAC TAT ATA GCA GGC ACC AAC AAC ATG GAC GGC	8462
Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly Thr Asn Asn Met Asp Gly	
310 315 320	
CAC ATC TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC AAC AAG	8510
His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly Asn Lys	
325 330 335	
AAA GTC ACG GA GTAAGCAGGG GGCACAGGAC TCAGGGGCGA CCCGTGCGGG	8561
Lys Val Thr Glu	
340	
AGGGCCGCCG GGAAAGCACT GGCGAGGGGG CCAGCCTGGA GGAGGAAGGC ATTGAGTGGA	8621
GGACTGGGAG TGAGGAAGTT AGCACCGGTC GGGGTGAGTA TGCACACACC TTCCTGTTGG	8681
CACAGGCTGA GTGTCAGTGC CTACTTGATT CCCCAG G GAG GAC TTC TAC AAG	8734
Glu Asp Phe Tyr Lys	
345	
CTG GTC AGT GAG TTC ACA ATC ACC AAG GGG CTC AGA GGC GCC AAG ACG	8782
Leu Val Ser Glu Phe Thr Ile Thr Lys Gly Leu Arg Gly Ala Lys Thr	
350 355 360	
ACC TTT GAT GTC TAC ACC GAG TCC TGG GCC CAG GAC CCA TCC CAG GAG	8830
Thr Phe Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro Ser Gln Glu	
365 370 375	
AAT AAG AAG AAG ACT GTG GTG GAC TTT GAG ACC GAT GTC CTC TTC CTG	8878
Asn Lys Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe Leu	
380 385 390	
GTG CCC ACC GAG ATT GCC CTA GCC CAG CAC AGA GCC AAT GCC AA	8922
Val Pro Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys	
395 400 405	
GTGAGGATCT GGGCAGCGGG TGGCTCCTGG GGGCCTTCCT GGGGTGCTGC ACCTTCCAGC	8982
CGAGGCCTCG CTGTGGGTGG CTCTCAGGTG TCTGGGTGTG CTGGGAAAGT GGTGCTTGAG	9042

TCCCCACCTG TGCCTGCCTG ATCCACTTTG CTGAGGCCTG GCAAGACTTG AGGGCCTCTT	9102
TTTACCTCCC AGCCTACAGG GCTTTACAAA CCCTATGATC CTCTGCCCTG CTCAGCCCTG	9162
CACCCCATGG TCCTTCCCAC TGGAGAGTTC TTGAGCTACC TTCCATCCCC CATGCTGTGT	9222
GCACTGAGAG AACACTGGAC AATAGTTTCT ATCCACTGAC TCTTATGGGC CTCAACTTTG	9282
CCCATAATTT CAGCCCACCA CCACATTAAA AATCTTCATG TAATAATAGC CAATTATAAT	9342
AAAAAATAAG GCCAGACACA GTAGCTCATG CCTGTAATCC CAGCACATTG GGAGGTCAAG	9402
GTGGGAGGAT CACTTGAGGT CAGGAGTCTG AGACTAGTCT GGCCAACATG GCAAACCCC	9462
ATCTCTACTA AAAATACAAA AATTATCCAG GCATGGTGGT GCATGCCTAT AATCCTAGCT	9522
ACTCAGGAGG CTGAGGTAGC AGAATTGATT GACCCAGGGA GGTGGAGGTT GCAGTGAGCC	9582
GAGATTACGC CACTGCACTC CAGCAGGGGC AACAGAGTGA GACTGTGTCT CGAATAAATA	9642
AGTAAATAAA TAATAAAAAT AAAAAATAAG TTAGGAATAC GAAAAAGATA GGAAGATAAA	9702
AGTATACCTA GAAGTCTAGG ATGAAAGCTT TGCAGCAACT AAGCAGTACA TTTAGCTGTG	9762
AGCCTCCTTT CAGTCAAGGC AAAAAGGGAA ACAGTTGAGG GCCTATACCT TGTCCAATCT	9822
AATTGAAGAA TGCACATTCA CTTGGAGAGC AAAATATTTC TTGATACTGA ATTCTAGAAG	9882
GAAGGTGCCT CACAATGTTT TGTGGAGGTG AAGTATAAAT TCAGCTGAAA TTGTGGAACC	9942
CATGAATCCA TGAATTTGGT TCTCAGCTTT CCCTCCCTG GGTGTAAGAA GCCCCATCTC	10002
TTCATGTGAA TTCCCCAGAC ACTTCCCTGC CCACTGCCCG GGACCTCCCT CCAAGTCCGG	10062
TCTCTGGGCT GATCGGTCCC CAGTGAGCAC CCTGCCTACT TGGGTGGTCT CTCCCCTCCA	10122
G G AGT GCC AAG ACC TAC GCC TAC CTG TTT TCC CAT CCC TCT CGG ATG Ser Ala Lys Thr Tyr Ala Tyr Leu Phe Ser His Pro Ser Arg Met 410 415 420	10169
CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT GCA GAT GAC ATT CAG Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His Ala Asp Asp Ile Gln 425 430 435 440	10217
TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG GGC TAC CGG CCC CAA Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro Gln 445 450 455	10265
GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC TGG ACC AAC TTT GCC Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala 460 465 470	10313
AAA ACA GG GTAAGACGTG .GGTTGAGTGC AGGGCCGAGG GCCACAGCCG Lys Thr Gly 475	10361
AGAAGGGCCT CCCACCACGA GGCCTTGTTT CCTCATTTGC CAGTGGAGGG ACTTTGGGCA	10421
AGTCACTTAA CCTCCCCCTG CATCGGAATC CATGTGTGTT TGAGGATGAG AGTTACTGGC	10481
AGAGCCCCAA GCCCATGCAC GTGCACAGCC AGTGCCAGT ATGCAGTGAG GGGCATGGTG	10541
CCCAGGGCCA GCTCAGAGGG CGGGGATGGC TCAGGCGTGC AGGTGGAGAG CAGGGCTTCA	10601
GCCCCCTGGG AGTCCCCAGC CCCTGCACAG CCTCTTCTCA CTCTGCAG G GAC CCC Asp Pro	10656
AAC ATG GGC GAC TCG GCT GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr 480 485 490	10704

