

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI0706424-1 A2**

(22) Data de Depósito: 05/01/2007
(43) Data da Publicação: 29/03/2011
(RPI 2099)



(51) *Int.Cl.:*
C12Q 1/68

(54) Título: **MARCADORES DE CALPASTATIN PARA FERTILIDADE E LONGEVIDADE**

(30) Prioridade Unionista: 05/01/2006 US 60/756,383

(73) Titular(es): Washington State University Research Foundation

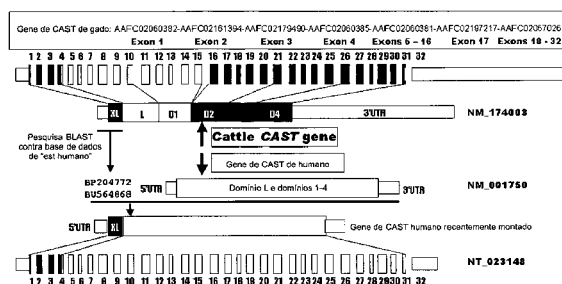
(72) Inventor(es): Jennifer J. Michal, Matthew D. Garcia, Zhihua Jiang

(74) Procurador(es): Dannemann, Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT US2007000252 de 05/01/2007

(87) Publicação Internacional: WO 2007/081788 de 19/07/2007

(57) **Resumo:** MARCADORES DE CALPASTATIN PARA FERTILIDADE E LONGEVIDADE. A presente invenção refere-se a composições e métodos baseados em novos marcadores genéticos calpastatin (CAST), tais como mutações sem sentido em exon 3 que resultem em substituições G48D ou P52L (NM174003.2:c.271G>.A e 283C>T), uma substituição G/T em intron 3 (A- AFC02060381 .1 :g.21 1OG>T) e uma repetição GAAA em intron 8 (A- AFC0206038 1.1 :g . 6700[(GAAA)4]+[(GAAA)5]. Aspectos particulares proporcionam novos marcadores para fertilidade (por exemplo, taxa de gravidez de raça, DPR) e longevidade (por exemplo, vida produtiva, PL) em, por exemplo, gado leiteiro. Aspectos adicionais proporcionam novos métodos compreendendo seleção auxiliada por marcador para aperfeiçoar fertilidade e/ou longevidade em gado leiteiro. Portanto, em concretizações particulares, uma combinação de seleção genética baseada em um ou mais dos novos marcadores CAST, e altos potenciais de PTA de traços de produção de leite, proporciona traços reprodutivos aperfeiçoados em associação com traços de produção de leite alta continuada. Aspectos adicionais desferem um domínio XL anteriormente não-reconhecidos no gene CAST humano, e, desse modo, proporciona o uso de mutantes/variantes de domínio de CAST XL como marcadores para fertilidade e longevidade humanas.



Relatório Descritivo da Patente de Invenção para "**MARCADORES DE CALPASTATIN PARA FERTILIDADE E LONGEVIDADE**".

INCORPORAÇÃO POR REFERÊNCIA

Este pedido reivindica o benefício do pedido de patente provisório dos Estados Unidos Nº de Série 60/756.383, depositado em 5 de janeiro de 2006. Referência é feita ao pedido de patente internacional PCT/IB2005/002983, depositado em 18 de julho de 2005, e publicado como WO 2006/097787 em 21 de setembro de 2006.

Os pedidos precedentes, e todos os documentos desse modo citados, ou durante seu prosseguimento ("documentos citados de pedido"), e todos os documentos citados ou referenciados nos documentos citados de pedido, e todos os documentos citados ou referenciados aqui ("documentos citados aqui"), e todos os documentos citados ou referenciados nos documentos aqui citados, juntos com as instruções do fabricante, descrições, especificações de produto, e folhas de produto para quaisquer produtos aqui mencionados, ou em qualquer documento incorporado por referência aqui, são, desse modo, incorporados aqui por referência, e podem ser empregados na prática da invenção.

CAMPO DA INVENÇÃO

A presente invenção refere-se à identificação de marcadores genéticos (polimorfismos de nucleotídeo simples (SNPs) e uma repetição em tandem curta (STR)) no interior de genes bovinos que codificam calpastatin ("CAST"), e suas associações com traços economicamente relevantes na produção de Leiteira. A invenção adicionalmente se refere a métodos e sistemas, incluindo processos baseados em rede, para controlar os dados de SNP/STR e outros dados relacionados a animais específicos e rebanhos de animais, cuidado veterinário, diagnóstico e dados de controle de qualidade, e gerenciamento de criação que, baseado na genotipagem, tem fertilidade prognosticável e traços de longevidade, condições de criação de gado, bem-estar do animal, informação de segurança de alimentação, exame dos processos existentes, e dados das locações de campo.

ANTECEDENTES DA INVENÇÃO

O declínio reprodutivo tem sido um desafio na indústria de leiteira ao redor do mundo por várias décadas, e tem tipicamente sido responsável na seleção de produção de leite aumentada [Sheldon IM, Dobson H. Reproductive challenges facing the cattle industry at the beginning of the 21st century. *Reprod Suppl.* 2003; 61:1-13]. Nos Estados Unidos somente, o primeiro serviço para taxa de concepção declinou em aproximadamente 65% em 1951 a 40% em 1996 [Butler WR. Review: effect of protein nutrition on ovarian and uterine physiology in dairy cattle. *J Dairy Sci.* 1998; 81:2533-2539], enquanto o número de serviços por concepção aumentou de aproximadamente 1,8 em 1970 a aproximadamente 3 em 2000 [Lucy MC. Reproductive loss in high-producing dairy cattle: where will it end? *J Dairy Sci.* 2001; 84:1277-1293]. Foi reportado que durante 1976 a 1978, o número médio de dias abertos (dias entre parto e a concepção subsequente) foi 122 dias para Jerseys e 124 dias para Holsteins [Washburn SP, Silvia WJ, Brown CH, McDaniel BT, McAllister AJ. Trends in reproductive performance in Southeastern Holstein and Jersey DHI herds. *J Dairy Sci.* 2002; 85:244-251]. Contudo, os dias abertos aumentaram para 152 dias para Jerseys e 168 dias para Holsteins por 1997 a 1999. No Reino Unido, a taxa de concepção para primeiro serviço de vacas leiteiras declinou de 65,4% (1975-82) para 44,3% (1995-98) a uma taxa de aproximadamente 1% por ano [Royal MD, Darwash AO, Flint AP, Webb R, Woolliams JA, Lamming GE. Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Anim Sci* 2000; 70:487-502]. Diminuições equivalentes na taxa de concepção de primeiro serviço foram também observadas no gado leiteiro em Ireland [Roche JF, Mackey D, Diskin MD. Reproductive management of postpartum cows. *Anim Reprod Sci* 2000; 61:703-712] e Australia [Macmillan KL, Lean IJ, Westwood CT. Os efeitos de lactação na fertilidade de vacas leiteiras. *Aust Vet J* 1996; 73:141-147]. A compreensão das causas genéticas de fertilidade reduzida é essencial para parar o declínio de fertilidade atualmente observado na lactação de vacas leiteiras.

Para se discutir as dificuldades de se alcançar níveis desejados de desempenho reprodutivo nos rebanhos de ordenha de hoje, um novo tra-

ço de fertilidade, a taxa de gravidez da raça (DPR), foi introduzido como um indicador de fertilidade de progenitor para seleção genética [VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD, Kuhn MT, Wiggans GR. O desenvolvimento de uma avaliação genética nacional para fertilidade de vaca. Dairy Sci 2004; 87:2285-2292]. A taxa de gravidez é definida como a percentagem de vacas não-grávidas que se tornam grávidas durante cada período de 21 dias. De fato, os dados para calcular a DPR são tomados a partir dos dias abertos reportados, que são calculados como data de gravidez menos data de parto anterior. A data de gravidez é determinada a partir da última geração reportada ou do parto subsequente menos comprimento de gestação esperado. Para cálculo de avaliações genéticas, os dias abertos são convertidos em taxa de gravidez da raça por transformação linear de taxa de gravidez = 0,25 (233 - dias abertos). As avaliações são expressas como capacidade de transmissão prognosticada (PTA) para DPR, e cálculos são gerados como um resultado direto de um desempenho de raças de touros [VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD, Kuhn MT, Wiggans GR. Development of a national genetic evaluation for cow fertility. Dairy Sci 2004; 87:2285-2292]. Em adição, a fertilidade da vaca é um componente maior de vida produtiva (PL) ou longevidade. O aperfeiçoamento de ambas DPR e PL conduziria a produtividade e utilidade aumentadas à indústria de leiteira.

O calpastatin (CAST) é um inibidor de protease endógeno que age especificamente em duas proteases independentes de Ca^{2+} , μ -calpain e m-calpain pela ligação e formação de um complexo inativo. O CAST é amplamente expresso em células mamíferas e tecidos, incluindo aqueles relacionados a reprodução. Por exemplo, o gene CAST é expresso na glândula pituitária humana [Kitahara A, Takano E, Ontsuki H, Kirihata Y, Yamagata Y, Knaagi R, Murachi T. Reversed distribution of calpains e calpastatin in human pituitary gland e selective localization of calpastatin in adrenocorticotropin-producing cells as demonstrated by immunohistochemistry. J Clin Endocrinol Metab 1986; 63:343-348], the human placenta [Thompson VF, Saldana S, Cong J, Luedke DM, Goll DE. The calpain system in human placenta. Life

Sci 2002; 70:2493-508], the human oocyte [Ben-Aharon I, Ben-Yosef D, Amit A, Shalgi R. Expression e immunolocalization of the calpain-calpastatin system in the human oocyte. Fertil Steril 2005; 83:1807-1813], the bovine corpus luteum [Orwig KE, Bertrand JE, Or BR, Forsberg NE, Stormshak F. Involvement of protein kinase-C, calpains, e calpastatin in prostaglandin F2 alpha-induced oxytocin secretion from the bovine corpus luteum. Endocrinology 1994; 134:78-83], as well as during spermatogenesis in the tests of humans [Liang ZG, O'Hern PA, Yavetz B, Yavetz H, Goldberg E. Human testis cDNAs identified by sera from infertile patients: a molecular biological approach to immunocontraceptive development. Reprod Fertil Dev 1994; 6:297-305; Li S, Liang ZG, Wang GY, Yavetz B, Kim ED, Goldberg E. Molecular cloning e characterization of functional domains of a human testis-specific isoform of calpastatin. Biol Reprod 2000; 63:172-178 e Wei SG, Wang LF, Miao SY, Zong SD, Koide SS. Expression of the calpastatin gene segment during spermiogenesis in human testis: an in situ hibridization study. Arch Androl 1995; 34:9-12.], mice [Li S, Goldberg E. A novel N-terminal domain directs membrane localization of mouse testis-specific calpastatin. Biol Reprod 2000; 63:1594-1600] e rabbits [Wang LF, Miao SY, Yan YC, Li YH, Zong C, Koide SS. Expression of a sperm protein gene during spermatogenesis in mammalian testis: an in situ hibridization study. Mol Reprod Dev 1990; 26:1-5]. Interessantemente, a proteína de CAST foi identificada como um dos antígenos alvos para anticorpos antiesperma encontrados em mulheres não-férteis [Koide SS, Wang L, Kamada M. Antisperm antibodies associated with infertility: properties e encoding genes of target antigens. Proc Soc Exp Biol Med 2000; 224:123-132]. In vivo, anticorpos CAST anti-BS-17 podem bloquear a capacidade de fertilização de esperma de camundongo para fertilizar ova pela redução significativa dos números de desenvolvimento de embriões [Koide SS, Wang L, Kamada M. Antisperm antibodies associated with infertility: properties e encoding genes of target antigens. Proc Soc Exp Biol Med 2000; 224:123-132]. Todos estes resultados indicam que o gene *CAST* desempenha um papel importante na biologia reprodutiva.

A estrutura primária da seqüência de amino-ácido de calpastatin

inclui quatro domínios internamente repetitivos (Domínios 1-4) e um domínio não-homólogo na extremidade amino-terminal (Domain L) [Murachi T. Calpain e calpastatin. Rinsho Byori 1990; 38:337-346]. Contudo, um novo domínio de peptídeo N-terminal, denominado domínio XL, foi identificado no gado

5 [Cong M, Thompson VF, Goll DE, Antin PB. O promotor de gene calpastatin bovino e uma nova região N-terminal da proteína são alvos para atividade de proteína kinase dependente de cAMP. J Biol Chem 1998; 273:660-666]. A região "XL" contém sessenta e oito aminoácidos, mas não compartilham homologia com outras regiões de calpastatin ou com quaisquer proteínas

10 conhecidas. Experimentos in vivo mostraram que a região "XL" é um substrato para fosforilação por proteína kinase [Cong M, Thompson VF, Goll DE, Antin PB. O promotor de gene calpastatin bovino e uma nova região N-terminal da proteína são alvos para atividade de proteína kinase dependente de cAMP. J Biol Chem 1998; 273:660-666].

15 Permanece vantajoso proporcionar SNPs/STRs adicionais que podem prognosticar mais precisamente os fenótipos de fertilidade e longevidade de um animal, e também um método de negócio que proporciona eficiências de produção aumentadas em gado de criação, bem como proporcionando acesso a vários registros dos animais, e permitem comparações com

20 objetivos esperados e desejados com relação a qualidade e quantidade de animais produzidos.

A citação ou identificação de qualquer documento neste pedido não é uma admissão que tal documento é disponível como técnica anterior a presente invenção.

25 SUMÁRIO DA INVENÇÃO

A presente invenção refere-se à identificação de marcadores genéticos (polimorfismos de nucleotídeo simples (SNPs) e uma repetição em tandem curta (STR)) no interior dos genes bovinos genes que codificam calpastatin ("CAST") e suas associações com traços economicamente relevantes na produção de leiteira.

30

A invenção envolve um método para subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo têm

um polimorfismo similar em um gene de *CAST* que pode compreender a determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de um SNP/STR no gene de *CAST*, e segregação dos animais individuais em subgrupos no qual cada animal em um subgrupo tem um polimorfismo no gene de *CAST*.

A invenção também envolve um método para subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo têm um genótipo similar no gene de *CAST* que pode compreender determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de um polimorfismo de nucleotídeo simples de interesse no gene de *CAST*, e segregação de animais individuais em subgrupos dependendo se os animais têm ou não têm o(s) polimorfismo(s) de nucleotídeo simples de interesse no gene de *CAST*.

O(s) polimorfismo(s) genético(s) de interesse pode(m) ser selecionado(s) a partir do grupo consistindo em mutações sem sentido na região de domínio XL, especialmente mutações sem sentido em exon 3 que resultam em G48D e P52L substituições (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição em tandem curta GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5]. A invenção adicionalmente se refere a um método para subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo têm um genótipo similar no gene de *CAST* que pode compreender determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de qualquer um dos SNPs/STR acima, e segregação de animais individuais em subgrupos dependendo se os animais têm ou não têm qualquer um dos SNPs/STR acima no gene de *CAST*.

A invenção também se refere a método para identificação de um animal tendo um fenótipo desejável conforme comparado à população geral de animais daquela espécie, que pode compreender determinação da presença de um of a polimorfismo de nucleotídeo simples no gene de *CAST* do animal, no qual a presença do SNP/STR é indicativa de um fenótipo desejá-

vel.

Em uma concretização vantajosa, o animal pode ser um bovino. Em outra concretização vantajosa, o *gene de CAST* pode ser um gene de *CAST* bovino.

5 A invenção também envolve métodos auxiliados por computador e sistemas para aperfeiçoamento da eficiência de produção para criação tendo fertilidade e longevidade e, em particular, o genótipo dos animais conforme ele se refere a *CAST* SNPs/STRs. Os métodos da invenção envolvem a obtenção de uma amostra genética de cada animal em um rebanho de
10 criação, determinação do genótipo de cada animal com relação aos traços de qualidade específicos conforme definidos por um painel de pelo menos dois polimorfismos de polinucleotídeo simples (SNPs)/STRs, agrupamento de animais com genótipos similares, e, opcionalmente, subagrupamento adicional de animais baseado nos fenótipos similares. Os métodos da invenção
15 podem também envolver obtenção e manutenção de dados relacionados a animais ou a rebanhos, suas condições de criação de gado, cuidado e condição de saúde e veterinária, história genética ou ascendência, e provisão destes dados a outros através de sistemas que são baseado em web, contidos em bases de dados, ou fixados ao próprio animal, tal como por um microchip implantado. Um aspecto vantajoso da presente invenção, portanto, é
20 direcionado a um sistema de computador e métodos auxiliados por computador para rastreio dos traços de qualidade para criação que possui predisposições genéticas específicas.

A presente invenção envolve vantajosamente métodos auxilia-
25 dos por computador e sistemas para aquisição de dados genéticos, particularmente dados genéticos conforme definidos pela ausência ou presença de um SNP/STR no interior do gene de *CAST* relacionado a traços de fertilidade e longevidade da geração de animal e associação daqueles dados com outros dados sobre o animal ou seu rebanho, e manutenção daqueles dados
30 em modos que são acessíveis. Outro aspecto da invenção envolve um método auxiliado por computador para prognosticar qual criação de animais possui uma diferença biológica na fertilidade e longevidade, e que pode in-

cluir as etapas de usar um sistema de computador, por exemplo, a computador programado compreendendo a processador, um sistema de armazenagem de dados, um dispositivo de entrada e um dispositivo de saída, as etapas de: (a) admissão no computador programado através do dispositivo de
5 entrada de dados que incluem um genótipo de um animal conforme ele se refere a qualquer um dos SNPs/STRs de *CAST* aqui descritos, (b) correção fertilidade e longevidade prognosticadas pelo genótipo de *CAST* usando o processador e o sistema de armazenagem de dados, e (c) fornecimento ao
10 dispositivo de saída da fertilidade e longevidade correlacionada ao genótipo de *CAST*, prognosticando-se, desse modo, qual criação de animais possui uma fertilidade e longevidade particular.

Ainda outro aspecto da invenção se refere a um método de se fazer negócio para controle da criação compreendendo provisão a um usuário de um sistema de computador para gerenciamento da criação compreendendo características físicas e genótipos correspondentes a um ou mais animais, ou um meio legível de computador para gerenciamento da criação
15 compreendendo características físicas e genótipos correspondentes a uma ou mais animais, ou características físicas e genótipos correspondentes a um ou mais animais, no qual a entrada característica física, crescimento ou mérito de carcaça em gado bovino e o genótipo é um genótipo de *CAST*.
20

É notado que nesta descrição e particularmente nas reivindicações e/ou parágrafos, termos tais como "compreende", "compreendido", "compreendendo" e similares podem ter o significado atribuído à lei de Patente dos Estados Unidos; por exemplo, eles podem significar "inclui", "incluído", "incluindo", e similares; e aqueles termos tais como "consistindo essencialmente em" e "consiste essencialmente em" têm os significados descritos na lei de Patente dos Estados Unidos; por exemplo, eles permitem elementos não explicitamente recitados, mas excluem elementos que são encontrados na técnica anterior ou que afetam uma característica básica ou nova da
25 invenção.
30

Estas e outras concretizações são descritas ou são óbvias a partir de e envolvidas pela seguinte Descrição Detalhada.

BREVE DESCRIÇÃO DOS DESENHOS

A seguinte descrição detalhada, dada por meio de exemplo, mas não pretendida para limitar a invenção somente às concretizações específicas descritas, pode melhor ser compreendida em conjunto com os desenhos
5 acompanhantes, em que:

A figura 1 proporciona uma anotação comparativa de ambos genes *CAST* humano e bovino. A anotação recuperou o domínio XL na proteína de *CAST* humana, mostrando a mesma estrutura de gene encontrada no gene de *CAST* bovino. Os tamanhos de cada exon (humano/gado) são conforme segue: exon 1, 30/30bp; exon 2, 63/63 bp; exon 3, 72/72 bp; exon 4,
10 60/60 bp; exon 5, 66 bp; exon 6, 42/42 bp; exon 7, 57/57 bp; exon 8, 114/114 bp; exon 9, 81/87 bp; exon 10, 69/69 bp; exon 11, 99/99 bp; exon 12, 81/81 bp; exon 13, 39/39 bp; exon 14, 93/93 bp; exon 15, 87/90 bp; exon 16, 102/102 bp; exon 17, 84/84 bp; exon 18, 48/48 bp; exon 19, 96/96 bp; exon
15 20, 96/93 bp; exon 21, 102/102 bp; exon 22, 84/84 bp; exon 23, 51/51 bp; exon 24, 72/72 bp; exon 25, 99/99 bp; exon 26, 105/108 bp; exon 27, 93/93 bp; exon 28, 45/48 bp; exon 29, 93/93 bp; exon 30, 72/72 bp; exon 31, 59/66 bp (incluindo seqüência de não-codificação parcial), e exon 32, 198/2045 bp (seqüências 3'UTR em ambas as espécies). A 5'UTR é 148 bp em humano e
20 128 bp de comprimento em gado. A distância de folga entre quaisquer dois exons não é proporcional ao tamanho do intron.

As figuras 2A-2E proporcionam uma seqüência de nucleotídeo (SEQ ID NO: 1) do promotor, 5'UTR (Negrito), exon 1 (*Itálicos e Sublinhados*) e seqüência parcial intron 1 do gene de *CAST*(AAFC02060382) bovino.
25 Cada limite exon-intron está em caixa.

As figuras 3A-3D proporcionam vide uma seqüência de nucleotídeo (SEQ ID NO: 2) do intron 1, exon 2 parcial (*Itálicos e Sublinhados*) e seqüência parcial de intron 2 do gene de *CAST*(AAFC02161394) bovino. Cada limite exon-intron está em caixa.

30 As figuras 4A-4C proporcionam vide uma seqüência de nucleotídeo (SEQ ID NO: 3) do intron 2, exon 3 parcial (*Itálicos e Sublinhados*) e seqüência parcial de intron 3 do gene de *CAST*(AAFC02179490) bovino. Cada

limite exon-intron está em caixa. Três mutações: G/A, C/T e T/G (foram colocadas em Negrito), que foram associadas com ambas fertilidade e longevidade em gado leiteiro.

As figuras 5A-5F proporcionam uma sequência de nucleotídeo (SEQ ID NO: 4) do intron 3, exon 4 parcial (*Itálicos e Sublinhados*) e sequência parcial de intron 4 do gene de CAST(AAFC02060385) bovino. Cada limite exon-intron está em caixa.

As figuras 6A-6E proporcionam uma sequência de nucleotídeo (SEQ ID NO: 5) do intron 5, exons 5 - 16 parcial (*Itálicos e Sublinhados*), sequência de nucleotídeo parcial de introns 5 - 15 e intron 16 do gene de CAST(AAFC02060381) bovino. Cada limite exon-intron está em caixa. Um polimorfismo de STR (GAAA) foi colocado em NEGRITO e associado com fertilidade e longevidade em gado leiteiro.

A figura 7 proporciona uma sequência de nucleotídeo (SEQ ID NO: 6) do intron 16, exon 17 parcial (*Itálicos e Sublinhados*) e sequência parcial de intron 17 do gene de CAST(AAFC02197217) bovino. Cada limite exon-intron está em caixa.

As figuras 8A-8H proporcionam uma sequência de nucleotídeo (SEQ ID NO: 7) do intron 17, exons 18 - 32 parcial (*Itálicos e Sublinhados*), sequência de introns 18 - 31 e 3'UTR do gene de CAST(AAFC02067026) bovino. Cada limite exon-intron está em caixa.

As figuras 9A-9B proporcionam uma sequência de nucleotídeo (SEQ ID NO: 7) e sequência de aminoácido (SEQ ID NO: 8) da sequência de codificação e *in silico* identificação cSNP. Os SNPs potenciais estão em caixa.

As figuras 10A-10C proporcionam polimorfismos de sequência de nucleotídeo no gene de CAST bovino. A, duas mutações sem sentido A/G e C/T em exon 3 (SEQ ID NOS. 10 e 11); B, a T/G substituição em intron 3 (SEQ ID NOS. 12 e 13); e C, uma repetição GAAA em intron 8 (SEQ ID NOS. 14 e 15). Dois haplotipos existem na população: G-C-T-GAAAGAAAGAAAGAAA (SEQ ID NO: 16) (série superior) e A-T-G-GAAAGAAAGAAAGAAAGAAA (SEQ ID NO: 17) (série inferior).

A figura 11 ilustra genotipagem de PCR-RFLP de uma substituição C/T em in exon 3 de gene de *CAST* bovino. TT = 308 bp, CT = 135 + 173 + 308 bp e CC = 135 + 173 bp, respectivamente.

5 A figura 12 ilustra um fluxograma da entrada de dados e a saída de resultados a partir da análise e correlação dos dados pertencentes à geração, histórias veterinárias e requerimentos de desempenho de um grupo de animais, tais como de um rebanho de vacas, e o fluxo interativo de dados a partir do dispositivo auxiliado por computador para um corpo de estudantes que aprendem o uso do método da invenção.

10 A figura 13 ilustra os relacionamentos potenciais entre os elementos de dados a serem admitidos no sistema. As setas unidirecionais indicam, por exemplo, que um celeiro é tipicamente pertencente a apenas uma fazenda, pelo qual uma fazenda pode possuir vários celeiros. Similarmente, uma prescrição pode incluir produtos veterinários.

15 A figura 14A ilustra o fluxo de eventos no uso do sistema baseado em computador portátil para entrada de dados na geração e criação de um rebanho de vacas.

A figura 14B ilustra o fluxo de eventos através das sub-rotinas relacionadas a entrada de dados concernente ao gerenciamento da fazenda.

20 A figura 14C ilustra o fluxo de eventos através das sub-rotinas relacionadas a entrada de dados concernente a dados específicos para uma companhia.

25 A figura 15 ilustra um fluxograma da entrada de dados e a saída de resultados a partir da análise e a correlação dos dados pertencentes à geração, histórias veterinárias e requerimentos de desempenho de um grupo de animais.

DESCRIÇÃO DETALHADA

30 O gene de Calpastatin (*CAST*) é amplamente expresso na reprodução de tecidos/órgãos. Contudo, como este gene é relacionado à fertilidade, permanece largamente indeterminado. No presente estudo, os inventores descobriram associações significantes anteriormente não reportadas de mutações sem sentido em um domínio XL recentemente identificado de

gene de *CAST* bovino com fertilidade (taxa de gravidez de raça, DPR) e longevidade (vida produtiva, PL) em gado leiteiro usando 652 progenitores derivados de sete antepassados. O alinhamento de ambos cDNA e seqüências de DNA genômico revelou três mutações sem sentido provisórias, mas duas delas (G48D e P52L) em exon 3 (NM_174003.2:c.271G>A e 283C>T), que correspondem à região de domínio XL, foram confirmadas por análise de seqüenciamento de duas reuniões de DNA e sete antepassados. Estas duas confirmaram mutações sem sentido mais uma mutação em intron 3 (A-AFC02060381.1:g.2110G>T), e uma repetição GAAA em intron 8 (A-AFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] formou somente dois haplotipos. Uma transição C/T foi então genotipada com enzima de restrição *MspI* e usada para uma classificação de associação inicial e uma análise compreensiva final. Através da família, a análise indicou que o genótipo individual era uma fonte significativa de variação ($P < 0,0001$) para DPR e PL, mas não para os traços de leite ($P > 0,05$). As hereditariedades realizadas foram estimadas para serem 0,55 para DPR e 0,66 para PL, indicando que o gene de calpastatin bovino, quando utilizado em seleção auxiliada por marcador, deve acelerar o aperfeiçoamento de fertilidade em gado leiteiro.

Aspectos particulares proporcionam quatro novos polimorfismos incluindo duas mutações sem sentido que formam dois haplotipos em um novo domínio N-terminal do gene bovino de calpastatin gene mostram associações significantes com fertilidade e longevidade em gado leiteiro.

Em aspectos específicos, duas mutações na região "XL" do gene bovino de *CAST* foram identificadas, que conduzem a mudanças de aminoácido (G48D e P52L) em ambas as posições. A genotipagem dos marcadores em 652 animais de sete famílias de progenitor descreveu uma forte associação com DPR e PL em gado leiteiro. Os resultados atuais indicam que formas diferentes de gene de *CAST* podem estar envolvidas em trajetórias diferentes de várias células/tecidos por expressão de um efeito pleiotrófico em funções diferentes.

Aspectos particulares proporcionam novos marcadores para fertilidade (*por exemplo*, taxa de gravidez da raça, DPR) e longevidade (*por*

exemplo, vida produtiva, PL) em, por exemplo, gado leiteiro.

Aspectos adicionais proporcionam novos métodos compreendendo seleção auxiliada por marcador para aperfeiçoar fertilidade e/ou longevidade em gado leiteiro. Em concretizações particulares, uma combinação
 5 de seleção genética baseada em um ou mais dos novos marcadores de *CAST*, e altos potenciais de PTA (capacidade de transmissão prognosticada) em traços de produção de leite, proporciona traços reprodutivos aperfeiçoados em associação com traços de produção de leite altos continuados.

Aspectos adicionais descrevem um domínio XL anteriormente
 10 não-reconhecido no gene humano de *CAST*, e, desse modo, proporciona o uso de mutantes/variantes de domínio XL de *CAST* humano como marcadores de fertilidade e longevidade humana.

A prática da presente invenção empregará, a menos que de outro modo indicado, técnicas convencionais de biologia molecular, microbiologia, tecnologia de DNA recombinante, e imunologia, que estão dentro dos
 15 conhecimentos de um técnico no assunto. Tais técnicas são explanadas totalmente na literatura. Ver, por exemplo, Sambrook et al. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Press; *DNA Cloning*, Vols. I e II (D. N. Glover ed. 1985); *Oligonucleotide Synthesis* (M. J. Gait ed. 1984); *Acid nucleic Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Animal Cell Culture* (R. K. Freshney ed. 1986); *Immobilized Cells e Enzymes* (IRL press, 1986); Perbal, B., *A Practical Guide to Molecular Cloning* (1984); the series, *Methods In Enzymology* (S. Colowick e N. Kaplan eds., Academic Press, Inc.); e *Handbook of Experimental Immunology*, Vols.
 20 I-IV (D. M. Weir e C. C. Blackwell eds., 1986, Blackwell Scientific Publications).

Antes de se descrever a presente invenção em detalhes, é para ser compreendido que esta invenção não é limitada a DNA particular, seqüências de polipeptídeo ou parâmetros de processo, como tais, podem variar. É também para ser compreendido que a terminologia aqui usada é para
 30 a proposta de descrever concretizações articulares da invenção somente, e não é pretendida para ser limitante.

A menos que de outro modo definido, todos os termos técnicos e científicos aqui usados têm o mesmo significado conforme comumente compreendido por um técnico no assunto ao qual a invenção pertence. Embora um número de métodos e materiais similares ou equivalentes àqueles descritos aqui possam ser usados na prática da presente invenção, os materiais e métodos preferidos são aqui descritos.

Na descrição da presente invenção, os seguintes termos serão empregados e são pretendidos para serem definidos conforme indicado abaixo.

10 O termo "vaca" ou "gado" é usado geralmente para se referir a um animal de origem bovina de qualquer idade. Termos permutáveis incluem "bovino", "bezerro", "touro novo", "touro", "novilho", e similares. Ele também inclui um animal individual em todos os estágios de desenvolvimento, incluindo estágios embriônicos e fetais. Os animais conforme referidos aqui podem também incluir indivíduos ou grupos de indivíduos que ocorreram de outra produção de alimento diferente, tal como, mas não limitado a, animais transgênicos para a produção de biofarmacêuticos incluindo anticorpos e outras proteínas, ou produtos de proteína.

20 Pelo termo "complementaridade" ou "complementar" é significativo, para a proposta do relatório descritivo ou reivindicações, um número suficiente no oligonucleotídeo de pares bases complementares em sua sequência para interagir especificamente (hibridiza) com uma sequência de ácido nucléico alvo do polimorfismo de gene a ser amplificado ou detectado. Conforme conhecido àqueles técnicos no assunto, um grau muito alto de complementaridade é necessário para especificidade e sensibilidade envolvendo hibridização, embora não seja necessário ser 100%. Desse modo, por exemplo, um oligonucleotídeo que é idêntico na sequência de nucleotídeo a um oligonucleotídeo aqui descrito, exceto para uma mudança ou substituição base, pode funcionar equivalentemente aos oligonucleotídeos descritos. Um

30 "DNA complementa" ou "gene de cDNA" inclui genes recombinantes sintetizados por transcrição reversa de RNA mensageiro ("mRNA").

Uma "reação mediada por polimerase cíclica" se refere a uma

reação biomédica na qual uma molécula gabarito ou uma população de moléculas gabaritos é periodicamente e repetidamente para criar uma molécula gabarito complementar, ou moléculas gabaritos complementares, aumentando, desse modo, o número das moléculas gabaritos com o tempo.

5 Pelo termo "porção detectável" é significativo, para a proposta do relatório descritivo e reivindicações, uma molécula de etiqueta (isotópica ou não-isotópica) que é incorporada indiretamente ou diretamente em um oligonucleotídeo, no qual a molécula de etiqueta facilita a detecção do oligonucleotídeo no qual ela é incorporada, por exemplo, quando o oligonucleotídeo
10 é hibridizado às seqüências polimórficas de gene amplificadas. Desse modo, "porção detectável" é usada sinonimamente como "molécula de etiqueta". A síntese de oligonucleotídeos pode ser acompanhada por qualquer um de vários métodos conhecidos àqueles técnicos no assunto. As moléculas de etiqueta, conhecidas àqueles técnicos no assunto como sendo úteis para
15 detecção, incluem moléculas quimioluminescentes, fluorescentes ou luminescentes. Várias moléculas fluorescentes são conhecidas na técnica que são adequadas para uso para etiquetar um ácido nucléico para o método da presente invenção. O protocolo para tal incorporação pode variar dependendo da molécula fluorescente usada. Tais protocolos são conhecidos na técnica
20 nica para a respectiva molécula fluorescente.

"Amplificação de DNA", conforme aqui usado, se refere a qualquer processo que aumente o número de cópias de uma seqüência de DNA específica por amplificação enzimática da seqüência de ácido nucléico. Uma variedade de processos é conhecida. Um dos mais comumente usado é um
25 processo de reação de cadeia de polimerase (PCR) de Mullis conforme descrito nas Patentes dos Estados Unidos N^{os} 4,683,195 e 4,683,202. Métodos, dispositivos e reagentes conforme descritos nas Patentes dos Estados Unidos N^{os} 6,951,726; 6,927,024; 6,924,127; 6,893,863; 6,887,664; 6,881,559; 6,855,522; 6,855,521; 6,849,430; 6,849,404; 6,846,631; 6,844,158;
30 6,844,155; 6,818,437; 6,818,402; 6,794,177; 6,794,133; 6,790,952; 6,783,940; 6,773,901; 6,770,440; 6,767,724; 6,750,022; 6,744,789; 6,733,999; 6,733,972; 6,703,236; 6,699,713; 6,696,277; 6,664,080;

	6,664,064;	6,664,044;	RE38,352;	6,650,719;	6,645,758;	6,645,720;
	6,642,000;	6,638,716;	6,632,653;	6,617,107;	6,613,560;	6,610,487;
	6,596,492;	6,586,250;	6,586,233;	6,569,678;	6,569,627;	6,566,103;
	6,566,067;	6,566,052;	6,558,929;	6,558,909;	6,551,783;	6,544,782;
5	6,537,752;	6,524,830;	6,518,020;	6,514,750;	6,514,706;	6,503,750;
	6,503,705;	6,493,640;	6,492,114;	6,485,907;	6,485,903;	6,482,588;
	6,475,729;	6,468,743;	6,465,638;	6,465,637;	6,465,171;	6,448,014;
	6,432,646;	6,428,987;	6,426,215;	6,423,499;	6,410,223;	6,403,341;
	6,399,320;	6,395,518;	6,391,559;	6,383,755;	6,379,932;	6,372,484;
10	6,368,834;	6,365,375;	6,358,680;	6,355,422;	6,348,336;	6,346,384;
	6,319,673;	6,316,195;	6,316,192;	6,312,930;	6,309,840;	6,309,837;
	6,303,343;	6,300,073;	6,300,072;	6,287,781;	6,284,455;	6,277,605;
	6,270,977;	6,270,966;	6,268,153;	6,268,143;	D445,907;	6,261,431;
	6,258,570;	6,258,567;	6,258,537;	6,258,529;	6,251,607;	6,248,567;
15	6,235,468;	6,232,079;	6,225,093;	6,221,595;	D441,091;	6,218,153;
	6,207,425;	6,183,999;	6,183,963;	6,180,372;	6,180,349 ;	6,174,670;
	6,153,412;	6,146,834;	6,143,496;	6,140,613;	6,140,110;	6,103,468;
	6,087,097;	6,072,369;	6,068,974;	6,063,563;	6,048,688;	6,046,039;
	6,037,129;	6,033,854;	6,031,960;	6,017,699;	6,015,664;	6,015,534;
20	6,004,747;	6,001,612;	6,001,572;	5,985,619;	5,976,842;	5,972,602;
	5,968,730;	5,958,686;	5,955,274;	5,952,200;	5,936,968;	5,909,468;
	5,905,732;	5,888,740;	5,883,924;	5,876,978;	5,876,977;	5,874,221;
	5,869,318;	5,863,772;	5,863,731;	5,861,251;	5,861,245;	5,858,725;
	5,858,718;	5,856,086;	5,853,991;	5,849,497;	5,837,468;	5,830,663;
25	5,827,695;	5,827,661;	5,827,657;	5,824,516;	5,824,479;	5,817,797;
	5,814,489;	5,814,453;	5,811,296;	5,804,383;	5,800,997;	5,780,271 ;
	5,780,222;	5,776,686;	5,774,497;	5,766,889;	5,759,822;	5,750,347;
	5,747,251;	5,741,656;	5,716,784;	5,712,125;	5,712,090;	5,710,381;
	5,705,627;	5,702,884;	5,693,467;	5,691,146;	5,681,741;	5,674,717;
30	5,665,572;	5,665,539;	5,656,493;	5,656,461;	5,654,144;	5,652,102;
	5,650,268;	5,643,765;	5,639,871;	5,639,611;	5,639,606;	5,631,128;
	5,629,178;	5,627,054;	5,618,703;	5,618,702;	5,614,388;	5,610,017;

5,602,756; 5,599,674; 5,589,333; 5,585,238; 5,576,197; 5,565,340;
 5,565,339; 5,556,774; 5,556,773; 5,538,871; 5,527,898; 5,527,510;
 5,514,568; 5,512,463; 5,512,462; 5,501,947; 5,494,795; 5,491,225;
 5,487,993; 5,487,985; 5,484,699; 5,476,774; 5,475,610; 5,447,839;
 5 5,437,975; 5,436,144; 5,426,026; 5,420,009; 5,411,876; 5,393,657;
 5,389,512; 5,364,790; 5,364,758; 5,340,728; 5,283,171; 5,279,952;
 5,254,469; 5,241,363; 5,232,829; 5,231,015; 5,229,297; 5,224,778;
 5,219,727; 5,213,961; 5,198,337; 5,187,060; 5,142,033; 5,091,310;
 5,082,780; 5,066,584; 5,023,171 e 5,008,182, podem também ser emprega-
 10 dos na prática da presente invenção. PCR envolve o uso de um DNA polime-
 rase termoestável, seqüências conhecidas de priméres, e ciclos de aqueci-
 mento, que separam o ácido deoxirribonucleico (DNA) de replicação, trança-
 dos e amplificam exponencialmente um gene de interesse. Qualquer tipo de
 PCR, tais como PCR quantitativa, RT-PCR, PCR de partida quente, LAPCR,
 15 PCR multiplex, PCR "touchdown" PCR, etc., podem ser usados. Vantajosa-
 mente, PCR de tempo real é usada. Em geral, o processo de amplificação
 de PCR envolve uma reação de cadeia enzimática cíclica para preparação
 de quantidades exponenciais de uma seqüência de ácido nucléico específi-
 ca. Ela requer uma pequena quantidade de uma seqüência para iniciar a
 20 reação de cadeia e priméres de oligonucleotídeo que hibridizarão para a se-
 qüência. Na PCR, os priméres são recozidos para desnaturar ácido nucléico,
 seguido por extensão com um agente de indução (enzima), e nucleotídeos.
 Isto resulta em produtos de extensão recentemente sintetizados. Desde que
 estas seqüências recentemente sintetizadas tornam-se gabaritos para os
 25 priméres, ciclos repetidos de desnaturação, recozimento de primer e exten-
 são resultam em acúmulo exponencial da seqüência específica sendo ampli-
 ficada. O produto de extensão da reação de cadeia será um ácido nucléico
 distinto dúplex com um terminal correspondendo às extremidades dos prime-
 res empregados.

30 Pelos termos "amplificar enzimaticamente" ou "amplificar", é sig-
 nificativo, para a proposta do relatório descritivo e reivindicações, amplifica-
 ção de DNA, isto, um processo pelo qual seqüências de ácido nucléico são

amplificadas em número. Existem vários meios para amplificar enzimaticamente seqüências de ácido nucléico. Atualmente o método mais comumente usado é a reação de cadeia de polimerase (PCR). Outros métodos de amplificação incluem LCR (reação de cadeia de ligase) que utiliza ligase de DNA, e uma sonda consistindo em duas metades de um segmento de DNA que é complementar à seqüência do DNA a ser amplificada, enzima QB replicase e um gabarito de seqüência de ácido ribonucleíco (RNA) fixado a uma sonda complementar ao DNA a ser copiado que é usado para produzir um gabarito de DNA para produção exponencial de RNA complementar; amplificação de deslocamento de trançado (SDA); amplificação de replicase QB (QBRA); replicação auto-sustentada (3SR); e NASBA (amplificação baseada em seqüência de ácido nucléico), que podem ser realizadas em RNA ou DNA conforme a seqüência de ácido nucléico a ser amplificada.

Um "fragmento" de uma molécula, tal como uma proteína ou ácido nucléico é significativo se referir a qualquer porção do aminoácido ou seqüência genética de nucleotídeo.

Conforme aqui usado, o termo "genoma" se refere a todo o material genético nos cromossomos de um organismo particular. Sua dimensão é geralmente dada como seu número total de pares bases. Dentro do genoma, o termo "gene" se refere a uma sequíssima ordenada de nucleotídeos localizada em uma posição particular em um cromossomo particular que codifica um produto funcional específico (por exemplo, uma proteína ou molécula de RNA). Em geral, características genéticas animais, conforme definidas pela seqüência de nucleotídeo de seu genoma, são conhecidas como seu "genótipo," enquanto os traços físicos do animal são descritos como seu "fenótipo".

Por "heterozigoto" ou "polimorfismo heterozigoto" é significativo que os dois alelos de uma célula diplóide ou organismo em um dado local são diferentes, isto é, que eles têm um nucleotídeo diferente trocado pelo menos nucleotídeo no mesmo lugar em suas seqüências.

Por "homozigoto" ou "polimorfismo homozigoto" é significativo que os dois alelos de uma célula diplóide ou organismo em um dado local

são idênticos, isto é, que eles têm o mesmo nucleotídeo para troca de nucleotídeo no mesmo lugar em suas seqüências.

Por "hibridização" ou "hibridizar", conforme aqui usado, é significativo à formação de pares base A-T e C-G entre a seqüência de nucleotídeo de um fragmento de um segmento de um polinucleotídeo e uma seqüência de nucleotídeo complementar de um oligonucleotídeo. Por complementar é significativo que no local de cada A, C, G ou T (ou U em um ribonucleotídeo) na seqüência de fragmento, a oligosseqüência de nucleotídeo tem um T, G, C ou A, respectivamente. O fragmento/oligonucleotídeo hibridizado é denominado um "dúplex".

Um "complexo de hibridização", tal como em um ensaio intercalado, significa um complexo de moléculas de ácido nucléico incluindo pelo menos o ácido nucléico alvo e um sonda de sensor. Pode-se também incluir uma sonda de âncora.

Conforme aqui usado, o termo "locus" ou "loci" se refere ao local de um gene em um cromossomo. Pares de genes, conhecidos como "alelos", controlam o traço hereditário produzido por um locus de gene. Cada combinação particular do animal de alelos é referida como seu "genótipo". Onde ambos alelos são idênticos, o indivíduo é referido para ser referido como sendo homozigoto para o traço controlado por este par de gene; onde os alelos são diferentes, o indivíduo é referido como sendo heterozigoto para o traço.

Uma "temperatura de fusão" é significativa à temperatura na qual dúpleces hibridizados dehibridizam e retornam para seu estado trançado simples. Do mesmo modo, hibridização não ocorrerá no primeiro local entre dois oligonucleotídeos, ou, aqui, um oligonucleotídeo e um fragmento, em temperaturas acima da temperatura de fusão do dúplex resultante. É atualmente vantajoso que a diferença nas temperaturas de ponto de fusão de dúpleces de oligonucleotídeo-fragmento desta invenção seja de cerca de 1°C a cerca de 10°C de modo a ser prontamente detectável.

Conforme aqui usado, o termo "molécula de ácido nucléico" é pretendido para incluir moléculas de DNA (por exemplo, cDNA ou DNA ge-

nômico), moléculas de RNA (por exemplo, mRNA), análogos do DNA ou RNA gerados usando-se análogos de nucleotídeo, e derivados, fragmentos e homólogos destes. A molécula de ácido nucléico pode ser de trançado simples ou de trançado duplo, mas a molécula pode ser de trançado simples ou de trançado dupla, mas vantajosamente é DNA de trançado duplo. "DNA" se refere à forma polimérica de deoxirribonucleotídeos (adenina, guanina, timina, ou citosina) em sua forma de ou trançada simples, ou uma espiral de trançado duplo. Este termo se refere somente à estrutura primária e secundária da molécula, e não se limita a quaisquer formas terciárias particulares. Desse modo, este termo inclui DNA de trançado duplo encontrado, entre outros, em moléculas de DNA lineares (por exemplo, fragmentos de restrição), vírus, plasmídeos, e cromossomos. Na discussão da estrutura de moléculas de DNA de trançado duplo, seqüências podem ser descritas aqui de acordo com a convenção normal de dar somente a seqüência na direção 5' a 3' ao longo do trançado não-transcrito de DNA (isto é, o trançado tendo uma seqüência homóloga para o mRNA). Uma molécula de ácido nucléico "isolada" é separada de outras moléculas de ácido nucléico que estão presentes na fonte natural do ácido nucléico.

Um "nucleosídeo" se refere a uma base ligada a um açúcar. A base pode ser adenina (A), guanina (G) (ou seu substituto, inosina (I)), citosina (C), ou timina (T) (ou seu substituto, uracil (U)). O açúcar pode ser ribose (o açúcar de um nucleotídeo natural no RNA) ou 2-deoxirribose (o açúcar de um nucleotídeo natural em DNA). Um "nucleotídeo" se refere a um nucleosídeo ligado a um grupo fosfato simples.

Conforme aqui usado, o termo "oligonucleotídeo" se refere a uma série de resíduos de nucleotídeo ligados, cujo oligonucleotídeo tem um número suficiente de bases de nucleotídeo a serem usadas em uma reação de PCR. Uma oligosseqüência curta de nucleotídeo pode ser baseada em, ou designada de, uma seqüência genômica ou de cDNA, e é usada para amplificar, confirmar, ou descrever a presença de um DNA ou RNA idêntico, similar ou complementar em uma célula ou tecido particular. Os oligonucleotídeos podem ser quimicamente sintetizados, e podem ser usados como

prímeres ou sondas. O oligonucleotídeo significa qualquer nucleotídeo de mais do que 3 bases de comprimento usadas para facilitar a detecção ou identificação de um ácido nucléico alvo, incluindo sondas e prímeres.

Uma "polimerase" é uma enzima que catalisa a adição seqüencial de unidades monoméricas a uma cadeia polimérica, ou liga duas ou mais unidades monoméricas para iniciar uma cadeia polimérica. A "polimerase" operará pela adição de unidades monoméricas cuja identidade é determinada por, e que é complementar a uma molécula gabarito de uma seqüência específica. Por exemplo, DNA polimerases, tais como DNA pol 1 e Taq polimerase adicionam deoxirribonucleotídeos à extremidade 3' de uma cadeia de polinucleotídeo em uma maneira dependente do gabarito, sintetizando, desse modo, um ácido nucléico que é complementar à molécula gabarito. Polimerases podem ser usadas, ou para entender um prímer uma vez, ou repetitivamente, ou para amplificar um polinucleotídeo por prímer repetitiva de dois trançados complementares usando dois prímeres. Uma "polimerase termoestável" se refere a uma enzima de DNA ou RNA polimerase que pode suportar temperaturas extremamente altas, tais como aquelas que se aproximam de 100°C. Frequentemente, polimerases termoestáveis são derivadas de organismos que vivem em temperaturas extremas, tais como *Thermus aquaticus*. Exemplos de polimerases termoestáveis incluem Taq, Tth, Pfu, Vent, "deep vent", UITma, e variações e derivados destes.

Um "polinucleotídeo" se refere a uma cadeia linear de nucleotídeos conetada por uma ligação de fosfodiéster entre o grupo 3'-hidroxil de um nucleosídeo no grupo 5'-hidroxil de um segundo nucleosídeo que, por sua vez, é ligada através de seu grupo 3'-hidroxil ao grupo 5'-hidroxil de um terceiro nucleosídeo e assim por diante, para formar um polímero compreendido de nucleosídeos ligados por um suporte de fosfodiéster. Um "polinucleotídeo modificado" se refere a um polinucleotídeo em que um ou mais nucleotídeos naturais foram parcialmente, substancialmente ou completamente substituídos com nucleotídeos modificados.

Um "prímer" é um oligonucleotídeo, a seqüência de pelo menos a porção do qual é complementar a um segmento de um DNA gabarito que é

para ser amplificado ou replicado. Tipicamente priméres são usados na realização da reação de cadeia de polimerase (PCR). Um primer hibridiza com (ou "é recozido") ao DNA gabarito, e é usado pela enzima de polimerase como o ponto de partida para o processo de replicação/amplificação. Os priméres aqui são selecionados para serem "substancialmente" complementares a trançados diferentes de uma seqüência de DNA alvo particular. Isto significa que os priméres devem ser suficientemente complementares para hibridizarem com seus respectivos trançados. Portanto, a seqüência de primer não necessita refletir a seqüência exata do gabarito. Por exemplo, um fragmento de nucleotídeo não-complementar pode ser fixado à extremidade 5' do primer, com o restante da seqüência de primer sendo complementar ao trançado. Alternativamente, bases não-complementares ou seqüências mais longas podem ser interespersadas no primer, provido que a seqüência de primer tem complementaridade suficiente com a seqüência do trançado para hibridizar com este e, desse modo, formar o gabarito para a síntese do produto de extensão.

"Sondas" se referem às seqüências de ácido nucléico de nucleotídeos de comprimento variáveis, usadas na detecção de seqüências de ácido nucléico idênticas, similares, ou complementares por hibridização. Uma oligosseqüência de nucleotídeo usada como uma sonda de detecção pode ser etiquetada com uma porção detectável.

Os seguintes são exemplos não-limitativos de polinucleotídeos: um gene ou fragmento de gene, exons, introns, mRNA, tRNA, rRNA, ribozimas, cDNA, polinucleotídeos recombinantes, polinucleotídeos ramificados, plasmídeos, vetores, DNA isolado de qualquer seqüência, RNA isolado de qualquer seqüência, sondas de ácido nucléico e priméres. Um polinucleotídeo pode compreender nucleotídeos modificados, tais como nucleotídeos metilados e análogos de nucleotídeo, uracil, outros açúcares, e grupos de ligação, tais como fluororibose e tiolato, e ramificações de nucleotídeo. A seqüência de nucleotídeos pode ser adicionalmente modificada após polimerização, tal como por conjugação, com um componente de etiquetação. Outros tipos de modificações incluídos nesta definição são coberturas, substitu-

ição de um ou mais nucleotídeos que ocorrem naturalmente com um análogo, e introdução de meios para fixação do polinucleotídeo às proteínas, íons metálicos, componentes de etiquetagem, outros polinucleotídeos ou suporte sólido.

5 Um polinucleotídeo ou polipeptídeo "isolado" é um que é substancialmente puro de materiais com o qual ele está associado em seu ambiente nativo. Por substancialmente livre, é significativo pelo menos 50%, pelo menos 55%, pelo menos 60%, pelo menos 65%, e vantajosamente pelo menos 70%, pelo menos 75%, mais vantajosamente pelo menos 80%, pelo
10 menos 85%, ainda mais vantajosamente pelo menos 90%, pelo menos 91%, pelo menos 92%, pelo menos 93%, pelo menos 94%, pelo menos 95%, pelo menos 96%, pelo menos 97%, mais vantajosamente pelo menos 98%, pelo menos 99%, pelo menos 99,5%, pelo menos 99,9% livre destes materiais.

Uma molécula de ácido nucléico "isolada" é uma molécula de
15 ácido nucléico separada e distinta a partir de todo organismo com o qual a molécula é encontrada na natureza; ou uma molécula de ácido nucléico destituída, no todo ou em parte, de seqüências normalmente associadas com ela na natureza; ou uma seqüência, conforme ela existe na natureza, mas tendo seqüências heterólogas (conforme definido abaixo) em associação
20 com esta.

O termo "polinucleotídeo que codifica uma proteína", conforme aqui usado, se refere a um fragmento de DNA ou molécula de DNA isolada que codifica uma proteína, ou o trançado complementar a esta; mas, RNA não é excluído, conforme é compreendido na técnica que timidina (T) em
25 uma seqüência de DNA é considerada igual à uracil (U) em uma seqüência de RNA. Desse modo, seqüências de RNA para uso na invenção, por exemplo, para uso em vetores de RNA, podem ser derivadas de seqüências de DNA, por timidina (T) na seqüência de DNA sendo considerada igual à uracil (U) nas seqüências de RNA.

30 Uma "seqüência de codificação" de DNA, ou uma "seqüência de nucleotídeo que codifica" uma proteína particular, é uma seqüência de DNA que é transcrita e transladada em um polipeptídeo in vitro ou in vivo quando

colocado sob o controle de elementos regulatórios apropriados. Os limites da sequência de codificação são determinados por um códon de partida no término 5' (amino) e um códon de parada de translação no término 3' (carbóxi). Uma sequência de codificação pode incluir, mas é limitada a, seqüências

5 procarióticas, cDNA de mRNA eucariótico, seqüências de DNA genômico de DNA eucariótico (por exemplo, mamífero), e ainda seqüências de DNA sintéticas. Uma seqüência de terminação de transcrição estará usualmente localizada 3' à seqüência de codificação.

"Homologia" se refere à percentagem de identidade entre dois

10 polinucleotídeos ou duas porções de polipeptídeo. Dois DNAs, ou duas seqüências de polipeptídeo são "substancialmente homólogas" umas as outras quando as seqüências exibem pelo menos cerca de 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, preferivelmente pelo menos cerca de 90%, 91%, 92%, 93%, 94%, e mais preferivelmente pelo menos cerca de

15 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9% de identidade de seqüência sobre um comprimento definido das moléculas. Conforme aqui usado, substancialmente homólogo também se refere às seqüências que mostram identidade completa (100% de identidade de seqüência) ao DNA especificado ou seqüência de polipeptídeo.

20 Homologia pode ser determinada por hibridização de polinucleotídeos sob condições que formam dúplexes estáveis entre regiões homólogas, seguido por digestão com nuclease (s) específica(s) de trançado simples, e determinação de tamanho dos fragmentos digeridos. As seqüências de DNA que são substancialmente homólogas podem ser identificadas em

25 um experimento de hibridização no Sul sob, por exemplo, condições estridentes, conforme definido para aquele sistema particular. A definição de condições de hibridização apropriadas está dentro do conhecimento do técnico no assunto. Ver, por exemplo, Sambrook et al. supra; DNA Cloning, supra; Acid nucleic Hybridization, supra.

30 Dois fragmentos de ácido nucléico são considerados ser "seletivamente hibridizáveis" a um polinucleotídeo se eles são capazes de se hibridizarem especificamente a um ácido nucléico, ou uma variante deste, ou

especificamente acelerando uma reação de cadeia de polimerase: (i) sob hibridização típica e condições de lavagem, conforme descrito, por exemplo, em Sambrook et al. supra e Acid nucleic Hibridização, supra, (ii) usando-se condições de lavagem de estringência reduzida que permitem ao menos
5 cerca de 25-30% dessemparelhamentos de par base, por exemplo: 2x SSC, 0,1% de SDS, temperatura ambiente duas vezes, 30 minutos cada; em seguida 2x SSC, 0,1% de SDS, 37°C uma vez, 30 minutos; em seguida 2 x SSC temperatura ambiente duas vezes, 10 minutos cada, ou (iii) seleção de primers para uso em reações de cadeia de polimerase típicas (PCR) sob condições
10 padrão (descrito por exemplo, em Saiki, et al. (1988) Science 239:487-491).