GAA AAC AGC GGC TAC CTG GAG ATC ACC AAG AAG ATG GGC AGC AGC TCC Glu Asn Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser 495 500 505	10752
ATG AAG CCG AGC CTG AGA ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC Met Lys Arg Ser Leu Arg Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr 510 515 520 525	10800
TAT CTG GCG CTG CCC ACA GTG ACC GAC CAG GAG GCC ACC CCT GTG CCC Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln Glu Ala Thr Pro Val Pro 530 535 540	10848
CCC ACA GGG GAC TCC GAG GCC ACT CCC GTG CCC CCC ACG GGT GAC TCC Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser 545 550 555	10896
GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val 560 565 570	10944
CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp 575 580 585	10992
TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro 590 595 600 605	11040
GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly 610 615 620	11088
GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC GCC CCC Asp Ser Gly Ala Pro Pro Val Pro Thr Gly Asp Ser Gly Ala Pro 625 630 635	11136
CCC GTG CCG CCC ACG GGT GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG Pro Val Pro Pro Thr Gly Asp Ala Gly Pro Pro Pro Val Pro Pro Thr 640 645 650	11184
GGT GAC TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala 655 660 665	11232
CCC CCC GTG ACC CCC ACG GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC Pro Pro Val Thr Pro Thr Gly Asp Ser Glu Thr Ala Pro Val Pro Pro 670 675 680 685	11280
ACG GGT GAC TCC GGG GCC CCC CCT GTG CCC CCC ACG GGT GAC TCT GAG Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Glu 690 695 700	11328
GCT GCC CCT GTG CCC CCC ACA GAT GAC TCC AAG GAA GCT CAG ATG CCT Ala Ala Pro Val Pro Pro Thr Asp Asp Ser Lys Glu Ala Gln Met Pro 705 710 715	11376
GCA GTC ATT AGG TTT TAGCGTCCCA TGAGCCTTGG TATCAAGAGG CCACAAGAGT Ala Val Ile Arg Phe 720	11431
GGGACCCAG GGGCTCCCCCT CCCATCTTGA GCTCTTCCCTG AATAAAGCCT CATACCCCTG	11491
TCGGTGTCTT TCTTTGCTCC CAAGGCTAAG CTGCAGGATC	11531

Таблица 1

Корреляция CEL последовательностей с хромосомами человека в 16 клеточных гибридах типа человек-грызун

ПРОЦЕНТ КЛЕТОК С ХРОМОСОМАМИ ЧЕЛОВЕКА

ХРОМОСОМА	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	CEL
ГИБРИД																									
GM09925	74	24	0	74	76	60	82	78	0	0	4	68	6	86	78	14	98	96	46	84	0	76	0	0	-
GM09927	69	83	75	77	0	93	79	73	0	82	0	0	77	79	90	0	81	73	87	89	0	0	0	0	-
GM09929	0	0	61	59	0	43	2	49	0	0	33	49	0	59	2	0	96	0	2	31	0	0	2	0	-
GM09930A	0	34	62	4	12	0	26	4	0	0	6	22	56	82	12	0	86	78	0	22	82	76	6	8	-
GM09932	0	0	0	68	86	46	0	80	0	2	28	26	0	0	0	0	96	0	2	0	92	0	0	0	-
GM09933	50	0	84	16	54	76	92	54	0	6	0	50	84	78	92	0	88	70	80	32	94	88	0	32	-
GM09934	0	50	0	0	83	79	4	87	0	0	77	87	0	2	89	0	90	89	0	91	89	2	0	0	-
GM09935A	0	0	52	10	28	12	0	0	0	8	0	22	74	72	0	0	93	59	0	9	91	71	0	0	-
GM09936	0	0	0	18	0	46	70	10	0	16	34	0	2	88	2	0	100	0	44	24	0	18	0	0	-
GM09937	0	0	54	38	0	62	54	70	0	4	0	42	0	70	60	0	96	66	0	0	0	0	0	0	-
GM09938	0	0	2	88	60	88	86	4	0	0	36	92	0	80	4	0	92	0	4	80	76	60	0	2	-
GM09940	0	0	46	0	0	0	84	62	0	0	0	0	0	0	62	0	100	0	0	0	0	0	0	0	-
GM10324	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	90	0	-
GM10567	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	0	0	0	0	0	0	0	0	-
GM10611	0	0	0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+
GM10095	0	0	0	0	0	0	0	0	94b	0	0	0	0	0	0	0	0	0	0	0	0	0	94b	0	-
Коэффициент	4	5	8	7	7	10	9	9	0	2	6	10	5	10	7	2	13	8	5	9	7	6	3	2	
набор клеток	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	

а) В целом, хромосома человека должна присутствовать в количестве 20-22% на клетку, чтобы ее можно было определить саузерн блот-анализом.

б) Содержит 9pter → 234 и Xq13 → qter.