O termo "capaz de hibridizar sob condições estringentes", conforme aqui usado, se refere a recozimento de um primeiro ácido nucléico a um segundo ácido nucléico sob condições estringentes, conforme definido abaixo. Condições de hibridização estringentes tipicamente permitem a hi-
15 bridização de moléculas de ácido nucléico tendo pelo menos 70% de identidade de sequência de ácido nucléico com a molécula de ácido nucléico sendo usada como uma sonda na reação de hibridização. Por exemplo, o primeiro ácido nucléico pode ser uma amostra teste ou sonda, e o segundo ácido nucléico pode ser o trançado de sentido ou de anti-sentido de um áci-
20 do nucléico, ou um fragmento deste. A hibridização dos primeiro e segundo ácidos nucléicos pode ser conduzida sob condições estringentes, por exemplo, alta temperatura e/ou baixo teor de sal, que tende a desfavorecer a hibridização de seqüências dissimilares de nucleotídeos. Alternativamente, a hibridização dos primeiro e segundo ácidos nucléicos pode ser conduzida
25 sob condições de estringência reduzidas, por exemplo, baixa temperatura e/ou alto teor de sal, que tende a favorecer a hibridização de seqüências dissimilares de nucleotídeos. Condições de hibridização de baixa estringência podem ser seguidas por condições de alta estringência, ou condições de estringência de meio intermediário para aumentar a seletividade da ligação
30 dos primeiro e segundo ácidos nucléicos. As condições de hibridização podem adicionalmente incluir reagentes, tais como, mas não limitados a, dimetil sulfoxido (DMSO) ou formamida para desfavorecer ainda adicionalmente a

hibridização de seqüências dissimilares de nucleotídeos. Um protocolo de hibridização adequado pode, por exemplo, envolver hibridização em 6 x SSC (no qual 1 x SSC compreende citrato de sódio 0,015 M e cloreto de sódio 0,15 M), a 65° Celsius em uma solução aquosa, seguido por lavagem com 1 x SSC a 65° C. Fórmulas para calcular hibridização apropriada e condições de lavagem para alcançar hibridização permitindo 30% ou menos de desequiparação entre duas moléculas de ácido nucléico são descritas, por exemplo, em Meinkoth et al. (1984) Anal. Biochem. 138: 267-284; o conteúdo do qual sendo aqui incorporado por referência em sua totalidade. Os protocolos para técnicas de hibridização são bem-conhecidos àqueles técnicos no assunto, e manuais de biologia molecular padrão podem ser consultados para selecionar um protocolo de hibridização adequado sem experimentação indevida. Ver, por exemplo, Sambrook et al. (2001) Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Press, os conteúdos do qual sendo aqui incorporados por referência em sua totalidade.

Tipicamente, condições estridentes serão aquelas em que a concentração de sal é menor do que cerca de 1,5 M de íon sódio, tipicamente cerca de 0,01 a 1,0 M de concentração de íon Na (ou outros sais) de cerca de pH 7,0 a cerca de pH 8,3, e a temperatura é pelo menos cerca de 30° Celsius para sondas curtas (por exemplo, 10 a 50 nucleotídeos), e pelo menos cerca de 60° C para sondas longas (por exemplo, maiores do que 50 nucleotídeos). Condições estridentes podem também ser alcançadas com a adição de agentes de desestabilização tal como formamida. Condições de baixa estringência adequadas incluem hibridização com uma solução tampão de 30 a 35% de formamida, 1 M NaCl, 1% de SDS (sódio dodecil sulfato) a 37° Celsius, e uma lavagem em 1-2 x SSC a 50 a 55° Celsius. Condições de estringência moderadas exemplares incluem hibridização em 40 a 45% de formamida, 1 M NaCl, 1% de SDS a 37° Celsius, e uma lavagem em 0,5-1 x SSC a 55 to 60° Celsius. Condições de estringência altas exemplares incluem hibridização em 50% de formamida, 1 M NaCl, 1% de SDS a 37° Celsius, e uma lavagem em 0,1 x SSC a 60 a 65° Celsius.

Os métodos e materiais da invenção podem ser usados mais

geralmente para avaliar uma amostra de DNA de um animal, geneticamente tipo um animal individual, e detectar diferenças genéticas em animais. Em particular, uma amostra de DNA genômico de um animal pode ser avaliada por referência a um ou mais controles para determinar se um SNP, ou grupo de SNPs, em um gene, está presente. Qualquer método para determinação do genótipo pode ser usado para determinação do genótipo na presente invenção. Tais métodos incluem, mas não estão limitados a, seqüenciamento de amplimer, seqüenciamento de DNA, espectroscopia de fluorescência, análise de hibridização baseada em transferência de energia de ressonância de fluorescência (ou "FRET"), classificação de alta produção, espectroscopia de massa, análise de microsatélite, hibridização de ácido nucléico, reação de cadeia de polimerase (PCR), análise de RFLP, e cromatografia de dimensão (por exemplo, cromatografia capilar ou de gel), todos dos quais são bem-conhecidos a um técnico no assunto. Em particular, métodos para determinação de polimorfismos de nucleotídeo, particularmente polimorfismo de nucleotídeo simples, são descritos nas Patentes dos Estados Unidos N^{os} 6,514,700; 6,503,710; 6,468,742; 6,448,407; 6,410,231; 6,383,756; 6,358,679; 6,322,980; 6,316,230; e 6,287,766, e revisados por Chen e Sullivan, *Pharmacogenomics J* 2003;3(2):77-96, as revelações dos quais sendo incorporadas por referência em sua totalidade. Dados genotípicos úteis nos métodos da invenção e métodos para a identificação e seleção de traços de animal são baseados na presença de SNPs.

Um "fragmento de restrição" se refere a um fragmento de um polinucleotídeo gerado por uma endonuclease de restrição (um enzima que cliva ligações de fosfodiéster dentro de uma cadeia de polinucleotídeo) que cliva DNA em resposta a um local de reconhecimento no DNA. O local de reconhecimento (local de restrição) consiste em uma seqüência específica de nucleotídeos tipicamente cerca de 4-8 nucleotídeos de comprimento.

A "polimorfismo de nucleotídeo simples" ou "SNP" se refere a uma variação na seqüência de nucleotídeo de um polinucleotídeo que difere de outro polinucleotídeo por uma diferença simples de nucleotídeo. Por exemplo, sem limitação, a troca de um A por um C, G ou T na seqüência total

de polinucleotídeo constitui um SNP. É possível ter-se mais do que um SNP em um polinucleotídeo particular. Por exemplo, em uma posição em um polinucleotídeo, um C pode ser trocado por um T, em outra posição, um G pode ser trocado por um A, e assim por diante. Quando se refere à SNPs, o polinucleotídeo é mais freqüentemente DNA.

Conforme aqui usado, um "gabarito" se refere a um trançado de polinucleotídeo alvo, por exemplo, sem limitação, um trançado de DNA que ocorre naturalmente não-modificado, que uma polimerase usa como um meio de reconhecer qual nucleotídeo deve em seguida se incorporar em um trançado em crescimento para polimerizar o complemento do trançado que ocorre naturalmente. Tal trançado de DNA pode ser de trançado simples, ou pode ser parte de um gabarito de DNA de trançado duplo. Nas aplicações da presente invenção que requer ciclos repetidos de polimerização, por exemplo, a reação de cadeia de polimerase (PCR), o próprio trançado do gabarito pode se tornar modificado pela incorporação de nucleotídeos modificados, servindo ainda como um gabarito para uma polimerase para sintetizar polinucleotídeos adicionais.

Uma "reação termocíclica" é uma reação de etapa múltipla na qual pelo menos duas etapas são efetuadas pela mudança da temperatura da reação.

Uma "variação" é uma diferença na seqüência de nucleotídeo entre polinucleotídeos relacionados. A diferença pode ser a retirada de um ou mais nucleotídeos a partir da seqüência de um polinucleotídeo comparada à seqüência de um polinucleotídeo relacionado, a adição de um ou mais nucleotídeos ou a substituição de um nucleotídeo por outro. Os termos "mutação," "polimorfismo" e "variação" são usados aqui permutavelmente. Conforme aqui usado, o termo "variação" no singular é para ser construído para incluir variações múltiplas; isto é, duas ou mais adições, retiradas e/ou substituições de nucleotídeo no mesmo polinucleotídeo. Uma "mutação de ponto" se refere a uma substituição simples de um nucleotídeo por outro.

Conforme aqui usado, os termos "traços", "traços de qualidade" ou "características físicas" ou "fenótipos" se referem a propriedades vantagio-

5 sas do animal resultante das genéticas. Traços de qualidade incluem, mas não estão limitados a, a capacidade genética do animal de metabolizar eficientemente energia, produzir carne ou leite, acumular gordura intramuscular. As características físicas incluem, mas não estão limitadas a, carnes marmoreadas, macias ou magras. Os termos podem ser usados permutavelmente.

Um "sistema de computador" se refere a meios de hardware, meios de software e meios de armazenagem de dados usados para compilar os dados da presente invenção. Os meios mínimos de hardware de sistemas baseados em computador da invenção podem compreender uma unidade de
10 processamento central (CPU), meios de entrada, meios de saída, e meios de armazenagem de dados. Desejavelmente, um monitor é provido para visualizar os dados da estrutura. Os meios de armazenagem de dados podem ser RAM ou outros meios para acessar meios legíveis de computador. Exemplos de tais sistemas são estações de operação de microcomputador disponíveis
15 de Silicon Graphics Incorporated e Sun Microsystems running Unix based, Linux, Windows NT, XP, ou sistemas de operação IBM OS/2.

"Meios de leitura de computador" se referem a quaisquer meios que podem ser lidos e acessados diretamente por um computador, e incluem, mas não é limitado a: meios de armazenagem magnética tais como dis-
20 cos flexíveis, meios de armazenagem rígidos e fitas magnéticas; meios de armazenagem óticos, tais como discos óticos ou CD-ROM; meios de armazenagem elétricos, tais como RAM e ROM; e híbridos destas categorias, tais como meios magnéticos/óticos. Pela provisão de meios de leitura de computador, os dados compilados em um animal particular podem ser rotineira-
25 mente acessados por um usuário, por exemplo, um operador de alimento.

O termo "módulo de análise de dados" é definido aqui para incluir qualquer pessoa ou máquina, que operam individualmente ou juntos, que analisam a amostra e determinam a informação genética contida nos mesmos. O termo pode incluir uma pessoa ou máquina dentro de um ajuste de
30 laboratório.

Conforme aqui usado, o termo "módulo de coleta de dados" se refere a qualquer pessoa, objeto ou sistema que obtém uma amostra de te-

cido de um animal ou embrião. Por exemplo, e sem limitação, o termo pode definir, individualmente ou coletivamente, a pessoa ou máquina em contato físico com o animal conforme a amostra é tomada, os recipientes que mantêm as amostras de tecido, o acondicionamento usado para transporte das amostras, e similares. Vantajosamente, o coletor de dados é uma pessoa. Mais vantajosamente, o coletor de dados é um fazendeiro, um criador ou um veterinário.

O termo "interface de rede" é definido aqui para incluir qualquer pessoa ou sistema de computador capaz de acessar dados, depositar dados, combinar dados, analisar dados, pesquisar dados, transmitir dados, ou armazenar dados. O termo é amplamente definido para ser uma pessoa que analisa os dados, os sistemas de hardware e software eletrônicos usados na análise, às bases de dados que armazenam a análise de dados, e quaisquer meios de armazenagem capazes de armazenar os dados. Exemplos não-limitativos de interfaces de rede incluem pessoas, equipamento automático de laboratório, computadores e redes de computador, dispositivos de armazenagem de dados, tais como, mas não limitados a, discos, acionamentos rígidos ou chips de memória.

O termo "história de procriação", conforme aqui usado, se refere a um registro da vida de um animal ou grupo de animais incluindo, mas não limitado a, a localização, geração, período de alojamento, bem como história genética dos animais, incluindo ascendentes e descendentes destes, genótipo, fenótipo, história transgênica, se relevante, e similares.

O termo "condições de criação de gado", conforme aqui usado, se refere aos parâmetros relacionados à manutenção de animais incluindo, mas não limitado a, temperatura do abrigo ou do alojamento, mortalidade semanal de um rebanho, consumo de água, consumo de alimento, taxa de ventilação e qualidade, condição do feno, e similares.

O termo "história veterinária", conforme aqui usado, se refere a dados de vacinação de um animal ou grupo de animais, incluindo, mas não limitado a, tipo(s) de vacina, número(s) de série da batelada de vacina, dose administrada, antígeno alvo, método de administração da vacina ao(s) ani-

mal(is) recipiente(s), número de animais vacinados, idade dos animais e do vacinador. Os dados relacionados a uma resposta sorológica ou imunológica induzida pela vacina podem também ser incluídos. "História veterinária", conforme aqui usado, é também pretendida para incluir as histórias da medicação do(s) animal(is) alvo(s) incluindo, mas não limitado a, fármaco e/ou antibióticos administrados aos animais incluindo tipo de medicação administrada, quantidade e taxas de dose, por quem e quando administrada, por qual rota, por exemplo, oral, subcutânea, e similares, e a resposta à medicação, incluindo efeitos colaterais desejados e indesejáveis desta.

10 O termo "dados diagnósticos", conforme aqui usado, se refere a dados relacionados à saúde do(s) animal(is) diferentes dos dados de detalhamento da história de vacinação ou medicação do(s) animal(is). Por exemplo, os dados diagnósticos podem ser um registro das infecções experimentadas pelo(s) animal(is), e a resposta destes às medicações providas para
15 tratar tais medicações. Dados sorológicos incluindo anticorpo ou composição de proteína do soro, ou outros biofluidos, podem também ser dados diagnósticos úteis para admissão nos métodos da invenção. Dados cirúrgicos pertencentes ao(s) animal(is) podem ser incluídos, tais como o tipo de manipulação cirúrgica, resultado da cirurgia e complicações que ocorrem do procedimento cirúrgico. "Dados diagnósticos" podem também incluir medições de
20 parâmetros tais como peso, morbidez, e outras características notadas por um serviço de veterinária, tais como a condição da pele, patas, etc.

O termo "dados de bem-estar", conforme aqui usado, se refere ao acúmulo coletivo de dados pertencentes a um animal ou grupo de animais incluindo, mas não limitado a, uma história de procriação, um história
25 veterinária, um perfil de bem-estar, dados diagnósticos, dados de controle de qualidade, ou qualquer combinação destes.

O termo "perfil de bem-estar", conforme aqui usado, se refere a parâmetros, tais como peso, densidade da carne, níveis de abarrotamento nos invólucros de procriação ou criação, comportamento psicológico do animal,
30 taxa de crescimento e qualidade, e similares.

O termo "controle de qualidade", conforme aqui usado, se refere

às características desejadas do(s) animal(is). Por animais não-domésticos tais como gado e ovelha, por exemplo, tais parâmetros incluem qualidade do músculo e densidade, teor de gordura, maciez da carne, produção de leite e qualidade, capacidade de procriação, e similares.

- 5 O termo "parâmetros de desempenho", conforme aqui usado, se refere fatores tais como produção de carne, produção de procriação, forma de leiteria, qualidade da carne, taxa de gravidez de raça (isto é, fertilidade), vida produtiva (isto é, longevidade), e similares, que podem ser as metas desejadas para a procriação e criação do(s) animal(is). Parâmetros de de-
10 sempenho podem ser, ou gerados dos próprios animais, ou aqueles parâmetros desejados por um cliente, ou o mercado.

O termo "dados nutricionais", conforme aqui usado, se refere à composição, quantidade e frequência de distribuição de alimento, incluindo água, providos ao(s) animal(is).

- 15 O termo "segurança de alimento", conforme aqui usado, se refere à qualidade da carne de um animal de criação, incluindo, mas não limitado a, tempo de preparação, lugar e maneira, armazenagem do produto alimentício, rota de transporte, registros de inspeção, textura, cor, sabor, odor, teor bacterial, teor parasítico, e similares.

- 20 Será aparente àqueles técnicos no assunto que os dados relacionados à saúde e manutenção dos animais podem ser variavelmente agrupados, dependendo da fonte ou intenção do coletor de dados, e qualquer um agrupamento aqui não é, portanto, pretendido para ser limitante.

- A menos que de outro modo definido, todos os termos técnicos e
25 científicos aqui usados têm o mesmo significado conforme comumente compreendido por um técnico no assunto de biologia molecular. Embora métodos e materiais similares ou equivalentes àqueles aqui descritos possam ser usados na prática ou teste da presente invenção, métodos e materiais adequados são aqui descritos.

- 30 Em uma concretização na qual o gene de interesse é *CAST* bovino, a seqüência de nucleotídeo de *CAST* bovino pode ser selecionada de, mas não é limitada a, a seqüência correspondente aos N^ºs de Acesso no

GenBank AAFC02060382 (SEQ ID NO: 1), AAFC02161394 (SEQ ID NO: 2), AAFC02179490 (SEQ ID NO: 3), AAFC02060385 (SEQ ID NO: 4), AAFC02060381 (SEQ ID NO: 5), AAFC02197217 (SEQ ID NO: 6), AAFC02067026 (SEQ ID NO: 7), ou um fragmento desta, um uma região do
 5 genoma bovino genoma que compreende esta seqüência.

A presente invenção, portanto, proporciona ácidos nucléicos isolados que podem especificamente hibridizar para a seqüência de nucleotídeo que pode ser selecionada de, mas não é limitada a, a seqüência correspondente aos N^{os} de Acesso no GenBank AAFC02060382 (SEQ ID NO: 1),
 10 AAFC02161394 (SEQ ID NO: 2), AAFC02179490 (SEQ ID NO: 3), AAFC02060385 (SEQ ID NO: 4), AAFC02060381 (SEQ ID NO: 5), AAFC02197217 (SEQ ID NO: 6), AAFC02067026 (SEQ ID NO: 7), ou o complemento desta, e que compreende o local polimórfico correspondente.

O(s) polimorfismo(s) de nucleotídeo simples de interesse pode(m) ser selecionado(s) a partir do grupo consistindo em mutações sem sentido na região de domínio XL, especialmente mutações sem sentido em exon 3 que resultam em substituições G48D e P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5].
 15
 20

O SNP/STR vantajoso na presente invenção está associado com certos traços economicamente valiosos e hereditários relacionados à fertilidade e longevidade nos bovinos. Portanto, é um objetivo da presente invenção determinar o genótipo de um dado animal de interesse conforme definido pelo SNP/STR de local de *CAST* de acordo com a presente invenção. É
 25 também contemplado que o genótipo do(s) animal(is) pode ser definido por SNPs/STRs adicionais dentro do gene de *CAST* ou dentro de outros genes identificados com traços desejáveis ou outras características, e, em particular, por um painel ou painéis de SNPs/STRs.

Existem muitos métodos conhecidos na técnica para determinação da seqüência de DNA em uma amostra, e para identificação se uma dada amostra de DNA contém um SNP/STR particular. Qualquer técnica co-
 30

nhcida na técnica pode ser usada no desempenho dos métodos da presente invenção.

Os métodos da presente invenção permitem que animais com certos traços hereditários economicamente valiosos a serem identificados baseados na presença de SNPs/STRs em seus genomas, e, particularmente, com mutações sem sentido em exon 3 do gene de *CAST*. Os métodos adicionalmente permitem, pelos métodos auxiliados por computador da invenção, correlacionar os traços associados ao SNP/STR com outros dados pertinentes à capacidade bem feita e produtiva dos animais, ou grupo de animais.

Para determinar o genótipo de um dado animal de acordo com os métodos da presente invenção, é necessário obter-se uma amostra de DNA genômico daquele animal. Tipicamente, esta amostra de DNA genômico será obtida de uma amostra de tecido ou células tomadas daquele animal. Uma amostra de tecido ou célula pode ser tomada de um animal a qualquer momento no tempo de vida de um animal, mas antes de que a identidade de carcaça seja perdida. A amostra de tecido pode compreender pelo menos, incluindo raízes, couro, osso, hastes de coleta bucais, sangue, saliva, leite, sêmen, embriões, músculo, ou quaisquer órgãos internos. Nos métodos da presente invenção, a fonte da amostra de tecido, e, desse modo, também a fonte da amostra teste de ácido nucléico, não é crítica. Por exemplo, o ácido nucléico teste pode ser obtido de células no interior de um fluido corpóreo do animal, ou de células que constituem um tecido de corpo do animal. O fluido corpóreo particular do qual as células são obtidas não é também crítico a presente invenção. Por exemplo, o fluido corpóreo pode ser selecionado a partir do grupo consistindo em sangue, ascites, fluido pleural e fluido espinhal. Além disso, o tecido de corpo particular do qual células são obtidas não é também crítico a presente invenção. Por exemplo, o tecido de corpo pode ser selecionado a partir do grupo consistindo em pele, tecido endometrial, uterino e cervical. Ambos os tecidos normal e tumoral podem ser usados.

Tipicamente, a amostra de tecido é marcada com um número de

identificação, ou outro índice que se relaciona a amostra ao animal individual do qual a amostra foi tomada. A identidade da amostra vantajosamente permanece constante através de todos os métodos e sistemas da invenção, garantindo, desse modo, a integridade e continuidade da amostra durante
5 extração e análise. Alternativamente, o índice pode ser mudado em um modo regular que assegura que os dados, e quaisquer outros dados associados, podem estar relacionados de volta ao animal do qual os dados foram obtidos.

A quantidade/tamanho de amostra requerida é conhecido àqueles técnicos no assunto, e, por exemplo, pode ser determinada pelas etapas
10 subseqüentes usadas no método e sistema da invenção, e nos métodos específicos de análise usados. Idealmente, o tamanho/volume da amostra de tecido recuperada deve ser tão consistente quanto possível dentro do tipo de amostra e da espécie de animal. Por exemplo, para gado, exemplos não-
15 limitativos de tamanhos/métodos de amostra incluem carne não-gordurosa: 0,0002 gm-10,0 gm; puro: 0,0004 gm-10,0 gm; raízes de pelo: pelo menos um e vantajosamente mais do que cinco; hastes de coleta bucais: 15 a 20 segundos de esfrega com pressão modesta na área entre lábio externo e goma usando-se, por exemplo, uma escova de citologia; osso: 0,0002 gm-
20 10,0 gm; sangue: 30 μ l a 50 ml.

Geralmente, a amostra de tecido é colocada em um recipiente que é etiquetado usando-se um sistema de numeração suportando um código correspondente ao animal, por exemplo, a etiqueta de orelha do animal. Conseqüentemente, o genótipo de um animal particular é facilmente deter-
25 minável em todos os tempos. O dispositivo de amostragem e/ou recipiente pode ser suprido ao fazendeiro, um matadouro ou varejista. O dispositivo de amostragem vantajosamente toma uma amostra consistente e reprodutível de animais individuais, enquanto simultaneamente evita qualquer contaminação cruzada de tecido. Conseqüentemente, o tamanho e volume de tecidos de amostra derivados de animais individuais seriam consistentes.
30

O DNA pode ser isolado a partir do tecido/células por técnicas conhecidas àqueles técnicos no assunto (ver, por exemplo, Patente dos Es-

tados Unidos N^{os} 6.548.256 e 5.989.431; Hirota et al. (1989) Jinrui Idengaku Zasshi. 34: 217-23 e John et al. (1991) Acid nucleics Res. 19:408, as revelações das quais sendo aqui incorporadas por referência em suas totalidades). Por exemplo, DNA de alto peso molecular pode ser purificado de células ou
5 tecido usando-se extração de proteinase K e precipitação de etanol. O DNA, contudo, pode ser extraído de uma espécime de animal usando-se quaisquer outros métodos conhecidos na técnica.

Em uma concretização, a presença ou ausência do SNP/STR de qualquer dos genes da presente invenção pode ser determinada pelo seqüenciamento da região da amostra de DNA genômico que se estende sobre o local polimórfico. Muitos métodos de seqüenciamento de DNA genômico são conhecidos na técnica, e qualquer método pode ser usado, ver, por exemplo, Sambrook et al. (2001) Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Press. Por exemplo, conforme descrito abaixo,
15 um fragmento de DNA que se estendendo sobre a locação do SNP/STR de interesse pode ser amplificado usando-se a reação de cadeia de polimerase. A região amplificada de forma de DNA pode, em seguida, ser seqüenciada usando-se qualquer método conhecido na técnica, por exemplo, usando-se um seqüenciador automático de ácido nucléico. A detecção de um dado
20 SNP/STR pode, em seguida, ser realizada usando-se hibridização de sondas, e ou usando-se métodos de amplificação baseados em PCR. Tais métodos são descritos em maiores detalhes abaixo.

Os métodos da presente invenção podem usar oligonucleotídeos úteis como priméres para amplificar seqüências específicas de ácido nucléico do gene de *CAST*, vantajosamente da região envolvendo um SNP/STR de *CAST*. Tais fragmentos devem ser de comprimento suficiente para capacitar recozimento específico ou hibridização para a amostra de ácido nucléico. As seqüências tipicamente serão cerca de 8 a cerca de 44 nucleotídeos de comprimento. Seqüências mais longas, por exemplo, de cerca de 14 a
30 cerca de 50, podem ser vantajosas para certas concretização. O desenho de priméres é bem-conhecido do técnico no assunto.

As moléculas de ácido nucléico da invenção incluem moléculas

de ácido nucléico tendo pelo menos 70% de identidade ou homologia ou similaridade com um gene de *CAST*, ou sondas ou prímeres derivados destes, tais como pelo menos 75% de identidade ou homologia ou similaridade, preferivelmente pelo menos 80% de identidade ou homologia ou similaridade, mais preferivelmente pelo menos 85% de identidade ou homologia ou similaridade, tais como pelo menos 90% de identidade ou homologia ou similaridade, mais preferivelmente pelo menos 95% de identidade ou homologia ou similaridade, tais como pelo menos 97% de identidade ou homologia ou similaridade. A similaridade, homologia ou identidade da seqüência de nucleotídeo pode ser determinada usando-se o programa "Align" de Myers e Miller, ("Optimal Alignments in Linear Space", CABIOS 4, 11-17, 1988) e disponível em NCBI. Alternativamente ou adicionalmente, os termos "similaridade" ou "identidade" ou "homologia", por exemplo, com relação a uma seqüência de nucleotídeo, são pretendidos para indicarem uma medida quantitativa de homologia entre duas seqüências. A percentagem de similaridade de seqüência pode ser calculada como $(N_{ref} - N_{dif}) * 100 / N_{ref}$, no qual N_{dif} é o número total de resíduos não-idênticos nas duas seqüências quando alinhadas, e no qual N_{ref} é o número de resíduos em uma das seqüências. Conseqüentemente, a seqüência de DNA AGTCAGTC terá uma similaridade de seqüência de 75% com a seqüência AATCAATC ($N_{ref} = 8$; $N_{dif} = 2$). Alternativamente ou adicionalmente, "similaridade" com relação a seqüências se refere ao número de posições com nucleotídeos idênticos divididos pelo número de nucleotídeos na mais curta das duas seqüências nas quais alinhamento das duas seqüências pode ser determinado de acordo com o algoritmo de Wilbur e Lipman (Wilbur e Lipman, 1983 PNAS USA 80:726), por exemplo, usando-se um tamanho de janela de 20 nucleotídeos, um comprimento de palavra de 4 nucleotídeos, e uma penalidade de folga de 4, e análise auxiliada por computador e interpretação dos dados de seqüência, incluindo alinhamento, podem ser convenientemente realizados usando-se programas comercialmente disponíveis (por exemplo, Intelligenetics® Suite, Intelligenetics Inc. CA). Quando seqüências de RNA são referidas para serem similares, ou terem um grau de identidade de seqüência com seqüências de DNA, timidi-

na (T) na seqüência de DNA é considerada igual à uracil (U) na seqüência de RNA.

Uma sonda ou primer pode ser qualquer extensão de pelo menos 8, preferivelmente pelo menos 10, mais preferivelmente pelo menos 12, 13, 14, ou 15, tais como pelo menos 20, por exemplo, pelo menos 23 ou 25, por exemplo, pelo menos 27, ou 30 nucleotídeos em um gene de *CAST* que são únicos a um gene de *CAST*. Como para PCR ou primers ou sondas de hibridização e comprimentos ótimos destes, referência é também feita a Kajimura et al., *GATA* 7(4):71-79 (1990).

As seqüências de RNA dentro do escopo da invenção são derivadas a partir das seqüências de DNA, por timidina (T) na seqüência de DNA sendo considerada igual à uracil (U) nas seqüências de RNA.

Os oligonucleotídeos podem ser produzidos por um processo de produção convencional para oligonucleotídeos gerais. Eles podem ser produzidos, por exemplo, por um processo de síntese química, ou por um processo microbial que faz uso de um vetor de plasmídeo, um vetor de fago, ou similares. Adicionalmente, é adequado usar-se um sintetizador de ácido nucléico.

Para etiquetar um oligonucleotídeo com o corante fluorescente, um dos métodos de etiquetagem convencionalmente conhecido pode ser usado (Tyagi & Kramer (1996) *Nature Biotechnology* 14: 303-308; Schofield et al. (1997) *Appl. e Environ. Microbiol.* 63: 1143-1147; Proudnikov & Mirzabekov (1996) *Nucl. Acids Res.* 24: 4532-4535). Alternativamente, o oligonucleotídeo pode ser etiquetado com uma radioetiqueta, por exemplo, ^3H , ^{125}I , ^{35}S , ^{14}C , ^{32}P , etc. Métodos de etiquetamento bem-conhecidos são descritos, por exemplo, em Sambrook et al. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Press. A etiqueta é acoplada diretamente ou indiretamente a um componente do oligonucleotídeo de acordo com métodos bem-conhecidos na técnica. Cromatografia de fase reversa ou similares usadas para proporcionar uma sonda de ácido nucléico para uso na presente invenção podem purificar o oligonucleotídeo sintetizado etiquetado com um marcador. Uma forma de sonda vantajosa é uma etiquetada com

um corante fluorescente na extremidade 3'- ou 5'- e contendo G ou C como a base na extremidade etiquetada. Se a extremidade 5'- é etiquetada e a extremidade 3'- não é etiquetada, o grupo OH no átomo de C na posição 3'- da extremidade 3'-, ribose ou deoxirribose podem ser modificadas com um grupo fosfato ou similares, embora nenhuma limitação seja imposta neste particular.

Durante a hibridização do ácido nucléico alvo com as sondas, condições estridentes podem ser utilizadas, vantajosamente junto com outras condições que afetam a estridência, para auxiliar na hibridização. A detecção por rompimento diferencial é particularmente vantajosa para reduzir ou eliminar hibridização de escorregamento entre sondas e alvo, e para promover hibridização mais efetiva. Em ainda outro aspecto, condições de estridência podem ser variadas durante a determinação de estabilidade complexa de hibridização de modo a determinar precisamente e rapidamente se um SNP está presente na seqüência alvo.

Um método de determinação do genótipo no local de gene polimórfico envolve obtenção de uma amostra de ácido nucléico, hibridização da amostra de ácido nucléico com uma sonda, e rompimento da hibridização para determinar o nível de energia de rompimento requerida no qual a sonda tem uma energia de rompimento diferente para um alelo, conforme comparado a outro alelo. Em um exemplo, pode ser uma energia de rompimento mais baixa, por exemplo, temperatura de fusão, para o alelo que propicia um resíduo de citosina em um local polimórfico, e uma energia requerida mais alta para um alelo com um resíduo diferente naquele local polimórfico. Isto pode ser alcançado onde a sonda tem 100% de homologia com um alelo (uma sonda perfeitamente equiparada), mas tem um desequiparação simples com o alelo alternativo. Desde que a sonda perfeitamente equiparada é ligada mais vedantemente ao DNA alvo do que a sonda desequiparada, ela requer mais energia para fazer com que a sonda hibridizada se disassocie.

Em uma etapa adicional do método acima, uma segunda sonda ("âncora") pode ser usada. Geralmente, a sonda âncora não é específica para qualquer alelo, mas hibridiza indiferente de qual nucleotídeo está pre-

sente no local polimórfico. A sonda âncora não afeta a energia de rompimento requerida para disassociar o complexo de hibridização, mas, ao invés, contém uma etiqueta complementar para se usar com a primeira sonda ("sensor").

5 A estabilidade de hibridização pode ser influenciada por numerosos fatores, incluindo termorregulação, regulação química, bem como controle de estirgência eletrônico, ou sozinho ou em combinação com outros fatores listados. Através do uso de condições de estirgência, em qualquer ou ambas da etapa de hibridização alvo ou a etapa de estirgência de oligo-
10 nucleotídeo de sensor, completação rápida do processo pode ser alcançada. Isto é desejável para alcançar hibridização corretamente indexada do DNA alvo para alcançar número máximo de moléculas em um local de teste com um complexo de hibridização preciso. Por meio de exemplo, com o uso de estirgência, a etapa de hibridização inicial pode ser completada em dez
15 minutos ou menos, mais vantajosamente cinco minutos ou menos, e mais vantajosamente dois minutos ou menos. Além de tudo, o processo analítico pode ser completado em menos do que metade de uma hora.

 Em um modo, o complexo de hibridização é etiquetado, e a etapa de determinação da quantidade de hibridização inclui detecção das quan-
20 tidades de complexo de hibridização etiquetado nos locais de teste. O dispositivo de detecção e método podem incluir, mas não estão limitados a, imagem ótica, imagem eletrônica, imagem com câmera CCD, imagem ótica integrada, e espectrometria de massa. Adicionalmente, a quantidade de sonda etiquetada ou não-etiquetada ligada ao alvo pode ser quantificada. Tal quan-
25 tificação pode incluir análise estatística. A porção etiquetada do complexo pode ser o alvo, o estabilizador, a sonda, ou o complexo de hibridização "in toto". A etiquetação pode ser por etiquetação fluorescente selecionada a partir do grupo de, mas não limitado a, Cy3, Cy5, Bodipy Texas Red, Bodipy Far Red, Lucifer Yellow, Bodipy 630/650-X, Bodipy R6G-X e 5-CR 6G. Etiqueta-
30 ção colormétrica, etiquetação bioluminescente, e/ou etiquetação quimioluminescente, podem adicionalmente efetuar etiquetação. A etiquetação adicionalmente pode incluir transferência de energia entre moléculas no complexo

de hibridização por análise de perturbação, extinção, transporte de elétron entre moléculas doadoras e receptoras, a última da qual podendo ser facilitada por complexos de hibridização de equiparação de trançado duplo. Op-
cionalmente, se o complexo de hibridização é não-etiquetado, detecção po-
5 de ser acompanhada pela medição de condutância diferencial entre DNA de trançado duplo e de trançado não-duplo. Adicionalmente, detecção direta pode ser alcançada por interferometria ótica baseada em silício poroso, ou por espectrometria de massa. Usando-se espectrometria de massa, nenhuma etiqueta fluorescente ou outra etiqueta é necessária. Preferivelmente
10 detecção é obtida por níveis extremamente altos de resolução de massa alcançada por medição direta, por exemplo, por tempo de voo (TOF), ou por ionização de pulverização de elétron (ESI). Onde espectrometria de massa é contemplada, sondas tendo uma seqüência de ácido nucléico de 50 bases ou menos são vantajosas.

15 A etiqueta pode ser amplificada, e pode incluir, por exemplo, DNA ramificado ou dendrítico. Se o DNA alvo é purificado, ele pode ser não-amplificado ou amplificado. Adicionalmente, se o alvo purificado é amplificado, e a amplificação é um método exponencial, ele pode ser, por exemplo, DNA amplificado por, ou DNA amplificado por amplificação de deslocamento
20 de trançado (SDA). Métodos lineares de amplificação de DNS, tais como círculo de rolamento, "runoff" transcricional, podem também ser usados.

Onde é desejado amplificar um fragmento de DNA que compreende um SNP/STR de acordo com a presente invenção, os priméres de avanço e reverso podem ter extensões contíguas de cerca de 8, 9, 10, 11, 12,
25 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, ou qualquer outro comprimento até e incluindo cerca de 50 nucleotídeos de comprimento. As seqüências as quais os priméres de avanço e reverso recozem estão vantajosamente localizados em qualquer lado da posição particular de nucleotídeo que é substituída no SNP/STR a ser amplificado.

30 Uma etiqueta destacável pode ser incorporada em um ácido nucléico durante pelo menos um ciclo de uma reação de amplificação. Meios espectroscópicos, fotoquímicos, bioquímicos, imunoquímicos, elétricos, óti-

cos ou químicos podem detectar tais etiquetas. Etiquetas úteis na presente invenção incluem corantes fluorescentes (por exemplo, isotiocianato de fluoresceína, Texas red, rodamina, e similares), radioetiquetas (por exemplo, ^3H , ^{125}I , ^{35}S , ^{14}C , ^{32}P , etc.), enzimas (por exemplo, horseradish peroxidase, fosfatase alcalina, etc.), etiquetas colorimétricas, tais como ouro coloidal ou gotas de vidro ou plástico coloridas (por exemplo, poliestireno, polipropileno, latex, etc.). A etiqueta é acoplada diretamente ou indiretamente a um componente do ensaio de acordo com métodos bem-conhecidos na técnica. Conforme indicado acima, uma ampla variedade de etiquetas são usadas, com a escolha de etiqueta dependendo da sensibilidade requerida, facilidade de conjugação com o composto, requerimentos de estabilidade, instrumentação disponível, e provisões dispostas. Etiquetas não-radioativas são frequentemente fixadas por meios indiretos. Polimerases podem também incorporar nucleotídeos fluorescentes durante síntese de ácidos nucleicos.

Os reagentes que permitem o seqüenciamento de produtos de reação podem ser utilizados aqui. Por exemplo, nucleotídeos de terminação de cadeia frequentemente serão incorporados em um produto de reação durante um ou mais ciclos de uma reação. Kits comerciais contendo os reagentes mais tipicamente usados para estes métodos de seqüenciamento de DNA são disponíveis e amplamente usados. Métodos de digestão de exonuclease de PCR para seqüenciamento de DNA podem também ser usados. Muitos métodos de seqüenciamento de DNA genômico são conhecidos na técnica, e qualquer tal método pode ser usado, ver, por exemplo, Sambrook et al. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Press. Por exemplo, conforme descrito abaixo, um fragmento de DNA se estendendo na locação do SNP de interesse pode ser amplificado usando-se a reação de cadeia de polimerase ou alguma outra reação de amplificação mediada por polimerase cíclica. A região amplificada de DNA pode, em seguida, ser seqüenciada usando-se qualquer método conhecido na técnica. Vantajosamente, o seqüenciamento de ácido nucleico é por métodos automáticos (revistos por Meldrum, (2000) *Genome Res.* 10: 1288-303, a revelação do qual sendo incorporado por referência em sua totalida-

de), por exemplo, usando-se um Beckman CEQ 8000 Genetic Analysis System (Beckman Coulter Instruments, Inc.). Métodos para seqüenciamento de ácidos nucleicos incluem, mas não estão limitados a, seqüenciamento de DNA fluorescente automático (ver, por exemplo, Watts & MacBeath, (2001) Methods Mol Biol. 167: 153-70 e MacBeath et al. (2001) Methods Mol Biol. 167:119-52), eletrophorese capillar (ver, por exemplo, Bosserhoff et al. (2000) Comb Chem High Throughput Screen. 3: 455-66), chips de seqüenciamento de DNA (ver, por exemplo, Jain, (2000) Pharmacogenomics. 1: 289-307), espectrometria de massa (ver, por exemplo, Yates, (2000) Trends Genet. 16: 5-8), piroseqüenciamento (ver, por exemplo, Ronaghi, (2001) Genome Res. 11: 3-11), eletroforese de gel de camada ultra fina (ver, por exemplo, Guttman & Ronai, (2000) Electrophoresis. 21: 3952-64), as descrições dos quais são, desse modo, incorporadas por referência em suas totalidades. O seqüenciamento pode também ser feito por uma companhia comercial. Exemplos de tais companhias incluem, mas não estão limitados a, a University of Georgia Molecular Genetics Instrumentation Facility (Athens, Georgia) ou SeqWright DNA Technologies Services (Houston, Texas).

Uma sonda específica de SNP/STR pode também ser usada na detecção do SNP/STR em seqüências de ácido nucleico amplificadas específicas do gene alvo, tais como os produtos de PCR amplificados gerados usando-se os priméres descritos acima. Em certas concretizações, estas sondas de SNP/STR específicas consistem em fragmentos de oligonucleotídeo. Vantajosamente, os fragmentos são de comprimento suficiente para proporcionar hibridização específica para a amostra de ácido nucleico. O uso de uma sonda de hibridização de entre 10 e 50 nucleotídeos de comprimento permite a formação de uma molécula dúplex que é ambos estável e seletiva. As moléculas tendo seqüências complementares sobre extensões maiores do que 12 bases de comprimento são geralmente vantajosas, de modo a aumentar estabilidade e seletividade do híbrido, e, desse modo, aperfeiçoar a qualidade e grau de moléculas híbridas particulares obtidas. Geralmente se preferirá designar moléculas de ácido nucleico tendo extensões de 16 a 24 nucleotídeos, ou ainda mais longas, onde desejado. Uma região de nu-

cleotídeo de etiqueta pode ser incluída, como na extremidade 5' do primer que pode proporcionar um local ao qual um primer de seqüenciamento de oligonucleotídeo pode hibridizar para facilitar o seqüenciamento de amostras múltiplas de PCR.

5 A seqüência de sonda deve se estender sobre a posição de nucleotídeo particular que pode ser substituída no SNP particular a ser detectado. Vantajosamente, duas ou mais "sondas específicas de alelo" diferentes podem ser usadas para análise de um SNP, uma primeira sonda específica de alelo para detecção de um alelo, e uma segunda sonda específica de alelo para a detecção do alelo alternativo.

 Será compreendido que esta invenção não é limitada aos primeiros particulares e sondas aqui descritas, e é pretendido envolver pelo menos seqüências de ácido nucléicos que são hibridizáveis à seqüência de nucleotídeo aqui descrita, o complemento ou um fragmento desta, ou são análogas de seqüência funcional destas seqüências. É também contemplado que um traço particular de um animal pode ser determinado pelo uso de um painel de SNPs/STRs associado com este traço. Vários traços economicamente relevantes podem ser caracterizados pela presença ou ausência de um ou mais SNPs/STRs, e por uma pluralidade de SNPs/STRs em genes diferentes. Um ou mais painéis de SNPs/STRs podem ser usados nos métodos da invenção para definir o perfil fenotípico do animal abjeto.

 Homólogos (isto é, ácidos nucléicos derivados de outras espécies), ou outras seqüências relacionadas (por exemplo, paralogos) podem ser obtidos sob condições de hibridização padrão ou estrigente com toda ou uma porção da seqüência particular como uma sonda usando-se métodos bem-conhecidos na técnica para hibridização e clonagem de ácido nucléico.

 Os marcadores genéticos, sondas destes, métodos, e kits da invenção são também úteis em um programa de procriação para selecionar procriação daqueles animais tendo fenótipos desejáveis para vários traços economicamente importantes, tais como fertilidade (taxa de gravidez de raça) e longevidade (vida produtiva). Seleção contínua e procriação de animais, tais como criação, que são pelo menos heterozigotos e vantajosamen-

te homozigotos para alelos desejáveis dos locais polimórficos de gene de *CAST* associados com traços economicamente relevantes de crescimento, entrada de alimentação, eficiência e/ou mérito de carcaça, e reprodução e longevidade, conduziriam a uma geração, linha ou população tendo números
 5 mais altos de descendência, com traços economicamente relevantes de crescimento, entrada de alimentação, eficiência e mérito de carcaça, e reprodução e longevidade. Desse modo, os SNPs/STRs de *CAST* da presente invenção podem ser usados como uma seleção de uma ferramenta de seleção.

10 Fenótipos desejáveis incluem, mas não estão limitados a, entrada de alimentação, taxa de crescimento, peso corpóreo, mérito de carcaça e composição, e reprodução e longevidade, e produção de leite. Traços específicos de carcaça com fenótipos desejáveis incluem, mas não estão limitados a, calor individual de carcaça (valor de carc adicional, \$), ganho diário
 15 médio (ADG, lb/d), espessura de gordura posterior (BFAT, polegada), peso vivo calculado (Ps Viv Calc, lb), grau de produção calculado (cYG), dias na alimentação (DOF, d), percentagem de adubo (DP, %), entrada de matéria seca (DMI, lb), entrada de matéria seca por dia na alimentação (DMI por DOF, lb/d), peso de carcaça quente (HCW, lb), valor do peso de carcaça
 20 quente (valor de HCW, \$), teor de gordura intramuscular (IMF%, %), classificação de aparência marmórea (MBS, 10 a 99), classificação de aparência marmórea dividida pelos dias na alimentação (MBS/DOF), grau de qualidade, menor do que ou igual para selecionar versus maior do que ou igual a escolha (QG, < Se vs, > Ch), área de "ribeye" (REA, pol²), área de
 25 "ribeye" por cem peso de HCW (REA/cwt HCW, pol²/100 lb de peso de carcaça quente (HCW) e profundidade de gorfura sbcutânea (SFD).

Um aspecto da presente invenção proporciona agrupamento de animais e métodos para gerenciamento de produção de criação compreendendo agrupar animais da criação tal como gado de acordo com o genótipo
 30 conforme definido por painéis de SNPs/STRs, cada painel compreendendo pelo menos um SNP/STR, um ou mais dos quais estão no gene de *CAST* da presente invenção. Outros SNPs que podem ser incluídos nos painéis de

SNPs incluem, mas não estão limitados a, SNPs encontrados no gene de *GHR*, gene de *FABP4*, gene de ghrelin, gene de leptin, gene de *NPY* gene, gene de *TFAM*, gene de *CRH*, gene de *UASMS1*, gene de *UASMS2*, gene de *UASMS3* e/ou o gene de *UCP3*. Os métodos de seleção genética e agrupamento da presente invenção podem ser usados em conjunto com outros métodos de agrupamento fenotípico convencional, tal como agrupamento de animais por características visíveis tais como peso, tamanho da estrutura, traços de geração, e similares. Os métodos da presente invenção proporcionam produção de gado tendo traços hereditários aperfeiçoados, e podem ser usados para otimizar o desempenho de rebanhos de criação em áreas tais como fertilidade, longevidade, procriação, entrada de alimentação, qualidade de carcaça/carne e produção de leite. A presente invenção proporciona métodos de classificação de criação para determinar aqueles mais prováveis de desenvolver uma condição corpórea desejada pela identificação da presença ou ausência de um ou mais polimorfismos de gene correlacionados com fertilidade e longevidade.

Conforme descrito acima, e nos Exemplos, existem vários traços fenotípicos com os quais os SNPs/STRs da presente invenção podem estar associados. Cada um dos traços fenotípicos e genéticos podem ser testados usando-se os métodos descritos nos Exemplos, ou usando-se quaisquer métodos conhecidos adequados na técnica. Usando-se os métodos da invenção, um fazendeiro, ou operador de "feedlot", ou similares, podem agrupar gado de acordo com cada propensão genética do animal para um traço desejado tais como taxa de crescimento, entrada de alimentação ou comportamento de alimentação, conforme determinado pelo genótipo de SNP/STR. O gado é testado para determinar homozigossidade ou heterozigossidade com relação aos alelos de SNP/STR de um ou mais genes de modo que eles podem ser agrupados tal que cada pequeno cercado para animal contém gado com genótipos similares. Cada cercado de animais é então alimentado e, de outro modo, mantido em uma maneira e por um tempo determinado pelo operador de "feedlot", e, em seguida, submetido à matança.

Os dados genotípicos individuais derivados de um painel ou pai-

néis de SNPs/STRs para cada animal ou um rebanho de animais podem ser registrados e associados com vários outros dados do animal, por exemplo, informação de saúde, ascendência, condições de criação de gado, história de vacinação, registros de rebanho, dados subseqüentes de segurança de alimentação, e similares. Tal informação pode ser enviada para uma agência governamental para proporcionar traçabilidade de um animal ou produto de carne, ou pode servir como a base para informação de procriação, alimentação e comercialização. Uma vez que o dado tenha sido ou não associado com outro dado, o dado é armazenado em uma base de dados acessível, tal como, mas não limitada a, uma base de dados de computador ou um microchip implantado no animal. Os métodos da invenção podem proporcionar uma análise dos dados de entrada que podem ser comparados com parâmetros desejados pelo operador. Estes parâmetros incluem, mas não estão limitados a, tais metas de procriação, gerenciamento de produção de leiteira, níveis de vacinação de um rebanho. Se o desempenho ou propriedades dos animais se desviam das metas desejadas, os métodos baseados em computador podem disparar um alerta para permitir que o operador ajuste as doses de vacinação, medicações, alimentação, etc., conseqüentemente.

Os resultados da análise proporcionam dados que são associados com o animal individual ou ao rebanho, no todo ou em parte, dos quais a amostra foi tomada. Os dados são então mantidos em uma base de dados acessível, e podem ou não podem estar associados com outros dados daquele indivíduo particular individual, ou de outros animais.

Os dados obtidos de animais individuais podem ser armazenados em uma base de dados que pode estar integrada ou associada com, e/ou equiparada a outras bases de dados. A base de dados, junto com os dados associados, permite informação sobre o animal individual a ser conhecido através de todo estágio da vida do animal, isto é, da concepção ao consumo do produto animal.

Os dados acumulados e a combinação dos dados genéticos com outros tipos de dados do animal proporcionam acesso à informação sobre a ascendência, identificação de rebanho, informação de saúde, incluindo vaci-

nações, exposição a doenças, localização de "feedlot", dieta, e mudanças de proprietário. Informação, tais como datas e resultados de diagnóstico, ou testes de rotina, são facilmente armazenados e alcançáveis. Tal informação seria especialmente valiosa para companhias, particularmente aquelas que

5 procuram linhas de procriação superiores.

Cada animal pode ser provido com um identificador único. O animal pode ser etiquetado, como programas de rastreio tradicionais, ou ter implante de chips de computador proporcionando dados armazenados ou legíveis, ou provido com qualquer outro método de identificação que associa

10 o animal com seu identificador único.

A base de dados contendo os resultados de genótipo baseados em SNP/STR para cada animal, ou os dados para cada animal, podem estar associados ou ligados a outras bases de dados contendo dados, por exemplo, que podem ser proveitosos na seleção de traços para agrupamento ou

15 subagrupamento de um animal. Por exemplo, e não para limitação, os dados pertencentes a animais tendo protocolos de vacinação ou medicação particulares, podem opcionalmente serem adicionalmente ligados com dados pertencentes a animais tendo alimento de certas fontes de alimento. A capacidade de refinar um grupo de animais é limitada somente pelos traços pro-

20 curados, e pelas bases de dados contendo informação relacionada àqueles traços.

As bases de dados que podem ser utilmente associadas com os métodos da invenção incluem, mas não estão limitadas a, dados científicos específicos ou gerais. Os dados específicos incluem, mas não estão limita-

25 dos a, linhas de procriação, progenitores, e similares, outros genótipos de animais, incluindo se ou não outros animais específicos possuem genes específicos, incluindo elementos genéticos transgênicos, locação de animais que compartilham características genéticas similares ou idênticas, e similares. Dados gerais incluem, mas não estão limitados a, dados científicos, tais

30 como aqueles genes que codificam para características de qualidade específicas, dados de associação de procriação, dados de alimentação, tendências à procriação, e similares.

Um método da presente invenção inclui a provisão do proprietário do animal ou cliente com equipamento de coleta de amostra, tais como hastes de coleta e etiquetas úteis para coleta de amostras das quais dados genéticos podem ser obtidos. Vantajosamente, o acondicionamento é codificado com uma etiqueta de código de barras. As etiquetas são codificadas com o mesmo índice de identificação, vantajosamente com uma equiparação da etiqueta de código de barras. Opcionalmente, o acondicionamento contém meios para enviar as etiquetas para um laboratório de análise. O acondicionamento opcional é também codificado com índice de identificação, vantajosamente com uma etiqueta de código de barras.

O método opcionalmente inclui um sistema no qual uma contagem de base de dados é estabelecida sob ordenação do equipamento de amostragem. Um identificador de contagem de base de dados corresponde ao índice de identificação das etiquetas e do acondicionamento. Após carregamento do equipamento de amostragem na execução da ordem, os índices de identificação são registrados em uma base de dados. Vantajosamente, o identificador é uma etiqueta de código de barras que é escaneada quando as etiquetas são enviadas. Quando as etiquetas são retornadas para a facilidade de teste, o identificador é novamente registrado e equiparado à informação anteriormente registrada na base de dados sob carregamento do frasco para o cliente. Uma vez que a genotipagem é completada, a informação é registrada na base de dados, e codificada com o identificador único. Os resultados dos testes são providos ao cliente ou proprietário do animal.

Os dados armazenados na base de dados do genótipo podem ser integrados com, ou comparados a outros dados ou base de dados para a proposta de identificação de animais baseada nas propensões genéticas. Outros dados ou base de dados incluem, mas não estão limitados a, aqueles contendo informação relacionada a teste de DNA baseado em SNP, vacinação, programa de précondicionamento Sure Health, estro, e resultados de gravidez, níveis de hormônio, segurança de alimentação/contaminação, contagens de célula somática, ocorrência de mastite, resultados de teste de diagnóstico, níveis de proteína de leite, gordura de leite, estado de vacina, re-

gistros de saúde, níveis de mineral, níveis de traço de mineral, desempenho do rebanho, e similares.

A presente invenção, portanto, envolve métodos auxiliados por computador para rastreo das histórias de procriação ou veterinária de animais da criação envolvendo o uso de um sistema baseado em computador compreendendo um computador programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada e um dispositivo de saída, e compreendendo as etapas de gerar um perfil de um animal da criação por admissão no computador programado através do dispositivo de entrada de dados de genótipo do animal, no qual o genótipo pode ser definido por um painel de pelo menos dois polimorfismos de nucleotídeo simples que prognostica pelo menos um traço físico do animal, admissão no computador programado através do dispositivo de entrada de dados de bem-estar do animal, correlação dos dados de bem-estar admitidos com o perfil fenotípico do animal usando-se o processador e o sistema de armazenagem de dados, e fornecimento de um perfil do animal ou grupo de animais ao dispositivo de saída.

As bases de dados e a análise destes estarão acessíveis a quem acessar. O acesso pode ser provido através de direitos de acesso, ou por subscrição a partes específicas dos dados. Por exemplo, a base de dados pode ser acessada pelos proprietários do animal, o local de teste, a entidade que proporciona a amostra ao local de teste, pessoal de "feedlot", e veterinários. Os dados podem ser providos em qualquer forma, tais como por acesso a uma website, fax, e-mail, correspondência de correio, telefone automático, ou outros métodos para comunicação. Estes dados podem também ser codificados em um dispositivo de armazenagem portátil, tal como um microchip, que pode ser implantado no animal. Vantajosamente, a informação pode ser lida e nova informação adicionada sem remoção do microchip do animal.

A presente invenção compreende sistemas para realização dos métodos aqui descritos. Tais sistemas compreendem dispositivos, tais como computadores, conexões de internet, servidores, e dispositivos de armaze-

nagem de dados. A presente invenção também proporciona um método de transmissão de dados compreendendo transmissão de informação de tais métodos aqui descritos ou etapas destes, por exemplo, via telecomunicação, telefone, vídeo-conferência, comunicação de massa, por exemplo, apresentação, tal como uma apresentação de computador (por exemplo, POWER-
5 POINT), internet, e-mail, comunicação com documento, tais como programas de computador (por exemplo, WORD), e similares.

Os sistemas da presente invenção podem compreender um módulo de coleta de dados, que inclui um coletor de dados para coletar dados
10 de um animal ou embrião, e transmitir os dados para um módulo de análise de dados, uma interface de rede para recebimento de dados a partir do módulo de análise de dados, e, opcionalmente, adicionalmente adaptados para combinar dados múltiplos de um ou mais animais individuais, e para transmitir os dados via uma rede para outros locais, ou para um dispositivo de ar-
15 mazenagem.

Mais particularmente, os sistemas da presente invenção compreendem um módulo de coleta de dados, um módulo de análise de dados, uma interface de rede para recebimento de dados a partir do módulo de análise de dados, e, opcionalmente, adicionalmente adaptados para combinar
20 dados múltiplos de um ou mais animais individuais, e para transmitir os dados via uma rede para outros locais, e/ou um dispositivo de armazenamento. Por exemplo, os dados coletados pelo módulo de coleta de dados conduzem a uma determinação da ausência ou presença de um SNP de um gene no animal ou embrião, e, por exemplo, tal dado é transmitido quando o regime
25 de alimentação do animal é planejado.