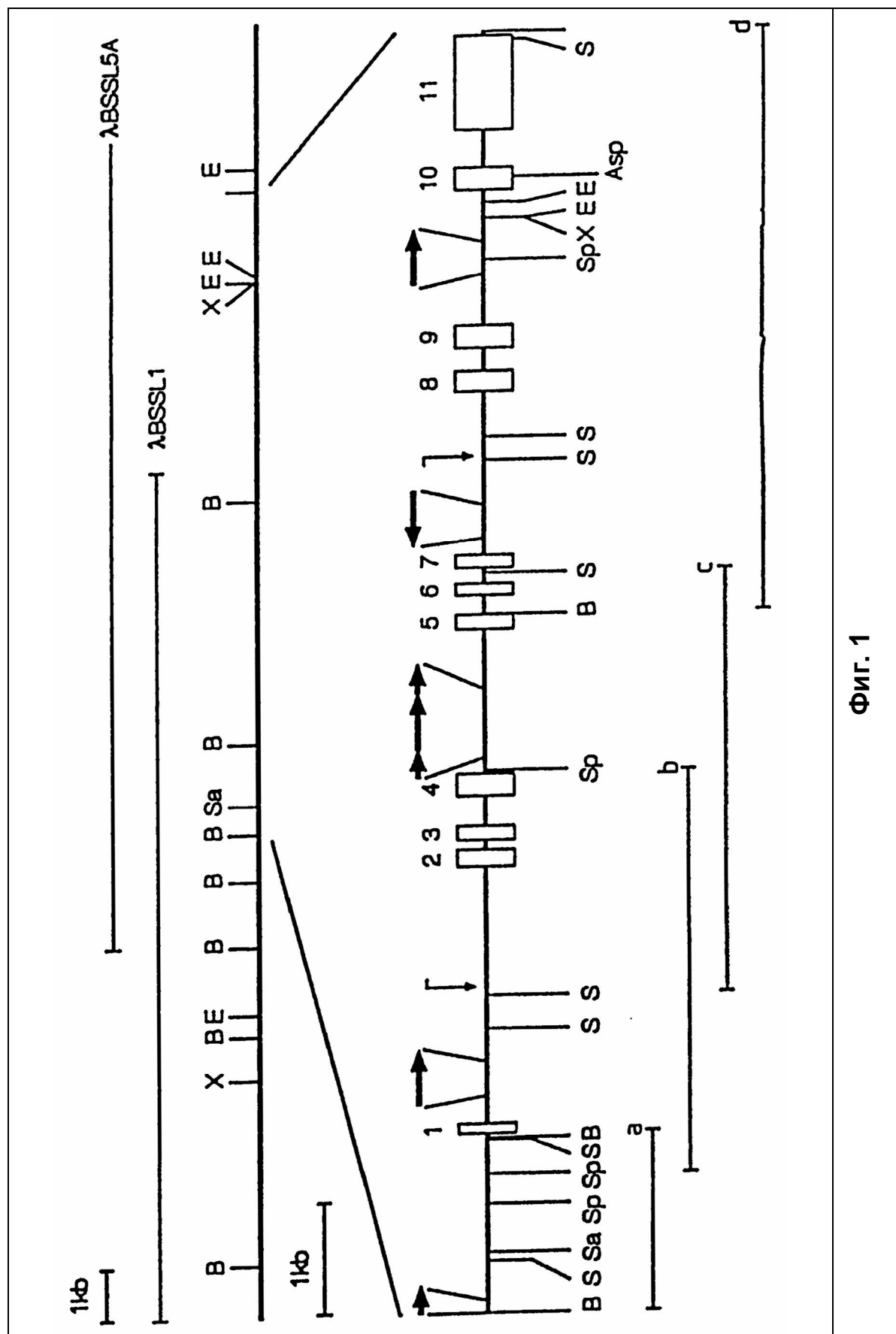
Применяемые для амплификации ДНК праймеры			Таблица 2
Олигонуклеотид	Положение нук-тидов ^а	Амплифицированная последовательность	
P1: 5'-AGACCTACGCGCTACCTG-3'	8492-8508	Экзон 10	
P2: 5'-TCCACCTAGCGGCAATGATG-3'	8646-8662		
P3: 5'-GACCGATGTCCTCTCTCTCTCTG-3'	7220-7239	Экзон 10 с праймерами из экзонав,	
P4: 5'-CAGCGGAGTCCGCGGATGTTG-3'	9016-9035	оказующих экзон 10 ^б	
P5: 5'-ACCAAGAGATGCGGACGACG-3'	9089-9109	Повтор в экзоне 11	
P6: 5'-GACTGACGCGGATCTGAGCTTC-3'	9722-9742		
^а Положение нуклеотида дается, как число оснований от начала первого экзона. Для сравнения по- ложения нуклеотида с ПОСЛЕД. № 1 следует добавить 1640 оснований к приведенной в столбце циф- ре.			
^б Для амплификации "экзона 10" из кДНК.			