Em uma concretização onde o dado é implantado em um microchip em um animal particular, o fazendeiro pode otimizar a eficiência de gerenciamento do rebanho porque o fazendeiro é capaz de identificar as predisposições genéticas de um animal individual, bem como tratamentos pas-
30 sados, presentes e futuros (por exemplo, vacinações e visitas veterinárias). A invenção, portanto, também proporciona acesso a outras bases de dados, por exemplo, dados de rebanho relacionados a testes genéticos e dados

realizados por outros, por ligações de dados a outros locais. Portanto, os dados de outras bases de dados podem ser transmitidos para a base de dados central da presente invenção, via uma interface de rede para recebimento de dados a partir do módulo de análise de dados das outras bases de dados.

A invenção se refere a um sistema de computador e a um meio legível de computador para compilação de dados em um animal, o sistema contendo dados admitidos neste animal, tais como, mas não limitados a, histórias de vacinação e medicação, teste de DNA, teste de tiroglobulina, leptin, MMI (Meta Morphix Inc.), diagnose de encefalopatia espongiforme (BSE), vacinação de brucelose, vacinação de FMD (doença de alimentação e boca), vacinação de BVD (diarréia viral de bovino), programa de précondicionamento de Sure Health, estro e resultados de gravidez, tuberculose, níveis de hormônio, segurança de alimentação/contaminação, contagens de célula somática, ocorrência de mastite, resultados de teste de diagnóstico, níveis de proteína de leite, gordura de leite, estado de vacina, registros de saúde, níveis de mineral, níveis de traço de mineral, desempenho do rebanho, e similares. Os dados do animal podem também incluir tratamentos anteriores, bem como tratamento provido sugerido dependendo do tratamento provido que depende da predisposição genética deste animal para uma doença particular.

A invenção também proporciona um método auxiliado por computador para aperfeiçoar a produção de animal compreendendo o uso de um sistema de computador, por exemplo, um computador programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada e um dispositivo de saída, as etapas de admissão no computador programado através do dispositivo de entrada de dados compreendendo dados diagnósticos de procriação, veterinários, de medicação, e similares de um animal, correlação de uma característica física prognosticada pelo genótipo usando-se o processador e o sistema de armazenagem de dados, fornecimento ao dispositivo de saída da característica física correlacionada ao genótipo, e alimentação do animal com uma dieta baseada na

característica física, aperfeiçoando, desse modo, a produção da criação.

A invenção adicionalmente proporciona um método auxiliado por computador para otimizar a eficiência de "feedlots" para criação compreendendo o uso de um sistema de computador, por exemplo, um computador
5 programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada e um dispositivo de saída, e as etapas de admissão no computador programado através do dispositivo de entrada de dados compreendendo uma história de procriação e veterinária de um animal, correlação das histórias de procriação e veterinária usando-se o pro-
10 cessador e o sistema de armazenagem de dados, fornecimento ao dispositivo de saída da característica física correlacionada ao genótipo, e alimentação do animal com uma dieta baseada na característica física, otimizando, desse modo, a eficiência de "feedlots" na criação.

A invenção adicionalmente compreende métodos de se fazer
15 negócio pela provisão de acesso a tal meio legível de computador, e/ou sistema de computadores, e/ou dados coletados de animais aos usuários; por exemplo, o meio e/ou dados de seqüência podem ser acessíveis a um usuário, por exemplo, em uma base de subscrição, via a Internet, ou uma comunicação global/rede de computador; ou, o sistema de computador pode ser
20 disponível a um usuário, em uma base de subscrição.

Em uma concretização, a invenção proporciona um sistema de computador para gerenciamento da criação compreendendo características físicas e bases de dados correspondentes a um ou mais animais. Em outra concretização, a invenção proporciona um meio legível de computador para
25 gerenciamento da criação compreendendo características físicas e histórias veterinárias correspondentes a um ou mais animais. A invenção adicionalmente proporciona métodos de se fazer negócio para gerenciamento da criação compreendendo provisão a um usuário de um sistema de computador e meios descritos acima, ou características físicas e histórias veterinárias a
30 um ou mais animais. A invenção adicionalmente envolve métodos de transmitir informação obtida em qualquer método ou etapa deste aqui descrita, ou qualquer informação aqui descrita, por exemplo, via telecomunicações, tele-

fone, comunicação de massas, meio de massa, apresentações, internet, e-mail, etc.

A invenção adicionalmente envolve kits úteis para classificação de ácido nucléico isolado de um ou mais indivíduos bovinos para variação alélica de qualquer um dos genes de fator de transcrição mitocondriais, e, em particular, para qualquer dos SNPs aqui descritos, no qual os kits podem compreender pelo menos um oligonucleotídeo que se hibridiza seletivamente a um ácido nucléico compreendendo qualquer um dos um ou mais dos quais são seqüências de *CAST* aqui descritas, e instruções para uso do oligonucleotídeo para detectar variação no nucleotídeo correspondente ao SNP do ácido nucléico isolado.

Uma concretização deste aspecto da invenção proporciona um oligonucleotídeo que se hibridiza especificamente à molécula de ácido nucléico isolada deste aspecto da invenção, e no qual o oligonucleotídeo se hibridiza a uma porção da molécula de ácido nucléico isolada compreendendo qualquer um dos locais polimórficos nas seqüências de *CAST* aqui descritas.

Outra concretização da invenção é um oligonucleotídeo que se hibridiza especificamente sob condições de estrigência altas para qualquer um dos locais polimórficos do gene de *CAST*, no qual o oligonucleotídeo está entre cerca de 18 nucleotídeos e cerca de 50 nucleotídeos.

Em outra concretização da invenção, o oligonucleotídeo compreende um nucleotídeo central que se hibridiza especificamente com um local de polimorfismo de gene de *CAST* da porção da molécula de ácido nucléico.

Outro aspecto da invenção é um método de identificação de um polimorfismo de *CAST* em uma amostra de ácido nucléico compreendendo isolar uma molécula de ácido nucléico que codifica *CAST*, ou um fragmento desta, e determinação do nucleotídeo no local polimórfico.

Outro aspecto da invenção é um método de classificação de gado para determinar aqueles bovinos mais provavelmente que exibem uma diferença biológica na fertilidade e longevidade compreendendo as etapas

de obter uma amostra de material genético de um bovino; e ensaio da presença de um genótipo no bovino que é associado com fertilidade e longevidade, o genótipo caracterizado por um polimorfismo no gene bovino de *CAST*.

- 5 Em outras concretizações deste aspecto da invenção, a etapa de ensaio é selecionada a partir do grupo consistindo em: análise de polimorfismo de comprimento de fragmento de restrição (RFLP), mini-seqüenciamento, análise de heterodúplex MALD-TOF, polimorfismo conformacional de trançado simples (SSCP), eletroforese de gel de gradiente de
10 desnaturação (DGGE) e eletroforese de gel de gradiente de temperatura (TGGE).

- Em várias concretizações da invenção, o método pode adicionalmente compreender a etapa de amplificar a região do gene de *CAST* ou uma porção deste que contém o polimorfismo. Em outras concretizações da
15 invenção, a amplificação pode incluir a etapa de selecionar o primer de seqüência avançada e reversa capaz de amplificar a região do gene de *CAST*.

- Outro aspecto da invenção é um método auxiliado por computador para prognosticar quais animais da criação possuem uma diferença biológica na fertilidade e longevidade compreendendo: o uso de um sistema de
20 computador, por exemplo, um computador programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada e um dispositivo de saída, as etapas de: (a) admissão no computador programado através do dispositivo de entrada de dados data compreendendo um genótipo de *CAST* de um animal, (b) correlação da fertilidade e longevidade prognosticada pelo genótipo de *CAST* usando-se o processador e
25 o sistema de armazenagem de dados e (c) fornecimento ao dispositivo de saída da fertilidade e longevidade correlacionada ao genótipo de *CAST*, prognosticando, desse modo, quais animais da criação possuem uma fertilidade e longevidade particular.

- 30 Ainda outro aspecto da invenção é um método de se fazer negócio para controle da criação compreendendo provisão a um usuário de um sistema de computador para controle da criação compreendendo caracterís-

ticas físicas e genótipos correspondentes a um ou mais animais ou um meio legível de computador para controle da criação compreendendo características físicas e genótipos correspondentes a um ou mais animais ou características físicas e genótipos correspondentes a um ou mais animais.

- 5 A invenção será agora adicionalmente descrita por meio dos seguintes Exemplos não-limitativos.

EXEMPLOS

Exemplo 1

Materiais e Métodos

10 Animais e preparação de DNA

- Amostras de sêmen de 652 progenitores que tinham procriado vários números de raças, mais sete antepassados, foram doadas pelo repositório de DNA de touro de Leiteria, Maryland, para análise. Os progenitores PTAs para seis traços foram obtidos a partir das avaliações de lista de progenitor do Laboratório de Programas de Aperfeiçoamento do Departamento de Agricultura Animal dos Estados Unidos para todos os touros Holstein (<http://aipl.arsusda.gov/>). Os traços analisados incluíam taxa de gravidez de raça (DPR), vida produtiva (PL), produção de proteína (PY), produção de leite (MY), produção de gordura (FY), classificação de célula somática (SCS) e mérito de rede em dólares (NM). O PTA médio, desvio padrão e faixa de variação para cada um destes traços em nossa população amostrada mais hereditariedade total para cada traço estão listados na Tabela 1. O DNA a partir do esperma foi extraído usando-se uma extração de clorofórmio fenol e protocolo de purificação anteriormente descrito [Ashwell MS, Heyen DW, Sonstegard TS, Van Tassell CP, Da Y, VanRaden PM, Ron M, Weller JI, Lewin HA. A detecção de local de traço quantitativo que afeta a produção de leite, saúde, e traços reprodutivos no gado Holstein. J. Dairy Sci 2004; 87:468-475].
- 15
- 20
- 25

Tabela 1

Parâmetros fenotípicos e genéticos de traços quantitativos em sete famílias de progenitor

Traço	N	Média	Desvio Padrão	Mínimo	Máximo	h ²
DPR	652	-0,53	1,06	-3,6	2,5	0,04
PL	652	-,078	1,37	-4,5	2,7	0,085
PY	652	6,29	16,28	-43	64	0,25-0,35
MY	652	28,18	618,31	-1915	2067	0,25-0,35
FY	652	3,81	22,05	-57	82	0,25-0,35
SCS	652	3,03	0,18	2,56	3,56	0,12
NM	652	-12,34	144,51	-380	538	N/A

Detecção de Mutação in Silico e Desenho de Prímer

- 5 Uma sequência de cDNA do gene bovino de calpastatin (NM_174003) consistindo no novo domínio XL foi usada como uma referência para pesquisar as sequências de DNA genômico do mesmo gene contra a base de dados de sequenciamento de genoma bovino (<http://www.hgsc.bcm.tmc.edu/projects/bovino/>). Um total de sete contigs
- 10 genômicos (figuras 1-8) foram recuperados e o alinhamento de ambos cDNA e sequências de DNA genômico descrever 11 polimorfismos potenciais de nucleotídeo simples (SNPs) na região de codificação (figura 9). Entre eles, três SNPs provisórios resultariam em mudanças de aminoácido na sequência de proteína, incluindo dois localizados em exon 3 putativo e um em exon
- 15 8 putativo do gene bovino gene. A nova sequência de DNA do domínio XL se estende nos exons 1 - 3 e exon parcial 4. Dois pares de prímer foram designados para cobrir ambos exons polimórficos parciais baseados nas sequências de DNA genômico. Para assegurar que cada região de exon foi completamente amplificada e sequenciada, pelo menos 100 bp de sequências de flaqueamento foram incluídas nos produtos. As sequências de iniciador para amplificação do produto de exon 3 foram: avançada, 5'-AAA TTT
- 20

GCG GTT GAC CAC ACT GTT A-3 (SEQ ID NO: 18) e reversa, 5'-TGT TAT GCC TGT TGC TTT GTA CCT C-3' (SEQ ID NO: 19) (número de acesso no GenBank: AAFC02179490). As seqüências de iniciador para amplificação do exon 8 tentativo total foram: avançado, 5'-GAT TCT TGC TGA ATT TGG
 5 AGG GAA G-3' (SEQ ID NO: 20) e reversa, 5'-GGG GTC TCA AAG AGT TGG ATA CGA T-3' (SEQ ID NO: 21) (número de acesso no GenBank: A-AFC02060381).

Validação de Mutação e Detecção por Seqüenciamento de DNA Reunido

O DNA de animais exibindo fenótipos extremos foi reunido para
 10 validar as mutações putativas descritas acima detectar novos polomorfismos em ambos os produtos que incluem exon 3 e exon 8 do gene bovino de CAST mais suas seqüências de intron de flaqueamento. Para alcançar os fenótipos relacionados de fertilidade, os animais foram classificados por PTA para DPR. Duas amostras de DNA reunidas foram formadas, utilizando os
 15 60 animais de topo com as PTA's mais altas para DPR e os 60 animais com as PTA's mais baixas para DPR. AS reações de PCR foram realizadas usando-se 25 ng de DNA genômico bovino como gabarito em um volume final de 10 µL contendo 12,5 ng de cada primer, 200 µM de dNTPs, 1,5 - 3 mM de MgCl₂, 50 mM de KCl, 20 de mM Tris-HCl e 0,2U de polimerase Platinum
 20 Taq (Invitrogen, Carlsbad, CA). Condições de PCR de "touchdown" foram efetuadas conforme segue: 95°C por 10 minutos; 8 ciclos de 94°C por 30 segundos, 71°C por 30 segundos e 72°C 30 segundos, seguido por 37 ciclos de 94°C por 30 segundos, 63°C por 30 segundos e 72°C por 30 segundos mais 1 ciclo a 72°C por 5 minutos, e, em seguida uma retenção estendida a
 25 4°C até uso. Os produtos de PCR foram examinados por eletroforese através de um gel agrose 1,5% com tampão 1X TBE de modo a determinar a qualidade e quantidade de seqüenciamento de DNA. O seqüenciamento foi realizado em um seqüenciador ABI 3730 no Laboratório para Biotecnologia e Bioanálise (Washington State University) usando-se um protocolo padrão.
 30 Os polimorfismos de nucleotídeo foram identificados por comparação de padrões de seqüência entre estas reuniões de DNA.

Detecção de Halotipo por Seqüenciamento Individual de Antepassado e Ge-

notipagem de Marcador

As seqüências de DNA reunidas de ambos os produtos revelaram polimorfismos genéticos na população de progenitores de leiteria derivados de sete antepassados. Para determinar halotipos potenciais entre estas mutações no gene bovino de CAST, os produtos de PCR a partir dos
5 sete antepassados foram amplificados e submetidos à seqüenciamento individualmente. Um total de quatro polimorfismos incluindo três SNPs e uma repetição em tandem simples foram observados, mas eles formaram somente dois haplotipos entre estes sete antepassados. Uma substituição C/T foi,
10 portanto, escolhida para genotipagem conforme pode ser descrito usando-se uma aproximação de PCR-RFLP (polimorfismo de comprimento de fragmento de restrição). Os amplicons de PCR que continham a região de exon 3 e seqüência de intron de flanqueamento foram digeridos a 37°C por três horas com 2U de MspI (New England Biolabs, Beverly, MA, USA). A visualização
15 de digestão de enzima e classificação de genótipo dos indivíduos foram conduzidas por eletroforese em um gel agrose 2% contendo brometo de etidium.

Análise de Associação de Gene de CAST com Fertilidade e longevidade em gado leiteiro

Três etapas foram empregadas para investigar a associação do
20 gene bovino de calpastatin com DPR e PL em nossa população amostrada. A primeira etapa foi para genotipar o marcador em 60 amostras altas e 60 amostras baixas de DPR. Teste exato de Fisher foi usado para examinar as diferenças em freqüências de alelo por classificação de associação inicial
25 entre estes dois grupos de animais. Na segunda etapa, o mesmo marcador foi genotipado na progenia a partir da família de progenitor com PTA mais alta para DPR e a família de progenitor com a PTA mais baixa para DPR. Teste exato de Fisher foi usado para examinar diferenças em freqüências de alelo para validar associações iniciais identificadas na primeira etapa. A última
30 etapa foi analisar compreensivamente os dados de todos os indivíduos para os traços de fertilidade de of DPR e PL, os traços de produção de produção de leite, proteína, gordura de leite, e mérito total de valor de rede para

todos os traços combinados. A associação entre as PTA's dos traços anteriormente descritas e o genótipo do filhote, CC, CT e TT foi examinado, usando-se a função de modelo misturada de SAS (Versão 9.1 Carey NC). O efeito da família de progenitor foi também incluído no modelo como um efeito fixo e genótipo de filhote como um efeito aleatório:

$$y_{ijk} = \mu + s_i + g_j + e_{ijk}$$

onde y_{ijk} é a PTA de traços examinados; μ é o valor médio total do traço; s_i é o efeito da família de progenitor ($i = 1, 2, \dots, 7$) e g_j é o efeito genotípico do filhote ($j = CC, CT, TT$). e_{ijk} é o efeito residual correspondente a y_{ijk} e foi

assumido para ser normalmente distribuído. O efeito de interação entre família de progenitor e genótipo do filhote foi também incluído inicialmente no modelo, mas foi excluído no modelo final porque ele não foi significativo. O modelo residual foi pesado por $1/r$ na análise, onde r é confiança para PTA. O nível de significância do modelo foi ajustado em $P < 0,05$. Seguindo um teste F significativo, meios individuais para os genótipos diferentes foram comparados usando-se comparações de par pré-planejadas usando-se a função pdiff de SAS (Versão 9.1, Carey NC).

Re-anotação Comparativa do Gene Humano de CAST

A base de dados do GenBank atual revelou que a proteína de CAST humana, mesmo em sua isoforma mais longa somente consistia em um domínio L N-terminal e quatro domínios de inibição de calpain repetitivos (domínios 1-4), necessitando, desse modo, do domínio XL que foi detectado na proteína de CAST bovino. Para validar adicionalmente esta observação, o gene de CAST humano foi re-anotado usando-se um processo de três etapas conforme segue: 1) pesquisar BLAST contra as bases de dados "est_humano" no GenBank usando-se uma sequência de cDNA de comprimento total do gene de CAST bovino como uma referência para recuperar todos os ESTs humanos (etiquetas de sequência expressas) que são ortólogos à sequência específica de domínio XL do gene bovino; 2) montar os ESTs humanos recentemente pesquisados com a sequência de cDNA isoforma mais longa atual para formar uma sequência de cDNA de comprimento total do gene humano; e 3) alinhar a sequência de cDNA recentemente

anotada e seqüências de DNA genômico para determinar a organização genômica completa do gene de CAST humano.

Resultados

Organização Genômica e Polimorfismos Funcionais no Gene de CAST Bovino

Uma pesquisa BLAST usando-se a seqüência de cDNA do gene bovino de CAST(NM_174003) recuperou sete contigs genômicos com uma seqüência total de 116,129 bp de uma montagem de seqüência de genoma bovino 6X (<http://www.hgsc.bcm.tmc.edu/projects/bovino/>) (figuras 2 - 9).

O alinhamento de ambos cDNA e seqüência de DNA genômico indicou que o gene bovino de CAST continha pelo menos 32 exons e 31 introns (figura 1). Contudo, exons 1 - 31 estão codificando exons com seqüência de codificação variando de 30 bp (exon 1) a 114 bp (exon 8) de comprimento. Exon 31 também continha 12 bp da seqüência não-traduzida 3' e exon 32 não possuía seqüência de codificação no todo. Entre os 31 introns, vinte e cinco tinham seqüências intrônicas completas. Intron 23 era o menor intron com uma seqüência de 173 bp, enquanto o intron 4 podia ser o intron maior no gene com mais do que 22 kb de seqüência (figura 1). O domínio XL da proteína se estendia dos exons 1 - 3 e parte do exon 4; o domínio L correspondia de parte de exon 4, exons 5 -9 e parte de exon 10; domínio 1 parte flanqueada de exon 10 e exons 11 - 15; domínio 2 estendido de 16 - 20 e parte de exon 21; domínio 3 foi codificado por parte de exon 21, exons 22 - 25 e parte de exon 26; e domínio 4 incluiu parte de exon 26, exons 27 - 30 e parte de exon 31, respectivamente (figura 1). Uma pesquisa de cDNA contra a base de dados de DNA genômico bovino indicou que somente uma cópia do gene de calpastatin existia no genoma bovino.

O alinhamento da seqüência de cDNA com a seqüência de DNA genômico do gene bovino de CAST também revelou 11 SNPs putativos na região de codificação; incluindo quatro transições G/A, três transições C/T, duas transversões T/A, uma substituição G/T e uma substituição CA (figura 9). Contudo, somente três cSNPs foram verificados alterar a seqüência de aminoácido da proteína de CAST bovino: G48D, P52L e I128K. Os SNP's

remanescentes eram mutações inativas. As substituições G48D e P52L foram localizadas no exon 3, correspondendo ao domínio XL. A I128L foi localizada no exon 8 que codifica para o domínio L. Todos os três SNPs de codificação de sem sentido alteraram a segunda base no códon, conduzindo a mudanças no aminoácido. As seqüências do DNA alto/baixo reunido e sete antepassados individuais confirmaram os SNPs de G48D e P52L SNP's no domínio XL (figura 10A), mas não a substituição I128K no domínio L. Como ambos os produtos de PCR continham seqüência intrônicas parciais, uma substituição G/T foi detectada em intron 3 (figura 10B) e uma repetição tetra-nucleotídeo GAAA em intron 8, perto da região de junção de exon-intron. A repetição GAAA apareceu em formas bi-alélicas: quatro repetições em um alelo e 5 repetições em outro (figura 10C). Entre estes quatro polimorfismos, somente dois haplotipos existem nos sete antepassados. O haplotipo 1 foi G-C-T-GAAAGAAAGAAAGAAA (SEQ ID NO: 16), e o haplotipo 2 foi A-T-G-GAAAGAAAGAAAGAAAGAAA (SEQ ID NO: 17), respectivamente (figura 10).

Comparações Alta-baixa Individuais e de Família para Estabelecimento de Associações Iniciais

Conforme indicado acima, somente dois haplotipos existem entre os sete antepassados examinados; desse modo, a genotipagem de um dos haplotipos foi suficiente. De três SNPs identificados nos produtos de exon 3, a transição C/T, que alterou o aminoácido na posição 52 de uma pró-linha para uma leucina resultou no ganho/perda de um local de corte de enzima de restrição para MspI. Um fragmento de 308 bp foi amplificado para a região, que continha somente um local de corte para a enzima de restrição. Portanto, digestão com MspI produziu três faixas de 135 bp, 173 bp e 308 bp, respectivamente (figura 11). Animais homozigotos para o alelo C têm um local de MspI, e após digestão completa exibiu duas faixas de 135 bp e 173 bp. Animais homozigotos para o alelo T perderam este local de MspI que resultou em uma faixa de 308 bp após a reação de digestão. Animais heterozigotos mostraram três faixas após digestão de MspI (figura 11).

A transição C/T foi então genotipada no topo e no fundo de 60

- indivíduos com valores de DPR mais altos. Entre os 60 indivíduos de topo, o número de genótipos de TT, CT e CC foi 24 (40%), 22 (36,7%) e 14 (23,3%), respectivamente. Contudo, dos 60 indivíduos de fundo, os genótipos de TT contaram 65% (39/60), enquanto somente três indivíduos (5%) mostraram
- 5 genótipo de CC e 18 animais (30%) eram heterozigotos (Tabela 2). A análise Chi-square revelou uma associação entre DPR e sequência de genótipo entre os grupos de topo e de fundo ($\chi^2 = 11,10$, $P < 0,01$). Esta associação entre polimorfismos de CAST bovino e DPR foi adicionalmente confirmada pela
- 10 genotipagem deste marcador em toda progenia de uma família de progenitor de DPR alta e baixa. Esta análise suportou a análise inicial e revelou uma associação altamente significativa entre DPR e genótipo ($\chi^2 = 92,91$; $P < 0,0001$; Tabela 2). Os testes exatos de Fisher de diferenças em sequências de alelo foram também altamente significantes entre os grupos de topo e de
- 15 fundo (Fisher's $p = 0,000146$), bem como entre a progenia de famílias altas e baixas (Fisher's $p = 0,000000$) (Tabela 2).

Tabela 2

Análise de Associação Inicial do Polimorfismo de CAST Bovino com DPR em Gado Leiteiro

Grupo	N	CC	CT	TT	Significância	C	T	Valor p de Fisher
Nível individual								
Topo	60	14	22	24	$\chi^2 = 11,10$	0,42	0,58	$P=0,000146$
Fundo	60	3	18	39	$P<0,01$	0,20	0,80	
Nível de família								
Topo	69	27	42	0	$\chi^2=92,69$	0,70	0,30	$P=0,000000$
Fundo	59	0	17	42	$P<0,001$	0,13	0,87	

- Associações Significantes de Gene de CAST com Fertilidade e Longevidade
- 20 em Gado Leiteiro

Inicialmente, um total de 659 filhotes de sete famílias de progenitores foi genotipado para esta transição C/T no gene de CAST bovino. Após verificação do genótipo baseada na análise da árvore genealógica, sete a-

animais foram removidos a partir do conjunto de dados devido a genótipos irregulares. Os 652 animais remanescentes incluíam 62 animais homozigotos CC, 378 animais homozigotos TT e 212 animais heterozigotos CT. As frequências de alelo C e alelo T na população foram 0,26 e 0,74, respectivamente. Através da análise de família para três genótipos de CC, CT e TT indicou que o genótipo individual era uma fonte significativa de variação ($P < 0,0001$) quando se examina DPR e PL, mas não era uma fonte significativa de variação quando se examina os traços de produção de leite ($P > 0,05$) (Tabela 3).

Nos valores de DPR, gado com genótipo homozigoto (CAST:c.283CC) tinha um adicional de 0,82 e 0,57 de unidades de PTA (3,28 e 2,28 dias abertos equivalentes) comparado aos animais homozigotos CAST:c.283TT e animais heterozigotos CAST:c.283CT ($P < 0,05$) (Tabela 1). A longevidade PTA foi 1,22 unidades maior entre CAST:c.283CC e animais TT e 0,89 diferente entre CAST:c.283CC e animais CT ($P < 0,05$). O aperfeiçoamento de ambos DPR e longevidade em animais CAST:c.283CC conduziu a um aumento do valor econômico por \$67,86 e \$51,14 por vaca comparado aos homozigotos TT e heterozigotos CT ($P < 0,05$) (Tabela 3). Os efeitos aditivo (a) e de dominância (d) de substituição C para T e suas unidades no desvio padrão (SD) são também listados na Tabela 3.

Tabela 3

Associações do Gene Bovino de CAST(NM_174003.2:c.283C>T) com Traços Reprodutivos e Produtivos em Gado Leiteiro¹

Traço	CC	CT	TT	Valor F	P	a ²	a em SD ³	d ⁴	d em SD ³
DPR	0,13 ^a	-0,44 ^b	-0,69 ^b	17,83	<0,0001	0,41	0,39	-0,16	-0,15
PL	0,22 ^a	-0,67 ^b	-1,00 ^b	23,41	<0,0001	0,61	0,45	-0,28	-0,21
PY	8,08 ^a	6,68 ^a	5,79 ^a	0,62	0,5402	1,15	0,07	-0,25	-0,02
MY	115,18 ^a	68,25 ^a	-8,56 ^a	1,73	0,1781	61,87	0,10	14,94	0,02
FY	5,39 ^a	3,72 ^a	3,52 ^a	0,18	0,8361	0,94	0,04	0,74	0,03
SCS	2,98 ^a	3,03 ^{ab}	3,03 ^b	2,03	0,1316	-0,02	-0,14	0,02	0,14
NM	43,63 ^a	-7,51 ^b	-24,23 ^b	6,14	0,0023	33,93	0,24	-17,21	-0,12

¹Subscritos diferentes indicam uma diferença significativa de $P < ,05$ dentro de cada série.

²Efeito aditivo.

³SD = Desvio Padrão.

5 ⁴Efeito de dominância.

A genotipagem de sete antepassados indicou que eram um antepassado CC, dois antepassados CT e quatro antepassados TT. O genótipo do antepassado foi mostrado ser uma fonte significativa de variação para ambos DPR e PL ($P < 0,0001$). Quando se examina DPR, o valor de PTA médio para a progenia do antepassado CC foi 1,50 (0,61 vs -0,89) e 0,94 unidades (0,60 vs -0,34) mais alta do que para a progenia de antepassados de TT e CT, respectivamente ($P < 0,05$). Os valores de PL PTA foram 0,62, -0,51 e -1,24 na progenia de famílias de antepassado CC, CT e TT, e foram significativamente diferentes entre quaisquer dois grupos de família de antepassado ($P < 0,05$). Considera-se as diferenças entre a progenia de antepassados CC e TT como diferenciais de seleção, e as diferenças entre indivíduos de CC e TT das famílias cruzadas como resposta de seleção, as hereditariedades compreendidas foram estimadas para serem 0,55 (0,82/1,50) para DPR e 0,66 (1,22/1,86) para PL, respectivamente.

20 Existência do Domínio XL em Gene de CAST Humano

A pesquisa BLAST contra a base de dados de "est_humano" em NCBI usando-se a sequência de cDNA do gene bovino de CAST(NM_174003) identificou dois ESTs humanos (BP2044772 e BU5644868) que mostraram similaridade de sequência alta para a sequência de domínio XL do gene bovino de CAST. A montagem destas duas sequências EST com a forma mais longa atual do gene humano de CAST(NM_001750) gerou uma sequência de consenso de 2,876 bp, incluindo sequência de codificação 151 bp 5'UTR, 2,331 bp e sequências 394 bp 3'UTR (figura 1). Além de tudo, a proteína CAST humana recentemente transladada é 10 aminoácidos mais curta do que a sequência de proteína CAST bovina, mas ambas têm o mesmo numero de aminoácidos (68 aminoácidos) no domínio XL. A similaridade do domínio XL entre o gene de CAST

humano e bovino foi 85% na sequência de nucleotídeo e 77% na sequência de aminoácido.

O alinhamento da sequência de cDNA humano recentemente montada com a sequência de DNA genômico revelou que o gene humano de CAST mais longo continha 32 exons e 31 introns (figura 1). Similar ao gene bovino de CAST, exons 1 - 31 estão codificando exons com sequência de codificação variando de 30 bp (exon 1) a 114 bp (exon 8) de comprimento. Exon 31 também continha 23 bp de sequência não-traduzida 3' e exon 32 compreende a sequência 3'UTR remanescente. Comparado ao gene bovino de CAST, o gene humano tinha sequência de codificação mais curtas em exons 9, 15, 26, 28 e 31 por 6, 3, 3, 3 e 18 bp, respectivamente. Contudo, exon 20 de gene humano de CAST tinha uma sequência de codificação mais longa 3 bp do que aquele observado no gado. No gene humano de CAST, intron 23 foi também o menor intron com uma sequência de 89 bp, enquanto intron 3 foi o intron maior no gene com cerca de 27 kb de sequência (figura 1).

Discussão

Foi reportado que o gene de calpastatin é expresso em tecidos reprodutivos múltiplos, incluindo, mas não limitados a, testículo, ovário, útero, pituitária, glândula mamária, células de germe e a glândula da próstata [Kitahara A, Takano E, Ontsuki H, Kiriata Y, Yamagata Y, Knaagi R, Mura-chi T. Reversed distribution of calpains e calpastatin in human pituitary gland e selective localization of calpastatin in adrenocorticotropin-producing células as demonstrated by immunohistochemistry. J Clin Endocrinol Metab 1986; 63:343-348; Thompson VF, Saldana S, Cong J, Luedke DM, Goll DE. The calpain system in human placenta. Life Sci 2002; 70:2493-508; Ben-Aharon I, Ben-Yosef D, Amit A, Shalgi R. Expression e immunolocalization of the calpain-calpastatin system in the human oocyte. Fertil Steril 2005; 83:1807-1813; Orwig KE, Bertrand JE, Ou BR, Forsberg NE, Stormshak F. Involvement of protein kinase-C, calpains, e calpastatin in prostaglandin F2 alpha-induced oxytocin secretion from the bovine corpus luteum. Endocrinology 1994; 134:78-83; Liang ZG, O'Hern PA, Yavetz B, Yavetz H, Goldberg E.

Human testis cDNAs identified by sera from infertile patients: a molecular biological approach to immunocontraceptive development. *Reprod Fertil Dev* 1994; 6:297-305; Li S, Liang ZG, Wang GY, Yavetz B, Kim ED, Goldberg E. Molecular cloning e characterization of functional domains of a human testis-specific isoform of calpastatin. *Biol Reprod* 2000; 63:172-178; Li S, Liang ZG, Wang GY, Yavetz B, Kim ED, Goldberg E. Molecular cloning e characterization of functional domains of a human testis-specific isoform of calpastatin. *Biol Reprod* 2000; 63:172-178; Li S, Goldberg E. A novel N-terminal domain directs membrane localization of mouse testis-specific calpastatin. *Biol Reprod* 2000; 63:1594-1600 e Wang LF, Miao SY, Yan YC, Li YH, Zong C, Koide SS. Expression of a sperm protein gene during spermatogenesis in mammalian testis: an in situ hybridization study. *Mol Reprod Dev* 1990; 26:1-5]. A strong association was found between the calpastatin e fertility in the present study. Foi também observado que animais com o genótipo desejável para fertilidade não exibe um nível diminuído de produção de leite: animais com valor de PTA mais altos para DPR exibem traços de produção de leite similares como animais com os genótipos menos férteis (Tabela 3). Esta é uma situação ótima, visto que foi inicialmente pensado que a seleção da fertilidade resultaria em uma perda de produção de leite, ou vice versa. Portanto, o envolvimento do gene de calpastatin em trajetórias fisiológicas e/ou bioquímicas diferentes que conduziu a várias funções deve ser adicionalmente avaliado.

Efeitos Pleiotrópicos do Gene de CAST Gene

Existe evidência que a atividade de calpastatin posterior à morte é altamente relacionada à maciez da carne em espécies diferentes [Koohmaraie M, Whipple G, Kretchmar DH, Crouse JD, Mersmann HJ. Postmortem proteolysis in longissimus muscle from beef, lamb, e pork carcasses. *J Anim Sci* 1991; 69:617-624], várias estudos foram realizados para pesquisar polimorfismos genéticos em calpastatin como uma fonte de marcadores genéticos que podem influenciar a maciez da carne. Por exemplo, Schenkel e colegas (dados não-publicados) identificaram uma substituição G/C em intron 4 do gene bovino de CAST que foi associada com força de cisalhamen-

- to ($P = 0,024$) em gado bovino. Em porcos, Ciobanu e outros [Ciobanu DC, Bastiaansen JW, Lonergan SM, Thomsen H, Dekkers JC, Plastow GS, Rothschild MF. Novos alelos em gene de calpastatin estão associados com traços de qualidade de carne em porcos. *J Anim Sci* 2004; 82:2829-2839] reportou que um haplotipo de CAST estava significativamente associada com força de cisalhamento inferior de Warner-Bratzler, perda de cozimento e suculência mais alta. No presente estudo, os resultados demonstram claramente que mutações funcionais da região de CAST bovino de domínio XL foram associadas com fertilidade e longevidade em gado leiteiro Holstein. O
- 5 haplotipo G-C-T-GAAAGAAAGAAAGAAA (SEQ ID NO: 16) é mais desejável do que o haplotipo A-T-G-GAAAGAAAGAAAGAAAGAAA (SEQ ID NO: 17) pelo aumento dos valores de PTA de 0,82 unidades em DPR e 1,22 unidades em PL (Tabela 3). Estes dados indicam que o gene de CAST desempenha um papel pleiotrópico em trajetórias fisiológicas diferentes e envolvidas
- 10 em funções diferentes. Em adição, o presente estudo reportou duas mutações de sem sentido no gene bovino de CAST, que estão localizadas na região recentemente identificada de domínio XL. O impacto potencial deste haplotipo de CAST benéfico para DPR e PL na qualidade da carne garante investigação adicionalmente.
- 15
- 20 **Benefícios do Gene de CAST para Seleção Auxiliada por Marcador**
- Declínios reprodutivos em vacas leiteiras foram largamente responsabilizados pela seleção intensiva de traços de leite. Por exemplo, baseado na regressão linear de valores de procriação para dias abertos nos valores de procriação para 3,7% de FCM (leite com gordura corrigida), Abdallah
- 25 e McDaniel [Abdallah JM, McDaniel BT. Parâmetros genéticos e tendências de leite, gordura, dias abertos, e peso corpóreo após parto nos rebanhos experimentais na Carolina do Norte. *J Dairy Sci* 2000; 83:1364-1370] estimou que para cada aumento de 1000 kg nos valores de procriação para 3,7% de FCM, os valores de procriação para dias abertos aumentou por 8
- 30 dias. Na Espanha, a produção de leite por vaca aumentou de 7800 kg em 1991 para 10.200 kg em 2000. Contudo, cada aumento de 1000 kg em uma produção de leite média foi acompanhado por uma diminuição de 3,2% a 6%

na taxa de gravidez, 4,4% a 7,6% na ciclicidade, e um aumento de 4,6% e 8% na incidência de ovários inativos [Lopez-Gatius F. Está a fertilidade declinando no gado leiteiro? Um estudo retrospectivo no noroeste da Espanha. *Theriogenology* 2003; 60:89-99]. Correlações genéticas negativas foram também reportadas entre produção de leite e outros traços reprodutivos, tais como intervalo de parto e primeiras taxas de concepção de serviço [Pryce JE, Veerkamp RF. A incorporação de índices de fertilidade nos programas de aperfeiçoamento genético. Na fertilidade na produção alta de vacas leiteiras, *British Society of Animal Science Occas* 2003; 26:237-249]. Os resultados a partir do presente estudo sugerem que a seleção de alelo/haplotipo benéfico no gene bovino de CAST não resultaria necessariamente em uma diminuição nos traços de produção de leite em vacas leiteiras. Portanto, antecipou-se que uma combinação de seleção genética baseada no marcador de CAST potenciais de PTA altos de traços de produção de leite aperfeiçoará os traços reprodutivos, enquanto permite traços de produção de leite continuados altos. Em adição, se os genes de CAST foram aplicados nos programas de seleção auxiliados por marcador, a hereditariedade de DPR podia ser dramaticamente aumentada. Comparado à hereditariedade estimada de 0,04 para DPR e 0,085 para PL (Tabela 1), o uso do gene de calpastatin na seleção auxiliada por marcador aceleraria, desse modo, o aperfeiçoamento de fertilidade em gado leiteiro, que declinou por várias décadas. A frequência baixa do alelo/haplotipo desejável (0,26) na população de leiteria atual estudada foi consistente com a baixa fertilidade exibida nesta população, indicando que a seleção auxiliada por marcador é urgentemente necessária.

25 Gene de CAST e Infertilidade Humana

A infertilidade é um dos problemas de saúde sociais e econômicos mais importantes em seres humanos. Por exemplo, embora a população mundial tenha aumentado notavelmente e é projetada para alcançar nove bilhões até 2050, é estimado que 50 - 80 milhões de casais no mundo permanecerão sem filhos devido à infertilidade [Montoya JM, Bernal A, Borrero C. Diagnostics in assisted human reproduction *Reprod Biomed Online* 2002; 5:198-210]. No presente estudo, re-anotação de gene humano de CAST re-

velou que o gene humano também tem o domínio XL como o gene bovino gene. Como a proteína de CAST foi identificada como um dos antígenos alvos para anticorpos de antiesperma encontrados em mulheres não-férteis [Koide SS, Wang L, Kamada M. Os anticorpos antiesperma associados com infertilidade: propriedades e codificação de genes de antígenos alvos. Proc Soc Exp Biol Med 2000; 224:123-132], seu envolvimento na reprodução humana necessário a ser adicionalmente explorado.

Exemplo 2

A figura 12 mostra um fluxograma da entrada de dados e da saída de dados de resultados da análise, e correlação dos dados pertencentes à histórias de procriação e veterinária e requerimentos de desempenho de um grupo de animais, tais como de bovinos. O fluxograma ilustrado na figura 7 adicionalmente indica o fluxo interativo de dados a partir do dispositivo auxiliado por computador para um corpo de estudantes aprendendo o uso do método da invenção, e a correlação de tais dados interativos para apresentar uma saída como um gráfico indicando o progresso da classe. O fluxograma adicionalmente indica modificações do método da invenção de acordo com a informação recebida dos estudantes para avançar o processo de ensinamento ou otimizar o método para satisfazer as necessidades dos estudantes.

A figura 13 ilustra relacionamentos potenciais entre os elementos de dados a serem admitidos no sistema. Setas unidirecionais indicam, por exemplo, que um celeiro é tipicamente possuído por somente uma fazenda, pelo qual uma fazenda pode possuir vários celeiros. Similarmente, uma prescrição pode incluir produtos veterinários.

A figura 14A ilustra o fluxo de eventos no uso do sistema portátil baseado em computador para entrada de dados na procriação e criação de um rebanho de vacas. A figura 14B ilustra o fluxo de eventos através de sub-rotinas relacionadas aos dados de entrada concernentes ao gerenciamento da fazenda. A figura 14C ilustra o fluxo de eventos através das sub-rotinas relacionadas à entrada de dados concernentes a dados específicos para uma companhia.

A figura 15 ilustra um fluxograma da entrada de dados e a saída de resultados a partir da análise e da correlação dos dados pertencentes às histórias de procriação e veterinária, e requerimentos de desempenho de um grupo de animais.

5 A invenção é adicionalmente descrita pelos seguintes parágrafos numerados:

1. Um método para subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo tem um polimorfismo similar em um gene de calpastatin("CAST") compreendendo:

10 (a) determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de SNPs/STRs no gene de CAST, e

(b) segregação de animais individuais em subgrupos no qual cada animal em um subgrupo tem um polimorfismo similar no gene de CAST.

15 2. Um método for subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo tem um genótipo similar no gene de CAST compreendendo:

(a) determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de SNPs/STRs de interesse no gene de
20 CAST,

(b) segregação de animais individuais em subgrupos dependendo se os têm ou não têm o(s) polimorfismo(s) de nucleotídeo simples/repetições em tandem curtas de interesse no gene de CAST.

25 3. O método dos parágrafos 1 ou 2, no qual os SNPs/STRs de interesse são selecionados a partir do grupo consistindo em mutações de sem sentido in exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e substituições 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5].

30 4. Um método para subagrupamento de animais de acordo com o genótipo no qual os animais de cada subgrupo tem um genótipo similar no gene de CAST compreendendo:

(a) determinação do genótipo de cada animal a ser subagrupado pela determinação da presença de mutações em sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e
 5 uma repetição GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5], e

(b) segregação de animais individuais em subgrupos dependendo se os animais têm ou não têm mutações sem sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e
 10 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] no gene de CAST.

5. Um método para identificação de um animal tendo um fenótipo desejável conforme comparado à população geral de animais daquela espécie, compreendendo a determinação da presença de um SNP/STR no gene de CAST do animal, no qual o polimorfismo é selecionado a partir do grupo consistindo em mutações sem sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição
 15 GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] no polimorfismo de nucleotídeo simples/repetição em tandem curta de gene de CAST é indicativa de um fenótipo desejável.

6. O método do parágrafo 5, no qual o fenótipo desejável é taxa de gravidez de raça (DPR), vida produtiva (PL), produção de proteína (PY),
 25 produção de leite (MY), produção de gordura (FY), classificação de célula somática (SCS) e mérito de rede em dólares (NM), ou qualquer combinação destes.

7. O método do parágrafo 5 ou 6, no qual o fenótipo desejável é fertilidade, longevidade e mérito de rede econômico adicionais, ou qualquer
 30 combinação destes.

8. O método de qualquer um dos parágrafos 1 a 7 no qual o animal é um bovino.

9. O método de qualquer um dos parágrafos 1 a 8 no qual o gene de CAST é um gene CAST bovino.

10. Um método interativo auxiliado por computador para rastreio da criação de bovinos compreendendo o uso de um sistema de computador compreendendo um computador programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada, um dispositivo de saída, e um dispositivo interativo, as etapas de: (a) admissão no computador programado através do dispositivo de entrada de dados compreendendo uma história de procriação de um bovino ou rebanho de bovinos, (b) admissão no computador programado através do dispositivo de entrada de dados compreendendo uma história veterinária de um bovino ou rebanho de bovinos, (c) correlação dos dados veterinários com a história de procriação do bovino ou rebanho de bovinos usando-se o processador e o sistema de armazenagem de dados, e (d) fornecimento ao dispositivo de saída da história de procriação e a história veterinária do bovino ou rebanho de bovinos.

11. O método de acordo com o parágrafo 10, no qual o sistema de computador é um sistema interativo pelo qual modificações na saída do método auxiliado por computador pode estar correlacionado de acordo com a entrada a partir do dispositivo interativo.

12. O método de acordo com o parágrafo 10 ou 11, compreendendo adicionalmente as etapas de admissão no computador programado de dados diagnósticos relacionados a saúde da vaca ou rebanho de vacas; e correlação dos dados diagnósticos para as histórias de procriação e veterinária da vaca ou rebanho de vacas.

13. O método de acordo com qualquer um dos parágrafos 10 a 12, no qual os dados veterinários compreendem um registro de vacinação para uma vaca ou rebanho de vacas.

14. O método de acordo com qualquer um dos parágrafos 10 a 13 no qual os dados de saúde são selecionados a partir do grupo consistindo em dados de condição de criação, história do rebanho, e dados de segurança de alimentação.

15. O método de acordo com qualquer um dos parágrafos 10 a 14, compreendendo adicionalmente pelo menos uma etapa adicional selecionada a partir do grupo consistindo em admissão no computador programado de dados relacionados ao controle de qualidade do bovino ou rebanho de bovinos, e correlação do controle de qualidade de dados às histórias de procriação e veterinária da vaca ou rebanho de vacas, admissão no computador programado de parâmetros de desempenho da vaca ou rebanho de vacas; e correlação dos parâmetros requeridos de desempenho do bovino ou rebanho de bovinos a um requerimento de desempenho específico de um cliente, correlação dos dados de vacina aos parâmetros de desempenho do bovino ou rebanho de bovinos, correlação do rebanho aos parâmetros de desempenho do bovino ou rebanho de bovinos, correlação dos dados de segurança de alimentação aos parâmetros de desempenho do bovino ou rebanho de bovinos, correlação dos dados de condição de raça aos parâmetros de desempenho do bovino ou rebanho de bovinos, admissão no computador programado de dados relacionados aos dados nutricionais do bovino ou rebanho de bovinos; e correlação dos dados nutricionais aos parâmetros de desempenho do bovino ou rebanho de bovinos, e alerta às mudanças indesejáveis nos parâmetros de desempenho do bovino ou rebanho de bovinos.

16. O método de acordo com qualquer um dos parágrafos 10 a 15, compreendendo adicionalmente as etapas de admissão no computador programado através do dispositivo de entrada de dados compreendendo um genótipo de um bovino; correlação de uma característica física prognosticada pelo genótipo usando-se o processador e o sistema de armazenagem de dados; e fornecimento ao dispositivo de saída da característica física correlacionada ao genótipo para um bovino ou população de bovinos, e alimentação do(s) animal(is) de uma dieta baseada na característica física, aperfeiçoando, desse modo, a produção de bovino.

17. O método auxiliado por computador de acordo com qualquer um dos parágrafos 10 a 16 para otimizar a eficiência de "feedlots" para criação de gado compreendendo fornecimento ao dispositivo de saída da histó-

ria de procriação e veterinária do bovino ou rebanho de bovinos, e alimentação do(s) animal(is) com uma dieta baseada em suas histórias de procriação e veterinária, otimizando, desse modo, a eficiência de "feedlots" para o bovino ou rebanho de bovinos.

5 18. Um método de transmissão de dados compreendendo transmissão de informação de tais métodos de acordo com qualquer um dos parágrafos 10 a 16, selecionados a partir do grupo consistindo em telecomunicação, telefone, vídeo-conferência, comunicação de massa, uma apresentação, uma apresentação por computador, uma apresentação em PowerPoint®, internet, e-mail, e comunicação com documento.

15 19. Um sistema interativo de computador de acordo com qualquer um dos parágrafos 10 a 16 para rastreamento de histórias de procriação e bem-estar of vacas compreendendo dados de procriação e veterinários correspondentes a um bovino ou rebanho de bovinos, e no qual o sistema de computador é configurado para permitir o operador deste mudar os dados com o dispositivo ou uma base de dados remota.

20 20. O sistema interativo de computador de acordo com o parágrafo 19, no qual os dispositivos de entrada e de saída são um assistente digital pessoal ou um computador de bolso.

20 21. Um método de se fazer negócio para rastreamento de histórias de procriação e bem-estar de criação compreendendo dados de procriação e veterinários correspondentes a um ou mais animais da criação compreendendo provisão a um usuário do sistema de computador do parágrafo 19.

25 22. Um método de se fazer negócio para rastreamento de histórias de procriação e bem-estar da criação compreendendo dados de procriação e veterinários correspondentes a um ou mais animais da criação compreendendo provisão a um usuário do sistema de computador do parágrafo 20.

30 23. O método de se fazer negócio de acordo com o parágrafo 21, compreendendo adicionalmente a provisão do proprietário do animal ou cliente com equipamento de coleta de amostra, tais como as hastes de cole-

ta e etiquetas úteis para coleta de amostras das quais os dados genéticos podem ser obtidos, e no qual as etiquetas são opcionalmente acondicionadas em um recipiente que é codificado com índices de identificação.

24. O método de se fazer negócio de acordo com qualquer dos parágrafos 10 a 16, no qual o sistema de computador compreende adicionalmente uma pluralidade de dispositivo interativos, e no qual o método compreende adicionalmente as etapas de recebimento de dados a partir dos dispositivos interativos, compilação dos dados, fornecimento dos dados para indicar a resposta de um estudante ou classe de estudantes a uma questão relacionada à operação do método auxiliado por computador, e opcionalmente modificação da operação do método auxiliado por computador de acordo com a indicação da resposta.

25. O método de qualquer um dos parágrafos 8 a 24 no qual os dados compreendem a presença ou ausência de um ou mais de um polimorfismo(s) de nucleotídeo simples/STR de interesse no gene de CAST.

26. O método do parágrafo 25 no qual o polimorfismo(s) de nucleotídeo simples/repetições em tandem curtas de interesse é selecionado a partir do grupo consistindo em mutações sem sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AAFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (AAFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] do gene de CAST.

27. Um método para o diagnóstico ou monitoramento de fertilidade e/ou longevidade em um indivíduo, compreendendo: obtenção de uma amostra biológica a partir de um indivíduo; e determinação, usando-se um ensaio adequado, a presença ou ausência na amostra de um ou mais marcadores de CAST de domínio XL, conforme descrito aqui.

28. O método de acordo com a reivindicação 27, no qual o indivíduo é bovino.

29. Um método para seleção auxiliado por marcador para aperfeiçoar a fertilidade e/ou longevidade, compreendendo classificação, como parte de um esquema de seleção, baseada em um ou mais marcadores de

CAST de domínio XL, conforme descrito aqui, para aumentar seleção para fertilidade e/ou longevidade.

30. O método de acordo com a reivindicação 29, no qual a seleção é para aumentar a fertilidade e/ou longevidade de bovino.

5 31. O método de acordo com a reivindicação 29, compreendendo adicionalmente seleção genética baseada em alto potencial de PTA de um ou mais traços de produção de leite, para proporcionar fertilidade e/ou longevidade aperfeiçoadas, em associação com alta produção de leite.

10 32. O método de acordo com a reivindicação 31, no qual a seleção é para aumentar a fertilidade e/ou longevidade de bovino, em associação com alta produção de leite.

Tendo-se, desse modo, descrito em detalhes concretizações preferidas da presente invenção, é para ser compreendido que a invenção definida pelos parágrafos acima não é para estar limitada a detalhes particulares colocados
15 na descrição acima, visto que muitas variações aparentes desta são possíveis sem fugir do espírito ou escopo da presente invenção.

REIVINDICAÇÕES

1. Método para identificação de um animal tendo fertilidade, longevidade e produção de leite desejáveis, ou uma combinação dos mesmos, conforme comparado à população geral desta espécie, compreendendo determinar a presença de polimorfismos de nucleotídeo simples/repetições em tandem curtas em um gene calpastatin("CAST") gene na presença de um polimorfismo de nucleotídeo simples/repetição em tandem curta no gene CAST do animal, no qual o polimorfismo de nucleotídeo simples/repetição em tandem curta é indicativo de fertilidade, longevidade e produção de leite desejáveis, ou uma combinação dos mesmos.

2. Método, de acordo com a reivindicação 1, compreendendo adicionalmente subgrupos de animais de acordo com o genótipo, no qual os animais de cada subgrupo têm um polimorfismo/repetição em tandem curta similar in no gene CAST, referido método compreendendo:

(a) determinar o genótipo de cada animal a ser subagrupado pela determinação da presença de um polimorfismo de nucleotídeo simples/repetição em tandem curta no gene CAST gene, e

(b) segregar animais individuais em subgrupos dependendo se os animais têm, ou não têm, os polimorfismos de nucleotídeo simples/repetições em tandem curtas de interesse no gene CAST.

3. Método, de acordo com a reivindicação 1, no qual o(s) polimorfismo(s) de nucleotídeo simples/repetições em tandem curtas de interesse é/são selecionado(s) a partir do grupo consistindo em mutações sem sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (A-AFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (A-AFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] no gene CAST.

4. Método, de acordo com a reivindicação 1, no qual o animal é um bovino.

5. Método, de acordo com a reivindicação 1, no qual o gene de CAST é um gene CAST bovino.

6. Método interativo auxiliado por computador para rastrear a

criação de bovinos de criação compreendendo o uso de um sistema de computador compreendendo um computador programado compreendendo um processador, um sistema de armazenagem de dados, um dispositivo de entrada, um dispositivo de saída, e um dispositivo interativo, as etapas de:

- 5 (a) admissão no computador programado através do dispositivo de entrada de dados compreendendo uma história de procriação de um bovino ou rebanho de bovinos e um genótipo de um bovino; correlação de uma característica física prognosticada pelo genótipo usando o processador e o sistema de armazenagem de dados, (b) admissão no computador programado através
- 10 do dispositivo de entrada de dados compreendendo a história veterinária de bovino ou rebanho de bovinos, (c) correlação dos dados veterinários com a história de procriação do bovino ou rebanho de bovinos usando o processador e o sistema de armazenagem de dados, e (d) fornecimento ao dispositivo de saída da história de procriação, a história veterinária do bovino ou rebanho de bovinos, e as características físicas correlacionadas ao genótipo para um bovino ou população de bovinos,

no qual a característica física é taxa de gravidez da raça (DPR) desejável, vida produtiva (PL), produção de proteína (PY), produção de leite (MY), produção de gordura (FY), classificação de célula somática (SCS) e

20 mérito de rede em dólares (NM), ou uma combinação destes, conforme comparado à população de bovinos, e o genótipo é um polimorfismo de nucleotídeo simples/repetição em tandem curta em um gene CAST.

7. Método, de acordo com a reivindicação 6, no qual o sistema de computador é um sistema interativo pelo qual modificações na admissão

25 do método auxiliado por computador podem estar correlacionadas de acordo com a entrada a partir do dispositivo interativo.

8. Método, de acordo com a reivindicação 6, compreendendo adicionalmente as etapas de admissão no computador programado de dados diagnósticos relacionados à saúde da vaca e rebanho de vacas; e correlação dos dados diagnósticos às histórias de procriação e veterinária da vaca ou rebanho de vacas.

30

9. Método, de acordo com a reivindicação 6, no qual os dados

veterinários compreendem um registro de vacinação para uma vaca ou rebanho de vacas.

10. Método, de acordo com a reivindicação 6, no qual os dados de saúde são selecionados a partir do grupo consistindo em dados de condição de criação de gado, história do rebanho, e dados de segurança de ali-
5 mentação.

11. Método, de acordo com a reivindicação 6, compreendendo adicionalmente pelo menos uma etapa adicional selecionada a partir do grupo consistindo em admissão no computador programado de dados relacio-
10 nados ao controle de qualidade do bovino ou rebanho de bovinos, e correlacionando os dados de controle de qualidade às histórias de procriação e veterinária da vaca ou rebanho de vacas, admissão no computador programa-
do de parâmetros de desempenho da vaca ou rebanho de vacas; e correlacionando-se os parâmetros de desempenho requeridos do bovino ou reba-
15 nho de bovinos a um requerimento de desempenho específico de um cliente, e correlacionando-se os dados de vacina aos parâmetros de desempenho do bovino ou rebanho de bovinos, correlacionando-se o rebanho aos parâ-
metros de desempenho do bovino ou rebanho de bovinos, correlacionando-se os dados de segurança de alimentação aos parâmetros de desempenho
20 do bovino ou rebanho de bovinos, correlacionando-se os dados de condição de criação de gado aos parâmetros de desempenho do bovino ou rebanho de bovinos, admissão no computador programado de dados relacionados aos dados nutricionais do bovino ou rebanho de bovinos; e correlacionando-se os dados nutricionais aos parâmetros de desempenho do bovino ou reba-
25 nho de bovinos, e alertando-se às mudanças indesejáveis nos parâmetros de desempenho do bovino ou rebanho de bovinos.