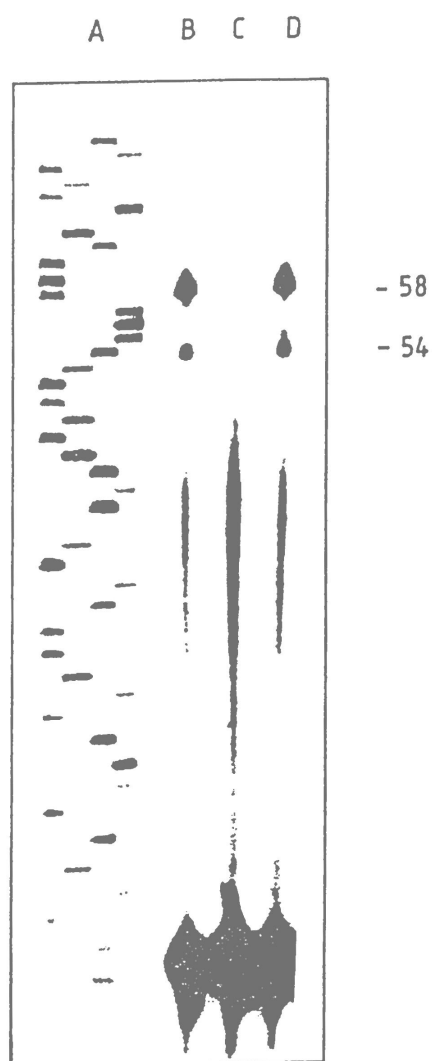
Таблица 3

Экзон-интронная организация CEL-гена

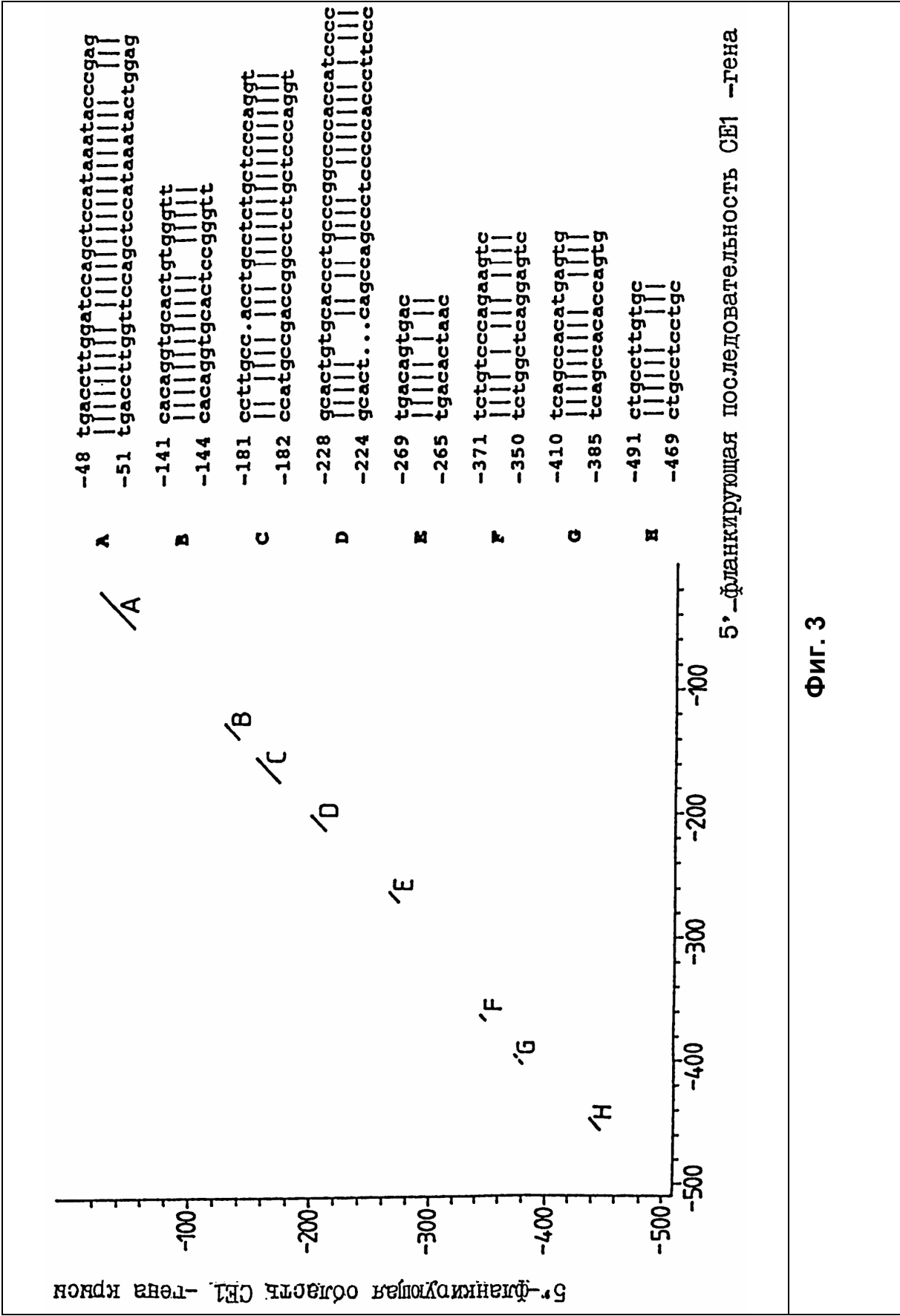
Экзон			аминокислота		последовательность у стыка экзон-интрон		Интрон	
№	Положение нуклеотида ^a	Длина (н-тип)	Полож. Число	5'-сплайс-донор	3'-сплайс-акцептор	№	Длина (н-тип)	
1	1- 87	87	1- 25 (25)	GCC GCG AAG gtaaga....gtgtctccctctgcaag	CTG GGC GCC	I	2343	
2	2431-2581	151	26- 75 (50)	TGG CAA G gtggga....tctgtgcacacctgcaag	GG ACC CTG	II	85	
3	2667-2789	123	76-116 (41)	AAG CAA G gtctgc....gtctcccccacatctcgaag	TC TCC CGG	III	277	
4	3067-3264	198	117-182 (66)	CTG CCA G gtgcgt....ctgcacctgcccccaag	GT AAC TAT	IV	1288	
5	4553-4683	131	183-226 (44)	TCT CTG CAG gtctcg....ttcttgggtccctgtag	ACC CTC TCC	V	177	
6	4861-4968	108	227-262 (36)	GCC AAA AAG gtaaac....tggtttctgcccccaag	GTG GCT GAG	VI	142	
7	5111-5228	118	263-301 (39)	CTG GAG T gtgagt....gggtctctcccaacccaag	AC CCC ATG	VII	1466	
8	6695-6881	187	302-364 (63)	GTC ACG GA gtaagc....acttgattctcccccag	G GAG GAC	VIII	197	
9	7079-7282	204	365-432 (68)	AAT GCC AA gtgagg....gtctctccctccctcag	G AGT GCC	IX	1201	
10	8484-8681	198	433-498 (66)	AAA ACA GG gtaaga....cttctctcacctctgcaag	G GAC CCC	X	328	
11	9010-9850	841	499-745 (247)					

^aПоложение нуклеотида указано в виде числа оснований от начала первого экзона. Для сравнения положения нуклеотида с ПОСЛЕД. № 1 следует добавить 1640 оснований к цифре в столбце.