12. Método, de acordo com a reivindicação 6, no qual o(s) polimorfismo(s) de nucleotídeo simples/repetições em tandem curtas de interesse é(são) selecionado(s) a partir do grupo consistindo em mutações sem
30 sentido em exon 3 que resultam em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (A-AFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (A-

AFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5] no gene CAST.

13. Método de transmissão de dados compreendendo transmissão de informação de tais métodos de acordo com a reivindicação 6, selecionada a partir do grupo consistindo em telecomunicação, telefone, vídeo-conferência, comunicação de massa, a apresentação, uma apresentação de
5 computador, uma apresentação com PowerPoint®, internet, e-mail, e comunicação com documento.

14. Sistema de computador interativo, como definido na reivindicação 6, para rastreamento de histórias de procriação e bem-estar de vacas
10 compreendendo dados de procriação e veterinários correspondentes a um bovino ou rebanho de bovinos, e no qual o sistema de computador é configurado para permitir ao operador destes trocar os dados com o dispositivo ou uma base de dados remota.

15. Sistema de computador interativo, de acordo com a reivindicação 14, no qual os dispositivos de admissão e descarga são um auxiliar digital pessoal ou um computador de bolso.

16. Método de se fazer negócio para rastreamento de histórias de procriação e bem-estar de criação compreendendo dados de procriação e veterinários correspondentes a um ou mais animais da criação compreendendo proporcionar a um usuário o sistema de computador como definido na
20 reivindicação 14.

17. Método de se fazer negócio para rastreamento de histórias de procriação e bem-estar de criação compreendendo dados de procriação e veterinários correspondentes a um ou mais animais da criação compreendendo proporcionar a um usuário o sistema de computador como definido na
25 reivindicação 14.

18. Método de e fazer negócio de acordo com a reivindicação 16, compreendendo adicionalmente proporcionar ao proprietário do animal ou cliente equipamento de coleta de amostra, tais como hastes de coleta e
30 etiquetas úteis para coleta de amostras das quais dados genéticos podem ser obtidos, e no qual as etiquetas são opcionalmente acondicionadas em um recipiente que é codificado com índice de identificação.

19. Método de se fazer negócio de acordo com a reivindicação 16, no qual o sistema de computador compreende adicionalmente uma pluralidade de dispositivos interativos, e no qual o método compreende adicionalmente as etapas de receber dados a partir dos dispositivos interativos,
- 5 compilar os dados, fornecimento dos dados para indicar a resposta de um estudante ou classe de estudantes a uma questão relacionada à operação do método auxiliado por computador, e opcionalmente modificar a operação do método auxiliado por computador de acordo com a indicação da resposta.

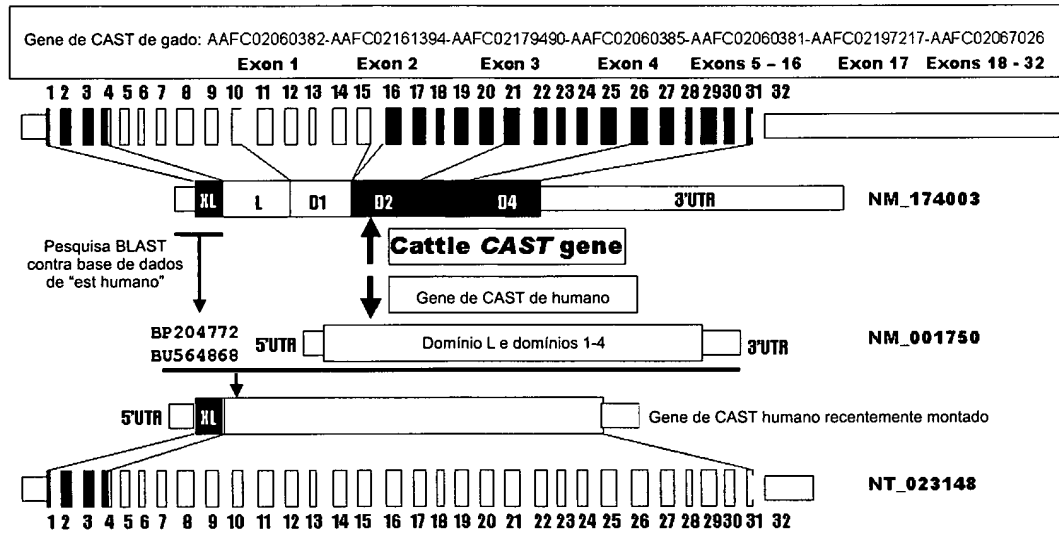


Fig.1

AACTGTCTGAAACACAGCACCGCAGAGAGCAGTGACCCCGAGGCTCACATTGTCAATGCCATGGCAGGGT
 TAACAGGTTTTCTTCTGTCTGATCTTCACAACTGGCAGCAACCACAGGCAGACAATCACTGCAGAGGAG
 GAGCCTTTTCTAGTCTGGCTTATAACATTTCACTTCAAGCTTGTATAATTACAGCCCCAAACCAGGAA
 GTGCTCTTTAAGATGAACCAAGACAAAAGCAATCTGGCCTCCTCGCAAACGAAATTCAACAGCCTCCTG
 AAAGGCAATGGGGTTTTAAATTTAGGACTGTTCTTCGGTTTTGATTGGAACCCAAGACTTGTTCACATGA
 AACATCATTGCACTTTCAAACAAGTAAAGCCGCACAAAACACACCCAGGCCCGTGTGTGTACACACACA
 CACACACACACACACACACACACACACACACACACAATCACAACACACGGAGTCACACACAGACCA
 ATTTGTTTCTGCAATCCGCTTCCTCATCCAGAGGGTCCAGGCCCGGGCCATCTGAGTTGGTAAATTCTC
 CTCCTAGTTAACTCAGAGCAGATTGCAGAAATGCTGTCTAATCTTGAGATTCTTAGGGAGTGGGGCTTG
 AGATGTAGACGGGGCCACGCCCCGGTTCCGCTCTCCTTGCACTGCAAGCCAAGTCTAGGCAAGTCCGG
 GGGGAAAACCCCGCCATCTCCAGCCCCCTCTCTGCGACCCACTGGGGCACCAAGGCGACCTCGGGTGG
 GGTGGGGTGTCCCTGGGAGAAGATACGGGACCCAGGGTGTGAGTTGCAACAGGCAGCCCCGGGCTGCTG
 CCGCCCGGGCGCTGCCAACTGCAGGCAGGAAGGGGAGGGCCCTGCCCGCGAAGGGGAGCTCTCGCGGGT
 CGGGGCTGGGTGCGAAAAGCTGCCCTCACAGGCGCGCCCCGCCAGCCCCCTCCGCGCCCTCGTCCCTCCAG
 CGTCCCCGGCTCCAGCCTCCCTTCCAGGCTCCGCGCGCGCCCGCCGAGGCAGCGCTCGCACCGGCCTC
 GCCATGTCCAGCCGGGCCCGAAGCCCGCCGCTCCCCGCGGCCCGCGCGCTGCCCGCCACACCCAGG
 AGGTGAGTGGCGCTCTACCGCCGGGGTTGAGCGCGGGGAGGATCTCGGGCCCCCGAAGCCTCGGGT
 CACCTTCAGAGGAGGGTTAGCCTTCTGGCCTTCTGTCCCTGGGCGCCACCCCTCCAGGCCCGGAGAAA
 CATCTGGAGAGAGACGGGTGTGGACTCAGAACCTGCCTCTGGGCAGGTGACCGGAGGGGCCGCGGATGG
 CCTGGGTCCCGCCAAGTTCAC'TGGGCGCGGCGCAAGCCACATTGGAGTGGGCTTCCCTGGCGGCTCAG
 CTGGTGAAGAATCCGCCTACAATGCGGGAGACCTGGGCTCGTTGGGTGGGAAGATCCCTGGAGAAGGG
 AACTGGCTCCAGTATCCTGGCCTGGAGAATTCCCTGCATCTGATAGTCCATGGGTGGAGAGTCAGGAC
 ACGACTGAGCGACTTTTGTCTCACTTGTCACTCATCGTTCGGTGGGAAAAGTTGGTGGATGAGACAGGGTGGG
 TGAGAAGCGGATAAGGGTAAAGGAGAATGAAATTCATTTGGGACCCCCCGGATTCCGGAGTCTGTAAG
 ATCAGCAGAAAAGTCTGGTCTTCCGCCAGTCAGGGCTGCAGGGCGTGGCTGCCTGGAGCAATAACACTT
 CTTCCCTAAGCAGCTTTGAGCCAAACCGGCAGGGCGGGGCGGGGCGGGGCGGGGCGGGGAGGGG
 CGGGGAGAGGGCGGGGTGAGCCGAGGGGCGGGGTCTGGGTAGGCCCGCCAGCTGTCTTGCCACCC
 CTGACAGCCTAGGTGCTTACAGAGTTAGTCCAGTCAGGTCTGCGGCAGGTGGAGTGCAGAACCCGTGGCC
 CTTTGTCTGCGCTGCACCCGTGTCTCGCCGGTCCCTCGGGTCTCTCGCTGCGTCTCTCGGAACACATCC
 ATCGTCGCCATGGCATTGCAAGCTGGTGGTACAAGACGCTAAATAGGAGTGATCGTCCCTGGGCAGGAC
 TGGGAAGGGAATGTGCTCTGTCTTCTGGGCTGTAGGCGAGGTCACTGTTAGATTTCGGGCGAGGGTT
 GGGGAGCAGGTGTGACCTCCTCTCAGACCTTCAGGGGAGGCTCCGGGAGAGGCTGAGACCCACCCCTGCT
 GTGGAATGTGGGAGCCAGCTCGGACGTACACGTGCTAGTTCGGCGTGAGTTCAAGGCTCACAAGTTGAATGG
 CATAGGGATTGTTGCTGGACGAATCTGCCCTCGGACGCGGATTCTGGATGAACGCGGACTCAGAAAGC
 GCTCAGCTCTAGAGTTTGTCTTTAACTCATAAGTAAAGCACAAAACCTTCAGAGGCGTGTGTTGGGCGT
 GTATTTTCCAGCGCCAAGTAGACAGGTAACATAAGATAGTGGGTGGGGCAAGGAAAAAATCCCAAACT
 CCAAGGAACATAAACAAGTGAATCAAGACCTTGTGAGAATGAGAGAGAGAGAGGGGAGACTCTATACT
 TTATTGCTGGATCGGGGGCAAAGATGTTCCATTTATTTAATGCACTGAATGCTTGTCCAACATATGTAAG
 TGTGATATCCAGTCAAGCTGGTACTTGTAAATGTTTTGAACCAGGAAGCATGCTTTAATATTAGGATGCTT
 TCCAAAATAATAACAATACATAAGTCTTGTGACAGAAAGGATAGAAATATTTATATCCATTGATT
 ATGATGAGAGTTGGAAAAAAGAAGTATACCTTCTCTTTCATCATGTGGTACTGGGAACAACAAGAGTGT
 TGTAAATCAGACACTCTTTAGAAGGAAACCTGTACTGGAACCAAGTAAATGTCCCTTTTCATCCCGTGGC
 CATTCAGTGAGCATCTACTATGTGTGACACTCAGTGTAGCCTCAAAGTGGAATAATGGACAGAATCTTG
 TTCTCTTTTTTCAACCGGTTTTCCACATCTATAAAACGTGGGTGATAATAGGTACCACATAAGGGTTGCTG
 AAGACTTAATTCCTGAATAAAAGTACCATAACCCATACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 GCTAAGTCGCTTCAGTCGTGTCGACTCTGTGCGACCCCATAGAGGGAAGCCACCAGGCTCCCCATCC
 CTGGGATTCTCCAGGCAAAGAACACTGGAGAAAATTTCCATTTTCTTGAGATGTTTCTTCATGAATGCT
 TCCTTGTACATTTTCTGGGTTTTTTTTGAATGTAGCATCTTATTGATGTGCTAAGAGCAAAGTTTACTTT
 GAGTCATTATCGTCTTAAGCTCTGAACACTGTTATTGATATTTGCAGATAATTCATATGGTGCATTTTG
 ATAAGAACATGAATTATCTGCAATATCAATAACAGTGGTTGAGAGCTAACATCACTTTTCTTCAAAT
 CATGATTCTTCAAAGGAAATTATGCCTGGATTATGGGATGAAAATAGTTTAAAGGTATGTTTTTAACCC
 TATTATTCTTAGGCACAGCTGACATCTTGTCTGAAACCTCTCTTATAATTCAGTGTGAGACATTTCTA
 AATTTCTTTTTCTGCTAGTAGGAACCTCTAGAATCAGTAACCTGCCAGTGATGAAGCGCTAACTTTTC
 CCCAACAGAGCTGGTCAAATATAGTGTGAGCCACATATGTGATTTTAAATTTCTAGTAATTACATTAG

Fig.2A

CTGACAGAGGAAATTAAAACGACATATTTTAATTAGCCCACTATATTCCAAATGGTACTATTTCAACATG
 GAATCAATGTGCAAATCATTGGGATATTGTAAATTTTCTTTTCATTTTCTTCAAAATCTGGTGTGTATT
 TTACACTTAGAGCACATCTCAGTTAACACGACATATTTCAAGTGCTTAATAGGCATGAGTGGCTAATGAC
 TACTGTATTGGACAGTATAGCTCTAGACATTGGTTTCTAACCCTTTCCTTTGGTGGCTGTGAGGGGATG
 ACCACAGCCATTTTGAGAATTAAATGAAAGTCGTGGACCACTTTCCCAACAACAAATGTACACACATACA
 GAAGACCATCTATAATCCCTTAAAGACCCGTGGAGCCAATTTATATATGTGTATACACACATATATATGA
 TTTATATTATATGCTATTAAATATAACATTTATGCAGTTTATTAATACAAAAATTTTATACAAAAACACG
 ATATAATACATATCTGCATATATGTATATCATGCTTCAGTAAGTCTTATAATTTAAATGTCTTTACTGAG
 ATCCTTTAATTACTGAAAAAGATTGTGAATGAGCTAACTTTAAAAATTTAGGCAGAATAAGGTAAAAATA
 TTGAAGGGTGTATCTTTTGTTTTATTTTGTTTAAGGGATTGCCACCTATAAAGCCAATCAAAGAAACGCT
 TCAGTATATATTTTCCAGTGAATACCATCAGTATTATGAATCTGATCTTTTGAAGAATATGTGGTTTT
 AGTAAAAAAAAAAAAAAAAAAAAATTTACAAAGCTAGATTTGAAGGACATCTGATGACCTTCATAATGATG
 AAGCATGACGGAGAGTAAATTTTGTGGCTGACACCAGTAATCATATCTGAGTCATAAGTTGACTCTA
 TTAGACCGTGTTAATTAAAGATGTTTCTTGGCCTAATTAGACCCATTGGTCCAACGGCAGAGTCAATCTAG
 GAGCAGTCATTTGGGAGGCCCGCCATTTGTGACCCAGCCTGGGAAGAGTGTCAAAACCTTCCCTGGCAGG
 ACAGACCCTTGGCTGCAGCGACTTTTAAGAATAAGGAATGTAAATTATAGATTCTGCAAAAAGGGCACA
 GCACCACCACCTTCTGCCTCGCCCCAACCCGCCCTCAACCCATCCCCGCCTCTGTCTCTGAATTGGT
 TTGCTTCCCATTCAGATAAGACTTAAGAGTACACCTGATGATATTATTCTTTAATCCATTTTGATACTT
 CCCAGAAATTTAATTGTATAAATACCTTTTAAAGAATGACGTAGCTTGCCGCTGGTGGTATAAATGGTGG
 GTGGCCACTGGAGGTTCCAGGTGAACAGTGGGTAATTAGTTCCTGAGTAGACTGTCTCAATGGAGTATA
 AAACCTCTAAGCAAACCATCTTTGTTTTGTTTTTCTTTGAGGTTTTAAAAAAATTTTATTGTGGTAAC
 ATACAGCAAATGTGCCGTCTTTACCATTTATAGCTGTAAAGTTTCAAGTGAAGTAAATATTCTAATCTT
 GTGCGACTATCTACCATCTAGCTTCAGAACAATTTTATCTTGCAGAACTGGGACTCTATACCTCATTA
 AACAAATAAATTTCTCATTCAATTTTCTATCTACATTTTAAACGCATGGTACTATATTCCAGTATCTGTGT
 CCTTATCATTTTTATTGTATTATGAAAGCATTTAACTGAAATGTACCCTTTTAAACAAATTTTAAAGTAT
 ACAGTGCATTATTGTTCACTGTAGGTACAATATTGTACAGCAGATCTCTGGAGCTTATTCAATTTCTTG
 ACCAAAATTTATGCCTGTTGATCAGTAACTCTGAATTTCTCTCTCCCTCAGTCCCTGCCAGCCATCAT
 TCTACTTTCTGTCTCTATGATTTTGACTATCCTAGGCACCTCACATCAGTGGAGTTTTATAGTATATGCC
 TTTTGTGACTGGCTTATTTCACTGGACTTCCCCGGTGGCTCAGATGGTAAAGCGTCTGCCTACAATGTG
 GGAGACCTGAGTTTGATCCCTGGGTGGGAAGATCCCCTGGAGAAGGAAATGGCAACCCACTCCAGTATT
 CTTGCCCTGGAAAATCCCATGGACGAAGGACCTCATAGGCTACAGTCCATGGGGTCACAAACAGTCCGAC
 AGGACTGAGCGACTTCACATAATCCCTCAAGATTATCCATGTAAAGCATATATCAAAATTTTCTTCTCT
 TTTAAGGCTGAAATACACACACACACACACACAGAGTATTGTGCTCATCCATACAACTACTGTGGAAA
 ACAGTAAGGTGGTTCTCTAAAAACGGCCATCCACGGAAGCAACCCAAGTATACTTGACGACGTCTCTCA
 GGTCTATAGGCTCTGTTTGTGCTGTGCTACTAGTCACTCAGTCGTGTCTGACTCTTTGCAACCTCATG
 GACTGTAGCCCGCCAGGCTCCTCTGTCCACTGGATTCTCCAGGCAAGAATCCTGGAGTGGGTTGCCATGT
 CCTCTCCAGGGGATCTTCAACCCAGGGATCAAACCTAGATCTCCTGCATGACGGCAGATTCTTTACCA
 TCTGAGCCACTTGGGACACCCTGCGGGCTCTATTTACTTTAATCATTTTCTTTCTGTTTCTCAGACTGT
 ATAATTTCAAGTTGTCTATTTTTGATTGCAATGATTCTTTCTTCAGCCTCCTCAAGTCTGCCTTTGAATC
 CCTCTAGTGATTTTTTTTCAATTTCAAGTTATTATATGTCTCCACTCTATAATTTTTTTTGGTTTCTTTT
 GATGATATATTTTTTATTGGTATTTCCATTTTGTATTATGCATTGGTTTCTTGACTGTCTCCACACTGTCT
 TTAGTTCTTTGAGCATCTTTAAGGCAGTTGTTTAAAGTTTTTATCTAGTGTATCTACTATCAGCTCTTA
 GGGACAGTTTCTGTTTTTATTTGTTTTTTTGTCTTCTCTTGATTGGGCAATACATCTCTGTTTCTTTG
 CATGCTTTGTGATTCTTTTTTTTTTTTTTTTGGTTGAAAACCTGGACATTTGAATCATTGAAAGTTTTGGA
 AATCAAATTTCTCTTTTCCCTAGGGTTTGTGTTTTTGTGTTTGGTAATGATGGGATCTTTTTGTGCCAA
 GGTTACAGCAAAAAGTGACACTGAAGAGCTTCTCAGGTCTTTTCTGATTCTGCACCTTTCCCTGGGCGTG
 AGCAATCTGTGATCAGAATACAGATCTGCAGTATTTGGAGGGCAGAGTCTTTTGCATGCTCCAGGGCAC
 ACTTGCAATGAGCTGTGTGCGCACACTTGCTGTGAGTTGTGTGTGCTCAGTCACTCAGTCTGTCTGACTC
 TGTGACCCCATGGACTACAGCCTGCTCAGCTCCTCTGACAACCTGGATTTTCCAGGCAAGAATACTGGAGT
 GGGTTGTCAATTTCTCTCCAGGAGATCTTCCCAACCCAGAGATCAAACCCACGTCTCCTGAGTCTTCTT
 TACCCCTGAGCCACCTGAGCTGTGAATTGGGGACGAGGAATGGGACGTGCTGCTGAGCAGAGCTAAAT
 GGATGCAAAATTAACCAAACTTACCATTTAAATGTTCCCTGGAAGTTGCAAGTCTTTGATAGACTCCAG
 AACTTACAGTAGTTACATCAGACAGATTCTGCCAGGGAGAATTTGTTCCAGATCAGGGGATGGATTCTCTG

Fig.2B

GTACCTCTTACTTCATCTTCCCAGAATCTTCCCTTTAAACTGTTTTAAGGATTGAAATCCTGTTAATGTG
 ATTCTAAACCTCTATTCTTCAAGTTGTCTCATTTTTCTTTTCATAAAAAATGACCCAGAAAGCTTTGCAGC
 TACCATAAAGAAAAAATATCCAGTGTCTTTCACTTAAGAGAAGTGTGAAATTACTTGATCCAATGAGTAC
 CATCTATGTAAGTTGCTTCCCTCAGATGGGAGCGTCATTTAACTCCAGAAAAATAAAGGGTCTATGGTTAGT
 TTTCCCTAAAATTTATGGAATTCCTTCTGACTCTGGAGAACTTACAGAATATAAAAGTGCTTTGGTAAAGG
 CCAAAGATGTTTACAGTTGCTTATAGCCACCTACTATGTGTATGTGTTGAACAGGGTGGCCTGGACAC
 AAAAATTGTGCTGATGGTCCATTCTGGTCTTCTCTTGCCTTTCTCATAGATAAAAAATCACAAGCACAGC
 CTAAAGACTAGTACAAAAGTAGCAGAGTAGCTATTGAAATTCTCATGGCGTCGTCTTTGACTCAGAGGAA
 ATTA AAAAAGCCTTACTACAGTGAACCCCTAAGCTAAACAAGGCAGAAAGGCAATCAATGTTACTTCTG
 AAAACTGGGCAAAAAGAAGGTGCACCCCTTACACTGCCACCTGGGCATATCCACCAGAATGTTAAATGTT
 AGAGAAAAGCAAAGGGAGTTCAGGGAAACATGAGGATTTCAGAATTATACATATGTTGCTGCTGCTGCTG
 CTAAGTCGCTTCAGTCGTGTCCGACTCTGTGTGACCCCATAGATGGCAATCCACCAGGCTCCCCATCCC
 TGGGATTCTCCAGGCAAGAACACTGGAGTGGGTGCCATTTCTTCTTCCAATGCATGAAAGTGAAAAGTG
 AAAGGGAAAGTCGTTTCAGTTGTGTCCGACTCTTTGCGACCCCATGGACTGCAGCCTACCAGGCTCCTCCAT
 CCATAGGATTTTCCATGCAAGAGTACTGGAGTGGGGTGCCATTGCCTTCTCCATATACATATGCATTTAA
 ACCTTAATTTGCGTCTTCTTTTTTATGAGAACCCCAAGAAGCTTATCCTGTATTTCTATAGTTTAAATA
 CAAATGTTTTTTCAGAGAAGCAGAAAGATACCATGCATCACACCTTTTCTTCTTTTAAATAATTATT
 CCCCCATAAAACCTAGCACACAGCAGCTGCTCACTGATTATTTTTTTTAGTAAATGAATTCTTCTCTCA
 TACTAGCCTTCTCCATTATTTTTGGGGTTTTTTCAGTATTA AAAAGCAAGTAAAAATGGCTAGGATATTCCCTT
 TGTGTGACTCTGTTGTTGTTCAACTGCTAAGTTGTATCCAACCTTTTTGTCAGCCCCATGGATTGCAGCCT
 GCCAGGCTCCTCTCTCCCTGGGATTTCCAGGCAAGAATCCTGGAGTGGGTGTCATTTCTTCTCCCCAGG
 GGAATCATCCCTTCCCAGGATTTGAACCCATGACTCCTGCATTGGCAGGTGGATTTTACTACTGAGTC
 TCCAGGGAACCCCTTTGTGTAACTTTAGGGAGATACAAGCTCTCAAGAAGATTCAAAAAATTTTTAACT
 TTTTTTTTTTAGCACAGAGGCAGAGAGAATAATATAACACCCCTCCTCTTGTACCCATCACCTTCTATAATT
 ATCAGCGTAACACTGTGCCTATCCTGCCTTATCTATCTTACTAGAAGCCCCACCAACACACACACACCC
 ACACCGTTGAAGTATTTTTATAGCATGGGAATTGCACAATAAGTGCTATCCCTGTGACCAGCTTGGTGT
 GTTGTGTGAAAACTGAGTTCTCTGTTCCCTGTGTACTTTGTTTCCCTCATCTGTACTATGGAATGTGAG
 ACTAACGTGACACAGTGTGCTGAGGTAGGCAAGTACTAAACATAAACTTAGCAGATGACTCTCGCTCCT
 TTGTACTTACGGCCAGTTAAATCTAAATTCCTGTCTATATTATTTATAGCCTTAGAAAATGAAAGCTTA
 AACCTCTGTTTCTTCTCATATTCATTTTATTGTCTCATAATTCAGCATGCCAGTGAAAAAGAAGGGAGTG
 AAGTGACCTTTATTGTGTTCTTAAATGAACGAGCCAATGGACCAGACTCCTCATTAAACAGTGGACTGGG
 CTCTCTCTCACTTGTTTGTGTACGTAAGCCAGCCCTGGGGTAGGCATCATGACCTGCATTTTACAAGGAG
 AAAACAGAAAGTTTAAATCATTGGCTCAAAGTCTCATAGCATTGCAGACCCACATTCTTCTGTCTTTATC
 CGAGTTTGGGGAATGGAATGGAATGCAGGGAGCTGGTTGCCAGGTAACAGAAAAGGCCGGAAGCCAACT
 GGGGTGGATGAGGCAACACAGGGGTAGAAATCTAGGAAGCTGCTACCATTCTAGACTGGAGGGGAAGC
 AGGATCACCTTGCAGAAGCTGAAACCACTGCTGGCCTGTCTGTCAGGAGGCTCCCATGGAGGCAGCCATC
 ACACATTCTTACTACACCAACTGACCGGGGTGAGGGCAGGGAGTCCACCCTAGGGCAAGCAGGTCCACTT
 CGGCCTGCACCAGGGAGCCTGTTTTAAAGCAAAACCAGCTTTCCATCTTGAGCCTTATGTTTGTGGCTG
 AACAGTCTTATTTTAGAAAGCAGAATACATGGAAAAATAAATTTTTAAGGGACTTTGCAGGAAATCCAGT
 TTGCATATTGTCTGTTAATTTTTCAAGCTATTCTGTTTTTTCAGCTTTAAAAAAGTCTTTTACCTAGTCT
 ACTAACCTGAAAAGAGCAAAAGTACTTCCCAAGGAATGTTAGAGGTTTGTGGAAATGCAGTCATCAGTCA
 AGAATTGTTTGTGCGACCGCTGCCACGGTACTAGGAAAAAAATTTAATCTACCATCTGGCTGAGCAGGC
 AAAATACACATTTATGAAGTAATCAGTGAGCAGGAGCGTATATAATAAAGTGCTGATTGGTTTGCTTTAC
 TAAGAAAATAGAGACTTACATGCAACATGAAGATCTGATGTGCCATTAGCAGGACAAGGCTCAGAACTAG
 TCAGGAAATGCTGGGTGAACACAGGGCACTGGGTTCCTGCCTTGAGCAGTTGGTTTACCCAGGAGGCC
 TGAGCTGCCCTGGCCTGTCTGACTTGGGGAAGCTCCCATGGGTGCTGACTCAGTGGGGTGATTGGAGGAG
 TAGAAACCTCCACTTCTGACATGGTGCCACAGAGCTCCTCTCCAAGCAGTGGTTTCAGTCTTAGCCGG
 GTGTTAGTTTTCAGCTGGGGAAGTTCAGGGATGTACATCTGTCCCAGCATCACCTGGATAAACTAAATC
 TGAATCCCCAGACTCTAGCTTAAGTGTTCCTCAAAAGTGCTTGGGCACTTGTGACCTGCACTGAAAGACGA
 GTTGTGTTTCGGTCACTAAGTCTTGTCTGACTCTTTTGGACACCATGGACTGCCACACACTAGGCTTCCC
 ATCCTTTTACCATCTCCCTGAGTTTGTTCAGGTCATGTCCGTTGAGTCAGTGATGCCATCCAATCATCTC
 ATCCTCTGTGCGCCACTTCTTCTCCTGCCCTAAGTCCCTTCAACATCAGGGTCTTTTCCAATGAGTTGG
 CCAAAGTACTGGAGCTTCAGCTTTAGTATCAACCCCTCCAATGAATATTCAGGGTGGTTTCTCTAGGA

Fig.2C

TTTACTGGTTGGAATCTCCTTGCAAGTCCAAGGGACTCTCAAGAGTCTTCAACACCACAGTTCAAAGGCAT
 CTATTCGTGGTGCACAGCCTTCTTTCTGGTCCAACCTTTACATCTGTGCATGACTACTGGAAAAACCGT
 TGCATTGACTATACAGTCCTTTGTGAGAAAAGTAATGTCTTTGCTTTTAAATATGCTGTCTAATTTGGTC
 ATAGCTTTTCTTTCAAGGAGCAAGTGTCTTTTAATTTTCATGGCTGCAGTCACCATCTGCAGTGATTTTGG
 AGCCCCCTAAAAATAAAATCTGTCACTGTTACCATTTTCCCGTCTATTTGCCATGAAGTGATGGAACCG
 GATGCCAAATTTTAGTTTTTTCGGATGTTGAATTTTAAGCCAACTTTCTCACTCTCCACATTCACCTTCA
 TCAAAAGACTCTTTAGTTTCTCTTTGCTTTCTGCCATTAGTGGTGTCTCCGCATATCTGAGGTTATTGG
 TATTTCTCCGACAATCTTGATTTTCAGCTTGTGATTTCATCCAGCCTGGTATTTACATGATGTACTCTGC
 ATATAAGTTAAATAAGCAGGGTGACAATACATAGCCTTGACATACTCCTTTCCCAATTTTGAACCAGTCC
 ATTGTTCCATGTCCAGTTTTAAACATTGCTTCTTGACCTGCATACATGTTTCTCGGAGGTGAGTAAGAT
 GGTCTGGTATTCCCCTCTCTTGATATGCTGTCTAGGTTTGTCTAGCTTTTCTTCCATGGAGCAACCTA
 CAGCTAAATGAAGGCTTTTGTATCCTTAGGGAGCACCCCTACCCACCCCTCCCGAGCAAGCCTTTGTCTC
 TGTGTGCAGCAGTGGTGCAGGGAGGGAGCTAAGCGGTGTGCACAGGCTCCAGAGGGCCCTGGGATTACG
 TATCATTTACTTCTTGATTTTGGCATAGCTGTCTCTGGACATGCAAACATTGATCTCCGTGAGCATTTCCCC
 AAATTGCAGAAGGAAGATCTTCCAGCTGTTTTACATAATCTTTATTATTTTCTTAAACAGCTTTATTTG
 GATATAATTCATAATTTACCTTTAAAATACACAATTCAGTATTTTTTAATATTTTAAATATTTTATTATTGA
 AAGTGAAGTGAAGTTGCTCAGTCATGTCCGACTCTTTTCAACCCCATGGGCTATACAGTCCATGGAAT
 TCTCCAGGCCAAAATACTGGAGTGCCTAGCTTTCTCTTCTCCAGGGGATCTTCCCCACCCAGGGATTGA
 ACCTAGGTCTCTGTCATGTTGGCGGATTCTTTACCAGCTGAGCCACAAGGGAAGCCCCATTTTATTATTA
 ACTCAGTATTATTATTTTATTAGTACTGAGTACAGAGTTGTCCAAGCATCACTTTAAGATCATTTTACAA
 TGTTTTTGTCAACCTTAAATGAACTCAGTACCTTTTGGGGGGGGTTTAAATCTAACAGAGGCATTTTCA
 GAAGCACACTAGTTTTTGTATGCTCTTTACAGCTGAGCTTAGGCTTGCAGAAGCTCAATTTTTGTCTGTC
 TTCTTTCTCTTCTTATATGTGTAGAAGTTTCGAGCTTTCAGCTCATCTGTGACACATACCTTTGTCTCTGA
 GTGTGTGAGGTTGATCTGGGTTGCCCTGGGGAGGCGACCCCAATGTGTCAGAGCCAAAGCCTGGCGCT
 GGCATACGTCCTTAGCTGGGTGGGAGCACAGTGCCAAGCTTCATGGCTCCTGTCTCTGTTGATTTCCTG
 AGACGTTTGTGAGATTGCTAGGGCAGAAAATGTATGCATATTAATGCTGTCTTCTAGAATGAGGACAAA
 GGACAGCTATGCTTACCAATATATAGCAATGCCAGGAGGTGGAACAGGTTGGAAAATTCTTTTTGAAG
 CAAGACCAAAAAACCCCTTTATCTGATATTTTCTCTGCCTCTAACTCCTGTCACTCTTGTTCTTTACAGT
 GTTTTCATCTTAGCTCCTGGCTTCCGGGGCCGTATCCTCTGAGTGAAATGTCTCCTACTTTAGGACCAGC
 ATCCTGCACTTCTGTCTTTGCTCCCGTCCAGCTCCTAAAACATCCTAAAACAGGATGACACTTTGGTAT
 TGTGTGCTGATTAATTTACACACACACAAAAAGAAGCTTTTGAATAGGAAGGTTTTATCTGGGTT
 TTGTTGATTTTTTTTTTTTTTTGAGTTGTTTTCTTTTAGAGCATTGGATTGAGACTTTCTTGACGGGCATC
 AAAAGTTTTTCAGCTAGCTTAACCAGCCCAGTCTCTGGCTTCCAGACTATAACCCACACATGTAGCTGACA
 TGTAGTGTGACTGAGGAGGATCTTACGTTTCACATTCTATTTCTTTATCAATAAAATGAGCTGCCAGTTG
 TGTCTCTTTTTTACTGCTATGGCAGTCAAGGAGACCACAGAGAGGAAATGCCAGTGCTGGGCTCACAAA
 GGAGATGCTTAGAACATGTTAGTTTCTTTCCACTCTTACAGAAAAGTCTTCTGGATATTTTCAGGTTT
 TGGTTTCAGTCCCTTTTCCATGTTTCTCTCAGCTCATCCTGACCCACCACTACATGTAGTGATAGTAGAC
 GAAGGTTTTCTTCTTCACTCTTTGAGATACTTGCTGCAGCTTCTTAAGCTGGATCAAATCTGAAATGTAT
 AATTTTCAAACATTTATTTTCATCAGCAGAATGGAAAGTATGTGGAAGTACACAGTTATTCCATGAATGT
 AGAAATCAGGCACTTTGAATCTCTCTCAATGAAAAGATAAAGTATTTTGTATCATTAAATAGCTACTG
 AATGGGGTTCCCTGGTGACTCAGATCGTAAAGAAATCCGCTGCAGTGTGGGAGACCTGGGTTTCGATCCCT
 GGGTTGGCAAGATCCCCTGAAGGAGGGCATGGCAACCCACTCCAGTATTCTTGCCCTGGAGAATTCCATGG
 ACAGAGGAGCCTGGCGGGTTACAGTCCATCGGATTGCAAAGAGCTGGACACGACTTAGCGACTAAGCACA
 GCACAGCACCAGCTACTGAATATGAATAAGCGCCATTTCTTGGTCAGGGCAATGATTCACTTATAAAGTT
 TCCATTTACTGGAAATGAATGGCATTAGGTAATCTTATTAATTCTATTAGAGTTTGTAGAGTGAGG
 ATGTATTGTAATTATAGAGGGAAAAGTCCCTCAGTTTAAGGACATGCTTGATAAAATGTTTAGGGGTAAAG
 TCTCAGGATATCTCAATTTACTTTCAAATGCTGGTACATATGCAAATATATGTAGATCATGTGTCTGTC
 CATAATATATGATGAGATGTGAGTATATGGCAACAATAGTTAATAGTTGTTGACTCTCAGTGATGGGACT
 ATGGATATTTTTCTTTCACTTTTCTGTATGTTAGAAAATTTTATAATAAAAAAGCCAGGTTAAAAAAC
 TAGAAAAATATAATCCTGATAGTTACTAAAAATGTCATTTCACTATGTGATTTTATTTATATTATTA
 TACTTACTCTATTGAGAGCAAATCTTTCTTTGAAGTTAAAAAATGACAATTTCTATAGTACATAATTTTT
 ATAATTTGTCTATTACAATTTCAACATTTTAAAAAGCATTTTCTTTTTTAAAAACCTTTTTCGTAAAAACATA
 TATTTATAATAGACATTTTAAGCAAAAAAGCATTTGTTATATATTTATATATATCCCTCTCAAACATTT

Fig.2D

ATTTGGAACTATGTAATTGTGAGAGATTTTGAGCATTTTCAGTGGAAGCAAGTAACTGATGATGTTGGTT
AATCTGGAGGACTGCAATTTTGAAATACCAGGGAATCAACTTTATAAACGTTTTATAGTAAGATATCTAT
CTACCTTTATTTTAAAGAGAAAAGGCTGTCTGCATGTTTCCTGTAGGGTAGATCGTAATGTATGTTTCCC
ACCTGAGCAAGAACCGAAATGGGACCTTGCGGAGAGCATGGAGCCTTGCCCAAGACAAGAGCGAGTTGTG
TGTCTGTATTCTCATTTCTAAAATGGAACCTCACACGTGGTGCTAGTGCAAAGCATCCGCCTGCCAATGC
AGAAGACTCAGGAGATGCGGGTTTCAGTCCCTGGGTGCGGAAGATCCCCTGGAGGAGGAAATGGCAACCCA
CTCCAGTATTCTTGCCTGAAGAATCCCATGGAGAGAGGAGCCTGGTGGGCTACAGACCTCGGGGTCACAA
AGAGTCGGACGCGACTGAGCGAGTAAGCACATTTCTAAAATGGAGTAATCATACTTCCTAAAAGTAAATG
AGGCTGTATGTTTCAGTGTGTATGCTCAAACAGTGGGTATTGTTGTGACTAGATATTACCTATATCTAACC
TGAAAAATATATAGAATTGGTGATAAAAAATGAATTTGATATCATGGTTCTAAATTTACCAGTTCTGAAAG
GAGACAAATTGAATAAATTTGTGTCATGTAGACCCATCCCTCTCAAATTTCTCTCTCCAAACACTGAATAA
TTCTTCTCTTTTGCTCACCTATTATTCTTAAAGCTTATCAATGGCCATCTAGCTCTGACATATATATGTG
GGCTTTTAAATGGCTCTTGTTTTTTCAAGATTCATTAATTACTCAATAGCTATTGCTCTCAACTGTGGAC
AAACTCTTTTTTTTTTAAACACACACCACATCTTCTTTATTATCATCATCTGTTGATGGATACAGGTTGCTT
CCATGTCTTGACAATTGTAAATAATGTTGTAAATGTAAACATGAACATTTTTGTGCTTGCTAGGTTGTGT
CCACCTCTTTGTGACCCTATGGAATATAGCCCGCTATACTCCTCTGTCTGTGCTTCCCTTGTAGCTCAG
TTGGTAAAGAATCTTCCTGCAATGCAGGAGATCTGAGTTCAAGTCCTGGGTGCGGAAGATCCCCTGGAGA
AGGAAATGGCAACCTGCTCCAGTATTGCCTGAAAAATCCCATG (SEQ ID NO: 1)

Fig.2E

GTCTACTATCACTACATGTAGTTGGTGGGTGAGGATGAGCTGAGAGAAACATGGAAAAGGGACTGAACCAAAAC
 CTGAAATATCCAGGAAGGACTTTTCTGTAAAGAGTGGAAAGAACTAACATGTTCTAAGCATCTCCTTTGTGAGCC
 CAGCACTGGGCATTTCTCTCTGTGGTCTCCTCGACTGCCATAGCAGTAAAAAGTGTCCCACCCAGCAGTTCAT
 GATAGGGATAACGTTTCTCTCTCCAGGGGTGGGCAAACTCTTGACTAGGTCACTGGGTGAAACAAGCCCTGTA
 GAGAGCTTCATTAACCAAGCTACTGCAATCTGCCTGCTTTTCTGTGGGCAGAAATTCATGAGAGGTTTCTGTTTG
 GGGGAAAGTTTTCATTTGTGAAAGTGATTTATTTCTAGTTTTTATACAAAGAACTACTGCACATGTCTAATATT
 GCTTAGCCTTTTATTTTACAAAAATATCTGTTGATATGATTTAAAGAAATATTATCTGTTTTATTTATCATAG
 TATGTGTGATATAGTTGGTAAGTCATGTCCAACCTTTTTGATCCCACAGAGTATAGCCTGTCAGGCTCCTCTGTC
 CACTGGATTTCCCAAGCAAGGATACTGGAGTGGGTTGCAATTTCTTTCTCCAGGAATCTTTCCAACCCAGGGAT
 CAAACCCAGGTCTCCTACATTGCAGGCAGATTCTTTACCGACTGAGCCATCAGGGAAGCTCTTATCATAGTATAG
 AGACAATAGATTTGCTGTGAGAAAACCAGCTTTGCAAGCTTTTCTCATTTACTGGGTGTATCTTTTTGTGAAAAG
 CGCATCACATGACTTTTATTCTTAATTTTTTTAATCCATGAACCTCAGGTGAGCATCTTTTTTGTATCAGATAA
 AATATATATGAAAGTTCTGTAAAGTAATAAAAAATGATAAACATTATTATTAATATTACTGTTATTGCCAT
 TCTTTTTAATTTGAATAAAGTGTGAGTGACATGCATATCCAGGTTTTGATTCCACTTTAGGGAACATGATGGGTTT
 TATGTTTCAATGTAACTAAAGATGATTGGGCTAAGACCCAATAGTGAGCTTTTCATTTTCTTATATTTAGGTTCC
 TTTTCCCACTTTCTTGCAATATCCCACTTGATATAAGAAATAATTCTTTATCCCTAATTTTCAGGAACATAAAATAT
 TTTTCTTTCTCAAAAAAATCAAGAATGAGAACTTCCCTGATGGTCCAGTGGTGGGAGTCTGCCTTTTCAGTGCA
 GGGGTGAGGGTTTGATCCCTGATGGGGGAAATAAGATCCCACATACTGTGGGCGACTAAGCACCACCCAGCA
 ACTAGAGAGGCCACACACCTCACCTGAGACCCAATGCCACCAAAAAATAAATAAAATAAATATTTTTAAAGTACA
 ATAATGTTTCATGGCTTCTTTTATCTTATCAAACCAATAGCGTGTAACCTACTTGACGCTTTTATATACCACATA
 AAAACTTCTCCTTGCCAACGCAGGAGATGTCAGAGAAGCATGCGAATTGGCTCCCTGGGTGGGAAGATCCCTTG
 GAGGAGGCATGGCAATCCCTCCAGTGTTCTTGCTGGAGAATCCCATGGACAGAGGAGCCTTCCCTGGCTGCAG
 TCCATAGGGTCGCAAAAGTGCAGACATGACTGAAGCAGCCACCATGCATGCAAAAAATACTCCCATTCATACCA
 GTGTTCTTGAGAAAAATAATTCTGATTTAGGACATAATGAAATCCTAAAGCCATAAATAATGAACGTTGGAGG
 TATGGGGTTCCTCAATGCACGCTTTATCTTTCTGTTTTGGTAAGTAGACAGAACTCCAGAGAATTTCTCCCTTGA
 GTTTATGCTGAACCTGAGGCTTAAGTAATGGCTTCAGCTCAACATGGTCCCTCAGACATCTCATCTCCATGAGGAAC
 CAAATGTCGCTTATGCTGGCAATTTGGGCTTTGAAATTCCAGGAGGCATTTTGATAGGAGACGAGATGCCGCTTG
 GCACCACCATCAGATCTATGTCTGGTTAATTACAGACTTACTTTCTGCATCCCAATAAATGTCTATAGCTTCCC
 TTTATTTTTGAAGGCCACCTATGTCAATGGAGAATTATTAACAGTTCTCAGGAGGAAGATTAAAGGTGCTCCTGT
 TTTTACAGTCTGTGATTTAGCCCTTTGGGATTAAAGTCACTATTATATAATCTGTATCACCTCCCTCTATTAA
 ATGTTATCTCTTTAATATTTTTTACAGCATGTCTAGTAAAAACAGTGGATCGCCTTCCAAGTCAGGAGAAAA
AAAGGATCAGATGAGGTTAATTTCTACAATACTGAATTTTCATTTTCTCCTGTTTTTAAATTGTGAATTCATCAT
 GAAATAAATAAGGTAGGTAAGCAATCTCCTACTTAAATTTGGCAAAATTTAGTTTTAACTAAGAGCTGCTTCAA
 TGCTGTAAGTGGAAATCATTAAATAAATAAAGTGTTCTCAGAGAGGAACATTTTCTCCAGAAATAGTTTTCA
 TTTAGGAGAACTCAGTTATTTAGGGTGTGTCTCTATAGGAATTGGTGAATCTTAAGTGTTTTTCTGTTTTTTT
 TTTTTTTTTTAAACAAAAACATACAGAAGGGAAACCAATGCTACTGTTGAAAAGGAAATCTTTTTTTATGAGAG
 TCATTGTGAAAGGAGAACAGTGATAACCTAGACTAAACAGTCATTTCAGGTTCAAAACAAAAGGGAGAGAAAA
 GACTATTTTTATTTGGTTCCATCCTGTGGAATACTAGTGGAAGCAAGCCTCACTTTCAGATGGGTTGTCTTTGA
 GTCCGTGAATTTCTCCTTTCACTGAACGATACCCCATATGTTCACTATCAGCATCGTCTGGTCTCCATGAATGG
 GTCGTGTGAAAGTGCTAGCCTCTGAGAGGGTCCTAGGCTTGCTTCTTTTGTGCTTTTTACATTTTAAATCTCA
 ATGGAATATTCATCTTCATTTCAGTAGAAATAATTATCATCATTATCCAGTTTGTAAGAGAGCCAGGTCCTGAA
 AGGTACCTGGAAATTTAAAAAGCATAGAGTAACCAATTTGGTTTTCTCTCTTATTAACATGCTTCCAGAAATCT
 AGCTTCTCTTTCTGCTTATTTTCTGCAGTTTTTGAAATAGTCCATATATTTCTGTATATTTCAGTATGAAGGGAGCC
 ATTATGACTAGATCTTGCTATGAATCTGCACGAAGGTTTAGTCCTCAGCTGTCTACTATTGTTTCCATTACCAA
 CACCCAAGTCAATATAATTTAGATTATTTCTTACCATGCAATCTGCAGTGAATATTTTGTTCATGGAAGAA
 AATTTTCTTTACAAAGTAAAAATAAACTTTTATTAAGGTTTTATCTGGCTCTTAGTTAAATTATTTAACTTTT
 TTTCCATGAGAACAGCCAACCTTTTACTAGCATGAAGATCTATGTACCTAATATCACTCAGTTCCCCTATCAATC
 CTTCTCTTTAAATATCTTCCAGATAAGTTATCTACCAATAGGTTTCTGTTGAATAAAGAGACAACATATAGTG
 TCTCTTTATTTGCCAGATCAGGAAATAATGTTTTTAGATGTTATTCTTTATGTTAATATTAATCAAGTTGAAAAC
 ATTTGCACAAGAAAGGAAATCTTCCAGTTCAATCAAGTGATAGTGACCTAATCTAAGAATCCTAGATCTTAGA
 GGATTATCTCAACAATTACGTTTTTATGTGGAAGCTTAAATGTTATGTCTTGAATGTACAGATATCATTTGGTTA
 AGGCAGCAGTGCTCCCCAGTGTTAGATCTCTTTAATATCTGGACTATAAAATGCCTCATTTTCATTTTTATA
 AAGGGCTGGGAAAAATTGATAAGGAGTCTGTTTATCAGTACAATCAGAGGAGAAATCAAGACTTCTGTGTCTGTG
 TGATGACTGTGCTGCAGTTCTTTGAAAGTGAAAGGCTGAGCTTCCATCTCAGGAGAGAAAGAAAGAAATTTTC
 TTTTCTTTATCAGCAGTTGAAGCAAGTCTGACTTTGTTTCATAACACGTCTGGGAAATTCAGATTACAGCTCTT
 GCGTTTGTGCCAGAACCTGGTCATCTGGGTCTG

Fig.3A

TTCTGCTCCACCCAGCTGGCTGACAGAGAGGAGAGCCAGGCTTCGCCCTGCTCCCGTGACATAAATCACTGCAG
 AGGTTTCAGATAGCCTGAGACTGAAGACTGCTTATTATCTCGCCAAATTCGCCCCGATCTATTTGATAGGGTGA
 CGGATGTGTAAGTTGAATCATCTAGTGGTCTAGTTATATTTCTAATGTAGTCTGTCTCCAAGGAGTCTCTCATTC
 CGTATGTCATGGATGACATATTATGGCCCTGGAGACAATTCAAGGAAATCATCCGTTTGGCATTATTTTACTCT
 ATGTTTCGTGATTATAAAACCACGTGCAGTCTGGACTTGTCTATCTGCCCTGCCCTACCATCTACAACCAGTTACTAT
 CGTCTTCAGATTACTAGGGAAAGATTTCTAAAAGTATTCTTTCCGACAAAATGCTTTCTTATCTGCTGTTCTGT
 TGAACAAGAAGTAAGATTAAGAAGCTTATTGGGGGGGTTTGGTTATCCATTTCAAAGTACATCTACCTCCAAA
 GAGTAAACAGCTTGTGTCCCCAAGCATATGCAAACATACTGGTTAAGCTTCCAAGGACTAAGCCAGGCCCTTGA
 AATAAAGTCAGTTTGAAAATGTGCTTTCTGCTCAGGGATACTTCTCCAGTGCCTTTTCAATTTGGAGGTGCAAGA
 AAAAGGGTACAGAGAAGGGCTTCGGGAAGATTTAACAGGATTGAAAATGATTAGGATCAGACAAGAACCATTAT
 GATATTGTCTAGGCAAACTACCTGGGGATGAATGCAGCTGTGTGAATTTCAGAGTCTTTCTCTTCTGCTGCC
 ATAAATGGTTTTTTATTTATCTTGAATATTACTCCTATACACCCACTAAGTGATCCTAGGATGTCTGCAGTTATTT
 TAATAGACTATTGCAACTAAAATATCATTTGTCTGTTGCATTTGCAGTATAAATGCTTGAGGATGGAGTTGAGT
 AAGTTTCTTTGAAAATGTATGTTTTGTTTTTGTGACAATTGTTAAGAACTTAAGTTCTTCTCAGGTGTAGAGGA
 ACTGTAAAGACCAAAAGTTGAATGGACTAAAAAAACCAAGATAAAATGACTACAGAGTAGGAAAGAGAAAAATGA
 AAACATTTAGCAGCACCTTTAAATACAGTGGAAACAAAGTTGTCTTTCCAAACCTGCAATCACATCTATTTAAAGT
 GTTTGACTGACACATCTGACCACCCTGGGGAGTGTGAGCATCTTAGCAGCATATTTCCGTGACGTCTCTGAGAAA
 TAGAATTTAAGACATGGGATGCACACGATCTCAGAGAGCCTGTTTTAAGGACTTCAGATTTAATGATGGATATAT
 TTTTATAGGTAAAGAAAGGGAGATGTATAGAGTCTAAACTGCTCCTTGAAGGTATTTGATAAATCTCTGTACCA
 TAACAGCTTTTCGTGGTATTCTTCCCTCAGTGTGTTTACCTCTCCTGCTTCTAAAGTCCATGAATTACTCCAACAT
 TTTAAAATTTCTTTCTAACATATTTTTCTCTTGATGAGATTAGTCTTTTTTAATTGCAAAAGTAATATGTACTT
 GTGAAAAGCAAAAGTATTTACCCTGTCTTACCCTTAAGTTATCACTGCTGTCTACAAAATTTGTGTGTACTCTC
 AAATCATTTGTGATGATATTATATGTTTCATAGTTATTATATATACTGATTTTTGTAGCTTTTAAATAATGATGCAG
 CTCTTACATATCATTATAGAAACGTATTTCAATATGTTTCTTCTAAAGAAAATTTAGAAGGTATCAACAATTTTA
 AAGAAATAAACACTGTCTATCTAGAATTCAGCCCTACAGAAATGGCCACTATTTCTATCTTTGAGTATTTTCTT
 CATCTTTTTTCTATGTGTGTTTCATGTTAGTCACTTAGTCTTGTCTCTACTCTTTGCGATCCCATGGACCACAGTCC
 ACCAGGCTCCTCTGTCTATGGAATTCCTCAGGCAGGAATACCGGGATGGGTAAACCATTCCCTTCTCCAGAGGATC
 TTCCCGACCCAGGGATCTAACCTGGGTGTCTTGCAATTGCAGGCAGATTTTTTACCATCTGAGGCACCAGGGAAGC
 CATATTTTATGTTTAAAGCTGAGATAACAAAGGAACATTACTAGACTAAAAGGGCCTAGACTAAAGATGATAACA
 ATGTGCTTTGTGATAGATAAATATTTTCAATTGAGAGAAGTCATCATTGTGTGTGGCTTTTAAATACATGGGGT
 TTTCCAGGTGGCGCTAGTGGTAAGAATCTACCTGCCAAAGCAGGAGATGTAAAAGATGCTGGTCCACCGAGTTG
 GGAAGATCCCCCTGGAGAAGGAATGGCAACCCACTTCAGTGTCTTGCCTGGAGAATTCCACAGACATGGAGCCC
 ACCAGAGGAGCCTGGTGGGCTACAACCCATGAGGTGCGAAAGAGTCGGACATGACCGAGCGACTAACACTTCCAC
 TTTTAAATACATAGAAAATCGTATTTAAGTACCTTGGCATTTACATAATATTTAATCTGGTTATTATTTTGTGC
 TGTGTGTTTGAGAGAGGAAAACAAGTTTCTTGTAAACAAGTTTTCTGGGACAGTGTTCGGGCAACAAGTGATAA
 AGCCGTCCTTAGCAAATTCCTGGAAGAGTCCACAGCAAATGTTATTTTAAAGGGGTGTTAAGTCATTACAACTCTCT
 TGTCCACTTTAATTGTGGGGACTTGATTGTGCTGGACAAGTTCATGTCCCAGTGCCAAGAAGGAGGCAACTCTGA
 AAACGTGCCCCAAGGGCAGTCCCCGTCTGTCTTACCTGGAGATCCAGAACTCTGCTGGGTGTCCGATTTGAAAGG
 CTAAACTCACTAAAGGAAGTTCCAAAATCCTTTTCAATTGCAGTTTCCAATTTTTTTTTTAAATCTAAATTTT
 ATTATTTTCCAATTAACATAAATATTTTAAAGATACTTATTTTCAATGAAGGCATATGGATGGTGATTTTCTG
 CATCTGTGTGGATCTAAGCATGCTTTTCTGCTGGCCTCACACAAGGTGCTACTACAAAATTTGTACGTTTTTGGGGT
 AACAGTTCTGTTCCCTTAAATATTTCTACCTGTTTCTCCATAGTTTCTGACTTTTTGTGCTATAGAAAAGATGT
 CTGAGGGAGCCTAATTGTTATTTATTTTATAGAGAACTGTTTTTCTGCTTTGGTGTATTTGTGATTTCTTTTCT
 TTTTATTTCTTACAATTTTAAATTTGCCAGGATGTGCCTACATGTGTGTTTTTCTATTAATTTTTTCTTAGAGTG
 TCTATTATGTAGTTATCGGTCTTATTTTCAAGTTTCTTTTAAACATACATATAGTTCTGAGATCAGACATTCATTA
 GGCACCTGCCCTCAGTTTTCTCTTTGACCTTCACCTTGTATTCTGCTAGAGTTTCTGAACTTTATCCCCCATATT
 TTTATTTCTGTCATTTAAATTTCTGCTCCTTGAAGTTCCTTCAATACAGGTTTAAATTTTGCCTTTTCCCTGTTGCA
 TTGTACCCCTCTCTAATCCAGTCTTCTGCCCCCTCTCTCTGTTTCTCTCTGTTATTTTCCCCTCTCATCTCTCT
 CTGTGCTGTGGCTCTCTACCTGGCTTTACGGGGGCAATGCCATCTCGTTTCTTATTAAGGATGCAGACAGTCT
 TCTGAAATTTTCTTCTGTAGTGATCATTTTATGAGATTTACTGTCTCAGTCTTCAGGAAAACGGCATCTTTCT
 TGACTGTTTTTCTTCTTCTCACTTATCTGATTAAATGTCTTTTTTGGAACTCAGGCAAGGCCTAGGAGGCCAAA
 GTTTTTCTGCAAACAGAGGCAGTGGAGGATGTGGTGAGTGGGTGGGAGGTGGCCTTTCCCGGAAGCCTCCACA
 GGGTCTGCTCAGTTACTGTGAGGGTGGAGATTTTACCTCAGTTTTGGCAATTCCTGACTTCTTTTGTGTTGA
 GTCAAATTCAGCCTCTTCTGTGAGAAGATGATTAGAAAACGCTGTGCGAGATTATTTGTAGTAGCCATTCCCG
 GAGTGGTCTCCTAATGACTCTCTTTCACGTCTTGAATCCAGTCTCCCCTTTTCTAGCCTCCTTCGGGTATAACCT
 GTGAAGTAGGCCTTCACTGCGGACTCTGTTCTG