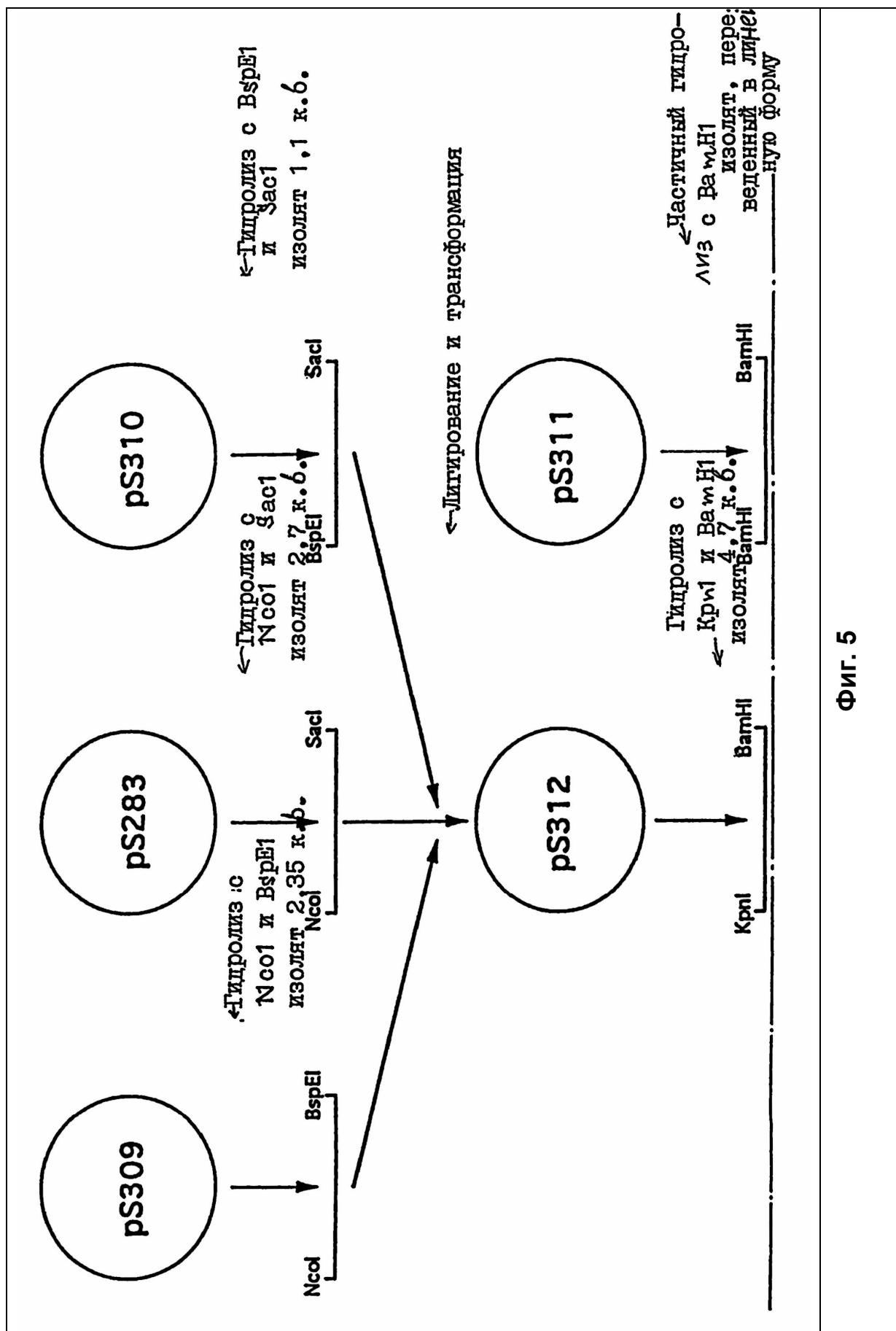


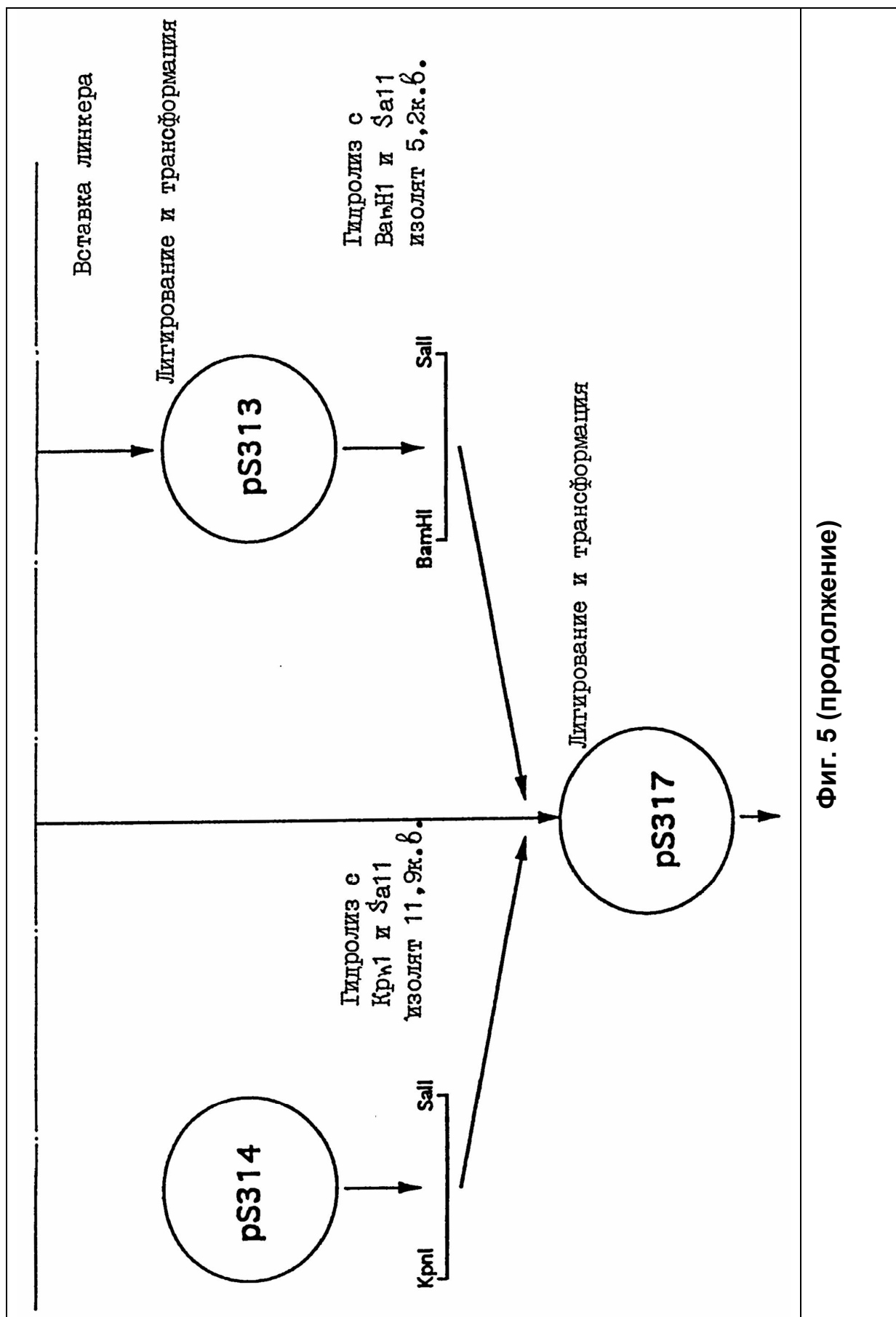
Фиг. 2



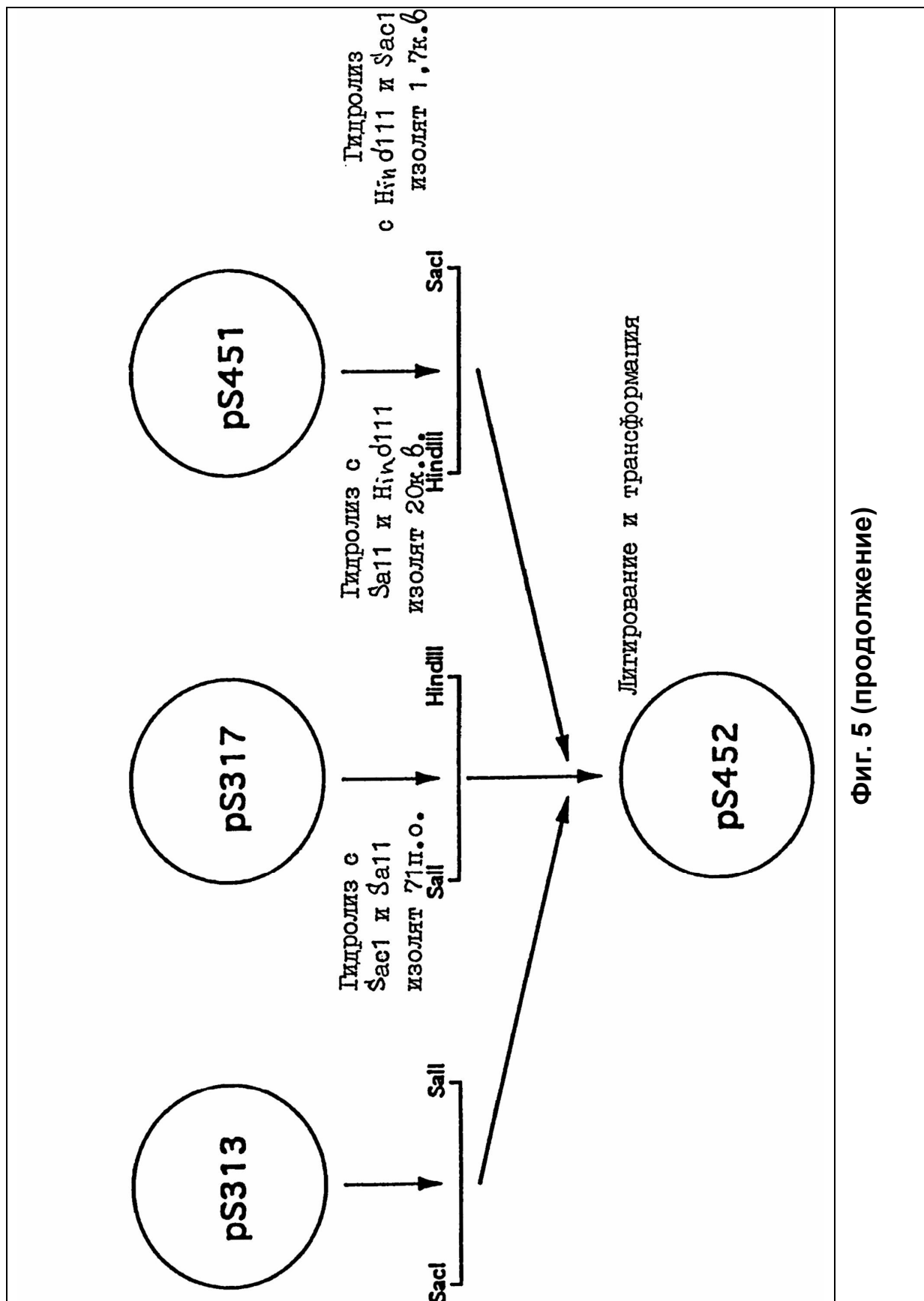
-1640 GGATCCCTCG AACCCAGGAG TTCAAGACTG CAGTGAGCTA TGATTGTGCC ACTGCACTCT AGCCTGGGTG ACAGAGACCC TGTCTCAAAA AAACAACAA
 -1540 ACAAATAAACC TCTGTGACT CCGGGTGATA ATGACATGTC AATGTGGATT CATCAGGTGT ACCCCCTGGT GGGGGATGTT GATAACGGGG
 -1440 GAGACTGGAG TGGGGCGAGG ACATACGGGA AATCTCTGTA ATCTTCCTCT AATTTTGTG TGAAACCTAA GCTGCTCTAA AAATGTACAT AGATATAAAC
 -1340 TGGGGCCTTC CTTTCCTCT GCGCTGCCCC AGCCCTCCCC CACTCTCTTC CTCTCCCTGC TGCCTCCCT CTGCCCTCCC CTTTCCTCCT TAGCCACTGT
 -1240 AAATGACACT GCAGCAAGG TCTAGGCAAA ATGCTTTTC CCTGGGGCGC CCAGCCACC TGCAGGCCCT TTATTTCTCTG TGGCCGAGCT CCTCTCTCCA
 -1140 CCTTCCAGTC CTTTCCCCAG CCTCCCTGCG CACTAGGCC TCCTGAATTG CTGGCACCGG CTGTGGTGA CAGACAGAGG GACAGACGTG GCTCTGCAAG
 -1040 TCCACTCGGT CCCTGGCAAC GCGCGCAGGG GTGGCAGAAC GGGAGTGTGG TTGGTGTGG AAGCACAGGC CCCAGTGTCT CCTGGGGGAC TGTGGGTGG
 -940 GAAGGCTCTG GCTGCCCTCA CCTGTTCCTC ATCACTGCAG AGGGCTGTGC GGTGGCTGA GCTGCCACTG AGTGTCTCTGG TGAGGGTGAC CTCACACTGG
 -840 CTGAGCTTAA AGGCCCAATC TGAAGACTTT GTTCGTGGTG TTCTTTCACT TCTCAGAGCC TTTCCTGGCT CCAGGATTA TACCTGTCTCA CAGAAATAC
 -740 GAGTCGCCCTC CTCTCCACA ACCTCACAGG ACCTTCCTCC TTCCCTCCCG CTGGCCTCTT TCCCTCCCT TCTGTCTCTC TGCTGGGCA TGCCCCAGGG
 -640 CCTCGGCTGG GCGCTTTGTT TCCACAGGGA AACCTACATG GTTGGGCTAG ATGCTCTCGC ACCCCCCAC CCACACCCCT TGAGCCTCTA GTCTCTCCTC
 -540 CCAGGACACA TCAGGCTGGA TGGTGACACT TCCACACCTT TGAGTGGGAC TCCCTTGTGC TGTCTTGGGA TTCCGACCCA GCTTTGGACTA CCGGCTCCAC
 -440 GGGCCCCAGG AAAAGCTCGT ACAGATAAGG TCAOCCACAT GAGTGGAGGG CCTGCAGCAT GCTGCCCTTT CTGTCCCAAG AGTCACCTGC TCGGTCCCT
 -340 CTGAAGCCCC TTGTGGGACC TAGGGGACAA GCAGGGCATG GAGACATGGA GACAAAGTAT GCGCTTTTCT CTGACAGTGA CACCAAGCCC TGTGAACAAA
 -240 CCAGAAAGCA GGCACCTGT CACCTTCCCC GCGCCCAACCA TCCCCCTTAC CACCGGCCAC CTTCGCCACT GCTCTCTCTC CCAAGTAAAT GGTAAACCTGC
 -140 ACAGCTCCAC TGTGGGTTG GGGAAACTG GATCTCCCTG CACCTGAGG GGTAGAGGG AGGGAGTGC TGAGAGCTCA TGAACAAGCA TGTGACCTTG
 -40 GATCCAGCTC CATTAATACC CAGGCCCCAG GGGAGGGCC