Fig.3B

TTTATTAAGTGTGTGTAGCTCCTCATGAGTTTCACCAGAAGGGCTTGAGGTGTAGGGCTTCGGGAGCATTCATG
 TCGCCTCTTTCTGGAATTTCTCTGACTGGTTTCCGGTGCGTGGGCCCTCGTTCTGTGCGTGTGTCTCATGACGT
 GCTTTAGGTACTGCTTGCGTCTCAGCGCCTGATACTTGTGTACGCACGTAATAGAGTTTGTATTTCTTATGTGG
 AGCCTGCCATATGGGTCTTCGCCAGCCATCTATGGATTTAAGGTCAAAACCTATGGTTAGAATACCAGCCATTAT
 GTAGCCATTCAAACAGCCTGTTTATCGGAAGACATGTTTTAGGTTGAATTTCTGAGAAAAACACAGGGAGTTTCT
 CAAATTCAGGACAAAAATTCTTTGCCGTCTTGAGACCACTGTAAGAGCATCTTCTTTCTCATGTGCCGTGGGAGC
 GCTGAAAAAGCTCCCAAGTTAGCCCGCCCCCAGTCCACAGTGCCTAAAGAGGCAGCACGCTGTAGTTGCTGAGAGC
 AGGCCCTGGAGTCCCAAGTGTGTGGACTTGTCCCAAGCTCTGCAACCTCTAGCTGGGTGTTCTTAGGCAGATTCC
 TTTGCCTCCCTAAGGGTACCTTTCTCATGTGTCCATGGAGACAGCGATAGGTTCTACATGAGTTAATGTGAAGAT
 CAATGAGACCAAGCAAATGAAGTGTCTAGAACAGTGCCTGGGACATGCACGGTGTAGCTGCTATCATGGTTAT
 TATTATTACTACTGTTATTACTTTTACTAGCCAATGAGGAGAACTTGTTTGGCAATAAAGGTAAGCCCCCTATAG
 CTTTAGAGCCAGGACTCAAATCTTTTGTGACTATATATTGCCAAGCCTTCCCATTAAAAAAGGAAGGAAGTT
 TGGTAGTTCATTAAAGCTTAGTAGTCTTTCTTTGTTTCAAAACAACCTCTGTTAACACCCAAAAATATATGCCAGTT
 GCTAGAAAAATACAGTCCCTCTGTAACCAAGAACATGAAAAATAATTTGCTGTATAAATTAAGAATCTGTGATATACA
 AAATGTTATCTAAGTAGTAGTGAGTGACTAATAAAAAATAGTTATTATTGCACTTCAGAGTAATGAGCCTAACTGC
 GTAAGATCTTAGGTCTACCAAAATTTGTCTCTGAAGACTATACCTTATAGCATAGGGAATTAGAATAAAAGTTAA
 AAAAAAATCAAAGGAATTGAATAGAAGTTACATTTCTTGGCTTTAGAATAGATCTTAAGTCTGGAAACCCGAG
 GGACAAATCAATTAAATATTAATCCTGAAATACACTGTTTAAACATTTATATTTTTTTCAGTCTTCTCTCTAGAC
 TTTATTAGCTTCTAAAAGATAATGGAATGACTGATTGAGAAGGGAGGATGACCTAAAATTGTGCCACAGCAAGA
 AACCTTTATCATTTAAGGACTGAATATGTATGATTCTTGTCTATAAAGTCAGTAATATTTTATCATATACCAATGT
 ACATAATAACTTGTAAAGACCTTCCAAAAACAGTTCACCTGTGAAGTCAAAATCTCTTAGCAAGTTGACATCATA
 AATTAGTCTTTTATATTTCAATCCACATTATTAACAAGCCCTTAGTAAATGTATAGAGCGCTGTTCAATTCAGTT
 CAGTTCAGTTGCTCAGTCGTGTCCAACTCTTTGTGACCCCATGGACTGCAGCACGCCAGGCCCTCCCTGTCCATCA
 CCAACTCCCAGAGTTTACCCAACTCATGTCCATTGAGTCAGTGATGCCATCCAACCAGCTCATCCTCTGTCTATC
 CCCTTCTCCTCCACCTTCAATCTTTCCAGCATCAGGGCCTTTTCAAATGAGTCAGCTCTTCACATCAGGTGGC
 CAAAGTATTGGAGTTTTCAGCTTCAGCATCAGTCGTTCGAATGAACACTCAGGACTGATTTCTTTTAGGATGGACT
 GGTTGGATCTCCTTGCAGCCTAAGGGACCTCAAGAGTCTTCTCCAACACCGCAGTTCAAAAGCATCAATTTGTT
 AGTATCATAAAAATTTATTTTAACTTTGTCTCTCATCTCCTAGGACATTACATGAGATTATGCAATATTTCTCA
 AGTTTAAGCATGCCTATGAATCACCTGGACATAGTATGAAAAATTCAGACTCCTGGGCCAGATCAGGGTGAGAT
 CAGGCTTTTCGGTATTTTTTCCAGTCACCTCAGGTGACTCTTCTGACGGTGGTCTTCTGAGCACTTTTCGAGTAACA
 CAGATAATTCTTACAGATCTCATAATGCAGTCTCCTGGAAAAATAGAGCTTTTCTTCTTTTCTTTCTTCTCATT
 TCCTTTTCTTTTGAAATGCCTGCCCACCCACCCCATCCCTCAGCCGTCTCCTGTCTCCATGTCTCCTGGGACAG
 CGTGGCTCTGAAGCACATTGAGGGCAGTGTGGTTTCAAGGAGTCATTGGACCAGATGGGGAGGCTTTGGTTGCAA
 GGGAGAAATGCAGAAATGCAACCCCTATTTCTTCTACTTCTCCTAGGACCTCGTTCCTTAACCTCTTCCCTCATT
 CCTTTTGTATTGTGTGTTTTGAGTCACTAAGTCGCATCCCATTCTTTTTTTTTTCTTTTACTCCTTATTTAAT
 GACTCCTCTTTAGACGACATTTTAGTTTAGTCTTCTCATCATGAACATTTTAAAATATTTCTGTTTGAAGCTGC
 TCAAGCTTCCATCCATCTTCTTTTTAACCCTCACTGCCAAAATTTCTGGTTGAAATTATTGAAGAAGCGAGTA
 GGGAGAAAGAGAGAGAAACCAAAGCAAATGTTAATAAATGGCTGCCTGAAGTAGGTAGGAAGGAGACACCAGT
 GAACGAGTCAAAGAAGCAGAAACAAAAGGGGTGGGGGAGCCTGGATGTTGCCGGGTCTAGGGTCGGAAGGCTGGA
 AGGAGCACCTAGGCTGAGGATAGACCCGGGACTGGGAGCTGAGGGAGCTTTCTTGTGTTGGAGACTGGGATATCGT
 CGGTTCACTTGAGAGCTGTTTCTGTTACGAGCCAGAGGCTGCTGAGTCAGTGAGATTTCGGCACCAGTACTCTT
 TTCCCTCCAGTGACGTGCATTTATAGTGAGCATCAGATAAAACAAAAGAAAAGAAAGTTTAAATCAGTGTTTA
 TTTACACAGAAAGAGATAAGTGTTGCCATGTGTACATGGTACAGATTTCTTGGATTGCTTGCCAGGTGTTTATA
 GGTAATGGATTCAATTTATCAATTGACCAATATCAAATAAGAAAGACATTATTTGGGAAGAGCTCAGGAATGAAT
 TCTACACATATGTTTATCTTAATCATTAGGACTGTGGGCTTCGTGGTTGGTTTTTACGTTAAGGTGGGCACCCC
 TGTTCTCTTTCTTTTCAAATCAAGTCTCTGATTCCTAAGCGATCTAATTTGAGCTCAAAGTACGTAATGTTTGG
 ACTTCAGCACCTTCCAGAAATATATCACTCTGAAGTCTGCAAAGTACAGTTTCAAAGAAAACACTTTTTTTTACAG
 GAAACAGAGAGCTAAAAGTGATCTCATGTTACATTTCCAATTTTGGTTTCACTTAGTGCTTGGGATAAAAAATTACT
 TCTCTGAATGCAATATTGTTACTTTTTTTTTTCTCTGCTTACCTTCTTAAACACTGAGTTTACTCAGTCAAGTCT
 GAAACCATGCATCTCTTTCCAGAAAAATGAAAGCTGACAAGAAAGTTGAATGACTAATTTTTAAAAATATTGCAT
 TTAATAAATAATGAGGGTTTCTTGTACAGCTATATGATGAGAGATTGCTGAAGTATTTGTTAGATAGGATCTGG
 TGATACCTGCCCTAACTTTAAAAAATAGGATTCAGGGGACACCCAGTGGTTAAAAGGAGGAAATTCCTCCTGAGTG
 CAGTGTTCAGTTAGAATAAGTATTTCTTGTCTCAGCAGTCAAAATGTTGATACATTGCTGATTTTGTCTTTCTGA
 ATAAGCATATGTTAGCTGTATGTGTGTGCTTTAAGAAGAAAGGGACTTGGTTTTGACATTTATATGAATTTTCTT
 CAAGTATTTGAAAACCTTACAAATTATAAGGTTAGAAATGTTAGAAATGTGATTAACTTTTCTTTTAAATGAAGAC
 ACAATAATATTAATAAATGAGAAATAAGCAACC

Fig.3C

[illegible]

Fig.3D

ACTCCCAACTTCATCTCCTCAGCCCAGACAGACAGACAGCCACCGCATAATCCAACAGCTGGAAGCTCCGCGGGC
 AGTAGACGGTGGCAATCACAGGGCATCCCTGGTTTGTCTTCCCTCTCTGGGGGACCCCGTCCCTTGCTGACTGGTG
 TCCAGCATCTGGAAACCGCTGTAGCATATAGTGGTCTAGGGTTTTCTTTAGTTATTTAAGGCAGAAACACTCCTT
 CCTTTTTAATCTCAATTTACTTACATAAGACTGTCATCAGTGTAAGGGGAAACAAACCTGTGTTTGCCAGGAGG
 CTGTGGCCCCAGCTTACAAACCGAAGTTCCCTTCTTTACTGTGAAGGGATTTGAGTTGTGACAGTTTTCTCATTTA
 TACGCCTGGTCACTGACTTGCTTCTTTGTTTATTTAGCAATGCGCTTGAGTGTCTGACAGGTATTTTACTAATAT
 TGAAATGTGCCACAAAGGCTGTTGAAATCGCCTATCCCATATCCTGTCTGACACCCTGTAGTACCTCAGAAGGA
 GGCTTGCCACTCCAGAAACACTGAGGTCATGTTTATTCTGCGGATCTTGATGACGCTGGATCACCCCTCCACA
 CACACACTCACACACCGTGTGAGGCCTAGGGGAGCGAGGTCCAGAGGGGCGCGCATCTGTACAGCTGACGCACG
 AGCACCAGGACCTGCTGCTCTGGCTCTGGAAGGCCACACTCATCACGGAAGCCTTTCAGCATTGACCCCGAAGT
 GATCCCAGGTTCCCTCAGCTAGCAGTTATCGGGCTCCCCAGGCGGCGCTAGTGGTAAAGAACCAGCCTACAATG
 CAGGAGACTTAAGAGACATGAGTTCAGTCCCTGGGTGCGGAAGATGCCCTGGAGGAGGGCATGGCAACCCGCTTC
 AGTATTTCTTGGCCTGGAGAATCCCATGGCCAGAGGGGCTGGTGGGCTAGTCCATGGGGTCACAAAGAGTCAGAC
 ATGACTGAAGTGACTTACCGTGCACAGCAGTTATTGAGTAATCACAGTATCTCTGAGTGTGGGCTTAGATGCACT
 GTATATATAAATTTTCATATCATCTCCAGAGTTCCTGTAAGACCAGTCACAGGACATGTTCTCACCAGTCAGAATCA
 CTGGGTACAGTACTCACTGTATTGAAAAAATTAGTTCTGCTCTGAAGGCTGCTCTTGTTCAACCTAACAAAGCT
 TAAAAGCAAACCTCAAAGGATCAAATTTTTCCAGGTAATAGAATCATGACCAGTCTTAAAAAATAAATAAAT
 CACTTTCTTTTCAATGAAAAAATATTCTTCTATTATATAAATGGGCTTCCCTGGTGGCTCAGGTGATAGAAAT
 CCGCCTATGATGTGGGAGACCTGAGTTCAATCCCTGGGTTTGAAGATCCCTTGAAGGAGAGCATGGTAACCCAC
 TCCAGTATTTCTGCTGGAGAATCTCCATGGACAGAGGAGCCTGGAGGGCTGCACTCTATGGGGTTGCAAAGAGT
 CGGACATGACTGAGTGACTAAGCACATGTATTATATAGATAACAAAGCTCCAAAAAATTCTCAGCAGCCACTAA
 GACCCATTATGAGTTAAAATTAGATGCACAGGGGTGCCCTGGGTTGATGAGTCTGTCCCAGTGCCACTGGTTCA
 AGGCCTTGGGCCCTCTCTCTCCCCAGAGCAGCAATTTCTAACCTGTGAGCCACAGACTTCAGGCCACACTTTCATT
 TCATACAGTCTCTCAGTCTCTCTCTGAGATACAAATAGGATTATTTTCATTGCTCAGGTGAGAAAGTGAGTGTT
 TAATAGTGCTCATAAACTTGACCGAGGTCTCATAGTTGGCATTGTTGGTGCAGCCAGGATTTTCTGCCACATGTA
 TCAAAGAGCAGGACCATTAATCTAGAGGATTTGGCACTTTTCTGTGTAGAGCCATCAGGTCTCAGGCTGGCATGA
 GGCCTCACTTAAGGGGCTTTCTGAGATTTTAGAAGATAGTGACACAGCAGCTTGCCTGGCTCTCGGAATTTCT
 ATTGTTACTTTGAAAACTTGAATTCTTATTTTGAAGTCTGAGTTCTGATTCAATTAATAAAACTTTTGTATGA
 ACCACAACAGACATCCAAGAAATGCATGCCGTTAATTCCTTTTCATGTCTCTCTGCAACAGCGGACAGGATGGATG
 TAATCTGTGCGTTCTTTATCCATCTTGTCTGCTCAAGACGAGGTGCCGTTTTATGGATCTCCTGGCCGGCAGACT
 CTGAACCTCTTTTCGAGATGGGAGAGGCAGCTTCTAGCAGAAGGTGTGGATTTCATGGTCCATGGAAACAGGGTCTA
 TGTCTTCTCAGGAGTCTGCACCTGAGCAGGAAAAATCACACAAATTCCTATAAATGGACCCGACCTGATGACC
 TGGAGTCAGAGTCCAGTAGAGGCAGATCTGGCATGGATGGTGACACTTCTGTCAACTGTTTTTAAGACTGAGGCC
 CTATGACAGGCTGGCCTGCCAGGAAGGTGACTCAGCCTTAGACGAGGGGCAGACTGTTCTGACCATTTCCTTT
 CTATAAACTTAAGTAAAAACAGAGACGTGCTCGTAAATGTTAGAAGGAATGCGGTTGAATGGCAAGGGAGAATCT
 GTTCAGTGGCCTCGAAGAAAAATTAGTTGAACCTCAAATTTGAATTGTCTGTTACCGGTGTTCTCACGTAACAG
 CTTAATCGTGTAGAGACCAGAGCGTGTCTGTGTGTGAAGACAAAGACAGAATCCTACAGGAGTGAGCTTGTTCG
 AGTTTTCCCTTACTGTGAGTAAATGAATTCATTCTATCCAGGGGTAGACCCAGCCCTGTGAGAGCTGTGAATGT
 GCAGACCTAGCTGGGTGAGAAATGTGTCCAGGCCAGACTGCGATTTCTAAGGACTGATGCTCATGGTTGGGGGA
 GAGAGGGGTGAGGATGAGGAAGACTGCACAAGTTAAGTTACGCTGACGTGAAGCCAAGGGACTTTGAGAACTT
 TTCCCGTGGAGACAAGAAAGACTCCCAAGGACTAGATGGTAGGTACCCAGAGAAGGCAGGAAGCCAGCCAGGCC
 TGCAAGAACTGGTGGTCCCAGCCACCTGACTGTGGGTCTGTGTCATGAACCCACCGCTTCGGACTCAGCATCTTG
 GTTCAGCATCCTGTATCACCTTGTTCCATGCTTGTGGTGCCACGTTGCTTCAGCGGCAGCAACTTCCCTCTGT
 TCCTCCAGCTAGCTAAACCCAGCCAGGACCTTATCCGTGCGGCCCTATCCACCTCGAGTGCCTCCCTTCCCC
 GTGCCGCACTTTATGCTTGATTTATTTATACCAGATCTCAGTCAAATGTCACCTTCTTGTCTATCTGTCTAAGG
 AACCCTTAGCCAGTTACTCTCTGTACATGACCCTGTTTTATTATCTTTACAACACTCTTCACTCTCTGAAACCA
 TCTTGATTTTTATTGACTTATTTAGAGTCCAAGTCTCTTACTGGAATGTAAGCAGTACAGGAGTAAGACCTGTT
 TTGTTCAATTGGTCATTATATACCACTGCCTAGAGGAGGACCAGGCTTCTAGCCAGGGTTTCAGTAAATGTCCAG
 TGAGTAAAGGGATGGTGGTGATGGGTGGTTTAGTCGCTAAGTTGTATCCAACACTTGCGACCCCATGGACTGTAG
 CCCGCCAGGCTCCTCTGTCCGTGGGCTTCTCCAGGCAAGAATACTGGAGTGGGTGCTGTTTCTCTCCAGAGG
 GGCTCTTCCGTGACCCAGGAATCGAACCTGGGTCTCTGTCATGTCAGACAGATCTTTACCAAAATGAGCTACAAGG
 AAGCCCGGAGTAAAGGGATGCATGAAGTCAAATACAGCTGAAGAGAGAACCAGTCAATGTATGTGTGGTTACAG
 TGAAAAAATGTGTAACAATACAAATCTGAGAATAAGGGGTTTGTGCTGTGGTATATGTGTGCATATCTAAGCTT
 TTGTGCTGTGTTTCAAGTGCCTGACATTGTTAAGAGCTCTATAAATATTACATAATGGAAGATCATTTTTTACAA
 ATTCTACTCCTGTTTTCAGCTAACTAAATTTGGTCTTTGTAAACTAGATAATTCAGTTGTCTCAAAGAACAAGC
 AAGGTCATACCCACAGGAAAGCTGAGATACCCA

Fig.4A

TCCTCAGTTGAGTTAAAAAGATACTGAATTTAATAATAATTTTAAAAATTACCTTTTAAAGTAAACGTTAAGTTTAT
 ACACAGGTAAAAATAAAATCATTATGTATGTTAAATTTATATTAAAAAAAAGGCATGTGTATCAAGTTATTTTT
 AACTTAATGGTTTTGTAACTCAGCCAAAAGTTTTTAATTTTTGTGGGCAAATCTTAGCTAACGGGTGAAGTAGAA
 TTCATTTACAGATAAGTCTGGGTCTGGTATGTGCATTCTGATGTGAGAATTGAAGCTGCAGTTAGTGCAGGATTG
 CAACGTGACCTTCCCTGATTACAGCTGGGACCACAAGTTAATAGCTGGAGCAGTGGACAGTGATTGGACTTTGCC
 AGGCTTCAGTCTCCTCACTGGAAATGAGGAGGTTGAAATGGAAAATGCAAGTATAATGAATAACCCCAATGCTAT
 ATGCTTGTCTTCTTTCTTATATCCAACCCCTTACACGGAATCCATGTTATGCACTATGATTATCAGAAGCTGCT
 TGGCTCCTCTGTCCATGGTATTCTCCAGGCAAAAATATTGGAATTTGGTAGCCATTCTCTTCTCCAGGGGATCGT
 CTCCACCCAGGGATCAAACCCAGGTCTCCTGCATTGCAGGCAGATTCTTTACCATCTGAGCCACCAGGGAGGCC
 TCTTTAATTCTATCACATTTATCTAGGATTTTTTCTTCTCTGGTTGATGGTAGTGCATGAATGCTTGGGAGAG
 ACTACTTAGTTTACAGTTGATACTTTGGAGTCTGGGGCCCCCTGCAGAGGCCCTCTGAAGACCCTCAGACACCCATT
 AAAACGTGTCCACCCCTGTTGATTACAGTGGGCACCCTGCAGGGCTCTGAGAATGGCCAACCGAGTGGACACATGACC
 ATTTGGACTTTGAGGACAGCAAGCTGAGTTTCATCTTTTCTACTTGCTAAGACAGTATTTGCCCTGAAGGGAAACA
 AGGCTTCAGCTCAAACTTCAGCTCAAATTTGTGGCAGACAGTTTCAAAAATTAAGCAATATAAATTTTGTGAAT
 ATCTTCTAAATGTTAAAGAATTCAACCAAATTTCTCTCAATTATATATTTTTCAAAAAGTGTCCAAAGTTGTGTG
 TAAGTGGAATGTAACCTTTACAGTGTATATATCACCTAATCGGCATGATAACCCCTCCTATGTATTCAATGC
 TTATAGTACATGCTTCATGTTTTATAGATACGGATCATAATCTCGCCATGTTGCATTATAAACTTTTATATGCTT
 GCTTCTCTGAAATACATAAGTAAATTCATAAGTAAAAGATACAAGATATCAAGTCCTTGTGGAGAAAGCAACTG
 TATATTTTATCTTCCCCCTTGTATTATATCCATTTGTCACTTAAGTAGTTATATGCCACTGCCGCTGCTAAGTC
 GCTTCAGTCGTGTCAAACCTCTGTGCGACCCCATAGACGGCAGCCACCAGGCTCCCCTGTCACTGGGATTCTCCA
 GGCAAGAACACTGGAGTGGGTGGCATTTCCTTCTCCAATGCATGAAAGTAAAAAGTGAAGTGAAGTCACTCAG
 TCATGTCTGACTCTTAGCAACCCCATGGACTGCAGCCACCAGGCTCCTCTGTCCATGGAGTTTTCCCGTCAAGA
 GAACTGGAGTGGGGTGCCATTGCCTTCTCCAAGTAGTTATATAAGAAAATAAAAAAAAAAAGAACTCTTAACCC
 CACGATCAGAAAATAACCCACATAAACTTCAGAGATCATACTATGCATGTTACTTTATAATGTGGCTTTTCTTT
 ATGTAAAATATAAGTTAAGTATTTTCAATTAATGCAGTGTCTTCTTAAACATGATCTTTAAACAACCACATGACA
 TTTTCATCCTGTAGCTGTTTATAGCATGACCTCTTATGTGGGAGGCAGTGTAGGTAGGTGAGAGAATAAGGTCTGA
 AATAGACATTTTGTCTTTTACAAGCTGTGACCACGGCCAAGTTATATACTCACCCCATTCCTCCATCTCCTTACC
 CATAAAATGGGGATAACAATAGTAACAGCCTTATATTGTGTAATTTTAAAGATTAACTAACTTGTATACATGATA
 CATGTTACATATTTATCATGTTATATGATGTTATACACGTTACACATGATGATGCTTAGAGCACTTACAATAGTA
 CTGTAATATATTGTATTATTAGAGCATTATTTATTATTATCTATAATTTAGTTATTCTGATGGAACAAATTACCT
 GATACACATTTGAGTAAGAACTCAATGTGCAGTGAAAATGAAGTGAAGTCCGTCAGTTGTGTCTGACTCTTTGAG
 ACCCATGGACTATACAGTCCATGGGATCTCCAGGCCAGAATACTGGAGTGGGTAGCCTTTCCCTTCTCCAGGGA
 TCTTCCCAGCCCAGGAATCAAACCTGGGTCTCCTGCATTGCAGGCAGATTCTTTACCAACTGAGCTATCAGGGAA
 GCCCTTATGGAGTGGGTAGCCTTTCCCTTCTCCAGGGGATCTTCCAGCCCAGGAATCGAACCTAGGTCTCCACA
 TTGCATTTCGATTCTTTACTGGCTGAGCCACAAGGGAAGCCCAAGAATACTGAAGTGGGTAGCCTATCCCTTCTC
 CAGCGGATCTTCTGACCCAGGCATCGAACTGGGGTCTCCTGCATTGTAGGCAGATTCTTTACCAACTGAGCTAT
 CAGGGAAGCCCAAGAACTCAAATAATTTATTATTATGCATCATTGTATTCTAACTTCAAAAAGCAAAAACAGTA
 TAGGTAATAAAGGACTCAATAAACATTCAATATTACCTATATTGTGTTTTGCTTTTTGAAGTCAGAATACCGTGA
 TGAATATCTGTGAAAAAAAAGGCTTTTGCTCTTGTACTATTATTTCCCTAAGATAAGTTTCTTAGACATAGGA
 TTGCATCCTTGCTGCTGAGTCACCTCTGCAGCCTGTGCTAACGCCCTGCTCTGCAGCGCTCAGGGTAAGAGAGAA
 AGAACGCACCCCTGTCCAGTGAGTGATGCTGTTCTGGGAATGCATGGCAGTGGAAAATGGCACCTGGCCATGTGA
 GGAGTTCCGGCAGTTTTCCAAAACATTGACTATTGAACAATTGCAATTTACTCAATGCACTCTTTCTTTTAATA
 TTTATATAGTCTTCAGGCTAAGATATAACAAATTA AAAA ACTGTTTCACTTTTTGTGTGCTTAGCATTGAGTATA
 TTAGTCTGAATTTTGCTTGAAAAGATACTTGATTGACAGAAAAATTTGCGGTTGACCACACTGTTAAGGTATCT
 TCCCCCATCGAAGCTAATATTTTCTTTTTTTTTTTTTTTTCTAGAAAAAGCAACAAGCCTTGGGAGCAGTCAGCTCT
CCAGAACTCAGGCTGATGAAAAAGCCCTGGTCCCAAGGTCAGTCATTTCTGACATGAAAAGACACCTACTTT
 ACAGATGTATCACTAGTGGGGGGTGGGACTTGTAGGGAAAAACAGACATGTCAAATATGTCACTGAAAATGGAACA
 AAGTATCATTTCTGGGCAGGTGGGGGAGGTACAAAAGCAACAGGCATAACATTTCTTAATAAGGATTGATGATTTA
 AGACCTAATAAAGACATTTGTGTTATTATCTTACTGGCATCTAATAGTAAGAGTCCAAAAATATTAATCATTAT
 TCTAAATGTTTATAAATAATATTTGATTTTTGCACATATTAGTAGAATAATAGTTTTTTCTTTCTTCAGAA
 AATAATTCAAATTATCTTTATGAGTATTAAGCATAGCAGGTTCTTGCTAAATGTGCTGTGTAATTATCTTATGTT
 ATCTGTCACAACTTATAATTATTTCTTTTACTTTGAGGAAACCTGGGGTCAAGAGGTTAAGTAACCTTGA
 CAAGATCACAAAGCCAGTAAGATGGAGGTAGGGTTCAAACCCAGGCTTCCAGTTCACAAAAGATCCTGGACCTTT
 TAAAGTCTGACCAAGCCTTCTTCGGCAATATAACACTACCAAGTTTTCAGCTTATAAAAACCACTGAAATTTCA
 TCTTTCAAGATAAGCTGTAACCTCAATTTAAGCTAATAGTTTTTAATATATAAAGTAAAAGCTATCTATGCTCCT
 TCCACTTCAGCTGGTGATCTTCCAGTATGTCCA

Fig.4B

1

CTTGAATATATTTTCGGTTACTGTTTATCGCCTTTTGCCCTCCAGGGAGGAGTCTGTTTTGCCCTCCAGTTCTGATCT
GCTAAGTTGCTTCAGTTGTGTCCGACTCTGTGCGACCCCATAGACAGCAGCCACCAGGCTCCCCATCCCTGGG
ATTCTCCAGGCAAGAACACTGGAGTAGGTTGCCATTTCTCTTCTCCAATGCATGGAAGTGAAAAGTGAAAGTGATG
TCACTCAGTCATGTCCGACTCTTAGAGACCCCATGGACTGCAGCCTTCCAGGCTCCTTCGCCCATGGGATTTTTCC
AGGCAAGAGTACTGGAGTGGGATGCCATTGCCTTCTCCGAGTTCTGATCTACTTGAGTCTATTTCCCAGAGCTGT
GCCTACTCTTCTGTCTGGTATAATACCTTAGAAAATAATGCCCTGGCAATTTTATGAGCGCCTAGAAATGTGCGGAA
GTTTTACCCACATATCTCAAAATACATGCAGCCCTGCAAAATAGGATTATTAGGCTCATTTTATATATTTGGGAA
AAGTTAAGACCCAAAGTCAAGAGCTAGTAATGGGGAAAGCTGAGTTTGAACAAAATTTCTTAAGTTTCCAAGT
CGATGTTTCTTCTGAAAATAATACAAAACAAAATTAGAATGATATTTTGTGTGTGTGTTCAGTCGCTCAGTTTA
TCTGACTCTTTGCGACCCCATGGACGGCAGCATGCCAGGCCCTCCCTGTCCATCACCACCTCTGGAGCTTACTT
AACTCATGTCCATTGAGTCAATAATGCCATGCAACCATCTCATCTCTGTGCTGCCCTTCTCCCGCATCTTTCC
AGCATCAGGGTCTTTTTCCAATGAGTTGGCTCTTCGAGTCAGGTGCCAAAGTATTGGAGCTTCATTTTCAGCATC
AGTCCTTCCAATGAATATTCAGGACTGATTTTCTTTAGGATGGACTGGTTGGATCTCCTTGAGTCCAAGGGACT
CTCAAGAGTCTTCTCCAACACCACAGTTCAAAGCATCAATTTCTTGGTGCTCAGCTTTCTTTATAGTACAACCT
TCATATCCATACATGACCACTGGAAAAACCATAGCTTTGACTAGATGAGCCTTTGTTGGCAAAGTAGTGTCTCTG
CTCTTTAATACGCTGTCTAGGTTTGTGAGCTTTTCTTCAAGGACGAAGTGTCTTTTAATTTTCATGCGCTGCAGTC
ACCATCCATAGTGATTTTGGAGACCAAGAAAATTTGTCAGTGTTCATTGTTTCCCCATCTGTATGCCATGAAGT
GATGGAACTGGATGCCTTGCTCTTAGTTTTTTGAATGTTGAGTTCTAAGCCATTCTCCACTTTCAGTTTCATCAA
GAGGCTCTTTGGTTCCTCTTTGCTTTCTGCCATAAGGGTGGTGTCTATCTGCATATCTGAGGTTATTGAGCACTTA
ATTACCAGGCAGCAATAGACTTAGAAGGTCTTTATTTCTATGGAAGAAAAATCAAGCAGAACTAAAGATCAGA
GTGTTTAAGTAGCTTGCCCAAGATTATTTACTATGTAGCACACCAAAATTTGAACCCATTGGAGTGATTACAA
AGGTTTTTACCACTGGACAACACCATGTACTGAACCTTTCATGGGAGTGTTACTCTCAGTTCAGTTCAGTTGCTCA
GTCGTGTCCAACCTTTTGGCAGCCCATGAATCAGACACGCCAGGCCCTCCCTGTCCATCACCACCTCCACAGAGTT
CACTCAAACCTCATGCTTCATCGAGTCAGGGATGCCATCCAGCCATCTCATCTCTGTAGTCCCCCTCTCTCTTGC
CCCCATCCCTCCCAGCATCAGAGTCTTTTCCAATGAGTCAACTTTTCGCATGAGGTGGCCAAAGTACTGGAGTT
TCAGCTTTAGCATCATTTCTTTCAAGGAAATCCAGGGCTGATCTCTCTCAGAATGGACTGGTTGGATCTCCTTG
CAGTCCAAGGGACTCTCAAGAGTCTTCTCCAACACCACAGTTCAAAGCATCAATTTCTCGGTGCTCAGCTTTCT
TCACAGTCCAACCTCTCACATCCATACATGACCACAGGAAAAACCATAGCCTTGACTAGACAGACATTGTTGGCAA
AGTAATGTCTCTGCTTTTTCAATATGCTATCTAGGTTGGTCATAACTTTCTTCCAAGGAGTAAGCATCTTTAAT
TTCATGGCTGCAATCACCATCTGCAGTGATTTTGAACCCAAAAATAAAATCTGATACTGTTTCCACTGTTTTCC
CCATCTATTTTCATGAAGTGGTAGGACCAGATGCCATGATCTTCATTTTCTGAATGTTGAGTTTTAAGCCAAC
TTTTCACTCTCCTCTTTCACTTTTCATCATGAGGCTTTTTGGTTCCCTCTTCACTTTCTGCCATAAGGGTGGTGTCA
TCTGCATATCTGAAGTTATTGATAATTTCTCCCGCAACTCTTGATTCCAGCTTGCTGCTTCTCCAGCCCACTGTTTT
CTCATGATGTACTCTGTATATAAGTTAAAAAACAGGGTGACAAATACAGCCTTGATGTACTCTTTCTTATTT
GGAGCTTGCTATTGCTCCATGTCCAGTTCTAACTGTTGCTTCTTAAGCTGCATATAGGTTTCTCAAGAACGAGG
TCAGGTGGTCTGGTATTCCCATCTCTTTCAGAATTTAACACAGTTTATTGTGATCCACACAGTCAAAGGCTTTGA
CATAGTCAATAAAGTAGAAATAGGTGTTTTTCTGGAACCTTGTCTTTTTCAGTGATCCAGCAGATGTTGGCAATT
TGGTCTCTGGTTCCTCTGCCTTTTTCTAAAAACAGCTTGAACATCTGGAAGTTCACAATTCAGTATTGCTGAAAC
CTGGCTTGGAGAATTTTGGAGCTTACTTTACTAGTGTGTAAGATGAGTGCAATTGTGCGGTAGTTTGGAGCATTCT
TTGGCATTGCCTTTCTTTGGGATTGGAATGAAAACCTGACCTTTTCCAGTCTGTGGCCACTGCTGAGTTTTCCAG
ATTTGCTGGCATATTGAGTGCAGCACTTTTCGAGCATCATCTTTCAGGATTGGAATAGCTAAACTGGAATGCCA
TCACCTCCACTAGCTTTGTTCTGTAGTGATGCTTTTCAAGGCCCATCTGACTTCACATTCAGGATGTCTGGCTCT
AGGTGAGTGATCAGACCATCTGTATTATCTTTGCTGCTGAGATCTCTAGGTGAGATAAAACCAAGGTAAAGAGTT
AAAAGTAAGTATGGTTAAAGGGACTTCCACCATGACTCATATGTTAAAGAATCTGCCTGCAATGCAAAAGACCC
AGCTTCCATCCCTGGGTGCGGAAGATCCCCTAGAGAAGGGAATGGCAACTTACTCCAGTATTCTTGCCTGGAGAA
TTCCATTGACAGAGAAACCCGGCAGGCTACAGTTTATGAGGTTGCAAAGAGTCAGCCACGATTGAGTGACTAACA
CTACTACACTGCTGTGGTTAAAGGAGGTGGAGGACAAAGCAGGGTGTAGGTTAACTAGGCTCAAACCTCTGTGT
TCTTTTCTTTCTAGCTAACTACTTCTCTGCGTCAGCCAGCAAGCTTCCAGTATGAATCCCACAGAAGCCAAG
GTATGGAGAGTCTTCAAGGGTCAACTTGGGTGAAAGCCTTCATTTTGTAAGCAAAATGAATAGAGACTTTGACA
GATAGCTTGTTGTCTCTAGGAGACTTACACTATGGAAGTAAATATTTCTTCTTAATTAATGAGGGTCTTTGAAA
ACTCAGCTTTCTTTTTCAGTATTTGCCTCAAACACCATCTATATACTTATACAGGATTTTAGAATATCTAGAAAAG
TGAGTCCGTGAGAGTTCTTTTGGTTCTAGCCTTGAGTTTTTTAAATTGTAAAAAGATATCTTAAGTACATTTA
TCAAGATCTAGATAGAAATATAAATTTCTTAATTTCTTACTAAACGGTTTTATTATCCTAACCATTTAAACGAA
ATCCAAGAATGAGAAAGTTTACACTAAATGAACATAGAAAGAAGTGACACAGACTCTGGTTATTTTCAGATGCTTT
CTGAAGTTTTTTTCTGACATTTGAGTCATGTGGGTCAACCATCTTATTGATAAACTGTGGTTGAGCTCTTGGGGTG
TATTCTGATAATTTCTTGGAAATTTAGAATAGA

Fig.5A

AATAAAAGAATGGCCTGCAGTGGGATTGCCACATAGAATAATAAGCCTACTTCTAAAGAGTTGACTGAGCGACCT
 AAACCCCTCTCACCGACTTACCATGACTCGAAGTCTGAGTTTCTCGGGAATGTTCTAGGTTGGAAGATTAAAGAGGAA
 GCATTTTTCAAGATGCTCTCAACTAATCTTATGAGCTAATATTCTTCAGTGCCTGTAGGAGCTACCCAGCTCC
 ATTATTATCAAATTGCCCTTGTTTAAATACCAACACACTTTTGCTTTGTGTTTCCCTTGTTACTTTACTGTCC
 CAAGTAGAATCCTATGCCATAATTTAATACTGGCACCATAATGTAGTACTGTTTAGTAATCCTGTGGCATCATCT
 TGAACGTGATTTAAAGGAGATTGAACTATCTTGAACACGCTGTCAAACCTAAGCCCTTATCCAATGTCATTT
 AGCACTCATTTTTATTTTTAAATGTAAATATGTTTTATTTAGTAATTAATTTCTTAATTTTTTCAAAATACA
 TGTTATCTGATTCTGTGATTAAACCACCTATTAGGTGTCACTACTTATTTATACCAAAGTGAAGTGAAGTCGCTCAG
 TCGTGTCTGACTCTTTGTGACCCCATAGACTGTAGCCTACCAGGCTCCTCTGTCCATGGGATTTTCCAGGCAATA
 GTACTGGAGTGGATTGCCATTTCTTCTCCAGGCTGATCTTTAAATGATCTTAAGACTATAGAAAAAACATTT
 AGCCCCTGCCCCGAAAAATTCTCATTGAACCTAGAGTCCGTGAAAATTAAGTAAGGAGATAAAAAAGGAGAGCC
 AGAAATATGAGATTTCTTTTTCTTATTATTTTGGTTTACATTTGACATGATGCCTCATTATAATTCATCCCTGC
 TGACACTCAGCTTAAGGGCCGATGCTTATTTCAAGTGTGTCCTGTTTTCCAGAGTAAATGTGAGCAAAATGCTTCA
 GCCATTGCGAGCGTGAATCATGAGAATCTTGGCACTCCTATAGACGTGGACTGAGGGGAGAGAGAGAGGAAAGT
 AGGAGAGGGAAGAAAGGGAGGAACAAACTAAGATGCGCCAGTCGATGACTCAGAACCCTCTTGGCCTTAATTC
 TCTGTCTTGGTGGATTTCAGTCATACCTGTTCTCATTAATCATCTTATCTTCGTGTGCTTCTTTAATTTTTTTTA
 ATTTTAATTAATTTTATTTAATTTAATTTGTATTTTATATTGTAGTAGGCTTCCAGGTGACTCAGTGGTAAAGA
 ATCTGCCTCCCAATGCAGGAGCTGTAGGCTTGATCCTTAGGTGAGGAAGATGCCTGGAGAAGGAGATGGCAACC
 CACTCCAGTATCTTGCCTGGATAATCCCATAAAGAGGGATAATCCACAGAGGAGCCTGGCGGGCTGCAGTCCG
 TGGGGTCACAAAGAGTCAGACACAACCTGAGCGACCCGAGCAGCCACACCCACAGTTGATTTACAGTGTATGTTAG
 TTTATGGTGTGCATACTTTTTCCATTTATTGCGAATATTGAGTAGAGTTCCCTGTACTATACAGAAGATCCTTG
 TTGGTTACTATTTTAAATACTAGTGTGTATATGTCAATCCCAACTCCAGTTTATCCCCCTCTTCCCCCTT
 TAATAACCATAAAGCTTCTTTCTGTTTCTGTTTTGTAAATAAGTTCTTTTGTATCATTTTTTTAAGATTCCCAT
 GTAAGAGATACCATGCTATTTGTCTTCTGTGTCTGAATTAGTTCAGTTAGCCAGTAATCTCCAGGGCTATCCA
 TGCTGCCACAAATGGCATCATTTTATTCTTTTTAATGGCTGAGTAATATTCATATATATATATATATTTAATCC
 ATTCATCTGTTGATAGACTTATAGGTTGCTTCCATGTGTTGGCTATTGTAAACAGTGCTTCAATGAACATTGGAA
 TACACGTATCCTTCTGAGCAATGGTTTTCTCTGGACCTGTTGCCAGGAGTGGGATGACTTGATCTTATGGTGGTT
 CTTTGCTTAGTTTTCTAAGGAACCTCCATCTGTTCTCCATCGTGGCTGTACCAATTTAATTTCCACCAACAG
 TATGGGAGGGCTCCCTATCTCACACCCTCTCTGGCATTAGTGTGTTGTAGACTTTTTGATGATGGCCATTCTGAC
 TGGAGTTAGGTGATATCTCATTTTAGTTTTCTTTGATTTCTCTAATAATTAGTGATGTTGCACATCTTTTGAT
 GTTCTCTTGGCCCATTTGTGTGTCTTCTTTAGAGAAATGGCTATTTAGGTCTTCCATTTTATACTTTTTTTT
 TTTAATATTGAGTTGTGTGAGCTGTTTGTAAATTTTGGATATTAATCCCTTGTGAGTCACATCATTAGCAACA
 TTTTCTCCTATTCTGTAGGTTGTCTTTTCACTTTGTTTGTGGTTTTTTTTTAATGTGAGAAAGCTTTTAATTAGGC
 CCCATTTGTTTATTTTCACTTTTACTCCATTTATCTGGGAGATGGATCCAAAAATATATTGCTGTGATTTATGT
 TGAAATGTGTTCTGCCTATGTTTTCTCTGAGAGTTTATACAAGTATCTGGTCTGTATTTGGGTCTTCAGCCC
 ATTTTGAGTTAATTTTTGTATGAAATTACAGAAATTTCTAGCTTAATCTTTTACACGTAGCTGTCCAATTGTC
 CTGGTACCCTTTTGGAGGAGACTATCTTTCTCCATGGTATAGTCTTGCCCTCTTTGTTGAGATTAATTGACCA
 TAGGTACATGAGTTTCTTCCAGAATTTCTATCCTATTCTGTTGATTTGTATTTCTCTTTTATATATTTCTGTT
 TTTATAGTTATTTGGTGATTGTAGCTTTGTAGTATAGACTGAAGTCTGGGGCCTGATTCCTCCAGCTCCATTTT
 TCTTTTCTAGATTGCTGTGGCTATTTGGGTCTTTTGTCTTCTCATAACAGTTTAAATTTTTTGTCTAGGT
 CTGTGAAAAATGCCGTTATTTTGTAGAGATTGCATGAGTCTGTAGATGGCCTTGGGTAGTTGAGTCATTTTGAC
 AATATTGATTCTTCCAATGCAAGAACGTGGTGTATCTTCCCATTTGTTTATGTATCTCAGTTTCTTTTATGAG
 TGTCTTACAGTTTTTGAAGTACAGATCTTCTGCCTCGTGAAGTAGGTTTATTCAGGTATTTTATCTTTTTCGA
 TGTGATGGTAAATGGGATTGTTTGCTTAATTTCTCCTGATCTTTTATTGTTGATGTCTTGAAATGCAACAGTTTT
 CTCTGTGCTGATTTGTATCCTGCAACTTTACTAGATTCAATGCTGAGCTCTAGTAGTTTTCTGGAAGCATCTTT
 TGGATTTTCTACATACAGTATTGTGTCACTATGCCAAAGCCTTTGACTGTGTGGATCACAATAAACTATGGAAAA
 TTCTGAAAGACATGGGAATACCAGACCCTGACCTTCTCTTGAGAAACCTATATGTAGGTGAGGAAGCAACAG
 TTGAACCTGGACATGGAAACAACAGACTGGTTGTTCCATGTCCAAATAGAAAAAGGAGTACGTCAAGGCTGTATAT
 TGTCACCCCTGCTTATTTAACTTATACGCAGAGTACATCATGAGAAATGCTGGGCTGGAAGAAGCACAAGCTGGAA
 TCAAGATTGCCAGGAGAAATATCAATAACCTCAGATATCCAGATGACACCCTTATGGCAGAAAGTGAAGAGC
 AACTAAAAAGCCTCTTGATGAAAGTGAAAGAGTAGAGTGAAAAAGTTGGCTTAAAGCTCAACATTGAAAAATGA
 AGATCATGGCATCCAGTCCCATCACTTATGGGAAATAGATGGGGAACAATGGAAACAGTGTGAGCTTTATTT
 TGGGGGGCTCTAAATCACTGCAGATGGTGATTGCAGCCATGAAGTTAAAGACACTCACTCCTTGAAGGAAAG
 TTATGATCAGCCTAGATAGCATATTCAAAGCAGAGACATTACTTTGCCAACAAAGGTCCTGTAGTCAAGGCTA
 TGGTTTTTCTGTGGTCATGTATGGATGTGAGAGTTGACTGTGAAGAAGGCTGAGCACCGAAGAATTGATGCTT
 TTGAGCTGTGGTGTGGAGAAGACTCTTGAGAG

Fig.5B

TCCCTTGGACTGCAAGGAGATCCAACCAGTCCGTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGA
 TGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTCATGCGAAGAGTGGACTCATTGGAAAAGACTCTGATGCTAG
 GAGGGATTGGGGGCAAGAGGAGAAGGGGATGACAGAGGATGAGATGGCTGCACGGCATCACTGACTCGATGGACG
 TGGGTCTCAGTGAGCTCCGGGAGTTGGTAATGGACAGGGTGGCCTGGCGTGCTGTGATTGATGGGGTCGCAAAGA
 GTCGGACATGACTGAGTGACTGATCTGATCTGATCTGATTGAAACAAAAATGTCTAGATGTGAGACAGCTACTAAG
 GTCAGGAACACACACTAACAGAAATACAATAATGGACTCTCCTCATCAGAATCCAGGGGTGCAGTGGGAATCACA
 GCTCATGGATATGGCAGGAATGTTTTTTTTTAATTAATTTTTTTTATTGAAGGATAATTGCTTTACAGAAATTTGC
 TGTCTTCTGTCAAACCTCAACATGAATCAGCCATAGGTATACATATATCCCAATCAAGCAGCTGAACCTTTCTTT
 CAGCTTAAATCCTTTTTACAATTGAACAACCAGTCCATGCTAAAGGAAATCAATCCTGAATGCTCATTGGAAGGA
 CTGATTTGAAAGCTGAAACTCCAATACATTGGCCACCCAATGGGAAGAACTGACTCATTGAAAAGACCCCTGATGC
 TGGGAAAGATTGAGGGCAGGAGGAGAAGGGGACGACAGAGGATGAGATGGTTGGATGGCATCACCACTGAACGG
 GCATGAGTTTGAATAAACTTCGGGAGTTGGTGAGGGACAGGGAGGTCTCGTGTGCTGTAGTCCATGGGGTCGCAA
 AGAGTCAGACACAACCTGAGCGACTGAACCTGAACCTGAAATGTATGAAATGCTCATCTCAAGCAATTTTGAA
 TGTAAGCAATTTTGCAGCCTCGTAATTGGGGTAGAGCAGTTTCTTCCCTGTGATACCTCTTTTGTGCAGCATAG
 TCGTACCCACACAAAGCCACAAACACCTCCTGTAACAAAAATGTGACTAGTGACCTCCCTCGTGCTAACACTCT
 TCTGGGCCCTGAGGAAAGATCAGTGAGCAAAAGAGACAAAGTCACCACTGCTCTATTAGTCAGACTTGGCTTAAC
 AGGTCAGCTTTACTGTGTGTGATTCCAGAGTCAGGGCTTCAGAGTTGATCTAACCTGTCTGTGATTCTGGCTCT
 GTGTGACCATAAACAAGTGTTTATATATCTCTTTATAGACACTATATATGTGTATATACATATATATATATA
 CTATATATGATATATATCATATATATGTTATATATATATATTATATATATATACCATATATTATATATGCTATATATA
 TATATATACCATACATTGTATATACTACATATATAATTTATTATATATGTGCTATATATAATACTATATATTATA
 CTACTATATATAATATATAACTATATATATATACACTATATATGTATATAAATATATATATTTATCACATCC
 AATCTATGGTGATTCTCTGGCTCTGCATGATCATGAAGACGTGTGTATATAGAAAGTTTATAGTAGACATGTACTGA
 ATAAAACACTGCATAAATCGTAACCATTATATTTATGAGCATTGTTGTAAGAAAAACCCCTTCTTCAAAGATGTTGC
 CATTTAAAGGAGAGCTGCTAAGTCGCTTCAGTCATGTCCGACTCTATGTGACCCACAGACAGCAGCCACCAGG
 CTCCCCGTCCCTGGGATTCTCCAGGCAAGAACACTGGAGTGGGTGGCCATTCTCTTCTCCAATGCATGAAAGTGA
 AAAGTGAAAGTGAAGTCGCTCAGTCGTGTCCGACTCTGTGGGACCCCATAGACAGCAGCCACCAGGCTCCCCGT
 CCCTGGGATTCTCCAGGCAAGAACACTGGAGTGGGTGGCCATTCTCTTCTCCAATGCATGAAAGTGAAGTGA
 AGTGAAGTCGCTCAGTCGTGTCCGACTCTGTGGGACCCCATAGACAGCAGCCACCAGGCTCCTCGTCCCTGGGA
 TTCTCCAGGCAAGAACACTGGAGTGGATTGCCATTTCTTCTCCAATGCATGAAAGTGAAGTGAAGTGAAGT
 CACTCAGTCATGTCTGACTCTTAGCGACCCCATGGACTGCAGCCTACCAGGCTCCTCCGTCATGGGATTTTCCA
 GGCAAGTGTAAGTGGGAGTGGGATGGCATTGCCCTTCTCCGTTAAAGGAGAGCAATATCCCCAAATACTAATTTAAG
 AGCTGAAATTTTACCCAGAGAGAGAGAGAAACGGACGGGTTTGTGAGGGCGGGAATGCGGAGGTGCCACCAC
 AGAGCACGGGGTGTGTGCTGGGTGGCCCTGTTCTGAGAGGGCAAGACTGCCAAGACCTGTGAGGAAGTGCCTTCA
 GACAGGTTGTGAAAGGGAGTGTGAAAGAGTGGAAAGAGAAAACTTGAGCAAGGAGATCCAAAGCCTCATAAAGA
 CAGCCTAAATGTGAACCTCAATCACACAGAGAGAAGGAAGCCCTTCTTCTAAGAATGTTGTGGAGACACAACCT
 GAGTAAAGGTGGTTAACTCTTTCTAGCCTAGATGCTCAAATCAGAGTTCAGATTTGCAAGAGGCATTTCTGGTA
 CAGGAAGAACTGCGAAAGGTCTGAATCATTTCTGAGCTGGGTCTTTATGCCTGTGTCCAGCATTCTGCCTCCTG
 GGTGTATTTGTTGGCTAAGTCTGCCATAACAACAGCTAGGGCTGGCTTAGAACAACTGAAATGTATTCTGTCTTA
 GTTCTGGAGGCTGAAGCCCAAAGCGGCGTGTGGCAGGGTCTCAGTCCTTTTGCAGATGCCTGTCTCAGGTCTC
 TCTCCAGCTTCTGGTAGTTTCTCAGCTTTCAGCAGCATAACTCCAGTCTGCACACAATGTCTACACTGTGTGTGC
 ATCGGTACCCCAATTTTTCTTTTATGACATCCATCCTGGATTAGGGCCCATCTAATGACCTTATTTTACC
 TTTTCTATTTCTATAAAACCCCTATCTCCAAATAAGGTCAAATCTGTTGTATGGGGTTAAGATTCCAATGTACCT
 TTTTGGAGGGACACAGTTAAACTCATAACACAGAAAAAAGGAAACAGACCTCATCATTAAAGTACTATTGGTTT
 GAATAGTACAAAGTGAAGTGAAGTGAAGTCGCTCAGTCGTGTCTGACTCTTTGCGACCCCATGGACTGTAGTCTA
 CCAAGCTCCTCCATCCATGGAATTTTCCAGACAAGCATTCTAGAGTGGCCTGCCATTTCTTCTCCAGGGATCT
 TTCCAACCCGGAGATCGAACCAGGTCTCCTGCACTGCAGACAGACGCTTTACCATCTGAGCCACCAGGAAAGCC
 CCAATAGTACAAAGCAATAACGTATTTGCGAATTCATATCAAGTAAAAATCCTATAGAAGCAAACGAGTTTACA
 ATGCAATGAAGCTTCTTTGCAAGAAGGCTGTTTTTACCACAGAACTACGCTGTGTTGATTGAATAACAAGTCAT
 GCTTAAAGCAAGATCTCCTGTTATCCATACTTTTAGAGTCTGAACCTGGGTCTAATTAACAAACAGTGTATGAAA
 CCAAAGGGTACTTTTGACCTGCGGATACAGAGAAGTGGCTGTTAGCAGAGTAGAGTTTCTTTACAGTTCTC
 TGTGTGTAACCTGGCAGCAACAAACACCCCTCACTAATCGCATTGCAGACATGACGGCTGCACCCAGCAGAAAAAG
 TCAGGCCACTTTACAGAAAAAATAATCATCCAAGTCTCAAAAGTATTCAGAAAAGAGGTGGCAGGACACGGGACA
 AACATATAGCCCTGTCTGCGCTCTTTTCTGGAATAGTTTTTTCAGTATTCAAAAAATTTTTACCTGCTGTGT
 AGTTTCATAGCAGAGAGGCTTCAACTTCACCCCTCCCAATTGTAATCACCTTTGTAAGTTTTATAAAATAGTTCAA
 AAATAAGAATAATGTTTTTATATTTTAAATAAATTTGAGTGTAAGCTGGACTTAATGACTTGCTTCCGAAGAATA
 GCGTATAAGAAGGAGAGACTAGAATCCTCAG

Fig.5C

TGGAGGAACCCAGCAAAACACCACCTTACTTAAGTGATGGAGATCACCATCGCTAGTGACAGAGCATGTGGATATC
 ATTGTACCTTCAAGAGGACGTGATGAGAAGCCACTCACTACTGTAGTATTCTTCCCCCAGACCCACAGCCCCAGC
 CTACTCACCAGAAAAACATCAGATAACGTCAAATTTAGGGGACGTTCTACAGAACACCTGACCAATACTCCTCAAA
 ACTGTCAAAGTCATGAATGCCTGTCAAGAAAAGACTGAAGAGCTGTACAGTGCACAGAGGAGTCTGAGAAGATG
 TGACGATAGTGTGAGATCCTGGATTACAGCCTCAAACAGAAAGAGGACATTAATGAGATAAATGGTGAAATAAAG
 TCTGGGGTTTGGTGAACAGTAATATACTAATAATGTTGGTTTCTTAGCTTTGATGAATGTGTACATAAGATGG
 TGACATTAGGACAACTGAAACTGGGAACCTTTATACTACCTTTGCAGCTTTTTCTAACTCTAAAATTATTACAAA
 ATAAAATTTTACTAGAAAAATGATTTCTATTTTAAACAAGAATTTTCTGAGAATAAACACCCCTAAAAATTTGAA
 ATAACTTTTATTGTAAATCACCATGTTTAAAGGTTTAAAGATAATCTTCCAGTCGCACATTTACTGCTTGAAGG
 AGGTAACCAACATACTACCTGAATCTGCCAGGTAACAAAAGCACTCCTGTGGTGAGTATCTCTTCCAAAATCCTT
 GAAGGAGCAGCAAAATAAGTGGGTAAACAGAAGTTTAAAGTGAAGACCTGTACTCCAGAAGGATCCTAAAGTGT
 TCTTTTGAATTCATAGATACCCATCGTGATGAAATAGATTTGTTTAAAGGAAGTATTTTTTAAAAAACAGCT
 TAATCGAAATATCTTTCACATACTACACAATTCCTCAATTTAACTATGCAATTCAGTGGTTTTTAGTATATTCA
 CGAACATGGAAAGCACTAGTCAATTTAGGATTTTTTTTATCACCTCGAAAAGAAACCCCAAAACCTTTAGCTA
 TCACCTTCTGTATAGGGCTTCCCTGGTAGCTCAGCTGGTAAAGAATCCACCTGCAATGCAGGAGACCCAGTT
 CTATTCCTGGGTGAGGAAGATCCTCTGGAGAAAGGATAGACCCACTCCAGTATTTCTTGAGCTTTCTGGTGGCTC
 AGCTGGTAAAAATCCACCTGCAGTGTGGGAGTGAAGACCTGGGTTCAATCCCTGGGTGGGAAGATCCCTGG
 AGAAGGGAACGGTACCCACTCCAGTATGCTGGCGTGAGAAATCCATTGACTGTATAGTCCACGGAGTCACAAA
 GAGTCAGACACAACCTGAGCGACTTTCACTTTCTCTGTCTTACAACCTTCTCTCCCCACCGCAACCCAGAAC
 AGTTGCTAATCACTTTTTGTCTCTATAGATTTGCATATTCTGAGCATTTTGTGTAAAAGGAATCATATATGACC
 TTTGATGACTGGCTAGCCATATTATTTTTATAATCTCCAGATGTACTCTGAAGAACAGATAGTTAAGGTACCTTG
 GAAAATTCACGTGTATCATTACAGATATTGAATCACCATTTTGAAATATAATCTCAATTTCCCAACCCACAAGAT
 ATATGTAGCCATCCACATTTCCACTCCCAGGCCATTATCAGTGTCTGTATGTAGCAAGAATAACCTTAGCATAG
 TGGTTAGGCACCCACATCAGGAAGCGAGACCTCCAGATTCAAATGCTAGGGCTGTCACTTAGCAGCTCTGCAACC
 CAGGGCAAATTTATTCAGTTTCTCTGGGCCTTAGTCTCCTCATCTGTAGCATGGGGGAAGTAACGGTCCCTACGTC
 ATAAGGGTGAACGAGCAAAATATGACCTAGTGTGGAGAGCCATTCCGGCGCAAATAGTACCACTGAGTGTTTAGAC
 AGTACTGCTGACTTCTGTATCTTATGTGATCAGAGACGGTTTCAGTCATAAGCGTCTTTTTAAATACCCAGTCC
 ACTCCTTTCATCTTTTGAAGACGGAAGTTTAGCTTGGAAAGACTCAACAGTGCATTCTTACTCATTCCATAAATA
 TTTTTTGAGGACATAGTCTGTGCCAGGGAGAGTCTCAAGTGTGTGCATAAGATTGAAATGGTCTGTTTCATCAT
 GGAGTCTAGTGGAGGATATGGATAAGCAGGCAAAAGGTGAATCAGAGAACATTAATATACACCTTTAGGAGGAAT
 TGCATATGGCAGAAGGAAACAAGGGGATTAGAACAAGAACATCTGGTTGATTTCTGAAGTGTATATAGAAAG
 AACTTCTGAGCAGGGGGTTGGGGGAACGTAGTTAAAGTAAGATACAGAGAGAAAGGATTTTAAAGATGTTTTGC
 TGAGTGTTCAGATTCTGTGCCCACTCAGGCACCTTCAACTGCAGTGAAGATTGAAATGCTTTTTAGTTCTTTTGG
 AGCACTGAGCCACTATTTCTTTAAATATAAGCCAAATTTAGGATTGAGCCGTGCTCTGGGCTGCAGTGTGGGGA
 AGGTTAGAGCCTTGGGGATGGGAAGGCAGGAGTATTTCAGAAGGAAAGTGAGTGAAGGGAGCAACTGGGAAGAGC
 GGGCGTCCCCCAGGACCTGAAGGCAGGAAGGCAAGGCTGGAGAATGGTGGCCTGAATGTACAAGGTTGCAGATAG
 GGCTGCCCTGGGACAGGACCATCCTGCAGGTCCTGGGAGGCCATGTTAAGGGATCACCCCATTTGGCACCAGGAA
 CCAGCAGGGACCGTAAACAAGAAAAGGAATGACTCACACATGGATCGTTAGCTGTTGGCCCAAGCATAGTGCA
 CTAATGTGCAAACTGAGCCCTTTTTAGGAACCCGTACTCTCAGGATGGAAGGTTGCAAACTTACTTTTTGGGGA
 AACTAGCCAGGAACTTCGGGAAATTAGCAAGGAAGGCAGAAGAAGCTGAAGGCTTAGAGCTTCGGGGCAACAA
 AGGATGTCTTTACCCACCTCCTCACCTGCTGTGAGCACTAAAGAGCAATCTTTTGTGTTGCTTGTGTTTAA
 TACTACGAGACAAACCTAGAAAAAAAAGGTTATTTCTCTTTATTTCTACCCGTTCTTAGCAATAACAGTGGTA
 AAATTGTTGTTTCCATGTATCTAGATACTAGTATTGACAGCATGGCTGCTCTGTGAAAACAAGGTGTTCTCTCTT
 TTGCCCTTCTGAAGACAACTCAACTGTAAAGGAGGAAAAATTTTAATTAATGGGTTATGCTTTCAAACCTAAT
 GTTCTGTTATATAGCAGTTTGACAGCACATATTGCAACATTTATTCTCTCTCCATCTTTTTTGCAAAGTTCTA
 TGGTGTCTACATAGCCTTGAGGGGTGTAGCAGATGTGATAAACTACCTGAATTCGCCCCAGGGATGAAAGAGCCCT
 ACCCAGGTTCTTACCCAGCTCCTTGAGTGTGCGAGCTGGTAACCCCTCATCAGCCCCCTTAGGGGTTCCCACC
 ACTTAAAGGAGACACCCCCCCCCCTTCTTAAAGTCACTCCCTCCCCCTTCAACTGCATCTAAGGACTGGTTG
 ATGCAGGCTTATTAAGGCTCAGCTCTCTCTTGCCAACCTTGACATCTCTAAAGGTTACCACATCTTCGTCTATC
 CCTGCAGGATTATGTGAGACTTCTGCCAGGTCTGCACTCCAGCTCAACTCTTCTCTTGGCCATCTTCTCTTACC
 TTAACAGGTATTGAGCCCAAGAGGAGTCCCTAATATCTCCCTGGAAACTCAGCCCTGTGTGTTCCACCGAGACA
 AGAAGCAGAACATTGTGTCCAGCTGAAGAACAAGGCTGAGGGGTATATGCGGGTGTCTTACTCCCTTCAATTCAG
 TTCAGTTGCTCAGTCATGTCCAACCTTTTGCAACCCCATGAACCACAGCACCCAGGCCTCCCTGTCCATCACTA
 ACTCCTGGAGCTTACTCAGACTCATGTCCACTGAGTCAGTGATGCCATCCAACCATCTCATCCTCTGTTGTCCCC
 TTCTCCTCTGCCCTCAATCTTTCTCAGCATCAGGGTTTTTTCAAATGAGTAAGCTCTTCACATCAGGTGGCCAA
 AGTATTGGAGTTTCAGCTTCAACATCAGTCCTT

Fig.5D

CCAATGAACACCGAGGACTGATCTCCTTTAGGATGGACTGGTTGGATCTCCTTGCACTCCAAGGGACTTTCAAGA
 GTCTTCTCCAACACCACAGTTCAAAAGCATCAATTCTTTGGTGCTCAGCTTTCTTTATAGTCCAGCTCTCACATC
 CATACTGACCACTGGAAAAACCATAGCCTTGACTAGACAGACCTTTGTTGGCAAAGTGATGTCTGACTCCCTGG
 AGGGACGTAAAGGCTGGGAAGTAAACCTGAGATCCCCGTGTACAGGTGTGTGTATGATTCAAGTTCTTGTGGT
 AGTAGTGAATACTAAAGCTATCTCATCTGCTCTTGATTCTCTTTCTGTCTTCCTCCTCTGACCCTCCATCTT
 CCATTTCTTTCTCCCTCCCCCTCCTCATCTGCCATTACGCACAAGCCCGGGGTGAGCTGAGAAGCAAGGTGTCTT
 AGTCCATTACAGGCAACTATAACAAAACACCACAGACTGCAGGGCTTAGGAACAACAGAAATGTTTTTCTCAAAG
 TTCTGGAGGCCAGAAGTCTGAAATCAGAGTGCCAGCCTGGTTGGGTGAGGGCTCTTCCAAGTGGCAGACTTCTT
 GTTGCATTTTCATAAGATGGAAAGGACAAGGGGCTCTCACAGGGATATTGCATAAGGGCATTAAATTTCAATTCAT
 GAGGGCTGTGTCTCAGGACCTAACACTTTGGAAGGTCTCACCTACTAACGCCACCAGCTTGGACACTGGCATCC
 CAATTTAGGAATTTGAGGGGACACAGACATTAAGGCAACGGCAAGCCTCCTGTGGTCTAAACATCGACTCCTCCA
 GGTGAGCACTCCCATTTCTTTATGCACAGAACTTCAGGGCTTCCCTGGTGGCTCAGCTGGTAAAGAATTCGCCT
 GCAATGCAGGAGACCTGGGTTTGATCCCTGGGTTAGGAAGATTCCCTGGAGAAGAGAATGGCTACCCACTCCAG
 TATTTCTGGCTGGAGAATTCCATGGATTGTATAGTCCATGAGGTACAGAGAGTACAGACAGACTGAGTGACTTC
 CACTTTTCACTTTTACCCCTGACGGCATTCTTCAAGATAATTCAAGACAGAAGCAGGAAGCTCCTACTCAGAT
 ACATGTTTGTAAAAGGGAAAGGAGACTCTGAGAGCCTTTGAATATTTTACTACTCAGTCTTTGCTGGTGTGCGC
 GCATGTTTATTTCTGGTTTACTTAGGCAAACTTACTATGGAGTGGAGCAAACCTTGGCTGTGAATGTTGAAATG
 GCCAACAGGACAGGCACGTTTTCAGTAGCATCAGAGGGCACTTGAATGCCAGCCAGAAGATCCAGGTTTTAGGGG
 TAAATTTCCCTCCGGGGGTGTAGAAAGAAATACCATCCAAAGTGAAGGGCACTGTGCATCCCTTCGAAAAATGCT
 CTGCCCTCCTTTAAGTTTTTACCCACCAGCTGCGTTACCAAAAACCTCTCAGCTGTGAGGCAATGGACACTCC
 CTGCTTTGTCTGACCAACTTTTGAAGTGGTTTATACTATGGGGTGGAGAAGGCACTGGCAACCCACTCCAGT
 ACTCTTGCCCTGGAAACTCCCATGGACGGAGGAGCCTGGTAGGCTGCAGTCCATGGGGTGCCTAAGAGTTGGACAT
 GACTGAGTGACTTCCCTTTTCACTTTTCCACTTTCATGCATTGGAGAAGGAAATGGCAACCCACTCCAGTGTCTTG
 CCTGGAGAATCCAGGGACGGAGGAGCCTGGTGGGCTGCCATCTATGGGGTCGCACAGAGTTGGACATGACTGAA
 GCGACTTAGCAGCAGCAGCAGCAGTATATACTATGGAGTGGGGCTTCCCTGTTAGCTCAGCTGGTAAAGAATCCAC
 CTGCAATGCAGGAGACCCAGTTTGTATCTTGGGTGGGAAGATCTGCTGGAGAAGGAGTAGGCTACCCACTCCA
 GTATTTCTGGCTTCCCTTTGTGGCTCAACTGGTAAAGAACCACCTGCAGTATGGGAGACCTGGGTTCCATCCCT
 GGGTTGGGAAGATCCCCTGGAGAAGGGAAGGCTACCCACTCCAGTGTCTGCTGGCTGGAGAATCCATGGAGTGT
 ATAGTCCATGGGATTGCAAAGAGTCAGACACGACCGAGCGACTTTCACTTCCCTTCCCTTCATTTACACCGGGG
 TCAGCCTTTCTTTTCAAGAGCCCTCCTCCCTGAATCTTAGGGCTCACTTTTCCACCATTCTCCTCTGCTCCTT
 GCCTGCACCTTCTCAGGCTGCCCTGGTTTCTTCTGTGACATCATCCTGCCTTATTTTCTTCTTCATGCATTT
 CCTGATGAAAAGTGATGGAGGAGGGGAGTGGTGGAGCTGGGGAGGAGATACAATGTGGAGATTGTTTCAAGCAGG
 TTTCAACTCAAGTCAGAGATGAGGGAAGGATGCAGCCCTGAGCTCTCTCCTGTCCCCAGACCACCTGTCCCTCAG
 TCAGAAAGAGAAGCCCTCCAGGGAACCAACAAGGCTCTTAGGCTGTCTCTGCTGGCAGCTGCCCTCATCGCATG
 GTCATTGTGGATTCAATTCGTCTGTCTTCCCTACTAGATTCTGAGATCCTTGAAGCTCCTGTTGCCCTCCCCAGAGCC
 TGACATAGAGCAGGCATCCTGTTACAGACTAAATAAAGGAATAAAACAGAGCCATGTGTCCAGGCTCTATTGCAA
 CCTCACTTTACTGTGCATTAAGGGCTTCCCTCATTAGCTTTCCAGTCTTTCAATCTGGGAAGACTGAGGGCAG
 GAGGAGAAAGGGATGACAGAGCATGAGATGGTTGGATGGCATCACTGACATGATGGACATGAGTTTGAACAACT
 CCAGGAGTTGGTGATGGACAGGGGAGACTGTTGTGCTGTAGTCCATGGGGTTGCAAAGAGTTGGACAAGACTAAG
 TGATTGAACTGAACAGAACTGGTTAGCTTAGACAGTAAGAATCTGCCCTGCAATGCAGAAGACCCAGGTTCAATC
 CCTGGGTTGAGAAGATCCCCTGGAGAAGGGAATGGCAACCCACTCCAGTATTCTTGCCTAGAGAACTTTGTGAAA
 AGAGGCTACAGACCATGGAATCACAACAGTTGGACATGACTGAGCAATGAATACACTTTACTGTGCATTACAAA
 GACCTTGTTTTTTACAAATTGAACATTTGCAGCAACCTTCTTGGATCAAGTCTATCAATGCTATTTTTCTAGCA
 GCATTTGCTCATATCATGTCTCATATTTTGCAAATTCACCAATTTACAGACTTTTTTATTATTATTATAGTTAT
 GGTGATCTATGATCATTGATCTTTGATGTTAATCATGATAATTGATTTTGGTTTATATTTTTAAATTAATACATAT
 ACATTACTTTTTTAGAAATGTTGTTTATTGCACACCTAATGGATACAGTGTAAACATAACTTTTTATCTTTAGTTT
 AGTCAGTTCACTCAGTCATGTCCAACCTTTTGTGACCCCATGGACTGTAGACGCAAGGCTTCCCTGTCCATCACC
 AACTTCCAGAGCTTACTCAAACCTCATGTCCATTGAGTCAGTGACACCATCCAGCCACCTCATCCTCTGTCTATCCC
 CTCTCCTCCACCTTCAATCTTCCCCACCATCGGGGTCTTTTCCAATGAGTCAGTTCTTACATCGGGTGGCCA
 AAGTACTGGAGTTTCACTTCAATATCAGTCTTCCAATGAATATTTAGGACTGATTTCTTTAGGATGGAGTGG
 TTGATCTTGCAGTTCAAGGGATTCTCAAGAGTCTTCTCCAACACCACAGTTCAAAGCATCATTTCTTCAAATCA
 TCAATCAGTTGTGGATGTGACTTGTGATGGAAGTAAAGTGTGATGCTGTAAAGAGCAATATTGCATAGGAACCTG
 GAATGTTAGGTCCATGAATGTTAGTTCCAGCTTTCTTTATAGTCCAACCTTAACATCCATACATGACTACTGGAA
 AAACCATAGCTTTGACTAGATGGACCTTTGTTGGCAAAGTAATGTCTCTGCTTTTAAATATGCTATCTAGGTTGG
 TCGTAGCTTTTCTCCAAGAAGCAAGCATCTTTAATTTTATGGCTGCAATCACCATCTGCAGTGATTTTGGAGC
 CCCCCAATAAAGTCTGTCACTGTTTCCATTGT

Fig.5E

TTCCCCATCTCTTTGCCATGAAGTGATGGGACCGAATGCCATGATCTTAGTTTTCTGAATGTTGAGTTTAAAGCC
 AACTATTTCACTCTCCTTTTTCACTTTCATCAAGAGGCTCTTTAGTTCCTCTTCACTTTCTGCCATAAGGGTGGT
 GTCATCTGCATATCTGAGGTTATTGATATTTCTCTCAGCAATCTTGATTCCAGCTTGTGCTTCATCCAGCCCAGC
 GTTTCCTCATGATGTACTCTGCATAGAAGCTGAATAAGCAGGGTGACAATACACATCCTTGACGTACTCCTTTTCC
 TATTTTGGAAACCAGTCTGTAGTTCCATGTCCAGTTCTGACTATTGCTTCCCTGACCTGCATACAGATTTCTCAGGAG
 GCAGATCAGGTGGTCTGGTATTCCCATCTCTTTCCAGAATTTTCCACAGTTTGTGTTTATGATCCACATAGTCAAAGGC
 TTTAGCATAGTCAATAAAACAGAAGTAGATGTTTTCTGGGAGTTTCTTGCTTTTTCAATGATCCAATGGATGTTG
 GCAATTTAATCTCTGGTTCCTCTGCCTTTTTCTAAATCCAGCTTGAACATCAGGAAGTTCATGGTTCATGTACTGT
 TGAAGCCTGGCTTCGAGGATTTTGAGCATCACTTTACTAGTGTGTGAGATGAGTGCAATTGTGCGGTAGTTTGAG
 CATTCTTTGGCATAGTCTTTCTTTGGGATTGGAATGAAAACCTGACCTTTTCCAGTCTGTGGCCACTTCTGAGTT
 TTCCAAATTTGCTGGCATATTGAGTGCAGCACTTTAACAGCAATCATCTTTTAGGATTTGAAATAGCTCAGCTGG
 AATTCATCACCTCCACTAGCTTTGTTCATAGTGATGCTTCCCTAAGGCCCACTTGACTTCACATTCCAGGATGTC
 TAGCTCTAGATGAGTGATCACACCATCATGGTTACCTGGGTCATGAAGATCTTTTGTATAGTTTTTCTGTGTA
 TTCTTGCCACCTCTTAATGTCTTCTGCTTCTGTTAGGTCCATACCATTCTGTCTCTTTATTGTACCCATCTTTGC
 ATGAAATGTTCCCTTGGTATCTCTAATTTTCTTGAAGAGATCTCTAGTCTTTCCCATCTATTGTTTTCTCTAT
 TTCTTTGCATTGATCACTGAGGAAGGCTTTCTTATCTCTCTGCTATTCTTTGGAACCTCTGCATCCAAATGGGT
 ATATCTTTCTTTCTCTCTTTGCTTTTCTCTCTCTTTCTAAGCAATTTTAAGGCCTCCTCAGACACCAT
 TTTGCCCTTGCTGCATTTCTTTGTCTTGGGGATGGTCTTGATCCCTGTGTCTGTACAATTTCAGGAGCCTCCGTC
 CATAGTTCTTCAGGCAGTCTGTCTATCTGATCTAATCCCTTGAATCTATTTCTCACTTCCACTCCGGGAGTTTGT
 GACGGACAGGGAAGCCTGACATGCTGTAGTCCATAGGGTTGCAAAGAGTCGGATATGACTGACTGAACCTGAACCTG
 AACTGATCCCCACTTTCTTCTAACCACATCCCACCAGCCAGTCAGCCAGCTCAAGCCTGCTTTCCCTTCCTATAG
 TCCCCCAAACCCACAGAACCCCAAATCCCAAAGTGCGTTTGTATTTATTCACCCACTGTGTGTCTTTCAATGCC
 ACCCAAATGATCATGGGTTACAAATGCTATATGACTTCAGCTAACTGTGAAGTCCCTGGAGGGCAGGAAATACG
 TATTTTAACCTTTATATCGCTCCGTCTACAGTCCCTTTTCATCGTGCTTTACATGGAAGCCTTTCAAATATGACAC
 TAGCCAGTTTAAAGCTCAGTTTTATTTTTATTTTTGGTTGTGCTGGGTCTTCACTGCTGCGCACAGGCTTTCTCTA
 GTTGGACGAGGGGCACTACTCTGTTGTGGTGCGTGGGCTTCTCATCTCTGCGGCTTCTCTTGCTGTGAAGCACAG
 GCTCTAGGTGCGTGGCTTCAGTAGCTGCAGCCTGTGGACTCAGTAGTCATGGCACAAGGGCTTACTTGCCCCGTG
 TAGCATGTGAAATCTTCTGACCAGGGACCAAACCTGTGTCCCCTGCATTGACAGGCAGATCATTATCCACTGT
 ACCACCAGGGAAGTCCCAAGTCCAATTTCAATACTCCATTAAAAAATGACTCAGACTTTAATGGGAAGTGTCTT
 TCAGAGTCAGAATAATGTACTCGGGGATACTGAATCTAGGCCCTTTAGGATCCTGGGAGAGGTGGTTCCAATTCC
 TGGGCCTGGAGGTAGGGCCATTATATCACTCCTTTCATTTTATGTTTGGCTTAAATTACTGTAGGTGTTTGGTCAC
 AATCATAAAAGCATTTTTTCAGGTTTATGACATATTTCTTTTCAATTGTGAATGTAAGTGAATGATGCCATCAATCAGTTT
 TTCTTGAATGGGCAATCATAGGCCCCATATTCAAGTTGATTCAAATTATGTATAAGTTGT (SEQ ID NO: 4)

Fig.5F

A C A C T C A C T C C T T G G A A G G A A A G T T A T G A T C A G C C T A G A T A G C A T A T T C A A A G C A G A G A C A T T A C T T T G C C A A C
 A A A G G T C C G T C T A G T C A A G G C T A T G G T T T T C C T G T G G T C A T G T A T G G A T G T G A G A G T T G G A C T G T G A A G A A A G C
 T G A G C A C T G A A G A A T T G A T G C T T T T G A A C T G T G G T G T T G G A G A A G A C T C T T G A G A A T C C C T T G G A C T G C A A G G A G
 A T C C A A C C A G T C C A T T C T A A A G G A G A T C A G C C C T G A G T G T T C T T T G G A A G G A A T G A T G C T A A A G C T G A A A C T C C A
 G T A C T T T G G C C A C C T C A T G C G A A G A G T T G A C T C A T T G G A A A A G A G T C T G A T G C T G G G A G G A T T G G G G G C A G G A G
 G A G A A G G G G A C G A C A G A G G A T G A G A T G G C T G G A T G G C A T C A C C A A C T C G A T G G A C A T G A G T C G G A G T G A A C T C C G
 G G A G T T G G T G A T G G A C A G G G A G G C C T G G T G T G C G G A G A T T C A T G G G A T C G C A A A G A G T C G G A A C A A C T G A G T G A
 C C G A A C T G A A C T A A G T G T G G G A A G A T G C A A G A A C T G G G C T C A C C T G C A A A C A A G G G C A T T T G T A G G T C T T C T T T
 T C C A A T T T G T A T T C T T T T A T T T C T T T T C T T C T G A T T G C C A T G T C C T A G A C T T C C A A A A C T A T G T T T A A T A A
 A A G T G G T A A A G T A G A C A T C T T T G T A G T G T T C C T C A T A T T A G A G A A T T G C T T T C A G C T T T T C A C T G T T G A G T G T
 G T T G T T A G C T G T A G G T T T G T C A T A T A T G G C C T T A T T A T T T T G G G G T A T G T T C C C T C C A T G C C C A C T T T G C T G T T
 T T T C C C T T G A T G C T T T C A A T A T T T T C C C T T G T C T T A A T T T T G C C A A T T T G A T T G T G A T A T T T G T T T A G T T T T
 T T G T G G C T G C T G G T G T T T G A T T T T T G T T G A T G G C T G T T C A G C G G T T A C T T G T G A T T T T A G T G T G C C G T G A G A C
 T C A T G T C C C T C C C A C T C C A C C A T C T G T C T C C A C C T C C C T T C A T A T A C T T C T G T A T T T G G C T T A T A A C A T T T A T A C
 T A G C C T C C C T G G G A T T T T T G G A A A T T G T A T G A G C A T A T T T C T A T A A A T A T T T T G A G C T T T T A G A G A A A A G G T
 G T T A A T G C A A G G T G A A G C C T G T A T C A A A T C T T T G T C A G G T T C G G C A C C T T A A A A T G A G A G T C A G T C T G G C T G T G
 G C A G T A T G T C T G G G T C T C A G G T C T C T C T T G T C C C T C A T A C T T G T T A G G C A C A C T G C T G T G T A G G G T G A A A G
 G T C A G A T C A T T G T T C A T A A G G G G A G A T C C C T T A G A G A T T A T G T A T A G T G G A C A C T G G A G C C C A G T G G C T T G G C T T
 T T A T T C C T G G T T C C A T G C T G G C T A T G A T A T T G G A A A A C T C C T C C T A A C T T A A T T T T A A A C T T A A C T G A A C T T A A T
 T T T T A C T T T A T G T T A G C G T A C A G T T G A T T G A C A A T G T T G T G T T A G T T T C A G G T G T A C A G C A A A G A G A T T C T T T G T
 G T C A T G A A A T G G G G A T G A T G G T G G T G A T G A T A A T A T G A T A C T C A C C A T A C C T T T C T T A C A G G A A T A T T G T G A G G
 A C T C C G A G A G T T A G A A C A G C C C T T G T A C T G A G G G C T A T C A A T A A G C G T A G T T A T T A T T G T C A T C A T T A C T A
 T G T T C A T C A G C C T C A C T G C C A C C A A A A T A A A C T A A A G A G A T G T G T T A G T G A T T A C T G A T T A A A A T G T A T A T T T C
 A T T A G A A C C T G C T T G A T T C T G T G A T T A T A C A T T C T T T C T T G T G A T T A T A C A T T C T T A A T T G T A T T G T A G A G T T C T
 T C C A T G C T A A A T A T T G C C A T A A C G C C A G A T T A C A A A T T A G A A G C T A T T G A T A G A C T C A A C C A C T A C A G T A C C A G
 C T G G C T T A T T T T C C T T G T T A A T C A T T T A A T A G A T T A C A T T T G C A T A A A C T A A A A A A A G C A G T G T T G C A A G T A T A
 G A A A C T C A T T A A C A A G G A T C C A G T T T G A A A C G A A C C A A A A T T T T A A A T T T T A A A A A C A T T T C T T T G T T C A T G A A C
 T T T G T T T C C T T C C C A A A G C A A T G T G T G C T G G C A T A A A A G T G T T G T T A T G T G A A G A G C A C A T G T A C C T T C A T T A C T
 G A A G C A T A A G A C C T G A T A A A T T A A G A G A G T T T T A A T C A G T G T C T T C A G A T A A T A T A G T A A A A T A T A G A A G T T T G
 T G C T G A T A C C A C C A A G C A C A T A A A A A G G A A A A A T A G C A G G A A G C A T C A G G C T T C A T T T T A A T C T C T T C T T T
 G A G A A G T T C A T A A A C T G T T T C A T C A A G T T T T A T T C T T G T T T T A T T T C C C T A T A T G G A A G G G A T G A T G C
 A C T T A A G G A A T T G T C T T T T T G A T A C A G G T A A C T T T G G A T T G G G T G A A A C T G C A A A T A C T C C A C T A A A G A A C C T G
 A A T C A C A G C C A G C T C T T C A G A A G C A T T C C T C A A A C T T T A G A T T T G A T C A C A G A T T T T T A T T G T G C A A C A A A G A A T
 C T A T G C A C T G T T T G A T C A A A C T G T C A T A A T A T T G T G C A G A C A G C A T T T C A G A A A A C T T T T G A T A A A A C C T T T A G C
 T T T G T T A A T T C G T G T T G C T T A C T T G A C T G T C A G A T A A A A T A A A G T C A C A C T T G G G G C T G T G C T C T T A T A T T A G G C
T A T T C C A G G C A G C A A A C A G C T G G A A G G A C C G C A T T C T C C T A A C A G A A A A G A C A C A A A A A C A G C T A T G T G G T T
 C A T T G T A C C A G G A A C A C G T A T G T T T A C T A A T G G T C A C T G C A G C T G C C T A G C A G A G C T A A T A A T T T C A G T T C A C
 C A T C T T C T G T A A C G C T T T G G C C C A T C A C A G A A T T C T A G T A G G C T T A G T T A C C C A T C A C T C A C A T T C T G G C A C T A A
 A A T A A A C T T C A C T G C T T C C T A T G T T C T C A A C A C T G T C A G C A T C C C A T T C T T T A C C T C A T C A G C T G A A A C C T G T
 G C T G A A T A T T C C A C T C A T A A A G A G G T C A A A C A T G T T A A T G G C T T T G A G G G A T G A G G G T T T T A T T T T C A T T T A T G
 T G T T A G C T A C A A T G T T C A T G G T A A T A A A T C T C T A C A T A T C T G T A A T A C T C G C A G A A G T A A A A T G C A G A T T G C C A T
 C T C A G T C G T A A G A G C T G A C T G T T A A T T G A T G T G C T T C T A T C T A A A G A G T T G C A A T T C C G G G A A T T C C C T G G C A G T
 C C A G T G G T T A G G A C T C T C T G C T T C C C C T G G A G G G A C A C A G A T T T G A C C C T G G T T G G G G A A C C A G G A T C C C A G A
 A G C C A T G T G G C A T G G C C A C A T A A A T A G C A G T T C C A A T T C C A A G A A A G C A T A A T A A T C T T A A C T C A C A A C A C T T A
 C A T T G A T C A A T A C A T A A A A C T T A T C T T T T A G C T C C C T T A G T T G A A G A A A T A T A T G A T A A A C T T T A T A A A C T C T G
 T T T A T T T T C C T T C C C T T C T T T T C T T T T T G C A T T T A G C T G T A A A A A C A G A A C C T G A G A A G A A G C C A C A A T
C A T C T A A G G T A A A T G A T T A A G T C G A C T A A T G A A T T A T T G G G A T T C T T T G A A A T A A T A A T G A A T A G C T A A T T T T C
 T A C A T T G A T T A C T G A A T G G T A A A A A A A A A A A A A A A T G G T C A G C T A A C T A A T T T T G G A T T C C C C T T T T G A T C A T G
 A T T A T A T A G A G T A T A T A A A A A A A A A T A A T G T T A T C T A T A T A T A T A T T G T T A A G A T G A A T A A T A T T T T A T A
 A A T G A A A A C A G T T T A C T T C A G A A A G C A T T T T G T T C A A A T C T A G G T C T T C G T T A T A G C T G A A G T A A T T A A A A G A A
 A A A A A A A A C A G A A A C A A A G G A C A A G A G A C A A A G G G T A A T G A G T A T C T G A T T A A T G A T G A G A C A T G G A G G A C A A A
 G G T G T C T A T T G A A C C C G G C A C C T A G A T A A C T A A A G A T G A C A G C C T C T C C T C A G A G G C C C T G G G T T T A C C G T G G T G
 G T C T C A C C T C C G T C G G T G G A C T A C T T G T C A G T T G C T C A G G A T A C A G C T C T C A C T T G C C C T T G G G A A C C G G C T T C C C
 T T T C T T T C C T G A T G C G T G C T G C T A G A G T A A A G C G A T T C C C G T C T G T G T G G A G A C G G A C C C C G G A A T G T C T G
 C A G G T G G G G C A A G T G T C T C T C T T C T G A A A G G T G C T G T G C G T T C C T T C T C T C A A G C A C A C T T G A G C A G T T G C T C A G
 A T A C G A T T G C A A G G T T T C C T G T C G G A A G C A C A

Fig.6A

AACGTGGGTTTGGGCTGGAGTACTTTAAGTGCAGGACAGGGGTGCGGGGCGCCAGGGTCATCTGGGTAAAAGATGACG
CTTATGTGATTCTGTGTAAGACCTCTTGGTGGCCCTATAGACGTATAAATTCAGCAGCTGGTGGAAGTTGTCTTC
TATTCCAAATTTAAATGTTTTTGGTTACCCAGCCTCAAGAGAATTGGGTTTTCTGGTGATACCTGTCTTCTCTACT
TGTAACGCAGAAAAGATCATATAAATCATTTAACTGAAAAAATATAGTGTTTCATCTCCTATGAAAGGAGTTTTT
TTTTTAATTAATTTAATTTCTAGAGTAGGATCCCCATATAAAACAACTCAGTGTTCAGAGTATCTTTTTGATG
TTCTGTGGATGGACCTTGCAGTTACCAAATGATCTCATGAATTTGTGAGAAACAGCCTATAATGTAGATACATAC
CTGTGTGAAAAATCATTCCAGATGAAATACTACATTTTCCATTATTTTTTTAATAATTATTGTTTTTCAGACTGTT
GCTAGGATTATTATCAACCAGACACCAACAGCCATTTCTCTCCCTAGGCCTTTCTCAAGAGAGGCAGAAATGGTCT
AATTGGGATCTTTGTTTCTGAGCTTATTTGCAAAAACAAAACAATGGCCTAGAATGTAATTGCAATTTTATCTTT
TGAAATTGACAGCCATCTGTGGTTCATGAGAAAAAAACCCAAGAAGTAAAGCCAAAGGAACACACAGAGGTAAG
TAATCATTATTAGGACTTGATATCATAAGATGAAGCCTTTTTTTTTTTTTCCCTATTTTTGTGAAGGATAAAATT
TTGAACTCTCATCTTTCAACACTTAAGTCCTACCTAGAATGGCAGTTATTTGTTTTTCTGTTAAACCGGCACCTC
TGTGTGGCATCAGCAGGTATTGCAATTTGCTTGTGTGATTCTTGCTGAATTTGGAGGGAAGGAATTGCATTGTTT
CAAAATTTCTACCCAAAGTGAATTTGTCCATCTGTAATCATATAATTTAAATTTCTCAAAATGACTACATAAA
ACAAAGTGTATTGAATTGCTTTCTACTCCTCAGAGAAAAGTAGCAATATGTGTCTATATTATTAACCCCATGGGG
TGTATGCGTGTTTTTAGCCAAAAAGCCTACCCAAGCACTCATCAGATACAGGAAGCAAGCATGCTCCTAAGGAAA
AAGCCGTTTCCAAATCAAGTGAGCAGCCACCATCAGAGAAATCAACAAAACCAAAGGTAATATAAGCAGACAGAT
GAAAAGAAAGAAAGAAAGAAAGTGAAGTCTCTCAATCGTATCCAACCTCTTTGAGACCCCATGGTTCCCCCAGGAAC
CTACCAGGTTCTCTGTCCATGGAATTTCCAGGCAGGAGTGCAGGATTGGATTGCCATTTCTTCTCCAGGGGA
TCTTCCCGACCCAGGGAGTGAACCCAGGTCTCCTGCATTGCAGGCAGATGTTTTACCAGACGGATAGATGGGAAA
ACGCTAATGTGAGCTAGGGAATAAGACACAGATCCTCTACCTACACTGATGTGTGTAGGTGACTCTTCGGAGCTA
TGCCAAATGTATCTAGTATAGTCAAGATTTGGTGCTGTCACTTCCCTGGTGTTGCAAAATAGGCTTGAGGGGACCT
GGGACTATGTAAGGCCTGAGGAGGGCATCAGACGGGTCTGGGGGATGTTCCGCCACGCCCTGTCACTGCCAGCA
GTACTCAGTTGTGACCTTGAGGGAGCAGGCTCCGGTGTCTTACAGAGAGGAAGTGGGAGTTGAGGCTGCCCTCT
GCACCCCTCTTCTGTGGCAGTTTAGTTGCCAAGCATGTCTGACGCTTGTCTGAGCCACCAGGCTCCTCTGTCT
CATGGGATTCTCCAGGCCAGGCCACTGGAGTGGGTCAACATTTCTTCTCTAGGGGATCTTCCCAACCCAGGGAT
TGAATCCGGGTCTCCTACATTTCAGGTAGATGATTTACCAACTGAGCTATGAGGGAAGCCCTCTCCTTATGGGCT
GTGGGAATTGAGATGCCCTGCGTGATCTGCTCCTTACCTGGAGGTCTTAGAGTTTTATGGCTCTTACTGGAGACA
TCAACTTACAGAGCCAAAAAATCAGAAATATAGTGATTTCCAGAACTAACAGGAGTCATATATAAGAAGAATA
AAATCTATGAGAAGGACTTCTTCGTTATGAGCTCCTGGAAGACCGTGTATCTTGTATACAACATGCCTTGTATCA
GGATTGCCGTGTGGCCCACTGACTGCAGCCACTTGTGTATGTTATTTGGCTTACTGGAATTAAATGAATTTTAA
TTAGCTGTACAACAATTACAGTCAGAATAGTTTCATATAAAATGCCAATTTCTAGCTTCTCTTGAAAAAATGGAATC
GCTGGCAGCAGTGGGCCCCATGGCAACACTGACTGGCGTCTTACAGCTGGTGAGCTCCCCGGGCCACC
TCACCTGCTGGCCCTCTGTAGGCTCTGGCGTCAAGCTCTGTGTGCCCCGCTTGCAATGAGTCCAGCAGAAGCTA
GTGACCATTTCCCTACAAGATGCTCACTAGGGGAAACGGGCTCCCTGGTAACATTTGAAATGTTTGATCTTGAA
ATGGTCATTCAAAGTTAACTTAGTCTGTGATCTTTTGGACCAAGTCACAGGACAAGATCTCCGGTGGTGAA
GAGCACTGTTCTCTGCTGCTGCTGCTGCAGCATCTGCCGAACAGCTGACAAGGTGAGCACACGTGAAAGATACGG
CGTGTGCGTGGGCCCTGACGCTTGTGAGGTCTGTTTATGAGAAACATGCTCAGCGAAGACCCCTTCTCCCTGGGT
TGGCCTCTGTTTAAACATTGAGTGTTCCAGCGCGTATACCCAAATAACCTGTCAATCGGTAAGAACTGAACAG
TCGTGTGTGACATTTGATGAAGAGGAGCGTGCGTCTAAAGGATAGAGAAGAGAGCAAGGCACCAAAAATGAATTT
AAAAATCCTGTATTGAGTACAGTTACCGGGCACCTGCCATGTGTAAGGCCTCGTGTCAATACCAAGAAAATTTAA
ATGCAGGTGGACTCAGACCCTGCCAGCAGGATTTATAGGTTTGAAATGAAGCAGATAAAACTGCCGTGAGCATTT
GAATTAGCAATTACTCTCAGCGTTTTCATATGAACACTACCGTCTGGAAGCAAGTAACCATGTGCGGTGACGAC
CTACATGAATTGATAGGGGAGGTTTTGACGGGGTCTCAAGACCCCTCTCAGGAGTATATGTGCCCTGCTGGCAA
ATTATGGCAGCTTTCGCGAGATCCTTTGAAGGGACTAGCCCCGAGAGTGAAGACAAGAAATGGTCTGCTGGTTCT
ACTCAGAGCAATACAGAGTTAGCAGTTGGTTAGTAATCAGTTTGAAGTCATCTTGTCCCAGAATTGCTAGTACT
TATTTCTCTTAAAGCCACAGGTAACATAATTGCAGTAACAATAAGATGGAATGAAAGCAAAAAAAAAAAAAAGTA
ACATGTTTATTTTCTGCTTCTCAAAATGGGAACTGTTATCCCCCTGCAGGCTAGATGGTAAGGAAAGTCATAGG
ATAAAAGTATGGCCAGTCTCCCTGAAACATCTGTTTTTCTTAGATTTTGAAAAGGATAAAGCATGCTTATTTTT
CTCCAAACTAATGCAGATACCAGGTGGCATGATGGCAAAATTTACATGAACTTTAAATGGAAATACTCAAATTTT
AAACTGTTTTGTTTTTCCCAGCAGACCCCCCTGTATGTCCCTTGAGACTCCCCCTCCCTACCCAGGGTACACCT
CACTCTCTTGTGTTGATGAGAGATACTTAGATGCTACAGATGGAAGAGTGGATAGCTTAAAGGTGGAGGATTGGA
TCTCTTGAATAGATTTAAAAATGCACTCAAAATTTGAAAATGTATGCCACACTTTATTAGCCAGGGTTATCTAT
AAAAATCAACAGTAAAGTCAGGCAAGGTTTTATATGGGATTCTGCTGTGCATTTGCTTGTGCTGGGGCTTGGGCA
AATTATTTCACTTTTCTTTGCTCTAATTTGT

Fig.6B

CTCC TAGAAAATGAGGTTGATAAGATCTGCTTTTCTTTTTCAGTCCACACTGTGTGTAAAGCACTGTGGACTGT
 GGTCTGGCTCTGCAGGACATTGAGCTGCAGGCTGGCCTCTCTTTTTTGGATGTGCTGTGCTCAGTCATGTCCAG
 CTGTTTGTGACCCCATGGACTATAGCCACACAGGCTCCTCTGTCTCCTCTTGCCCTGGAGAATCCAGGCAAGAATA
 CTGGAGTGGGTTGCCATTTCTCTCTCCAGGGCTTCTTCTGGACCCAGGAATCAAGCCCCAAGTCTCTTATAGCTCC
 TGCATTGGCAGGCGGGTTCTTTACCACCTGAGCCACCTGGGTATCCCTGGCCCCCTCTAATGTGCCTGGGGCTGCAT
 TCAGGCAGATTGAAACAGGCCCCCACAGGGAGCTGTTTGTGGAGACAGGGATTCTGTCTGATTTGGAGAACCCGG
 GAGGACTTTGAGTGCATATAGTCTGAGTAAGTCATTAAAGGAAAGCTGGGATTGCAGGCTTGCAAACAAAAAAG
 AGAGAAAAATAGAAGTGTGATCTGAGTGCAAAGGGCATTCCAGTGTGCGTGAAGCACAGGAGACATGAGTGGGAG
 TGTAAGACCAGAAAGACAGGTGAGGCGCTGCTGAGGAGCACACTGAGATGCATTAATCCCATGGTAATGATAAT
 AGCTGGCCACGTAACACTGGGGATCAGAGGACAGATGGAGCGATGTGATGGGATTAAAACAGGCTGCAGGAG
 GATTAAGTGACTTCAACGGCTACGATGAGACAGCATAGAGATTGGAAGAGGGAGGCTGGTCAGGGGCTGTTCTG
 GGGAGCCGAAGTCAGCCTGAATTAGAGCTAGGATAGAATCAGTGAGTAGGGAGCAAGGCAGGAGGTGTTGCAGAG
 GTGCAAGCTAAGAGTTGGACACGACTGAGCGACTTCACTTTCACTTTCCACTTTCATGCATTGGAGAAGGAAATG
 GCAACCCACTCCAGTGTTCTTGTCTGGAGAATCCCATGGACGGAGAAGCCTGGTAGGCTGCAGTCCATGGGGTCA
 CACAGAGTCCGACACGACTGAAGCCACTTAGTAGTAGTAGTAGTAAGATCTCAGATGACCCAGGGCCTCAGTGTC
 TGAATGGACTAGGTTGGCAACCCTAAGGTGAGACAAGAGAGCCGGGTATCACTGCTAGGTTTGTGCCAGTGCAGA
 GTCAGACCCCATGAGGCGAGCATCTTCTTAAGCGGGAGCCCGGCCCTGGTTACCACTGACGGGTGGACAGAGGT
 CAGAAGGAGGGGAATTGGTGGAGGTGGAAAACGGATGCCTGCCTTTGGACAGGCTGAAGTTTCTGAGGTGTCCCG
 AGAGAGGTGATTGGAACATCGCTTTGGAGCCCAGGTGACGATGGAGAAGGAGAAGCGGGCTGGGGGGAGAGAGTG
 ACAC'TGAGTGAACACCTGCGCAAGGCTGTGGATGAATCGAGGAAC'TGAATTCAGTTGCCAGGACTTGGGTTCAAC
 CTGGTGCCCTCCACAGTATCTTTAATACAACATCACAGCAACTGCTTATGTAGTACTTCGGAAGAGCTTAGCATG
 TGCCCAACCACAGCTTTCTCACTGTGATCGATTAATTAGTTGAACCCCTACAACCCCTCTGCAGTAGGTGCATTG
 ACTGCCCTGGTTCTCACACACAGAGGATACAGAGGTTTCAGAAGGAGACGTGACCTGCTGGCAGCCTCCTGGAGCCA
 GTGTGTGTTGGGGCTGATGTGCAACCCAGAGCATCTGAGCCTGTACGCTCCGCCCTCCAAGCTCTGCTGCGTGG
 ATTTTCAGAGACCTTCAAATAAGACAGTTTCTATTGTCTTCAACCCAATGACCTCACCATTTAGTTCTCTGCCTGG
 CCCCAGAAAGCATCTGGCCCTGATCCTAGAGAATGATTGCCTGGGCCCAGCAGCACAGTAGAAACACACAGAGTG
 TGCGGGGCGAGGAGCAAAGCAGAAGATTTGCCAGACGGGGCAGAGAACAGAGGCCACAGGTGCCAGTGTGGAACC
 AGGGTGGTGTCTGTGAGAGATAAGAGTCAGGAGACGGGTGCTGGGCAGAGGAATGGAGGGAAGATTATGATGCTG
 GGTAAGGCATGGAAAATCGCAGGTTAGAAAGTTGTCAGCAGTTACTGACAGAGCAACTGACGCAGTGGTGGGAAG
 GGCATTCAATTCACGGAATGAAAGAAACAGTGTGTGATGATTGACAAAACAAGGGGTGTGGAGGTCTTGATGT
 GTGGGTACAAAGACTGAGAGAACAGCGCCTAGAAGCAGTGGCTGCCTTGTGATAAGAAAGAAATTCATGAGCATTG
 CAAAGACTAGCTGTGGGTCTGGGAGAGGCAGCCAGAGCAATGAAGAGATGGGTAGACCTGGATTTCATATCCTA
 GTGCTATCCTTTACTGAGTGACCTGAGAAGTGACTCAGCCTCTCTGAGCCTTGTGTTGTGTCATCGCTAAATGAGA
 TTAATATCATTGTTATTATTACTTCTGCTGTTCTGAAAAGTATATAACACGTATAGTGTCTGGCACAGAGCAGAT
 GTCCAGTGTATGAGCGATATTATTTATGTTACAGCGTATTCAATTTCAAGTATGTAATATTATTAAGCTTGATGA
 AAGTATATGAGGAGTTTTTATCTCTTTCTCAACCCACCCAGTAATAAGAAAAATAAATTGTTAAACATCGGCCGTAC
CAGCTGAATCTAAACCAAGTAAACCATCTGGAAAGGTATGAAGACAGCAGGGCATTTTGCATAGGTATTTCTTC
 ACAGTCTCTGTGTTTGTGTCATGATCCGATTTCTCGAGGGCAAACAATATGAGGTGAGGTAGCTTAAAGGACCTTT
 TTAATGAATTGTTGTTTGAAGCTAATGAAACTGGGTGGAAAGTGTGACTTTCTTCATGTATGTCTATAACATTTT
 ATATTGTAACCCCTCTGTTCTGAAAACCTCAAATTATATATCGAGGAGTTTACATTCCAGTTTCCCTCCTGCAATG
 TACTTTCTTTAAACAGAGGTTTAAAGGAATCTAAAGAATGCCAGGCACCTTTTATTTCTGAAATTTCCATTGAA
 TCACCCACAGAGCTAACAGACCCAGATTCTATTACACAGTCTGGTGAGTTCCGTAAGTTGTAGTAGATATATGGT
 AGTTGCAAAAAAAAAAGTAAGCGATCATTTTTAAAGCTGAATTAATAGTAAAGGGTCTTTCTTCAACTTGTATT
 TTTCCCTTTGAAAGCAGCTTAGTCTTCACAATGAATCATTAAAAAATAATAATAAACAAGTTGTTTTATTAT
 TTGGCCATACTGTGAGACATGTGGAACCTTCTCTCACCAGGAATGGAACCCCTTGCCCTTGCAATGCCCCCTGCCGT
 GGAAGGGTGGAGTCTTAACCACTGGACTATAGGGGAAGTCTCACAATGTATGATTCTTAACCTCAAAAATAGGGT
 GGAGTCACTGTTTCTTTTACCATTTTAAAGCCATATCGACAGAATAAACTGTGCTTCTTATAAATTTAACATCC
 CTTTCAGGAAATTTCTATTTGTAAACATCCTCAGACAAGCAAGTGAACAGAAAACCAAAATGAATAAAAGAGCA
 CAGGGCAATCCGTTTCATGAGATGCATTTTATTTGGAAGAGGTGGAACAACCTGATAAAAAAACCAAGCATCTTT
 CCTCTTTTCATATAGCAACATGGTCATCATTCTTTGCAATTTGTTTTCTTCTTTCTTTTGACTATCGGTACAGTGG
 TCTGGCAGCCATAATGAACCATATGCTTGTTCCTTAGTTAATAATGCTGCTTATTTTGTGCTGCTGCTAAGTC
 GCTTCAGTGTGCTCCGACTCTGTGCGACCCCATAGACGGTAGCCACCAGGCTCCCCCGTCCCTGGGATTCTCCA
 GGCAAGAACACTGGAGTGGGTGGCATTTCTTCTCAATGCGTGAAAGTGAAAAGTGAAAAGTGAAGTGAAGTGCCTCAG
 TCGTGTCCGACTCCTAGCAACCCCATGGACTGCAGCCGACCAGGCTCCTCTGTCCATGGGATTTTCCAGGCAAGA
 GTACTGGAGTGGGGTGCCATTGCCTTCTCCA

Fig.6C

AATAATGCCTTTAGACTTACTGAAATTGTGATTCTTACTTGAATGTGGCTTTCTATCTTGCCAACTAGAGTCTG
 TACTCTGGGACTTTAATAAGTTCTATTACAGGCCAGATTTTAACCATTTTGATAGCTGAAGAACTTTTGGTGTTT
 GCTATTTTTCTACATAGAATTTCTTTTGATGATTAAAATGTCAAAGTGAAGGATGTGCAGCAAGTAGCATTTACA
 TATCTGCTCTGTTTTCCAGTCAGACATGGACACTGCTCTGGATGACTTAATAGACACTTTAGGAGAACCTGAAGA
GATGAAAGAAGATAACAACAATATACCGGACCGGAAGTGTGCTACGTGACCTTGATGTCTCTAAAGATTTCGTA
 CCTAGTGAGTAAAGAGGGCTGGTAGGTCATGGCCCTGATGCACGCAAGGATGCACCGTGTTCTGTCACTTGAAAGAG
 AAAGAGGATGCTGTAGCGTCCGTGGTAGCAAATGAACATGAGTGTGTGGTTGGAAAAGTGTCCAGGGAGCTAT
 GCAAGCTGTGCTGATCAGCAAACCCCGATTGCTTGTCTTCTGAGACTACAGCTGGGATGTACTCTGTTGAGTC
 ACAAAGACTTACTTTCTGTCTTCGAGGTTTACATCCTCATGGTTTCCAGATATTTCCATCCTGTGACTTGAGCTT
 ATTCTGACGTCTAGATATTTCTGCCCCGAGTTCATTTCTACTCCTGATTCATAAAAAGCAATTGTATATTCTTT
 GGTATTAAGAAAGATGTTCCATAGATTGGAAGAAAATATTGCAACAGACACAACAAAGGACTTGTTTCTAGAGTT
 TATGAAGAATTTTGAAGTGAAGAAAAACAATCCAGTATAAAATTAGGAGAAAATTTGACCAGTAGACACATGA
 AAGTCTCCCTAGTAATTGGGAAAATGATACTAAACCGCAGAGGGATAGCACTTCACACTTACCTGATTGGCAGAA
 GTGTAGACGTCTGACACGTGACAGAGGCTGTGGTTGTACCTGAATATAAGCATTCCTCTAACCAGCACTTCCA
 TGTCTAGACAGATTCCCGGGAGGGGGCCCTTGCAAGTGTGTAAAGAGTTATAAGAACAGGTGAAACATTTGGAA
 ATAATTGCATTCAACAGGAAACAAGATGATGACAAGTGAATGCTTTATAGTGGTTAAAATGGTTTGTTCATGTA
 TCAATATGGTTTCATTACAAAAGCATAATTTTGAATGAAGAATAGTACATTGCAGAATGATAGTAGAATTTGAT
 CTGATACCATTATGTAACACTTGAAAATAGGCAAAATAATGGTGTATATTGTTTATGGATACTTATATGTGGAA
 TAATTGTTTTTACAATATACTCAGCTTTTACTGATGGATGTGATAAAGTTTGGGGTGGTGGCTATCTCTGGAAAA
 CAAGGAAGAAAATTTGGATTGGAGAGTGTCTAAAACAGAGATATCACTCTGACTTTTTTTTATTACATAAGAAATA
 GCTAGATCTATTGAGTTATATAGAAATTCACGTGTACCAACTACTCTTTTATAGGTATATAAATATTAGATAATG
 CAAATTATTTAAAAATAAAATGTATTTTATATAGGAATCAAATTTAATATATTGGTTTGAATATTTTATAGCATT
 ATTTATGGGTTAAGCCAATTAATATAGGGCAAGTATCTGTAGAAACATAAAATTTTCAAGTTGCATAAACTGCTG
 GCTTCTTAATGATTTGTATTTAATCTCAATAGTTCCATAAGTTAATAATGCAATAGTACTTTGATTTAAGAGTTA
 TGTATACAACATTCTCAGCTAAAATCCCTATACGTACTTTTTTTTTTCCCTAGTATCCAATGAGTTCTACCTACA
TAGAGGAACTGGGTAAAAGAGAATCCACACTTCTCCAAAATATAAGGAACTTCTGAATGTAAGTTAAACAGTTA
 TGTATTTACACCTTATTGTTTTCTCTTTGGAATTATAATTTTATCACATTAAATATATCCAATTATTTTAACTTA
 GATAAAATAAGGGGCTTGTGAAACTATCATTACTCTGAAGCTCACCGTATTAGTTTGTATGGGGCTGCCATGTCAG
 AGTGTTCAGATTGGATGGCTTAAACAACAGAACTTACTATCTCACAGTTCTGGGGCTCGAAGACTGGAATCA
 AGATGTGGGCAGAGTTGGTTCTGTTTCAGGCCCTCTGCTTGGTCTCCACATGTGGATACCTCCGTGTGGAATAA
 CTCTACATCTTCTTCCATGTGTACATATCTGCATGTCGTCTTTTCTCTATATGGATCGCTGGGTCACTTCCCCC
 TTTTATAGGAACACTGGTCATATTGAAATAAGACTGATCTTAATGTTGTCACTTACCTTAATTACCTCTGTAAT
 TAAATTCAATACCACTTCATGGCAAATAGATGGGGAACAGTGGAAACAGTGGCTGACTTTATTTTTCTGGGCTC
 TAAATCACTGCAGATGGTGATTGTCAGCCATGAAATTAAGAACGCTTACTCCTTGAAGGAAGATTATGACCAA
 CTTAGATAGCATATTTCAAAGCAGAGACATTATTTTGCCAAACAAAGGTCCATCTAGTTAAGGCTATGGTTTTTCC
 AGTGGTCACTGTATGGATGTGAGAGTTGGACTATAAAGAAAGCTGAGCGCCGAAGAATTGATGCTTTTGAAGTGTG
 GTGTTGGAGAAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAACCAGTCCATCCTAAAGAGATCAGTCTT
 GGGTGTTCATTGGAAGGACTGATGCTAAAGCTGAAACTCCACTACTTTGGCCACCTCATGAGAAGAGCTGACTCA
 TTGAAAAAGACTCTGATGCTGGGAAAGATTGAGGGCAGGAGGAGAAGGGGACAACAGAGGATGAGATGGTTGGAT
 GGCATCGCCAACTCGATGGACATGGGTTTGGGTGGACTCCGGGAGTTGGTGATGGACAGGGAGCCTTGGCGTGCT
 GAGGTTTATGGGGTCGCCAAGAGTCAGACACGACTGAGCGACTGAACTGAACTGAAAGGCCCCATCTTCAAATAG
 AGTCCCATTCTGAGGTTTTGGGGGATGGGACTTCAGCATGTGAGTTTGGGGAGGATACAGTTTCAAGGCTCCCAT
 TCCTGAGGTATGCATTTTGTGCTGACTGTAAGGATTAATCTTGCCTTCTCAGCATCATCAAATAATATTGTTTC
AGAAAGAAGAAGGGATCGCAGGGCCTCTCCAGACTCTTGGTGAGTTTACATAAATGTCTTCCAAGATCAGAT
 TTAACATTTTTTCAATCTTTTGGTTGGGGGAATAGTTATCAATCTGTATTTTGGATTATGACTATAGTACTCTC
 TGAATCAGATGAAGTTTCCATCTATTCAAAGGAAGAGGGTGGGGATGGGGGTCACTTTCAGTGCTTGTATTATTTCC
 TAATTAGTTCTTTTTTGGAGGAAAGAAGAATGGGACAAGAAAGTTTTAGGACAGAATAAATATTATTAATAAATAA
 GATGAGAATATCCTACCAGCATAATAACAAGTACGTTGAGAGAAACAGAAAAATAATATGCACTGAGCATCCAAT
 TTCTCCCAAACATAGGCTTTTCTGAGTTATTTTCAATTAAGTATCTCCTTTGGTAGTCAAAGACTTGCACGAGT
 GCATTACCACTTTGCTGGGTTTGTAACTGTTCTGCTCACTTAAATGGTTCCCATATGTGACCCGTGCTGCTGCT
 GTTTCTTTAGAAACCCCTGGGGCCCAATGATGCCATCGATGCCTTGTCATCCGACTTCACCTGCAGTTCCCTTAC
AGCTGATGCAAAGAAAAGTGAAGAAAGGTATGGTTTTTAATGCCCTTAGGGAAGCTTGTTAGAAACTACCTCCC
 ACTTTAAGACAACAACTTTTTTTTTTAACTTCATTTTCACTTCACTGCGTCTTCATTGCTGTGTTCCGGGCTTTC
 TCTAGTTGGGGCAAGCGAGGCCTGTTCTCTAGTTGCGATGTGTAGGCTTCTGCAAGGGGCTTCTCTTGTGCTGG
 CGCGGGGCTCTAGGTGCACAGGCTTCAGTTGT