Фиг. 4

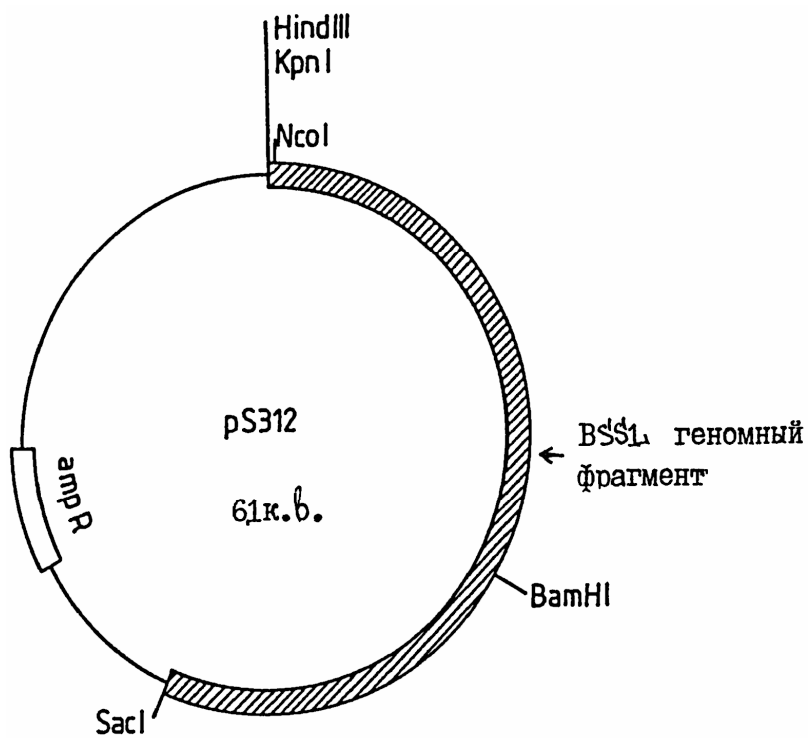




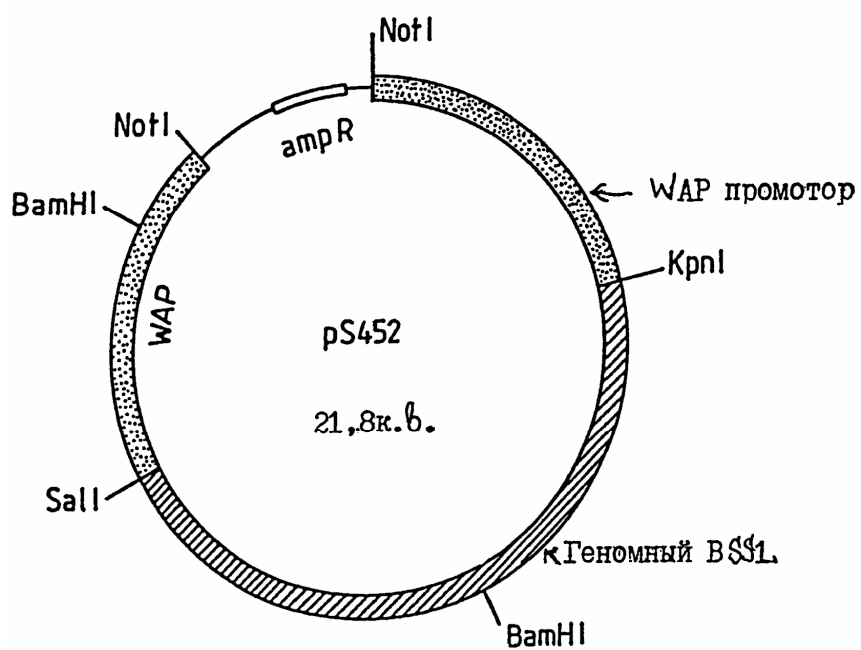
Фиг. 5 (продолжение)