Fig.6D

TGTGGCTCGAGGGCTCTAGAGCACAGGCTCAGTGGTCTTGGCGCACGGGCATGGTTACTCCAATGCATGTGGGAT
 CTTCCCTGGCCAGGGAGCGAACCTGTGTCCCCTGCATTGCAAGGCGGCCTCTTAACCGCTGGCCACCAGGGAAGC
 CCCAAGATGCCAAGGCTTTTACTTCTGGTCTTACCGTTTGGTTTCATATGTTTCCTTCATCTGCCAGTCAAACC
 TTCTTCTGTATTTTATTTTTCAGAAATCTACAGAAGAGGCTTTAAAAGCTCAGTCAGCTGGGGTGATCAGAAAGTG
 CTGCTCCACCCCAAGAGAAAAAAGGAAAGTGGAAAAGGTATCAATAATTACTTCTTTGAACTTCAGCACGGTGC
 CCTGGATAGCAGTTTGGTTTCTGAGGCTGATCCAGCTGACTGGGGGGGAGGTCTCAATAGTGCATTATACCCGT
 AGACCTCCTTACTCCCTCAGGCCTCGGGCCTGCGTGGCATAGTGAAACCAGTCAGGCTTATGTTGGTGGGCA
 GGCCTTCTCTGTTTCTACAGATGCTTATGGGTGATTTTGGCACATGTGCCTTGGTTGTGGGAAAAGACTGTCTCA
 CTGTTACATTACACAGAAGGGCAGGACCCAGAGACACCACTCTCTGGGGACATCAGATCTTTGACTTGATTTTC
 TAAAGGGCCTAGAGTCCTGTGACCCTCAAAGGGAGCGTTTCTGCAGTTACTCTGTTCACCTTGACCACAGTCAGC
 TATAACCTGCTTCTCAGGAATCCTGAAAGGCGGCCCTGCAGCTCTGGACTGGATGAGAAGACCTGCCCTGTCTGC
 AGATCGCCAGATACTGCAGGGAGGAGTTTTCTGTGGGTGCTGCTCTGCCCTCGGCTCGTGGAGTGGCCAGTTTA
 TTTTTTATCTGGAGGAGCAGAGAGCACCAACCCCTGAGTGTGGGAATTCCTACAGCTGTGAGCCACTTCAGGGC
 CCCCTGCTTCTGGGCTCCGCTGGGCTACAGGACAGGGTTATCAGAGTCATTTTAAGGGACCTAAGTTTAGAAA
 AAGAAACACAGTACTTTGATACGAGTATTTTACAATTATTACTCTACGTGCACCTCTCGCCTCCTAAGAGGTGAA
 GTGCACACCCGTACAGGAGCCCCGTACTGTGCAACCGGTCTACCCCATGTTTCTGTACCCGTAAGTTCTT
 ATTTCTTTCTCCTTTTTTGTCTGGTTTCAAGTAGCACTGCCTTCTTAGAGTCTTCTCAAGCCACAGCTCATT
 CCTAGAGATTTACTGACAAAATGTGAATGTCATGAAGAAGTAGGTGGCTTATCGAGGTTGAGGAGAGCGTCTCGT
 TTTTCATGTACAGAATTCTTTGCTTCTCCGCAGATGCCATGACTGAGCACGCCCTGGAGGCCCTGTCAGCCTCC
 CTGGGCACCCGGAAGCCGGAGCCGGAGCTCGACCCAGCTCCATTAAGGAGGTCGATGAGGTACTGACCTGGGGG
 TTCACGTACAAAGCCTGTTAGTCTGCAGCTTACAAGGTTACGAAACTAATGCCCAACATATGTTTAGAACTTTTG
 TACCCTTTTGTTTTATTAGAGATCTCAAATGAACTGTGCTTTGTTCTGCTGTACTTTCAGGGTCTGGCTAATT
 CAGGTATTTTATTGCCAGGGAGTTAACAACCAGGAAGGATATTTGAGTGTGTTGCTTATAAACCATCTAGAATTG
 TCATACTTCAAATGAACGACTAATCTCCTTTGCCAATTCCTAAGTGACACCTGGCTAAACAAGATAACATAGAGG
 TGGTGGCCTGATGAGAAACAGGTTGTCTGCTCGTATGGAAAATTAGAAATCTGTCTTTGCAAAAAGGTTCTAGGTG
 CTTCTATTTATAAAGAAAAGAGAGGAGATGGCTGCAGATGGGGGTGGGAGATCCTTTATATAACTCTGGGGAAAAG
 TGTGAAACCTTGTGCAATTAGAATAAAATCAAGCAATAATTGCTTTATAACTAAAAGGAATGCACGTTTAGA
 GTATGGAGCTGGGGAGTTGCTCCTGTTTGGTAGAGGAGTTTTCATGTTAAGTTTCAGTTTTTAAGGGATTGCTGA
 GGAGTCTCTTGGGGGAAAAGAGGAGGGGAAGCAGAGGGATGTCTGGGGCCACCTTGGCCAGGCAGTGGTGGG
 GAGACCCTCCGAGCGCCCGCTCCAGCTCAGGTGGGACCTGGTGGCTGTCTGTGTCACAATGCTGCAGCCTCT
 GCTGTCACTTCTGTTCCGGGTCTCCGAGGAAGGTGGGTCTCCGTCACTTTTCAGCAGCTCCCCCGTCTGGGACC
 ACCCGTCCCCAATGGGCCAGTTTCTATCGTCAACCTTTTTGGAGGACTCACTGGGTGCCGTATGGCACGAAAGG
 CTTTACGAAGGCGAAGGCAAGGGAGCGGAGAAAGCCGCTTTGTCTCCAGGACCAAGTAATGTACAGGGAGAAG
 GGAACCTTGATTGCAGCAGGGTGCAGGGGGCACAGCAT (SEQ ID NO: 5)

Fig.6E

ACCACAGTTCAAAAGCATCAATTCTTTGGCACTCAGCTTTCTTTATAGTCCAACTCTCACATCCATACGT
 GACTCCTGGGAAAACCAAAGCTTTGACTAGATGGAACCTTTGTTGGCAAAGTAATGTCTCTGCTTTTTAAT
 ACCAGGGAATATGTTAAATTTCTCTAGAAAAGCTAGCAAACTCTTAAATTTAAACAAAAAGAACAGTTC
 CTTTTCACTCTTCCACTAGTTTTCTTTTAAATTATATATAAAGTATGAAGATTTTCTAACCTCTGGA
 TCCTTACATTGAGTTTGTAGTACCTGATCATGAATTCCTACTTTGGTTAGTCTGCTGTGATAACCCATA
 AAGATTGATATGGTATAGGAGTCCATGGCCACTCCAGCACATTCTACGCCAGGGGATGCTTCAGAGAGGA
 GGTCCATGCCCTGCTTGTCTCTATAAATAATGTACCATGGCATATTTATTCTGTAGTGATACCTCATT
 ATTAATCTCCCCAGATTCTAGGGCTTCGCTGGTAGCTCAGTTGGTAAAGAATCTGCTTGCAATGCAAGAA
 ACCAGGGTTCTATCCCTGGGTCTGCAAGATCCCTTTGAGAAGAGAATGGCAATCCACTCCAGTACTCTTG
 CCTAGAAAATCCCATGGACAGAGGAGCCAGCGGGCTACAGTCCATGGGAGCTGCAAAAGAGTCAGACAT
 GAGTTAGCGACTAAACCACCACCATTTTTAGTGTGTTAACTAGTCTGGAACATATGTCAATTCAGGGATG
 TTTGTTTTTCATTGCCTATTGAATACAGACCACAGTGCAGACATTTAAGATATATAAAGATCAAAGAGGAA
 TAGAAAAGGTATCTAGCTGTGCATATGACTCACTGTTTCAAGCCCAAGTCGGGGCACGGTCCATATATGGG
 TGACTCTTCCATGAGGACTGTTTTATAGCTGGAGACCTCACAGATCCCTGAACTGGGGAGAAGTACGGTG
 TAATAGGATGTCCACTATATCTGAAATAGAATATAACTGGAAACATTTTTAAAAAAAATCCTTTCATCC
 CTTAATGACGGCTATTATTCTATAGAAAAATAAACTATAATACCAAAGCCTAGAAAAAGTACCTGGCTT
 ATTAAGAGTGCTCATTAATATTTTTAAATTAATCAACTAACTAAGCATTCAATGGAGAACCACGATCTT
 AGGCTGAGACGTAGGTGGGAGCAATAAGGTGAGCTAGAATCAGCTAAATACCACCCACCCACAAACACG
 CACACATACATACACACTTTTTGTGCCAGTAGCTCCTTGGCTTGTGATTTAAATCCATGTCTTGATAC
 TGATACAGATTATTACATAGCATGAACTTCACATTACCCACCAGTAACCTTAAATGTAGACGTTCCGTC
 CTGAAAGCCACTGATCTGCTGCGGCAGTGCACCTTCTATGTACGTGCATTGCGTTGCTCGTGGTTTCAGTAA
 TCCAGTGAATCTTGGTGTTACTGAAGGGAGTTCTGTAAAAGAGGTAAACCACTTATCCATGTATCAAC
 CAGATTTGTGATGGGATGTCTCTCCATTTCTCATCTGTTTAACTGAACGAGTTTAAACACTCTTGAGC
 ATTGTTTTTTTTTTTTCTCCCCCTTAGATAAGTACCTCATTGAGTATGACAGATGAAGGCTGTAAATCA
 TTGCAACTAACTTGTGCACTCTTCTTAACCTTCTTCCAAGAATATTTATTTAATTTTCTCTGCTGGTCTT
 TACAAATCAAGAGAGCTAGGTAAGGGTTCTTAAAGAAGTAGCCCCGTTTTACAGCTGAAGAACTGAGGC
 AAAGAAAGTCTGGGTGGCTTCCCAGGGTCACGGAGACAGGGCAGGGACGCCTCCCCGGGACTCCTGGCTG
 GCTTGTTTTTCTCCACCTCCAGTCTCCCTACACCACCTCCTCCCGGCCGCTCACTAGGCCCTTTAATCT
 GATTGGGTCTTCTCCTCAACACCCATGGAGTACATAAAGTGATGTCTTATTTTACCTGTATTATCCAATT
 TTTTTTTTTCTCTCCACTTTCTTTAGGCAAAAGCCAAAGAGAGAAAGTAAAGAAATGTGGTGAAGATG
AGGAAACAGTCCCATCGGAGTACAGATTAACCCGGCCACAGTACGTTGCTCACCTCGTTGCGTGCCAGC
 GGCTCGCTGCTCCTGCCCATGTGTCCCGTGGCGCCCTGTCTGTCCCAGGAGGCTTCAGGGCATGCTGG
 CCAAGTTTCAGATAACAACTCTGGAACCCCTGAACTTTTATTTTCTGTCTGAACAGTTGAACCCCTTTAA
 AAGAGGGTTTTCTAAGTAGCGACCATGTGTGGTCACTAAAATTCCACAAATTCAAGTGTTCAGCTATT
 AACCATACTTGACATAGATCCTTAACCTTAAAGTGGAAGGAATAATTTGGAGTTTATATTTCTCTAACT
 GCCAATGGATAAAAGCACATAGATCTTGTAAGTGGTAGAGAAACATATAATGATTTACCAGTGTCTTCAG
 TGAAAGTGGGTTTTGGTTTGGTGTGAGGAGACTGGACTTATCCCTGTAATTATATTTGTCAATTTATCTTTG
 AAAGAATGCACATAAATGTATGATGGGAAATAGTTGTAATGCTTATATATGGAAGTATAAAGTATTTTA
 AGAAATATTCTAGAATGTGAATGTCCAAACAACTAAAACAGTTTTTA (SEQ ID NO: 6)

Fig.7

TCACTCAAACATATCAGAAAAAGACAGTCTCTTCTCTCATATCCTGGAATTTCAAAAACAAACCTTAATGGAATAA
 GAAAGCAAATTTTTTGTTCATTTTAAATGAATTTTCAATTTTCAATTCACAGGATAAAGATGGAAAAACCACTCTT
 GCCAGAGGCTGAAGAAAAACCAAGGTGAGGAAATAAGTTTTTTTCTGATACCTAAAAAAAATACAACCTGTAA
 GATATAAACAGATAATTTTCATAGACCAGATCTCACTGTGGAGACTCATATGCCAGCCTAGTATTAAGGAGAAAG
 GAACCTTTGACAGCAGCATTGGCCATTTACGCATTTTAATGATTATTTACCTCATGAATTATTACCTATTATA
 AGTTTATAGGTGATATTCATTGTACATAGGAGCTGATATCAAGAGGGCCCTCAAATTAGGTCATTAGCATTTACT
 TTGATCCTATTCTTATCATCAAATGCTCTGATTCTACACTACCTTTACTTAATTTGGTATTAGCCAGCATAGTGA
 CCTGTCTATGATGATACTAAGCAACTACATTTATTAGTCACTTACTGTGTGCCTGACACCACACTAAATGGTGT
 GTGTGTGTTAGCATTTTGAGAACTCGCTGAATGAGGAAGGCACTATTATTATCCCACTTGACAGGTGAGGAAATTA
 AGACATAAAAGGATCCTGTAAATAGCCCAAAGTCAAATAGCTAGCAAGTAGCAAGGCCAGGGCTCAAAACTCAGC
 CTGTCCCTCTCTGGGGCCAAGCTTCTGAATTACATGCCATCTGTCTGTCTTAGATTAGGGCATGTGGAGATTTG
 ACCAATTACAGTAAATAAATAATACTACTTCACCAGAAGGAAGTAGGGATAACAGAAGCTTTCCTGCCAAACCCAC
 GTTCTCATAGAAAGGGAACATATGATCTATGTTTCTCTGGCTTATTATTACCGAAGTCAATAGGAAGTAGAAA
 AAGCAATCACATAAAAAAACCATGCAGCTCTCAGAAATTTACTTTGAAATTGAAGTATCAGGAAAACCTTTTAGC
 TGAATAAACCTTTAAACATTTTTTATGGTACTCGGTGTTTCAGTATTCATTACTTGTGTGTGACATTTATCTTATA
 TTGATAAATGAAATGTATGCCCTGTATAAATTAGCTTTGGGTTTTAGATCCTAAATGCAACCTAATATATGTTA
 ACTAGAAAACTTTTTCTCTAGGAAAGAATTTTTCTTCTATCAAAATAAAAAAAAATATTTTTTTTCTACTATA
 TATAATATTCCTTATATTACAGCCCTGAGTGAATCAGAACTCATCGATGAACCTCTCAGAAGATTTTGACCAGTC
 TAAGTGTAAGAAAAACAATCTAAGCCAAGTGAAGAAACAGAGGTATGTTTTCTAAACTTAAATCTCCAGTTTT
 GGTCGTTTTTTCAGTGTGTTGTCTATCATCTTGAAATGCCAGGACCAACGTAATTTAACAGGAAGCAGCTGATGA
 AAAGGCCTGTTTTGCCATATTTTTTACATGCTGTGTTTCATCTTGCCCTGTTTATGGATGTGCTTTGTTTTCCCAA
 GCTGACGCTGATACTCTTGACGTTAAGTCCGTGAATCAAGTATTCGGTTCCAGGAGAGCCTTCACACCAATGCTT
 TTCTTTCCCTTCTGCTTCTGTAACCTGACCTACTTAAAGCTGTCTCATCTCATCACCTCAAGACGGCCGTGGCCT
 CTGTCTTTGGTCTCTTCCAGCCCAACCCATGTGTCCACTGGCATTATTCTTTCTCAGCATGTTTGTAGCCACAA
 CATCCCCCTGTTCCAATGCCAAGAGCTTCCCAAAAGAGTTCCTGCTTGTATTTGGGTCTGGAGAAGGAAATGG
 CAACCCACTCCAGTGCTCTTGCCCTGGACAGTCCCATGGACAGAGGAGCCTGGCGGGCTACAGTCCACGGGGTCGC
 AAGAGTCGGACACGACTGAGCGACTAAACACCACCACCAGCTTCCCTCCCTACCAAGATGTGTCCCCATCA
 CTAAACTCTGATAGTAAGTACTTATGTATCTTCAAGCTTTTCTCATGTGGTTCTTTCTGCTTGGAAATACCTCTC
 CAAATTCTAGATCAGCTAAAATTGTAGTCAGCCCTAAAAGGCTCTGCTAAAATATGGCCTTTGTCTCCCGCATGT
 CCTCCCCAAACCTTACCACACTTCCAACCTTTAAGAAGCAGAATACAACACGCCAGGGAAGTCTGTCTGCTTT
 CTTTTTATATCCCTGTTGGGACATTTGTTGCAACAAGCTTTTAGCTGCTTTGGTCTTTACTTCTCACATGTGTA
 CATATTTTTACATAGTAAATTTCTTGAAAGAAGAGATCATACACTATTAATGATCATCTGGTAAGCACCTTGAT
 ATAGGATTTGCTTCCATAAATGTCTGGAATAAGAGCAACCTGTTTGAAATTTACTCTTGTGACAAAATGATAGT
 GACTCTCAAACCTCAGTGTGTCATCAGGATGACTGGGAAGGTTTATCAAAGTACAGATTACTGAGCCCCACCCCA
 GAATTCCTTGATTAGTATGTTAAGGTGGGGCATAACGAATTTGAATTTCTAACAAGTTCCCTAGCAGTGCTGAGGT
 TGGTGGTTCGACGACCTCATTTTGAGAACCCCCAGGATAGTCTTACCACCTTGACATTTGTACCTAGCTATTTACA
 ACTGACAGGGTGCTGTTAGATATATGAACATGATATATTTAAATATTGCATAATTGCAATATATATATATATA
 CTACAGCTATATAATATTATGTTGTATAACATTATGTATTATATGATGGCAAATCATCTAAATAATGACAGTTGA
 TAACAATCTTTGCGCTTACTGTGTTCCACGTAATGCAAAAGCACTTAGCGTGTGTTGTCTCATTAAAGGCCTT
 GAGATGTAGGTACTGTCTATTGTTCTCATTTTGCAGATGTGGAACTGTGTCTCAGAGAGGTTAAGTAACCCACCT
 AAGGTAAGAGGTAAAACCAAGTGACGTAGTGGTAGCTGCAGAGGTTTAAAGTACCTTGAGTTATTCACTGCTCAA
 GATCCTTAGGCATTCAAGAAAATCATGCTCACAGCGGTAGGGTAGCAGACCGTATTGTTGGTTTCATTGTTGTCC
 CATTTTAATAACTGGGTATTTTCAAGATCCTTCGTTGTGCCTGATTCTTTCTGCGAGCATCCCCGGCCGTGCCC
 CTGTACCCGTGGCAGAGGACGTGCCTCGGACCTCTATGTGTTCCGTGACGTGCGCTCCGCCACAGCAGCTCCAAG
 TGGTGAGTGACCTCTGGGCCTTGGGAAATGTCTTGAGAAAGCAGTGTTCTTCTCCCTGCCCTTCTCTCTGTA
 TTACCAATGTATTACCTTGTCTAGCAAGATAAAAAATTAATCTTAGATAGCTCTATTATCTCCAGGGAGGATA
 ATACAGCTCCAGGGAGGAGCTGGAAACTCAGGCTAGATTTTAAGGTTTCCACTATTTTCAAGGTGGTTTTGTTTTTA
 ACCTCTTGACATGCTTTCTAGCTCCTCAGCTACTGTGGATAATCATGCTTCTTCACTTTCTTGTGTCTTTCAAGC
 TGATTTGGAGCCCCACCAACAGCCAGAAGCACTTTTTATTGACTGTCTGATGGAAAAGTCTGCACCTTTATCAT
 CAGTTCAGTCACTCAGTTGTGTCTGACTCTTTGCGACCCCATGGACTGCAGTATGCCAGGCCTCCCTGTCCATCA
 CCAATTCCCGGAGTTTACTCAAACCTCATGTCCATCAGTCTGATGCCATCTGACCATTTTATCCCTTTTCTCT
 CCTGCCTTCAATCTTTCCAGCATCAGGGTCTTTTCCAATGAGTCAGTTCTTCGCATCAGGTAGCCAAAGTATTG
 GAGTTTCAGCTTCAACATCAGTCTTCCAGTGAATATTCAAGACTGATTTCTTTAGGATGGACTGGCTGGATCT
 CCTTGAGTCCAAGATATTCTCAAGAGTCTTCTCCAACACCACAGTCAAAGCATCAATCATTTAGCATTCAGC
 TTTCTTTGTAGTCCAACCTCTCACATCCATAC

Fig.8A

ATGACCACTGGAAAAACCATAGCTTTGACTCTACAGACCTTTGTTGGCAAAGTAATCTCTCAGCTTCTCAATATG
 CTGTTTAGGTTGGTCATAGCTTTTCTTCAAGGAGCAAGTGTCTTTTAATTTTCATGGCTGCAGTCACCATCTGCAG
 TGATTTTGGAGCCCCCAAAAATAGTCTGTCACTGTTTTTATTGTTACCCCATCTATTTGCCATGAAGTGATGGG
 ATCAGATGCCATGATCTTAGTTTTGTGACTATTTTAAGCCAACTTTTTCACCTCTCTCTTTCACTTTCAACAAGA
 GGCTCTTTTAGTTCTTCCCTTTCTGCCGTAAGGGGTGGTGTCTGTATATCGGAGAAGGCAATGGCACCCCACT
 CCAGTACTCTTGCTGGAAAAATCCCATGGATGGAGGAGCCTGGAAGGCTGCAGTCCATGGGGTCGCTGAAGGTCG
 GACACGACTGAGCGACTTTACTTTCACTTTTCAATTAAATGCATTGGAGAAGGAAATGGCAACCCACTCCAGTGT
 TCTTGCTGGAGAATCCCAAGGATGGGGGAGCCTGGTGGGTGCCATCTATGGGGTCGCACAGAGTCAGACACGA
 CTGAAGCAACTTAGCAGCAGCAGCAACAGCAGCATCTGTATATCTGAGGTTATTGATGTTTCTCCCGGCAATCTT
 GATTCCAGCTTGCTTCCCTCCAGCCAGCGTTTCTCATGATGTACTCTGCATGTAAGTTAAATAAGCAGGGTGA
 CAACATACAGCCTTGTTGTACTCCTTTTCCCTATTTGGAACAGTCTGTTGTTCCATGTCCAGTTCCTAAGTGTG
 TTCTTGACCTGCATACAGATTTCTCAGGAGGCAAGTAAGGTGGTGTGGTATTCCCATCTCTTGAAGAATTTTCCA
 CAGTTTGTGTGATCCACACAGTCAAAGGTTTTAGCGTAGTCAATAAAGCAGAATTAGATGTTTCTCTGGAACCT
 TCTTGCTTTTTTCCATAGTCCAACAGATGTTGGCAATTTTATCTCTGGTTCCCTGCTCTTCTAAATCCAGCTTG
 AACATCTGGAATTTCTCAATTCATATACTGTTGAAGCCTGGCTGGAGAATTTTGAAGCATTACTTTACTATCGTG
 TGGGATGAGTGAATTTGTGTGGTAGTTTGAACATTCTCTGGCATTGCCTTTCTTTGGGATTGGAATGAAAACCTGA
 CCTTTTCCAGTCTGTGGCCACTGCTGAGTTTTTCCAACTTTGCTGGCTTATTGAGTGCAGCACTTTTACAGCATC
 ATCTTTTAGGGTTTGAATAGCTCAACTGGAATCCATCACCTCTACTAGCTTTGTTTCATAGTGATGCTATGGCC
 CACTTGACTTTCATTTCCAGGATGTCTGGCTGTAGGTGGGTGATCACACCATCATGATTATCTGGGTATGCAGA
 TCTTTTTTGTACAGTTCTGTGTATACATAGAATTCTGTGAATACACAGAAGAACACAGAAGAACCTTTATCATCA
 TGACTCCTTTTTTCTCCCACTGAATGGTGACAGATGGACTCACTTCCCTTTCCCTTTCTTCTTTTGTGTCTCTG
 ACTTTTTGACTTTTGTATGAAGTGTGATATTTAACTGGTTGAATCTGTGTTCTTCTGTTTTCTATAGCAAGTAGA
 TTCTGACTCAGAGGATACCAAGTGTGGAACCTATGTTTCTTACTATAGACTTAGTGAATCTACAGGCTAAATAA
 CTATCTTCTAGATTCTTTGCAGTATCTTCTCATTAAATATTTCATGAGGACACTGTGAAATCTGTTCTGATCATGT
 CTCTCTTATTAATGAGGAAAATGAACCTTTGAATGGACCTTGAGCAATTCACCTTCCAGATCTCAGCTAG
 ATTTAAATAGGGTCTGCACTCAGTCCCCCTCAAGCCTAGTGTGGATATGGGCCAGGTCTGACACTCACAGCTC
 AGTTGGATCATTCTCTGTGGAAAAATTAAAGTGGGTGAAGAACTCCTCTCTGTTGGAACTTGGAAGGCTTG
 CATGTGGAGGACAGGATCCTGCTGAGAAGTGGGAACAGCACTGTTTCTGGGGAGACCAGTGAGCCAGGTGAGG
 CCAGCACAGAGGTGTGGAGCCCTTGGGATACTGGGGTCTGTGCATTTAGTTGTTTGAAGAGAAGAGGGCCTTGGGG
 AAGGTGTGGGACGTGTGTCTGGGTACCCGAGTGGGTGCTTCTGCCAGGGAATCCTGGGTGCCATCGGCAAAAT
 CCCGTACAGCCTTCTCTGCTTGGACAGGACGTGACCTCCTTTTTGGCGTGGGAGATTCTCAGGGTACCCTC
 CAAAACTGTGAGGTTGTGGGAGTCAAGCCACAGGCCAGAGTGGTAAAGAACAGAGTGGGAGAGAGAATCTTGG
 CAGAGACTGTGCAGGGAGTAGGCACAGAGTCAGTGTGAGTCAAAAAGACAGTAGCCAACATTCCAGTGCAGA
 ACTCTCGTACCTAAACATGGCCACTACCTGTCTGCTGAGTGTGACTGCCAGGGCTTAAGAGCAAGTTAGCTCCT
 TTCCAACCTTTACCTTGACTGTAGAGACTCAAGCATTTGAGAGCGGGAAAAAAGAGCCATTCTCTTACCC
 CCTCTCTGGGCCCTTACTCCACCAACTCTGGTTACCACGGTTGCGGAGTTTCTCACCCTGAGGTGAGTTGGCA
 TTTCTGGGGTTGTGACGCCCACGGGAGCCAGAAACAATAGGCGCAGAATGTGATAAAGGTGAGAGAACTCCAGAG
 AGGCTGGAGGATGAACGAGCAAAAGCAGACAGTAAGCCAGGAGGGCGGGGATGAGTGTCTGCTGTCTGCATTCTT
 GTGGTAGAAAAAGAAAGTTTCCGGAGATTGAGTTCTAGTTCTGAACCTTGACAATGGAGATAGAAGCTTTGTCC
 GTAGTCTGTTTCTAGATCCATTAGTTATTAAGACTGCTGCTTCTCCTAGCTCTTCTAATGCTGTTTCACTTATT
 TCTTGAATCAGAGGACACAGAATAAGGCAGCGTTGGTTTCCAAAGATTCTGCAGCAGAGACTAGAAAACAAAGT
 GTGAGCAAGTGACTGGGATCAGGTAGCAATGAGATTGTTTGGCTAAGACTTGGGGACCCAGCAACTCTTGGTTA
 CAATCCCAAGCAGGTTGAAGACATCAGGGGTGTTGTCCCTGGGGGGGAAAAAATCTATGGAGTTTCCAACCACT
 CAGTTCAGTCCAATTGCTCAGTCGTGTCCAATTCTTTGTGACCCCATGGACTACAGCACTCCAGGCCTCCCTGTC
 CATCACCACCTCCAGAGTTCACTCAGACTCATGTCCATTGAGTCACTGGTGCCATCCAGCCATCTCATCTCTG
 TCGTCCCCTTCTCTCTGCTTTCAATCCGTCCAGCATCAGGGTCTTTTCCAGTGAGTCAGTTCTTCGCATCAG
 GTAGCCAAAGTATTGGAGTTTTCAGCTTCAGCATCAATCCTTTTTCCAATGAATAATCTAACCCTAGATGGTGATA
 AATCCCTTCCCACTTAACTGCCCTTTATTCCTGGTGATCTTTTGTGGTGCAGCTTAGCAGCAGAACAAAGTTTTC
 TTACTTAGGAGGGTAGCTAAGCTATAGAAAACAGAAAGACGCTTTTACTCCAATATGTACTCCATGTGAGGTCAT
 ATATCTTTTACGATAAGAATAATCTTTGAGACTGTTTAACTCAGGAGCCTTTACTTTGGATCAAGACCAACTGTG
 GGAATTCATCAATTCTGCAACCCACACAGTAGAAACAATAGTTTACTCAATGGTACTTATCTTGTATGATGCATG
 ATGTGTTTCAGTTTATATTCTTCATTAATAAAGCCTGACCAGCCTGTAAATTGATTTAAGTGCTGCTCAC
 GTATGGATCGTGACCTATGGTTTGAAGAAAGCCAGCTCAAGGCAATGGCACCCCACTCCAGTACTCTTGCTGGA
 GAATCCCATGGACGGAGGAGCCTGGTAGGCTGCAGTCCATGGGGTCACTAAGAGTCGGACACAACCTGAGCAACTT
 TACTTTCACTTTTCACTTTTATGCATTGGAGAAGGAAATGGCAACCCACTCCAGTGTCTTGCCTGGAGAATCTC
 AGCGACGGCGGAGCCTGGTAGACTGCTGTCTAT

Fig.8B

GGGGTCGCACAGAGTTGGACACGACTGAAGCAACTTAGCAGCAGTAGCAGTCTATCTTTTAGCCCAAACATAAT
 TTATTCCTTTATTTTCCCTTTATGCCATTTACATTCTGTGTTTCATCCAGTTTTGTTTTGTTTTTTTTTAATAACAC
 TGTGTCCGGTGTATTACTGTATCTTAGGGAGCATGTTTCATTTTCATCAGCTGGACCACAGGCAGACTGATCTC
 ACTGTATCACTCAGGTTAGTAGGTACACAGACCGTCATCTGAAGTGGCGTGAGTCCACTCTGACCAGTAGAAGC
 CACACAGAGAACGCAGGCTTCTGTGCCTACAGGCTCAGCCAGAGTCATCGCCTCCAGGGCAGGTCTGGCTGTTTC
 AATGGGAAAAACGTCTCTCGAGGGCTTTTCTGGGCTCTTTCTGTGCCATCTCTAGGACAGTTGCCTGAAGAT
 CATGTCTTTTTCTCTTTCTGGCTTTCTGGCTTCTCCGTGGTCTCTCCCATTCAGTGAAGTGCCTTATTCCTGAA
 CCAGAGTGTCTGGTTAGATCTTATAAGCTTCTCAGTGCAGCTCACAAATTGCACACTCTTAATGTTTAAATGTAAG
 AAGCTTGGCCCTCATTTTAAACAGTCAGGTTTTTAAACCTAGATTCTAATTCCTAGATTTCCTCAAATTATTA
 AAGATGACCTGTGCTTCTTATCTAGGCCTGAAAGAAGGATGTGCTCATCTGAGTTGTTGCTTGTAGTCTCTTTTA
 AAAGGCATGGTTAACCATCTACTTAGTATACCCGCTGCTGTGCTTTTCTAGGGCATGGTGCCAGACGATGCTGT
 TGAAGCCTTGGCTGGAAGCCTGGGCAAAAGGAAGCAGATCCAGAAGACGGAAGCCTGTGGAGGATAAAGTCAA
 GGTAAATGGCAGCTCAGAACTTCTAGAAAGGAGCTATCATTGGAGCTATGATTGGTCACCCCAACTCCCCAAATA
 AATCAGTTTAAAAAATAAATTCCTGCAGCACATGGCCGGCTGGACCCCTCTGTTACGATGCATGTCTGCTTA
 ATAAGTCAGCAGAGCGAGCGCCAGTGTTTAGTGTGGCAGGTGGGCGGTACAGAGATCGGGCTTCTGTAG
 TCTCATGTTGTCCACCCGGTTTTCCATTGCCAAGGACCAAGGCGGAAGTGGTACTGCTGCGCTCTCACAGCT
 GCCCCTTAACC CGCAGGCTGAAAACAGTGGTCTCTGAGCCTGCTAACCAGGAGGATACAGCGTCAAAGATACA
 GTTGTGTGGGAACCTCACAGTGGAGTCACTGACCTCGGATTTAGTCACTCAGGCATCCACTTTATCAGAGAACG
 AGGTACTAACACTGAAAACGGTTTAATTTTTATTCAGGAAAAAGCCAAAGAAGAGGATCGTGAGAACTTGGTGA
 AAAAGAAGAAACGATTCTCTCTGATTACAGATTAGAAGAAGCCAAGGTAAACAGGCCGGGATCTTTTTTTCTAAC
 TCATTTTTCATTCATATTGAATTTTCATATACATAAAGAAGGGACATGGAACAGACCTGGGATATAAAAGTCATGTT
 TCTTAGTATTTCAAACATCAAGTAGTATTACTTAACCTCACATCCTGTTAGCACACCCACTGTAAAGCTAAATAT
 TTGTCACTGGCCTGTCACTCCTTCTTAGGGGAAAGAAGTCTAAGATTTTCTGTCTGATACGTGGTTTGTACA
 ATACCAGAAATAAACAACCTCAGGAAATCTTTATATGGATAGATAGGAAACCAAAGGAAAGTAAAGTCGCTCA
 GTCGTGTCCGACTCTTGTGACCCCATGGACTGTAGCCTACGAGGCTCCTCCGTCCATGGGATTTCCAGGCAAG
 AATACTGAAGTGGGCTGCCATTTCTTCTCCAGGGGATCTTCCCAACCCAGGGATCGAACCAGGTCTTCCGCAT
 TGCAGGCAGACGCTTGACCGTCTGAGAACCAATTACTTCTTAAAGAAGGATTTACAAAACCTTAGATAAAGTTGAA
 AAACAGCCCATTGACATTGTATTATAGTGAGCAATATGGCCTTCTTGGTGGCTCAGATGGTAAAGAATCTGCC
 TGCAATGCAGGAGACCCAGGTTCAATTCTGGGTGGGAAGACCCCTGGAGGAGGAATGGCATCCCACTCCAGT
 ATTCTGCTGGGAGATTCCCATGGACAGAGGAGCTGGCAGGCTACAGTCTATAGGGTTGCAAAGAGTCAGACA
 TGACTGAGCAACTCAATATAGCCAGTGATAAATTGACCCATTTCTTTATTTTGTAGTATACAGATTGTACAGAA
 AAGTGATATGGGAGAGACCACTACTTAAGAGAATATTCTACAGACTTTATGAATAATGCAGGAAAATCAAATTG
 ATACTCTTTGTATTTCATCGATTTGCACTAGTCCATTTATGTGGCAAAGATAGGAACATACTGAGCGCTCTCCCT
 GATCGAAAGTTTATTATGCTGGGTCCCTTTAGTCTGCTTCTTTCTGCATTTACAATTCACTCGAAACCATGGC
 ATTTCTATTTGTTTGGGCTGTATGTTTATAAGTTCTTAGACCAGTGGCCGTAAGTAGCTGTACACTTCGAT
 AAAGCTGGTCAAGTGCTGTAACTCTCTTTAGAGTAGGAGGTTGGAGGTTGGTTACTTTCTCTCTCCCTTGTCTG
 GATTTTTCATCGTAACTTCTAAGAACGTGGCCAGTTAGCTTTCTGACCCTTTATTTCTGTGGTTGTCTCATCCTC
 AACTATGGATCTCACCCATTATGAATTACCTATGGCTTCATAACATGTCCAGGTGCTGCTTGATTTCCGGGTTAA
 ATCTTGTGTTGTTCTTCAACTGCTCCCATCAGGCCTCCTCAGGAGAGTCAGTCCCCAAGGTAGACTGCAGGTGT
 CCCCTTCCAAACTTGGGATGATCTCAGTGTGAGTGTGGAACAAGTGTGCCCCCTGTCCCCACTCCTGTGTGCCCC
 TTGGGAGATTTGCTGCCTTGGCAGCTAAGGGGTGGCAGAGAGCGCACCCCTGGGCTCTGTTCCCTCTCTAACTGGA
 CAGTGATTAGAAGGTCCATGAGGACGCAGAGGACGTGGAAATGGGGGAAAGCACTTTGAGAAAAGGGAAGCTGGT
 GATACTCTTGACCTCTGCATCTCCTCTAATGTTAAAGCAAATGAAAACACCTTGATCAGTGTATCAGGGAACT
 GTGCAGTTGACCTTGAGTGACACGGGTTTGAATTGCCTGGGTGCACTTCTGTGGGACTTTTCTCAGTAGTAA
 CACTACAGCATCTCACCTTCTGTGGTTGACTGAATTCACAGATATGGAATTGAGGATACTGAGGGCCAGGTGTAA
 ATTACACATGGATTTTTTTCAGCTGCTTGAGATGTCCGTGCCCTAACCTCTGCAGTTATCCAAGGGTTAACTGTGT
 GTACAAACAGAGAATCTTGTGCTTACACATTTTGTGATGTGTAGATAAATGGGATTTTACCATTATTTATAAAT
 CACCTTTTACTAAAGAGCCCCAAATGCTTTATTTTGTCCATTACTCAGAAGTGGCACTTAGGGTCATTTGAAGG
 TTGGCATCATCTGAAGTGAAGTTTGTAGCATCATACCCCTAGGGTGAGTTTCCAGTGCTTGGTATCGGTGCT
 TGGCACAGAGTCGGTGACAGTGGTTGGATGCACTGCCTCTAAAGTTACCCGGCTTTTCCAGGCAACTCGTGCA
 AAGCCCCCTCCCCAGGCCAGCGAGGGTCTCATCCCGTGTTCCCTCTGTGTCTGTGCACGCTGTGCCACACACC
 ACAGTAGCCATTGTCCGTGGGCTTCTGGAGTTGGCAAGCCCCATTAGCAGCCCCACGCTAGGGGATGTCTGTC
 AGCTGTTTCTTGACATGAGCTCATGAGAGCCTATTTCTGGTGAGAGAATCTGTTTCTGGAATGAGGAATCGAC
 TTCATAATGTACATCATCAGCTATAACCTATCAACCTCTGATGCTATTTACAGAGTGACACTTAGAATACATGT
 GAAATGACTCAATGCCATGTTTTTCTCCACC

Fig.8C

TGCACAGGATAAAGACGGAACCACTGCTGCCAAAAGAGGTCAAGGAACCGCTCCCA GT AAGCAAGCCAGTTT
 TTCTCTGAGTGTATCTTTCTCCTTTAGTCCTGATGGTTAATTTAGGAAGGGCAGAGCTTTCTGTGTAATGAGCTG
 ATCAGCATATATGCAGTGACAGTACTTTGGAGAGGAGACTGACGAGGTAGCGTTTGCTGACACGTCTCCTGGTTC
 TTTGCGAGCCCTTGAGTGAAGACGTCTCTCGATGCTCTGTCCAAGGACTTCACTGTCCCCTCAGACACATCATC
 GCCTGT AAGTCTCCTGAGAGTCCTGGTTTTAGTGCCTCACTTTTTAGGGTAGCAGAAATAAGTGGAACCTGTGA
 CTTAGAATCCGACATGAGAGATGAGGAAACAGTCATGAAATTAGCGGCCCTCAACCCATTGTGGCCATTAAAGAT
 GTATTCCATCAGGAAATATGGTCCAGCACACTGGATGTTTTCAACAGTGTAACTAATAGAGCCAGTGGCCGTGCA
 CTCTGTCAAGTATATGCATGGAAAGAGAATTTAAACCGAGCTGTCCAATATATTTGAAAAATAAATGCGTAAAG
 ATATGGAAAAAAGTGAAGCATTATTTATTTTAGGGAAGATTAAGCCAATAAGTGCCAGGATTTATTATACAGT
 AATGTTAAAAAGTAAAGTTCTGATTAGAAAGTGAAGCCAAAAGCCAATGGGAAAAATAACCCATCGATGACAGAT
 TTAGTATATCTTTAAAGTGAGAAACGTCATTATTTAGAAAGATTCTAAAATAATACTACTTAACATTAAAGAA
 AAAATGATACAACTGAAACCTCATCTGTGTGCCAAATATTAAAGCAATAGTAGCACTTATTCTGGAGAAGGCAA
 TGGCAGCCCACTCCAGTGTCTTGCCTGGAGAATCCCCGGGACAGGAGAGCCTGGTGGGCCGCCGTCTATGGGGT
 CGCACAGTCAGACACGACTGAAGCGACTTAGCAGCAGCAGCAGCAGCACTTATACAGAATTAAAAATAGGATAAG
 TATCTAATTTAGGCAAGCAAGGCTACTGGAGACAGGCATACCCATTCACTGCTTTTAGCAGCATGATTTTCTAGC
 AGCTTTTAAAGAGAAGTTATTTGGCGCAATATAGCTAGCACCATACAGAGAGAGAGCCACTTCTAAAATTTACTC
 AGAGCAACTCATTACAAAGAGCAAAGCTAAAACACTTCCATCAGCCTAAATGTGCAACAAGAAATGTATAAGCCA
 TAAACTCTTAAATACTTAGCTTATTGAATATTTAGGATAATAAGCAGGAAGAGAAAAATAGGAAAGTGATAAAAA
 TAATGCTTGGAGTAGAAAGTAGAATAAAATGGCTCTTAAGCTGCTGTGTTGCTGTGCTTCCATTCCGATCAAGAG
 AACAACCATCATAAAGCTGTCCCGAGGGTTTTATTTAAGTAATGGCATTGTGGGCATAATTTGCTACAACATATT
 ACGGGTCAAACTCTTTTTCTGTGATTGTTTTCTTACAAAAGCAATTTGAAGATGTAAACTTTTCAGCTGTCTCTC
 TGAAGTGGTTTTCCCAAACCCAGCTCCAACCCAGGCAGCCGGTCCACCCCCAGCACTGCG GT AAGCAGCAT
 GGCTTATAAGTACTTGGTTTTTGGAGCAGCACTAGAGGGTCAACAAAATTGTTACCACAGTCACCTTCTGTTTCA
 GAAGAGGTAAGACTGGCAAGATAGGGAGATGACAGAGATTATCTAAAATGATTAAATAAAACAATTCCTGTGAAA
 GTATTTTACACACACTGAAGGAGCTTAATATATTTGTTGCCTTATTAGAATTGAAGTGCAATAATGCATATTGTGT
 ATATATTAGATTTTAAATTTGTTTCGCTTTCTTGACAATGTTTCCTTCTTATATTTTTCTGATACATGCT
 GTAAGTACTACTCATTATTTAACCTCAAGTTTACTTTAAGGAATTGTTCAAGTGTGTTGAGCTTCTCTCTGCCACC
 AAATCACAGATTCTGGTAGAAAAGCATTGTTTTCTCCAGTCATACCAGTAAAAACTTGTCAATTTTCATGCTA
 GTAATGGAAAAACCTCAGTGTCAATATCGTTTTAATTATTTACATGTTCTGAGGCAAGACTGTGGGACCAAAGCT
 TTTCTGACTTCTTGATTAATTTCTTTACAATACCAAATTATACATGATGCAGGAGACCACGGCACATCAAAA
 GATAAAGAGATTGCCCTAAACCAAACCCCACTCCTTGGGCTGAGATAGAAGTGCTTTATGAGCTAGTTGGAGA
 CGAAGTAGCTTGAGTGAAGTCAGCTAGGCTCCCTTGCCCTTTTGATGTAGCATCAAAGCTTGCTTCTCTGCTGCACT
 CTGAGATTCTGAGTGAATTTTCTCTGTGGGTGCTGCTGGGATCATCTGAAAAATATCCAGAGGCACAAATGAC
 AAAGCCCAGATCTGGCTGTGGGCGGCCACCATTGCTGTCTATAATCCTGGGCGCTCATGAAGCAGTTAAACC
 AGTAAAATGCCCTCTCTCCATGTGTCTGTGTATGTCTGCACACACACCACAGTGAACAAAGTCCCTGTGCA
 CAGAAGGAAGACTCATTAAATAACAGGCTGCCGTTTACACTTCAGTTCTTAATCTCGGATTTTAGGAGTCCCTA
 GAATAGTGATCGAGGCTGTTAAGAAACATCCACAATCAAAAAACCTTTTTGGCCTTCCACAGTATGTTTTAGAAT
 GATCCCTTTCCCATCACTTCTCCTCGTCAGAGTTTTTAGTTTTGTTAATATTTTATTTGAAGCTGTCAATAAAAAC
 TAGGTATAAAGCTTTATTTACAACAATACACTTGGGAATCAGAGAGTAACTTAGTGATTTTGTGGGGAACATAA
 TGTTACATGTATTGAAGGCATCATAAGACTTCTAGAAATCGGGCCAAGAACCCCATCTTTTAGGAAGGACCTAT
 TTTCTAGGAAGAAAAATGCATCTAGGGGTGGGAGGGTAGACAGGTGCAGAGAGGTGCTGAGTGAAGCGAGGGAC
 AAAGAGCCGGCTCTCTGCCACCTGCCCCCTCCCTGGGGAAAAACCCACGAGCTTACGGAAGTCTCAGAGGCACTG
 TTGGGGCTGCAGGTGACCTAAGGTGGCTAACCACTGACTAATTCAGATTTACTTTTTAGCAGGTGACAACAAAG
 AACTTGACGATGCCCTGGATCAACTTTCTGACAGTCTCGGGCAAAGACAGCCTGATCCAGATGAGAATAAACCCG
 TAGAGGATAAAGTCAAG GT CAAGAAAAAAGTCATTTAAATTGAAAGTTTTTATAGCCCTCTTTTTTGGGGGGG
 AGGGGGGTATTAGTTCTAAAAGCTCTGTAGTCTTACAGACCCGTTCACTTCAGCTTCTTCAGCGTTACTGG
 TTGGGGCATAGGCTCACCGTGATATTGAATGGTTTGCCCTTGGAACGAACAGAGATCATTCTGCTGTTTTTGGAGA
 TTGCATCCAAGTACTGCATTTTCGACTCTTTTGTGACCATGATGGCTACTCCATTTCTTCTGAGGGATTCTGTC
 CCACAGTAGTAGATATAATGGTCATCTGAGTTAAATTCACACATTCCAGTCCATTCTAGTTTGCTGATTCTAGA
 ATGTGATGTTCACTCTTGGCATCTCCTGTTTGACCACTTCCAATTTGCCCTAACATTCCAGGTTTCTATGCAAT
 ATTGCTCTTTACAGCATCGGACTTTGCCTCTATCACCAGTCACATCTACAACCTGGGTATTGTTTCTGCTTTGGCT
 CTGTCCCTTCATTCTTCTGGAGTTATTTCTCCACTGATCTCCAGTAGCATATTGGGCACCTACCAACCTGAGGA
 GTTCTCTTTTCAGTATCCTATCATTTTTGCCCTTTCATACTGTTTCATGGGGTTTCAAGGCAAGAATACTGAAGTGG
 TTTGCCATTCCCTTCTCCAGTTTTTCATTCCAATCCCAAAGAAAGACAATGCCAAAGAATGCTCAAACCTACCGCAC
 AATTGCACTAATATCACACGCTAGTAAAGTAA

Fig.8D

TGCTCAAAATTCTCCAAGCCAGGCTTCAGCAATATGTGAACTGTGAACTTCCAGATGTTCAAGCTGGTTTTAGGA
 AACGCAGAGGAACCCAGAGATCAAATTGCCAACATCTGCTGGATCATGGAAAAAGCAAGAGAGTTCCAGAAAAACA
 TCGATTTCTGCTTTATTGACTATGCCAAAGCCTTTGACTGTGTGGATCACAATAAACTGTGGAAAATTCTGAAAG
 AGATGGGAATACCGGACCACCTGACCTGACTCTTGAGAAACCTATATGCAGGTGAGGAAGCAACAGAACTGGACA
 TGGAAACAACAGACTGGTTCCAAGTAGGAAAAGGAGTACGTCAAGGCTGTATAATGTCACTCTGTTTTATTAACTT
 ATGTGCAGAGTACATCATGAGAAATGCTGGGCTGGATGAAGCACAAGTTGGAATCAAGATTGCTGGGAGAAATAT
 CAATAACCTCAGATATGCAGATGACACCACCTTACGGCAGAGAAGGCGATGGCACCCCTACTCCAGTACTCTTGC
 CTGGAGGGTCCCATGGATGGAGGAGCCTGATGGGCTGCAGTCCATGGGGTCGCAAGAGTCCGACATGACTGAGT
 GACTTTACTTTCACTTCTTACTTTTCATGCATTGGAGAAGGAAATGGCAACCCACTCCAGTGTCTTGCCTGGAGA
 ATCCCAGGGACAGGAGGAGCCTGGTGGGCTGCTGTCTATGGGGTGCACAGAGTCCGACACAACCTGAAGTGACTT
 AGCAGCAGCAGCAGCAGCCACCCGTATGGCAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGGTGAAAGTGAAAAG
 GGAGAGTGAAAAAGTGGGCTTAAAGCTCAACATTCAGAAAACGAAGATCATGGCATCTGGTCCCCTCACTTCATG
 GGAAATAGATGGGGACACAGTGGAACCTGTCAGACTTTATTTTTGGGGGCTCCAAAATCACTACAGATGGTGATT
 GCAGCCATGAAATTAAAAATACGCTTACTTTTTGGAAGGAAAATTATGACCAACCTAGATAGCATATTAAAAAGCA
 GAGATATTACTTTGACAACAAAGGTCTGTCTAGTCAAGGCTATGGTTTTTCCAGTGGTCATGTATGGATGTGAGA
 GTTGGACTGTGAAGAAAGCTGAGCGCCGAAGAATTGACACTTTTGAAGTATCATGCTGGAGAAGACTCTTGAGAG
 TCCCTTGCAAGGAGATCCAACAGTCCATCCTAAAGGAGACCCGTCTGGGTGTTTATTGGAAGGATTGATGCTG
 AGGCTGAAACTCCAATACTTTGGCCACCTCATGCGAAGAGTTGACTCATTGGAAAAGACCCTGATGCTGGGAGGG
 ATTGGGGACAGGAGGAGAAGGGGATGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATGCACATGAGT
 TTGGGTGAATCCGGGAGTTGGTGATGGACAGGAGGCTGGCGTGTGCGGTTTATGGGGTCACAAAGAGTCAG
 ACACGACTGAGCGACTGAACTGAACTGATAGCCCTGACTTGGCTGCAATCGGTCTTCCCTAGTGGCTCAG
 CTGGTAAAGAAATCCACCTGCAATGCAGGAAACCTGGGTTTGATCCCTGCGTTGGAAAAGATCCCCTGGAGGAGGA
 ATGTTAACCCAGTATTCTTACCTGGATAATTGCTCATTGACTTGATGGACATGAGTTTGAGTAAACACCAGAAGT
 TGGTAATGGACAGGGAAGCCTGGCATGCTGCAGTCCATGGGGTCGCAAGAGTTGGACATGACTGAGTGACTGAG
 AATAATTACTTAAAAGCTTATTTTCACTTTTAAATACAAGTAGAAAATTATAGTGTGGAATGACAAACAGAAGC
 ACTGGGTTTGTCTTTTTCATCAACTGGAAATCTGTTGACATTGTTGCTCTAAGTTACTTATAAACTTAGAGG
 GACTATGTTTTGCTCAGAAAATATCCAATAAATGCTGAATGCTGCATTTTCGAATGTGCTAGATGAAGGTAGAAC
 AGAATAGTAATTCTGTATGTTCAATGTTTACAGGAAAAAGCCAAAGCTGAACACAGAGACAAGCTGGGAGAAAAG
 GATGACACCATCCACCTAAATACCAACATCTTTTGGATGACAACAAGGAGGTAAATGAAGGTGGTGTCCAGGTT
 GGATCTATACTCCAAAGCTTTTGGATTCAAACCTCACTTGAACAGAACACTGTAAACAATAGCATAAAAAATGTAA
 CATGACTGACACATTTAATCTGTCCCCACTGAGCAGGCTGTCTAGAATTCTGTTAATATTTATTTAGGCCAGTA
 TCCTTGCATGTTTACTGTGCTGGGCACTGTGGCTAGAGCAATTGAAGAATGTTCCATCACTCGCTGATGCTTAA
 ACAGTGATTCCGAGGCAGGTACACAGAAGAATTCTCTGCAAGAAAAGGCTGTAATAACAATGAATGATATTGTA
 ATCTTACCTGACACAGAGTCTCTCTGTTGAAGCAGAACAGAAAAAGAGCTAGGATGGAAAATAATTATGGGGCAA
 ATACGATCCAAGGCCTGCTGTCTCTCTTTCTTCCCCAACACCACCACCGGTGCTGTTGAGAACGAAGACCGG
 CAAAACCATATTTCCAATTGAAAGCTCTGTTGCTCTATCTCAGTACAGCAACAGTATCGTCAAGAAGTTTATCTG
 TCTTTGTGTCTCTCTGCTCAGCCTGCTGTGTTCCCTGCCCCCTTCTGTTTGGCTTCTAGATATTAAATAACCTT
 TTCCAGATCTGTTTCATCAGATAAAAAACAAGGAAGTCCAAATGACAGTTTCTGTCTCTGCACTGATGATTTGTGAA
 GGGTTTGAAGTGTATAGAATACTATCAGCTTCACTGTCTTCTGTTACAGAAATCAGTTCCCCCTCGCTTGCAATCAG
 CAGTCATGACAGCGCTTGACACCCTGCTAGAATGCCCTGCTCTGATGGTCCCCTGCTGACCTGTTGTGTTCTT
 TGGGAAATGTGTTTATTAGTGTCTTGGCCTTTGGCGATGGGAAAATCCCTGATGAGTTGCTCTTGGAAATCCTCAC
 TGAAAAAAAATCACCTACAGACCATTTCAGTGCCTCAAATCCTTTTGAAGTGAATAAGCTGTAGTGATACATA
 TATAGTAAAACAGCAAATAAGACATAAAAAACACAAGTTACCTTTTTGTGATGTTTAACTAGAGTTAAGCAGTGAC
 AGAGTTGCTAAAATAGTATTTCTTTCACTTCCAGGAACAGATTATCCTTTGCGTTTTCCCCAAGGCTATTGAAA
 TAGTTATGAGATAGTCATTAAGCAAAGCTGGTGAAGAGATAATACGGACCCAGGTTTATCAGTTATGCAAAGTAT
 ACAGTTGTCCCTTGGTATCTAGGGGGGTTAGTTCCAAGACCCCAAGAGACCAAATCTGAGGGTACTTTAGATC
 ATCTCTGCATTACTCAGAATACCTAATACAGTGCAAGTGCTACATAAGTAGTTGCTAGCATGTGGCAAATTCAG
 ATTTGCTTTTTGGAAATTCCTAGAATTCTGTATCCCCAAATAGTTTCAATCTGCTGGTTGGTTGAATCTGCCAAT
 GCAGAAGCCACAGATGCAGGGGGATGACTGAACTCCACTGGATAAAGATCATGTAAATACTGACTTACACCTGAG
 TAAAACCTTTACAATTTGTTTAAACCAACGGTCTGCTTTGTTTCTAAGGGCACACCCGGGAAGCCAAAGGCATCAG
 AGAAGCCCAAGGCATCAGAGGTAAATATCATAGCTGTGATTTCCAGAAACAATTTTTTAACTCTAGCTGCAG
 CTTTCACATTTTAGAAATCTAAATTCAGTGAATGCCTACCTTGTGCTTGGTCTGATGCCAAGGCATGTATATTA
 TATATTATTTCTGTCTTCCCAAGGCCACTCAACATAGGCAGTTCCATTTCCAGATTAGAACATCATAGCTCAG
 AGAATACTCAGCACCTTGATATAACAGAGTGGAAGAAAATGGAGCTGTCAAGATTTTCTGTTGTAAGATCTCTTT
 TATAATAGTTTGTAGGAAAAGAATGCAGCA