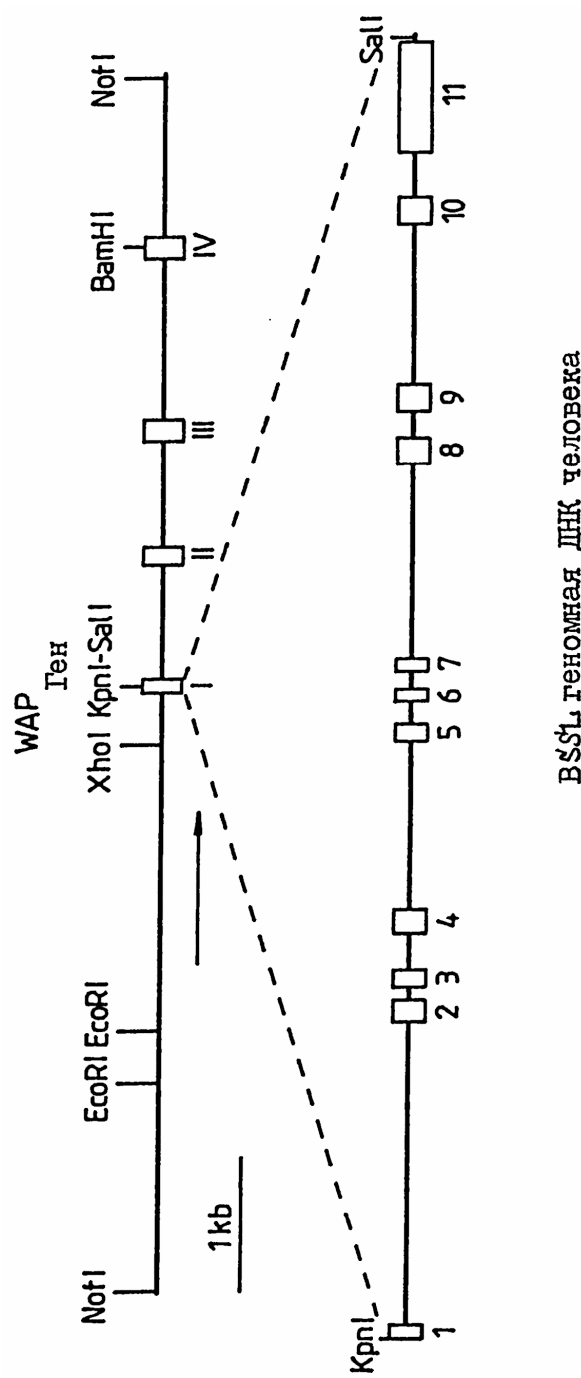
Фиг. 5 (продолжение)



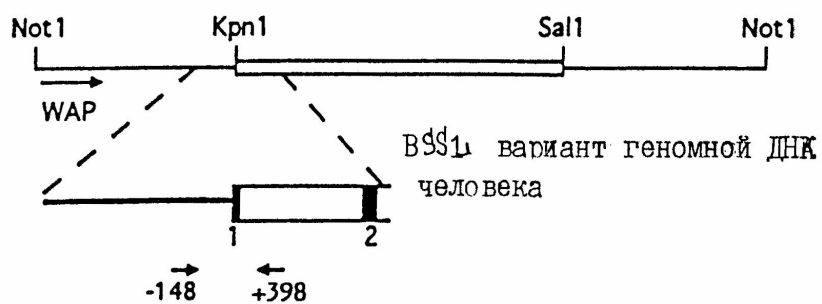
Фиг. 6



Фиг. 7



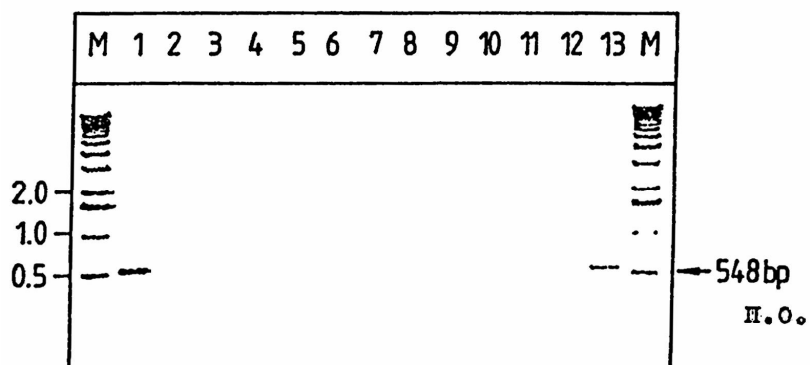
Фиг. 8



Фиг. 9А

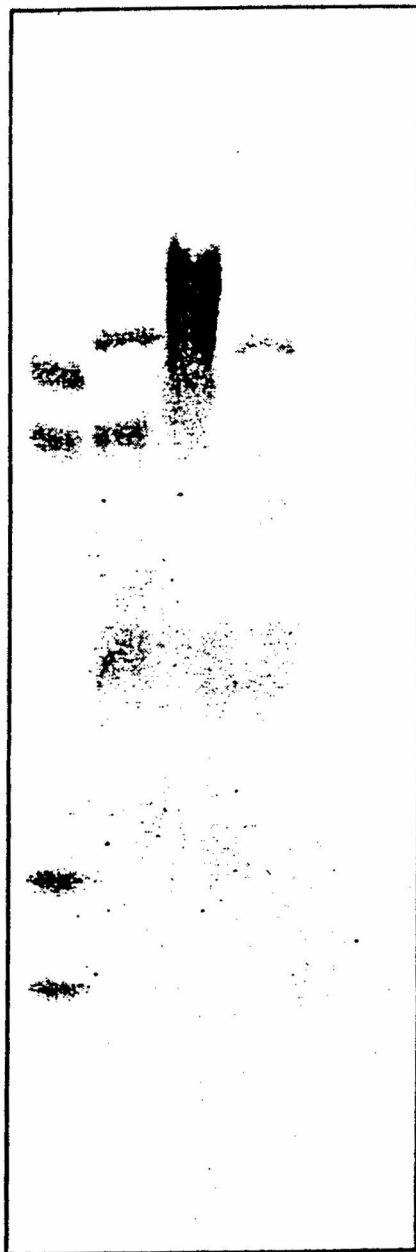
Праймер	Последовательность (5'-3')
5'-праймер	СТГТГТGGCAAGAAGGAAGTGTGT
3'-праймер	СAACTCCTGACCTCAAGTGATC

Фиг. 9В



Фиг. 9С

1 2 3 4 5 6



Фиг. 10

ДП "Український інститут промислової власності" (Укрпатент)
Україна, 01133, Київ-133, бульв. Лесі Українки, 26
(044) 295-81-42, 295-61-97

Підписано до друку _____ 2002 р. Формат 60x84 1/8.
Обсяг _____ обл.-вид. арк. Тираж 50 прим. Зам. _____

УкрІНТЕІ, 03680, Київ-39 МСП, вул. Горького, 180.
(044) 268-25-22