Fig.8E

AGGCTTTCTTTGAACACATTTTCAGGGCTGTGGGGAAATAACCCACAAAGCATCTTCCCTTGCTTCTCCAAAA
 GCAAATCTTGATGGTTATTTCTGTTGATATCTCAGAAACCTGCAGGTGCCAGGACCCCATTTGATGCCCTCTCAG
 GGGACTTTGACAGCTGTCCCTCGACTACAGAAACCTCGACAGACACACCAAAGGTACTGGGCTTTTTCTTCTTGG
 CTTTTGTTTTTAATGTATCTCTGTACAAACAACACACACACATTATATAATTTTATACCGTAAAAATAAACTGA
 TTAATGCCCTAGGAACTTGAAAAAATTCTATTTCTCCCAAACACCATTATTTGGTGTTATTTCTTGTCTTTT
 TGCATATGTAGGTTTACATAATTATACCTTAGATTCTATATATATATATATAGAATATATATATTATATATAGA
 ATATTCTATATTCTCTACTTAGATTCTATATATAAAACATTTCTTAAGAGTATATCATAAATGTTTTCCATGTTG
 CTTTATAAGCTCCATTGCCACCATTTTTGATTGCCACACTATATTGGCTTGTAACAGTACGACTTTATCTGTGGT
 CATTAAAAAGGAGAGCAGTCCCTCACTGTATACACCATGCCCCCTTGCCCTCTGGGCTTTCCTCCCTTTTCTGCC
 CTGAGCAGCTCTCCTCCTTACCCCTCAGAGACCAGTTCAGCATCAGCTGCTCCAGGAAGCTCTCCCTGTCTGCT
 TCTCCCTTCCCAGGTGAGGGCTTCTCCTCTGCCATCCCTGCGTTGTGAGTTACCTGTCCCATGGGCCCTCACTCC
 ACTGAGCTGTGGGCTTTAGATAGCAGTCATGTCTGGTAGCTCTGTGGCCAGCATCAGCAAAATGCCTGGTGCA
 TAGCAGAGGAGATCGAATGTGAATGGGGAAGGTGAATCCACCAAATAGGAGGAAAGTGAAATGAGTGTCC
 TCCGTTGTTATTTTCAAGATAAAATGCATGTGGTTTCAGCGGGTATTAAACATATTTCCAGTCACTGAATTGCG
 AGTCTTCCCTATCTTGCGTGATTTAATTTCTAGATCACATCAGTATCCTATTTTCAACTCCTAGAACCTCATCA
 ATAATGCCTTTTGGTGAATGGACTTTACTTTTGTATGTGACTGGCCCTAGCTCACTCTGGTTGCTTCCACTGTT
 AGATTATTTCTCCATAATTCTGAAAGCAGTGTCTCTTGGCACATAGAAACCAGTCTGATAGAACAGGCTCTA
 TTTTGATAAGCAAAATACAGTCTGAGTGCATCGTGAAATTTGGCAGGGCGTCAACGTGAGTGCCTTTTCTCAT
 TTTTTTCCCTTCTTACCTCCCACTTTGAACACCTGTGTGTACATTCACTGTGGTTCAAAAAGTCATGACTGTGAA
 ATACTCAATTTTTCATTATTTAACTGGAGATTCCAGGCTCCTAAAAAGAAGAGTCTGGATCGATTGTCTGCAG
 CAGCAGTTAATTGCTAGATGGAGTGTGACTGCAGCAGCTTCATTACATCTCCTCTGCTCTAGGACAAAGACAAG
 AAGCCTGCTTCCAGTGCCGAAGCACCTAGGAATGGCGGGAAAGCAAAGGATTCCACAAAGGTAAAGCTC
 GGTCGTGTCCGACTGTTTGTGACCTTCATGGACTGTAGCCCCCAGGCTCCTCCATCCATGGGATTTTCCAGGCA
 AGAGTACTGGAGTGGGTGGCCATTTCTTCTCCAGGGCATCTTCTGACTCAGGGATCCAACCAGGGGCTCCTGC
 ACTGCAGGAAGACTCTACCATATAAGCCACCAGGGAAGCCCCCACCAGAGCAGTAAATAGATGACAAATGAGCTT
 GGCAATTACAGGGAACGGCACGTTCTTAGTCAATTTACTTTCTGCTTGGGGTCTATTTAGATAAGACCATTTTA
 AGTACTATCTATACCTTAGAACATCTTGCCAAAGAACTGAAGTTTTAAAGATTTTTTTTTTAATTATTTATTTT
 AATTGGAGGTTAATTCCATTACAATATTGTAGTGGTTTTTGCCATACATTGACATGAATCAGCCATGATATATAT
 TATTTTTTAAAGTAATCGATACTTAAGTATCAATATTACTTATCAATAATCAGTATTCAGTAAGTATCACTTAC
 TGAAGGAACAAAGATTCGGAACACAATTGAAGCGTGTGCATTGCCAGCCCTTGTACACAAAGTCCCAGTGTCA
 GCTGTGCTGGTCTGGGGTGAGCGGAGGCCCTCTCCCTTCATCAGCCCCAAACGCCCTTGCACTCCCTTGTGTGCTAC
 TCTCCCTTACCGTGTGCAGAAAGCCTCGCCTTCTACTTAAAGAGAAAAGTGAAGGTGCCTTGCAAGAAATAACCA
 AGACCCCTTCCATCTTCTCCTAATAAAACCACTCTCCCTCCTTGTCTGCAGCCCTTCTCTCTCTTTCTCAA
 GGGGTCACTAACTCTCTCCAGGCTCCTTGGTTGCTCAGCTGGTAAAGAATCCACCAGCAATGCAGGAGACAGG
 GTTTGATCCCCAGGTTGGGAAGATCCCCTGAGAAAGGAAAGGCTACCCATTCCGGTATGCTGGCCTGGAGAATT
 CCATGAACGTGTAGTCCATGGGGTCCCAAAGAGTCACTCCTGAATATTCAGTGAAGGACTGATTCTAAAGCTG
 AAACCTCAATACTTTGGCCACGTGATGCAAGAACCAACTCATTGGAGAAGACCCTGATGCTGGGAAAGATTGAA
 GGAGCGGGGAGAAGGGGACAAACAGAGGATGAGATGGTGGGATGGCATCACCAACTCGATGGACATGAGTTTGA
 AGGCTTCGGGAGTTGGTGTGATGGACAGGGAAGCCTGGCGTGTGCTGAGTCCATGGGGTCGCAAGAGTCCGACATGG
 CTGAGCAAGTGAACGTGAACGTGTCTCCAGTATCTCTGTCTCAATTCTCTCTTGTAGCACAAACATATTCAGC
 TCTTCAACAACCTCAAAGCCCTATTCCCTCCATCCCTTCACTCCTTGGAAGATCTGTCTGTGACCCCTGCTCCCC
 CTCCTTGGGCTCCCCCTGAGAACAGCAGTGACATGCTGGTGTGTGCTGTTTCTTATCAGCCCTTAGTTTGCT
 TGTTCTCTCGGGGTGTGCTTTGGGACGTTAGATTTGATGCTGTAATGACTGAGACTTCGGGACGGGTGAATATTT
 TATAACTTCACTGCCCAAAGACCCAGACTAACACTTCTACCTCACATTTCCACTTGGATAAACATTGACAGTA
 TAGATCCGTATACATTATTAGATCAAGTCTAAATCGCCACCAATTATTAAGAAAAAGTCAACAGTTTGATTACG
 ATGTTAGTCATCATAACAGCAAATTAAGACATGTTAAATATGTGTGCGTGGCAAAGGAGATGACACCACCAGAG
 AATGTAGTTTCTGTAAATCTTGGAAACAGGAAAGTAGTATAGAGAACTAGAGAAGCTCTACAGTCTCTGAGATCC
 AATATACTCATATTAATACTAGTAGTCTTCTCCCAAGTGGTACCTCATTTTTTCACTTGTATAGAAAGCAG
 AACTGAGACTAATGAGCAGATCTGGTTTCTTAGAGCTGGGCGTTAATGCAGCATGGCACTTTGAAAGTTTCTCTAT
 TATTAAATGACTCAAACGTGTGGTTAGGAAAGGGATGGAAGTTGGATTGGATAAACAACTTCTGGGTCCCCCTCCA
 ATTCTAAAACCATAGTATTCAAATGGTTATGACTGTTAATTCGAATTATAACGAGAAATAAAAGATACAGGTCCA
 GAAACAATAATCCAGCTGCATCCAACTCTTAAAAAGCAAGATAGTGAGTTCTGGTTCAAGGTGGTGTGACAGAA
 AAGATCAGCTCGCTCGTACCATGGAAACACCAGATCTACAGCTGCATACACAACAAATTTCTCTAGGAAAAAAC
 TCCTACAGTGATATGAATGACTCCTATCCATCAGGCCAATAAAAGAAGGCCAGCATTAAGAAGGTAGAAGAGGC
 TGGGACAGACTCTTGACATAAATCCAGTCCC

AACACAGTCAACTCACCGTCAGAAGGAACTCAAACTCAGAGCTTCTCCTGAGGAGCAAATCAGGTTGAACCCC
CAAGGCAGCACCCCAACATTTAAAGACCACCTGAGAAATGAGCACACAAAACCTCTAGCTTTGGGAAGCCAATGG
AGTTTGTGTCCAGAAGGCCATAGCAAACCTGAGGAAAGCTTCTTCAAGAGCTCATGAGCTCAGACGCACCCGCCCC
CCCAACCAAGCAATCAGAGCTACAGTCTTAGCCGGCCACTATCCCCAGAGTGTGCACAGACCACAGACCAAAA
CACGCCCCAGCTTTCTGTGTAAGAGGCCCTCCTTGCTTGTCTGAAGCTCCGGTGTGAGAAGCAGGCTTCCGATTT
GGCAGAAGATGCCCAGATGTGCTCTTAGGAGACATCAGCTGGGTGCTCACCACCACTGAGCTATTTCTCCATCTC
ACTCCAGGCCACCAGTATCTCCAGAAAAGGGGCTTTGACGCTTACCTAGACGCCAAATTTTTGTAAGTCCCCC
CAGAGGGATACTTCTGATCACCTGCTGTGGCCCTTCAAGACTATTATGTTTGTATATTGAACCAATGAAGATAA
AATAGGCTGCTTACAGTGTGGAGATAATCAAGAGTCAGGGCAGGGTTTAAACAGTAAAGTTCATCTGCTACACG
AGGCTACTCCTTCAAGACTGGAAGAGGTGGCTGTTTTATCTAATGCACACAGACCAACACAGAGAGTCAAAGAAA
ATGAAGAAATGGAAGAATACATTCCAAACGATTAAACCTCAGGATAGAAACCTTAATGAAGCGGTTTACCTGAT
GAGTTCAAAGTAACAGTCATAAAGATGCTCACCAACTTTGGGAGAAGAATGGATGAACACAGTGAAGCTTCAAC
AAAGATATAGGAAAAATACAAGAAAGTAACAGATAGTCACAGAGCTGAAGAAGACAGTAACGGAACCAAAAAACA
CACTGGAGAGGTTCAACAGCAGACTAGAAGCAGAAGTACGGATCCAGGAACCTCAAGGACAGAGCAACGGGACTCA
CCCAGTCAAAGCAGCAAAAAAAGATAGCCAAAAAGACTTATCAGACAATATCACGCAGACCCACACAGATGAAC
ATTTGCATTATAGGAGTCCCAAAAGGAGAAGCAAGAGAGAAAGAGACCAAGAACTTATTGACATAAATAATGGCT
AAAACCTTCCCTAACCTGAGGAAGGAAACGGACATTCATAACCAGGAAGCCAGATGCTTTCAAATAAGAAGAAAT
CCAAAGACACCCACACCAAGACACATTAATAATGTCAAAAGTTAAAGACAAGGAGAGGAATTTGAGTATACCAAA
AGAAAAACAACCTTTTTACATACAAGGGGCCCTCTGTTAAGACTCTCAGCAGATTTTCAACAGAAATTTTGCAGGCC
AGAAAAGGAGTATCATGATATATTTTCAAAGTGCTGAAAGAAAAAACGCCAAGGATATTCTCTTCTTAGAAAAA
CTGTCTTTCAGGAAAAAATTTTCCCGATAAGCAAAAGCTAAAGGAGTTCATCACCATTAAACAGTTTTACAAG
TAATGTTAAAAAGACTTCTTTAAGCTGAAAAGGGTGCTAATTAGTAATGGGAGAACAAATCAAAGTACAAATCTCA
TTGGTAAAGGTTAATATACAGCAAAATTCAGAATAATTAAATGTAAAGGTGGTGATTAATCACTGATAAAGCTA
GTAGGAAATTTAAAAAGACAAAAGTAATTTAAAAACAGCTAACTATAATAAGGGATACACAAGATTAAAAAGACA
TAAATGGACATCAAAAACAAAACATGAGGGTGATGAGTAAACTGTAGAGCTTTGGAATGCATTTAAACCTAAG
TGATTGTCAACTTAAATATAACTGTAACCTTATGGTAACCACAAAGTAGCTGCACGAAAACCTTTGGTAGATA
AACGAAAGAGAAAGGAACCTACTGTAGGAAATCATCAATCACAAAAGGAGAGAGCCAGAGAAAAGAGAGCAAG
AACTACAGAACAGTCAGAAACAGTTCACAGAATGGCAATACGTACACACTTATTAATAAATACTTTGAAAATAA
TGAGCTGAACCTCCCGAGAATACAGAGAGGCAGAATGGATTAAAAAACATGACACATCTATGTATACTGCCTATA
AGAACTCGCTTCAAATTTAAAGACAGACTGAAAGTGAAGAGATGGAAAAATAAGCCTTGCAAGTGGAAACCAAT
GATGTTGGCCTTCAACAAAAAATGAAAGCTAGAGTAAACATTTAGAATCACTTAAGGTTATTTTGTGTTTAT
AAACTTCAAGGGGAAGAATTCAGTGTGTTGGACTAAATAAAGTAATCTCAAAAAGGACTATCTTATTTTCCCTGCTA
ACTGCCATCTTTTACCTTTTTTAAATCTGAAAAACATTTTTCTCATAAAGCTCTTCTGGTTGATCAGTTTGGGGT
ACTCAATACAGAAAAATAAAGAACTCATTGATGGTTTTTGTCTGCTTTTTAGGCAAAGGAGGAACTTCCAAGCC
AAAAGCTGATGGAAAAAGTACAAGTTAAAGTTCACTATTTGTTAAGTTGGGTGTTTGTCAAAATGAAGACCA
CTGTACCTTCTACCAGCCCGTGTGTTATAGGTGAGAGAACTAAGTCCATTTCTGAGTATTTTATGTTCTCTTT
ATGCTATTGGCAATGAAAGGGACACTGAAATCTGTAGAGTAATTCCTGAAAAGCAATAAATGGTGCTTCCAGAA
TGTTTTAAAAACATTAATCTTGTAATTAAGAGCTCTTAAATATGAGGCCAGGCTTTAAACAAAATGACCATCT
AAGTCTAGTTGACTCTTTGAAACCCAAATGCTCATGACTGCAGGTTACTGTGGCAAAGCAAAATTGACAAGTTAC
ATCTTATGGTGAATGGCCTTACGATCTAGATATTACCCGAAACACTTGTGGGGTGAGGAGAAGTGCACGCCCTAC
GGTCTGACGGTAATTCAGGGTTTTGAGAATAGGGCAAAGGAAAAGACTGATAACTTGTGATCATCTTAGTTTTA
GAAAGTAAAGCCCTGTGTTTGAATCTACTTTACAGTAACCGAAGAGCTGGTTTGATGAGGGAGACTCTGACATC
TGTCTTGAGAGCTAGGCCCTTATTAATGTCAAAGAGAAGAGCCCTTATTTCTAAATGACAAAAGCATATAACTACT
CGCTGCAAAAAATCTAAAAAGAAATAACAATGTGGATTAAATTGGATTGGCTCTCTCACCGTGATTCACTTTTCAT
TTCCTGTAGAGGAAGGTATACTGATTAGAATGCTCCAAAGATTTTCTCAAAAAAAGGTTCTTCTCTTAATT
TTCTACTCTAATATCTGCTTAAAGTAAATGCCTCCTTAGATTTATTATGCCTGTAATAAGAAAATAACCTAGCAA
ATGGTTCACTGGATTTTCTTCTTTGAATTTTTCAAGTATCTGCATATAAAATCTTCAGCGGGTAGATGGTGACT
TCTGAAGAAGAAAAGGCTTTGATAACAGAAACAATTTCTGGGTGGCTTGGAGACAGTGGTATTTGCTGAGTCTTT
TGACCTCCTAAACATTGTCTGTTATTCTTTTCTGAAAAGAACTGAATTTGTCTGGTTCACCTGTGTTATTCTA
CTGAGTATTGATAAACTTTAAATTTTTTAAAAATTCGCTTTCAGTTGGGAGAGAAAGGAACTTTATATTTCTAAGAG
ATACATTTGATAGTTTCTTAAAGCAGCACACAAAAAGGAAAAACCTTTGCAAACTTTTGACATTCTCCCCACA
GTGCCTGTAAATCTCATTAGTATTTTCGATTGCACTTATTTTGTGTTAGCATTGGAAAACGATGCCTCAGC
TGTTCTTTCAGTGTCTGATTCTCATGACCCCTTCTCTTAGACTTGTGGACTGTGTTTGTGTTTCTTTGGGT
TGTTGTTTATAAGTCAGTCATAAAATACTGTGCATTGGGCACATGTCTCCTCTTGAGCTGCTAATCGTAGAGACC
CTGGACAGACCAGGAAGCGCCGACCCCCCT

Fig.8G

TTCAGGTTGAACTTCTCTTTGCCAGAGGTTAGGACTTCCGCACCTTGTCCGCAGAGACCCAGCGGGACTTGACTC
 AAGTCAGAAGAGCAAAACGCAGACTCTGTCACTGCATAGCAAGTTTCAAGAATAACAAAACCTAATGCTTAGAAT
 ACTCACATGGCTTGATCTCTGCCACAGTACTGCCTTAAGCTGGGCTTGCCATCCTCTCTCTGTGAAACCGAGCAG
 CCGCTGTTAGCGCAGAGCCATTGCCACGTCACTCTTGCGCTGTACAGCATGTGGCTTAATGGAAGTTGGGTTTAT
 TGCCAAAAACAAAAACAAACTTCTGCTTAAAAATAAAGAAAAAAGATCTTTTTAACATTAAAAAAAATAATCTGT
 TTTTTTTTCCCCCATAGCCATTTAAGAACTAGGGAGGGTCTGATATCCGACAACCTTGAAAGGTTTCTCATATTT
 TTCTGAATGGTTGATGAACATAGCACTGATGAAATTAATATCAAAATCCCATCTTCATTTTCTGATAGTGAAC
 GCAGGAAAAGATATCAGGTACACATCATAAACTGATTTGGAATTCCTGCTTTCAGAGAGTACAGCCTTCCTTAG
 CATCTGTTTAACTATTTCTTAGGACTTGGCAATTTCACTAAAAATTAGTAGTATCACTGATACTGTCTAATGGG
 GTAGGAGGCAAGCAAAATAACAATACTATTCTAATGGGTATCCAATTCTATTAATACTAAACAATTCTCTTTAAA
 AATCATGCATATGCCCTGAGATTGCAAACCTTCATCCACTTCAAATGCTCTGTCTTCATCTCATGAGAGGTAAAGGC
 ATTATTAGTCTAAATCTGGGAAAGATAAGCCAGGATTATGGTAATTTAAACAGACCTTTTCAAGGCACTAACAG
 AAAATAAACACTTTTCTGAGGGATTTTAACTATGTCTGTCTGTAAGTAGTACAAAGTTGTTTTAAAAAAAATGAAA
 CCTTTGACATTGACAGATGTGTTTGCAAAGATATGACTCCAGTATTACTAATTTACATGCATATATCAAGTGT
 TTTTAATTACATACCAAATCCAGAGATTTTAAGAAATGCCTGTAAAAGTAAACATTTATTTAAGGTTCTCTGAAT
 ATGCCTTTTGTATTCAAAATGTCTTCAAGACTTGGGTGTCTGCTGGTAATTAATGATGAAGTGAAGTAC
 AGGATCTAAATCAGTCTTTTACCCATCTAATTCACAGTAAATAATTGCAAATACTTGCTTAGAGGTAAATAAC
 ATTTTTTATTACCAAAAAGGGTTTTGCAATTGTTGCGGTGATGTGTGTCTCTTTAGAAAACGGACTTCTCCAAA
 AGCAAATGAAAATAGGTTCTCAGGTGACCAAAACGGAAAACCACTATGTCCATGTTTCATTACTCAAGATACAA
 GATGCAATTAAAGGAAAGTATTTTACATTAAATCTTGGCTTTGTATGTTTTTTAGAAAGGAAAACACCTGGGGGG
 TAGAAATGACCAGTGATTTCTGTTTTTTCGTTCTTAAACCATGAAGTCAGGCAGTATATATAAAATGGAATTAT
 AAACACATTTTATGCTCGAAAAGACCTTCTATTCCATAGTGTGTTAGAACACATTGACATTTTTCATATCCTATT
 CCTTTATCTTAAGTTTGTCAATATATTAAACTCATAAAAAGATTTCAACTCCAACAACACCATACCTGGTCTTCGC
 TTAGGCTGATATAAGCAAGTCGTTTCCATTTACAGTAGCTTACATTTAGTAGAGTCTGCAGTGCTCACTCGGAG
 AAGGCAATGGCACCCCACTCCAGTACTCTTGCCTGGAGGATCCAGGGACAGAGGAGCCTGGTGGGCTGCAGTCC
 ATGGGGTCACTAAGAGTCAGACATGACTGAGCGACTTCACTTTCACTTTTCACTTTTCATGCACTGGAGAAGGAAA
 TAGCAACTCACTCCAGTGTTCTTGCCTAGAGAATCCAGGGACGGAGGAGCCTGGTAGGCTGCAGTCCATGGGGT
 CACACAGAGTCAGACACAACCTGCAGCGACTTAGCAGCAGCAGCAGCAGTGGCCACTGGCAATAAACACAAG
 TTATTTAGGGATTATAAACAGACCATCTCATCCTCTGTGAGCCCCATCTCCTCCTACCTTCAGTCTTTCCAGC
 ATCAGGGTCTTTTCAAATGAGTCACTTCTTCAACGGTTGGATGGCATCACCGACTCAATGGACATGAGTTTGAGT
 AAATTCGGGAACTGGTGA (SEQ ID NO: 7)

Fig.8H

1 atggcattttgcaagctgggtggtacaagacgcatgtcagtagaaaa
 M A F A S W W Y K T H V S R K
 46 accagtggatcgcttccaagtcaggagaaaaagaaaggatcagat
 T S G S P S K S G E K K G S D
 91 gagaaaaagcaacaagccttgggagcagtcagctctccagaact
 E K K A T S L G S S Q L S R T
 136 caggctggtgaaaaagccccgggtccccaaggtaactacttctct
 Q A G E K A P V P K V T T S S
 181 gcgtcagccagcaagtccttccagtatgaatcccacagaagccaag
 A S A S K S S S M N P T E A K
 226 gctattccaggcagcaaacagctggaaggaccgcattctcctaac
 A I P G S K Q L E G P H S P N
 271 aagaaaaagacacaaaaaacaggctgtaaaaacagaacctgagaag
 K K R H K K Q A V K T E P E K
 316 aagccacaatcatctaagccatctgtggttcagagaaaaaaacc
 K P Q S S K P S V V H E K K T
 361 caagaagtaaagccaaaggaacacacagagccaaaaagcctaccc
 Q E V K P K E H T E P K S L P
 406 aagcactcatcagatacaggaagcaagcatgctcctaaggaaaaa
 K H S S D T G S K H A P K E K
 451 gccgttttcaaatacaagtgagcagccaccatcagagaaatcaaca
 A V S K S S E Q P P S E K S T
 496 ataccaaaagaccaaaccaggacaagatctccgggtggtggaag
 I P K T N S Q D K I S G G G K
 541 agcactgttctctgctgctgctgctgcagcatctgccgaaccagct
 S T V P A A A A A A S A E P A
 586 gacaagaataaagaaaaataaattgttaacatcgccgtaccagct
 D K N K E N K L L T S A V P A
 631 gaatctaaaccaagtaaaccatctggaagtcagacatggacact
 E S K P S K P S G K S D M D T
 676 gctctggatgacttaataagacacttttaggagaacctgaagagatg
 A L D D L I D T L G E P E E M
 721 aaagaagataacacacatataccggaccggaagtgtcggatcca
 K E D N T T Y T G P E V S D P
 746 atgagttctacctacatagaggaactgggtaaaagagaatccaca
 M S S T Y I E E L G K R E S T
 811 cttctccaaaatataaggaacttctgaataaagaagaagggatc
 L P P K Y K E L L N K E E G I
 856 gcggggcctcctccagactccttgaaaccctggggcccaatgat
 A G P P P D S L K P L G P N D
 901 gccatcgatgccttgtcatccgacttcacctgcagttcccctaca
 A I D A L S S D F T C S S P T
 946 gctgatgcaaagaaaactgagaaagagaaatctacagaagaggct
 A D A K K T E K E K S T E E A
 991 ttaaaagctcagtcagctgggtgatcagaagtgtgctccaccc
 L K A Q S A G V I R S A A P P
 1036 caagagaaaaaaaggaaagtggaaaaggatgccatgactgagcac
 Q E K K R K V E K D A M T E H
 1081 gccctggaggccctgtcgcctccctgggcacccggaagccggag
 A L E A L S A S L G T R K P E
 1126 ccggagctcgacccagctccattaaggagggtcgatgaggcaaaa
 P E L D P S S I K E V D E A K
 1171 gccaaaagagagaaagtaagaaatgtggtgaagatgaggaaaca
 A K E E K V K K C G E D E E T

Fig.9A

1216 gtcccatcggagtacagattaaaaccggccacagataaagatgga
 V P S E Y R L K P A T D K D G
 1261 aaaccactcttgccagaggctgaagaaaaacccaagcccctgagt
 K P L L P E A E E K P K P L S
 1306 gaatcagaactcatcgatgaactctcagaagattttgaccagtct
 E S E L I D E L S E D F D Q S
 1351 aagtgtaaagaaaaacaatctaagccaactgaaaaaacagaggca
 K C K E K Q S K P T E K T E A
 1396 tccccggccgctgcccccggtgccctgagcagaggacgtgcctcgg
 S P A A A P V P V A E D V P R
 1441 acctctatgtgttccgtgcagtcggctccgcccacagcagctcca
 T S M C S V Q S A P P T A A P
 1486 gcgaagggcatggtgccagacgatgctgtggaagccttggtgga
 A K G M V P D D A V E A L A G
 1531 agcctgggcaaaaaggaagcagatccagaagacggaaagcctgtg
 S L G K K E A D P E D G K P V
 1576 gaggataaagtcaaggaagaaagccaaagaagaggatcgtgagaaa
 E D K V K E K A K E E D R E K
 1621 cttggtgaaaaagaagaacgattcctcctgattacagattagaa
 L G E K E E T I P P D Y R L E
 1666 gaagccaaggataaagacggaaaaccactgctgccaaaagaggtc
 E A K D K D G K P L L P K E V
 1711 aaggaaccgctcccacccttgagtgaagacgtcctcctcgatgct
 K E P L P P L S E D V L L D A
 1756 ctgtccaaggacttcactgtcccctcagacacatcatcgccctcaa
 L S K D F T V P S D T S S P Q
 1801 tttgaagatgctaaactttcagctgtcgtctctgaagtggtttcc
 F E D A K L S A V V S E V V S
 1846 caaaccacagctccaaccacccaggcagccgggtccacccccagc
 Q T P A P T T Q A A G P P P S
 1891 actgcgcagcgtgacaacaaagaacttgacgatgccctggatcaa
 T A Q R D N K E L D D A L D Q
 1936 ctttctgacagtctcgggcaaagacagcctgatccagatgagaat
 L S D S L G Q R Q P D P D E N
 1981 aaaccctagaggataaagtcaaggaaaaagccaaagctgaacac
 K P V E D K V K E K A K A E H
 2026 agagacaagctgggagaaagagatgacaccatcccacctaatac
 R D K L G E R D D T I P P K Y
 2071 caacatcttttgatgacaacaaggagggcacacccgggaagcca
 Q H L L D D N K E G T P G K P
 2116 aaggcatcagagaagcccaaggcatcagagaaacctgcaggtgcc
 K A S E K P K A S E K P A G A
 2161 caggacccattgatgccctctcaggggactttgacagctgtccc
 Q D P I D A L S G D F D S C P
 2206 tcgactacagaaacctcgacagacacaccaaaggacaaagacaag
 S T T E T S T D T P K D K D K
 2251 aagcctgcttccagtgccgaagcacctaggaatggcgaggaaagca
 K P A S S A E A P R N G G K A
 2296 aaggattccacaaaggcaaaggaggaaacttccaagccaaaagct
 K D S T K A K E E T S K P K A
 2341 gatggaaaaagtacaagttaa 2361 (SEQ ID NO: 8)
 D G K S T S * (SEQ ID NO: 9)

Fig.9B

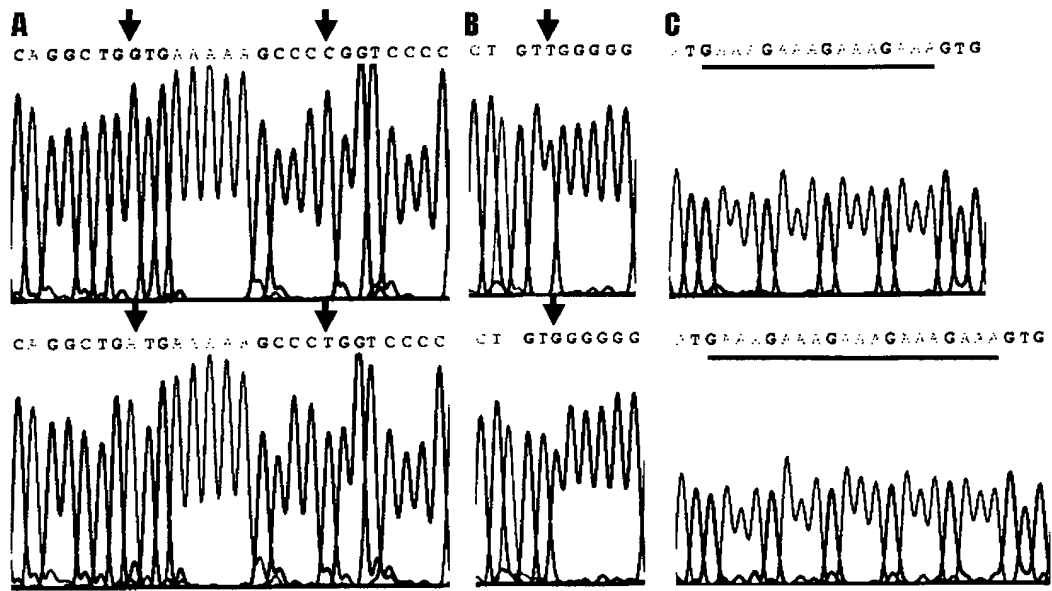


Fig.10

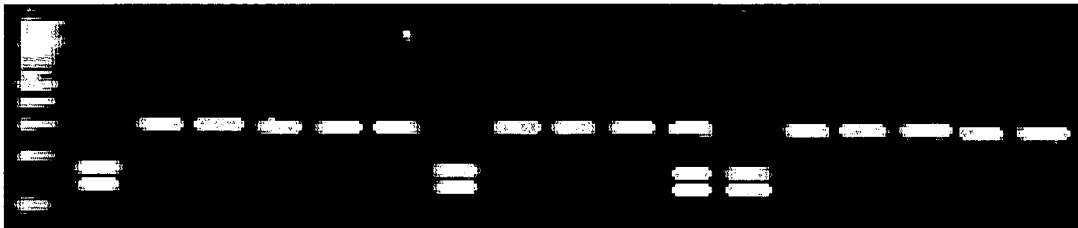


Fig.11

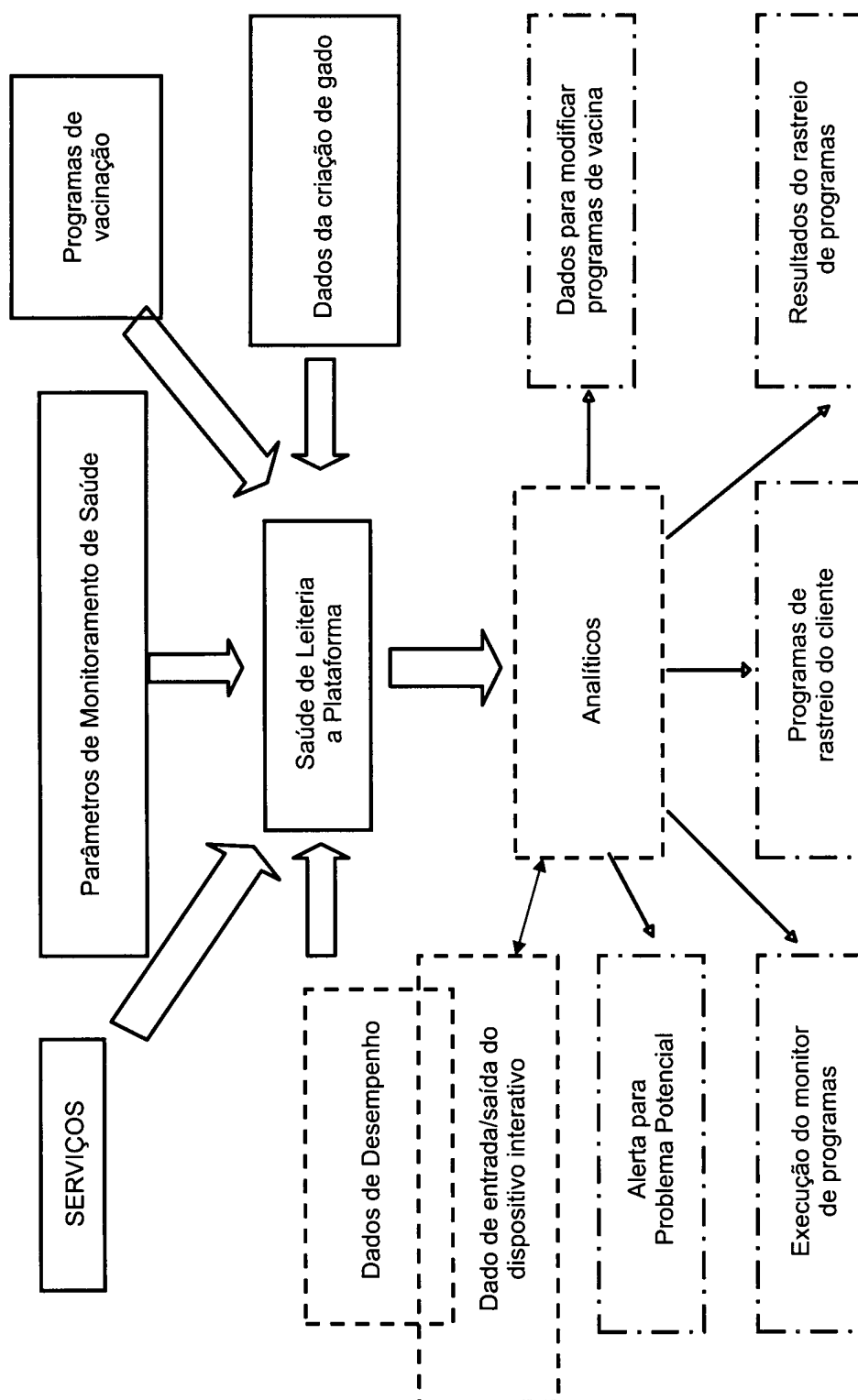


Fig.12

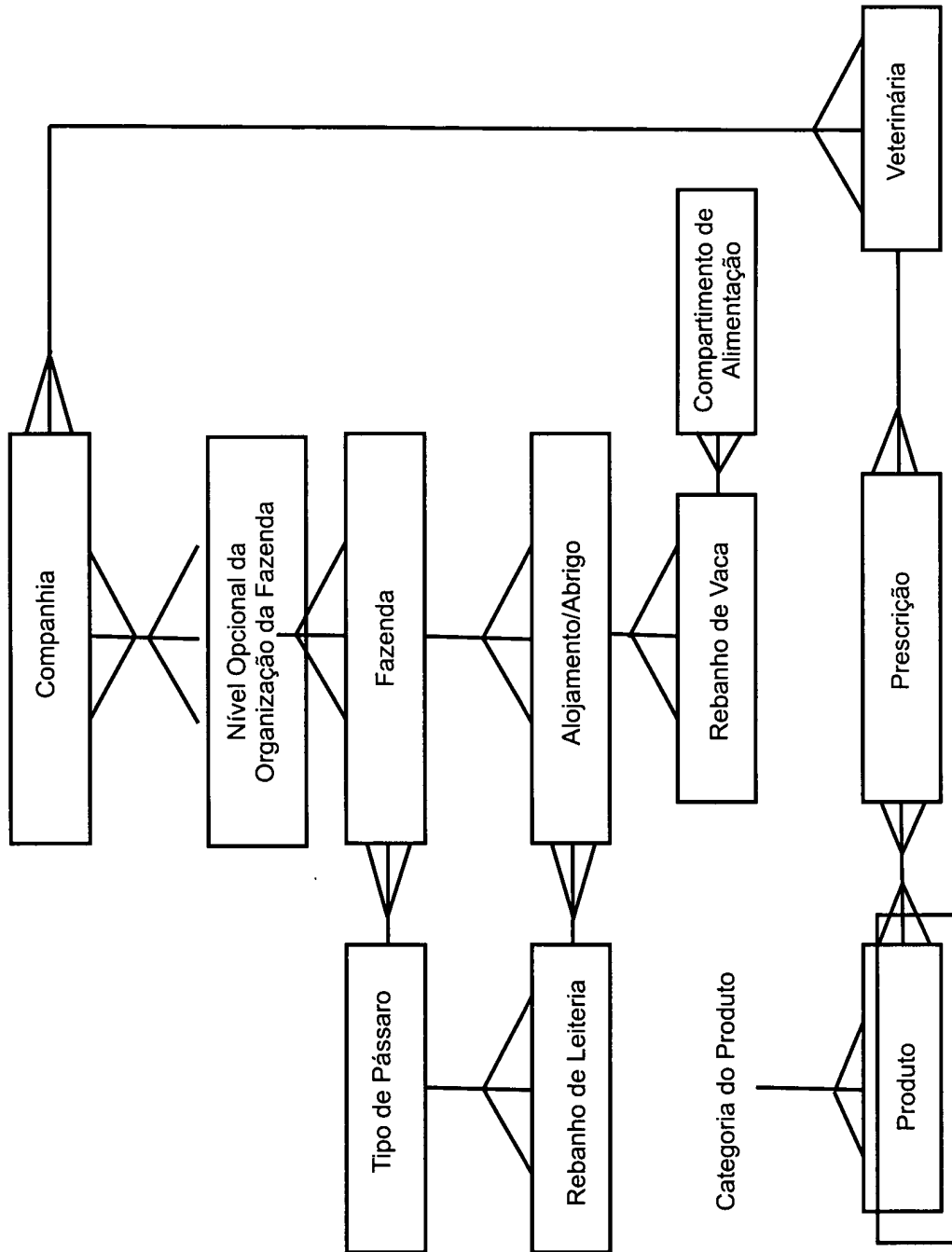
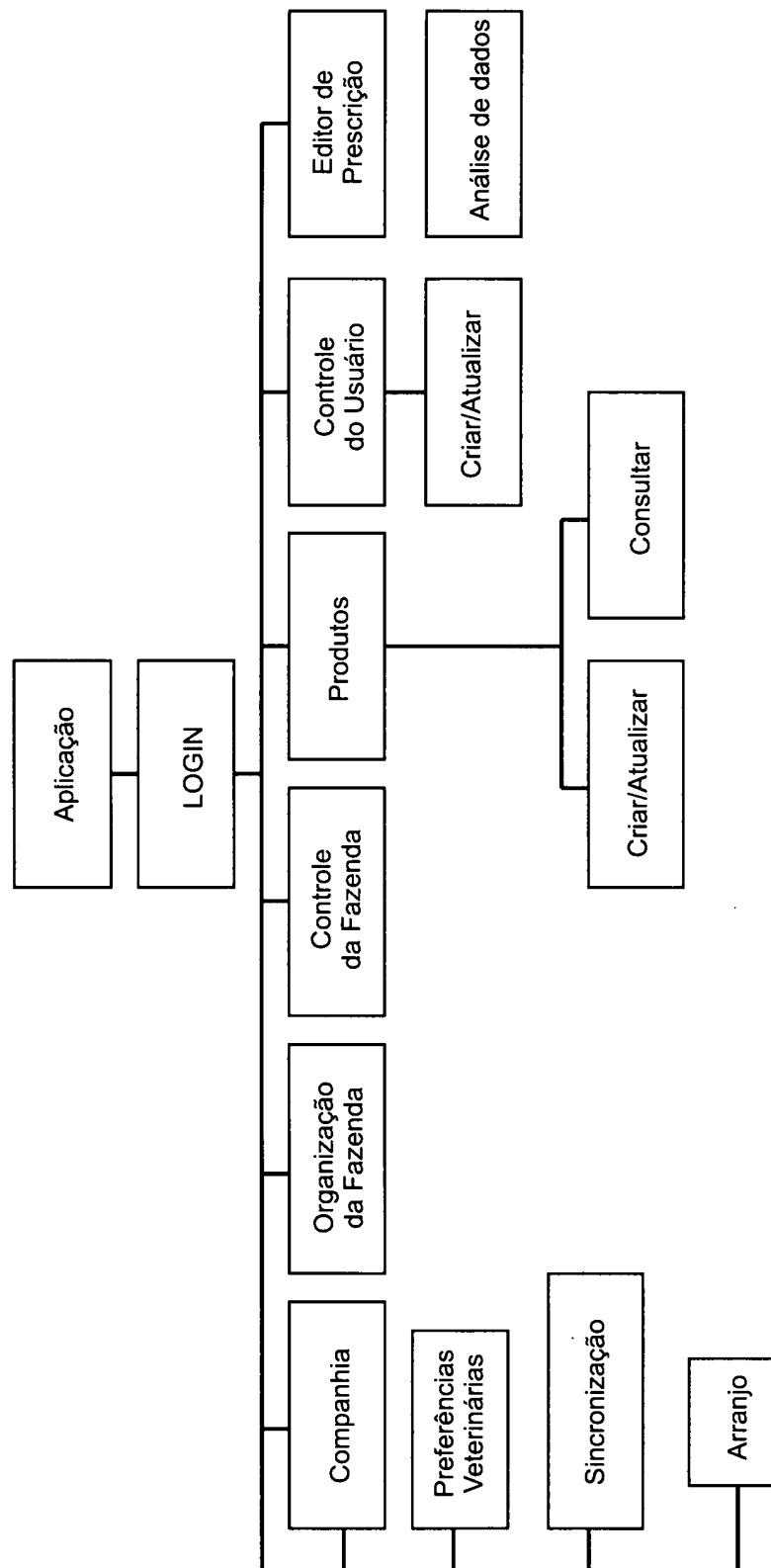
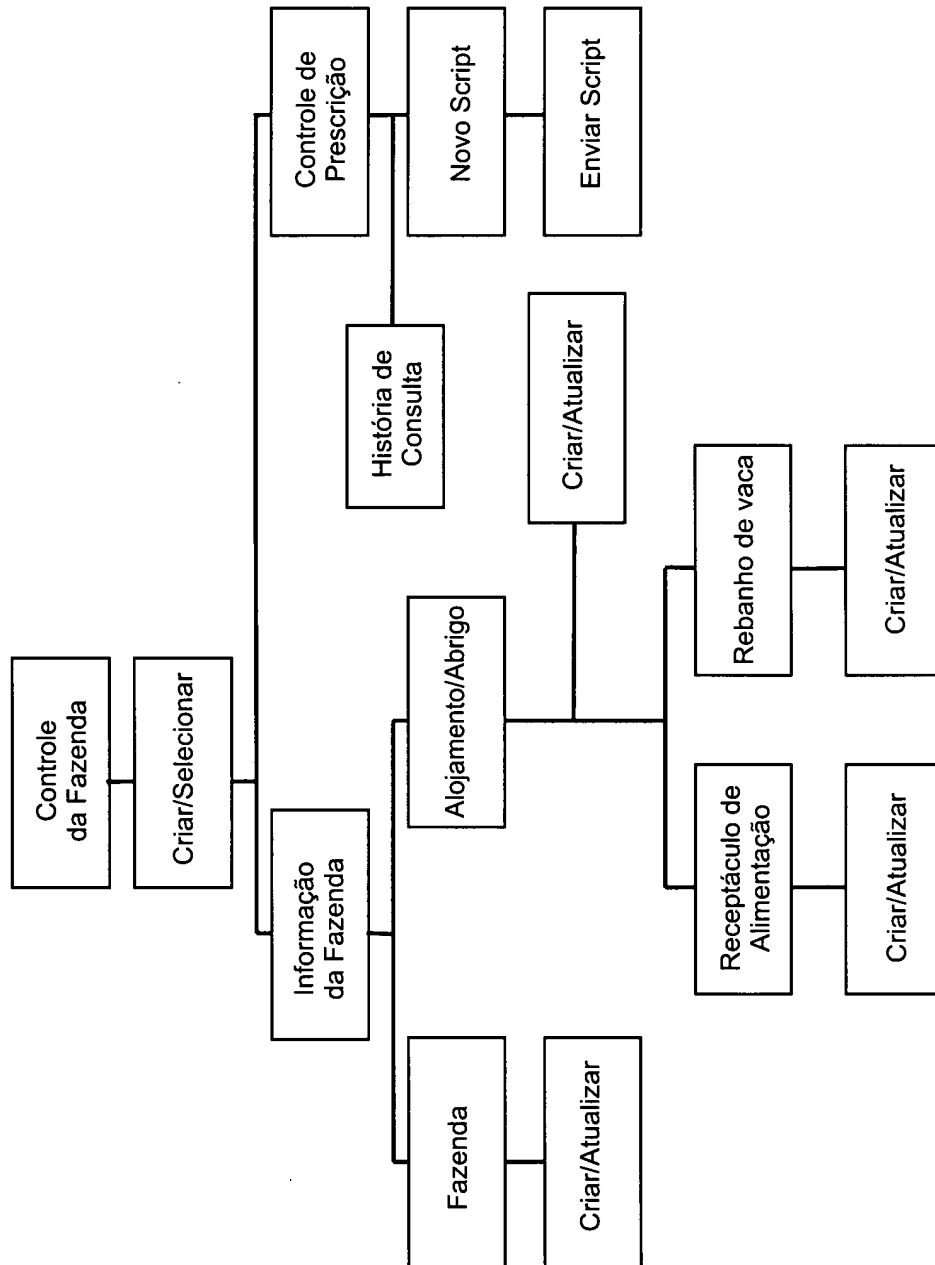
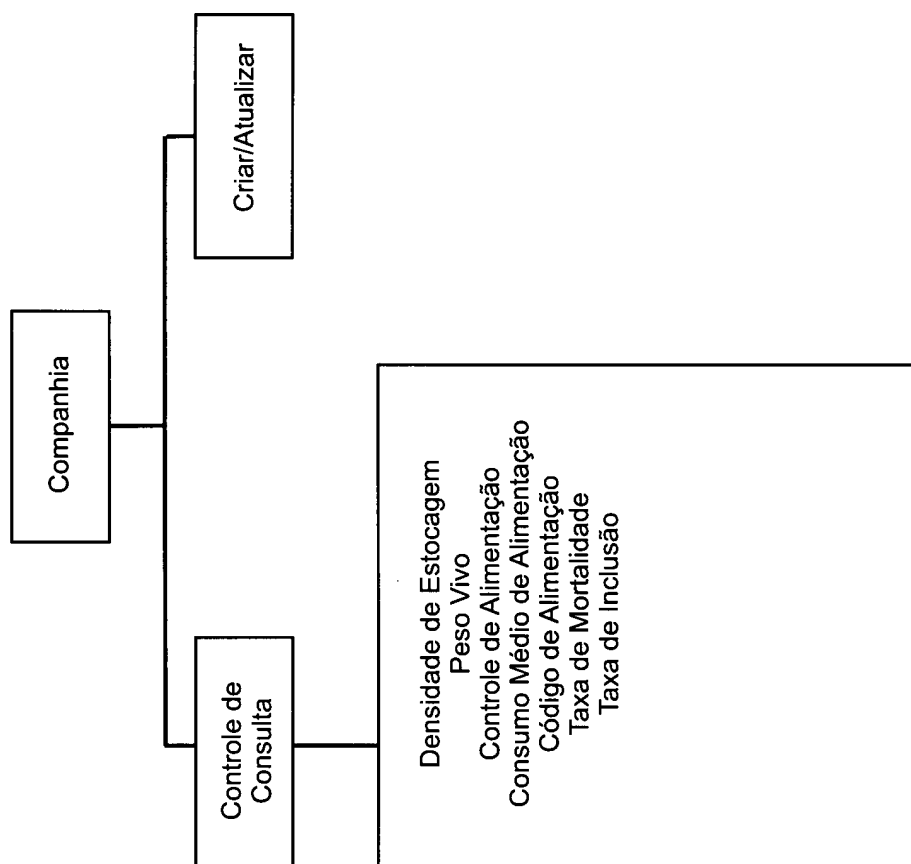
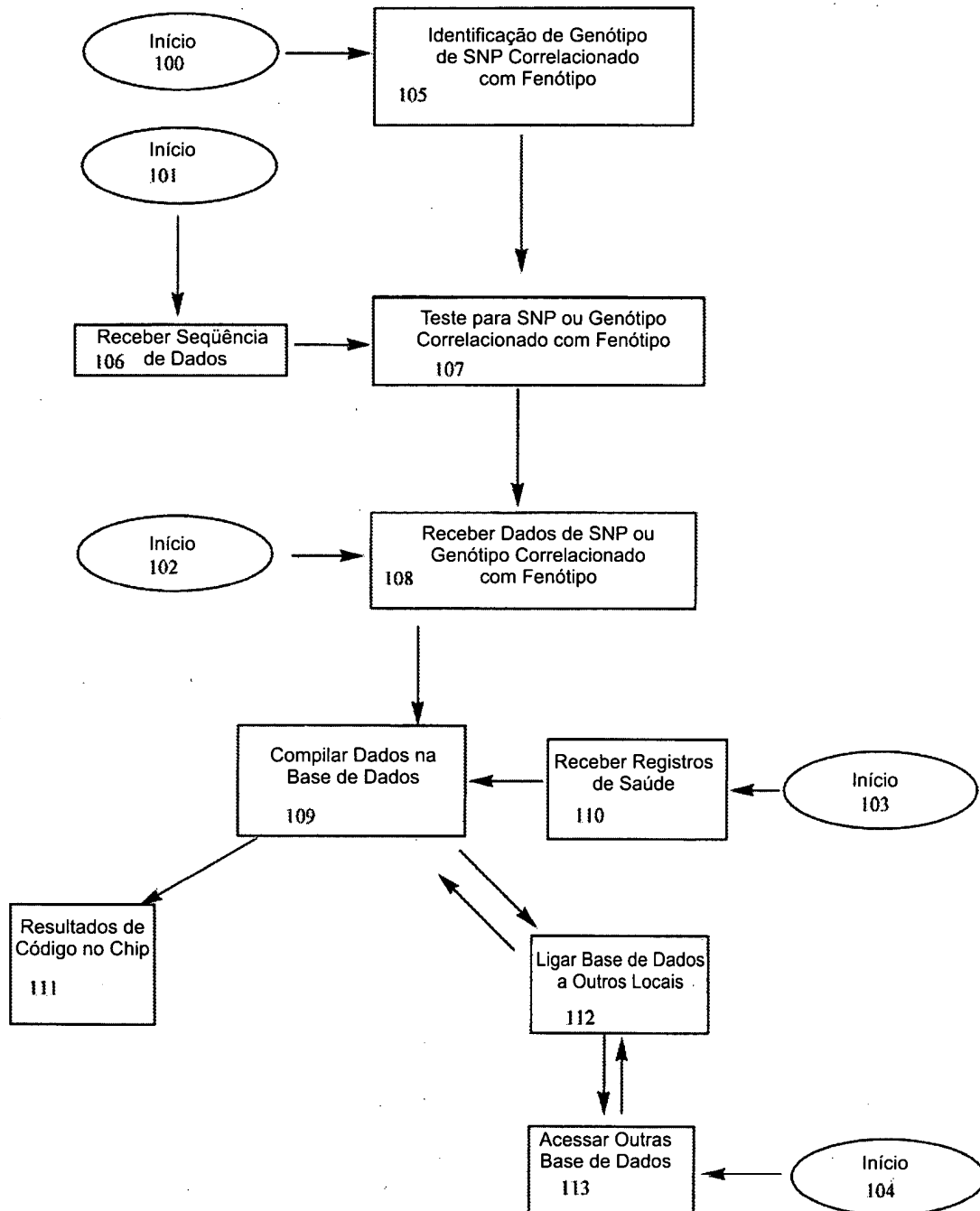


Fig.13

**Fig. 14A**

**Fig.14B**

**Fig.14C**

**Fig.15**

RESUMO

Patente de Invenção: "MARCADORES DE CALPASTATIN PARA FERTILIDADE E LONGEVIDADE".

A presente invenção refere-se a composições e métodos baseados em novos marcadores genéticos calpastatin (CAST), tais como mutações sem sentido em exon 3 que resultem em substituições G48D ou P52L (NM_174003.2:c.271G>A e 283C>T), uma substituição G/T em intron 3 (AFC02060381.1:g.2110G>T) e uma repetição GAAA em intron 8 (AFC02060381.1:g.6700[(GAAA)4]+[(GAAA)5]. Aspectos particulares proporcionam novos marcadores para fertilidade (por exemplo, taxa de gravidez de raça, DPR) e longevidade (por exemplo, vida produtiva, PL) em, por exemplo, gado leiteiro. Aspectos adicionais proporcionam novos métodos compreendendo seleção auxiliada por marcador para aperfeiçoar fertilidade e/ou longevidade em gado leiteiro. Portanto, em concretizações particulares, uma combinação de seleção genética baseada em um ou mais dos novos marcadores CAST, e altos potenciais de PTA de traços de produção de leite, proporciona traços reprodutivos aperfeiçoados em associação com traços de produção de leite alta continuada. Aspectos adicionais desgeram um domínio XL anteriormente não-reconhecidos no gene CAST humano, e, desse modo, proporciona o uso de mutantes/variantes de domínio de CAST XL como marcadores para fertilidade e longevidade humanas.