



(12) PATENTSKRIFT

Patent- og  
Varemærkestyrelsen

---

(51) Int.Cl<sup>7</sup>: C 12 N 15/40 C 07 K 14/18

(21) Patentansøgning nr: PA 1986 03230

(22) Indleveringsdag: 1986-07-07

(24) Løbedag: 1986-07-07

(41) Alm. tilgængelig: 1987-01-09

(45) Patentets meddelelse bkg. den: 2004-09-27

(30) Prioritet: 1985-07-08 US 752981

(73) Patenthaver: REGION WALLONNE REPRESENTEE PAR L'EXECUTIF REGIONAL WALLON, Bruxelles, Belgien  
CHIRON CORPORATION, 4560 Horton Street, Emeryville, California 94608, , USA

(72) Opfinder: Andre Renard, 17 Rue T. Gerkens, B-4950 Beaufays, Belgien  
Dina Dino, 1245 Washington Street, San Francisco California 94108, USA  
Joseph Martial, 13 Rue Vignoble, B-4050 Esneux, Belgien

(74) Fuldmægtig: Internationalt Patent-Bureau A/S, Høje Taastrup Boulevard 23, 2630 Taastrup, Danmark

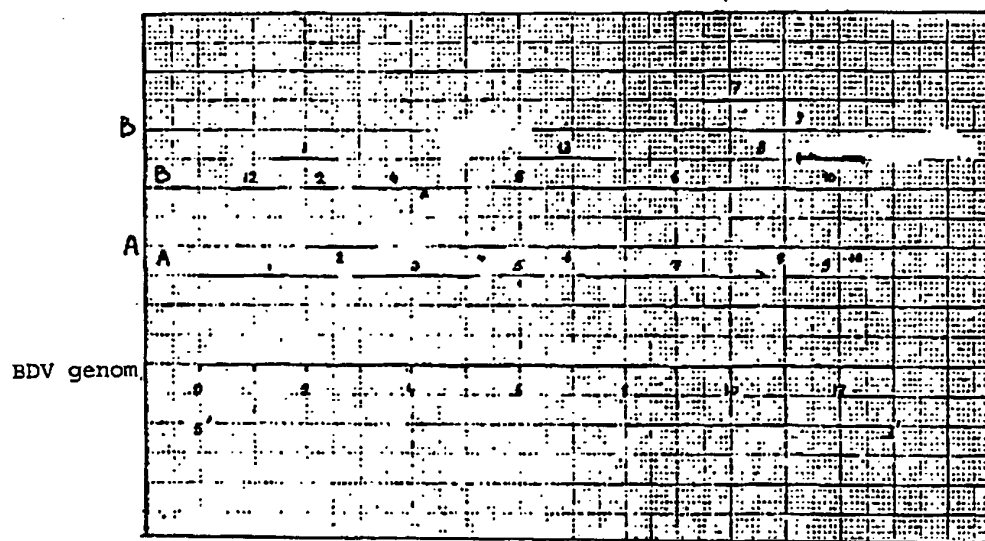
---

(54) Benævnelse: Nucleotidsekvens kodende for et polypeptid relateret til kvægdiarré, rekombinant ekspressionsystem for polypeptidet, vektor og værtscelle omfattende systemet, selve polypeptidet, vaccine mod kvægdiarré omfattende polypeptidet, fremgangsmåde til fremstilling af polypeptidet og af vaccinen og endelig anvendelser for polypeptidet og nucleotidsekvensen

(57) Sammendrag:

Nucleotidsekvenser af genomet for kvægdiarré-virus (BDV) er angivet. Sekvensen tillader skitsering og konstruktion af vacciner, der er effektive mod BDV.

Figur 1



- A: Afbildning af 9 BDV cDNA kloner, der omspænder hele genomet. Klonerne blev fremstillet fra oligo dT primet cDNA (DT kloner) eller fra tilfældigt primet cDNA ved brug af kalvebrissel oligonucleotider (CT kloner). Navnene på klonerne er som følger: 1 - pCT63; 2 - pCT36; 3 - pCT183; 4 - pCT70; 5 - pDT28; 6 - pDT17; 7 - pCT174; 8 - pDT65; 9 - pCT185; 10 - pCT40.
- B: Anvendte cDNA fragmenter til konstruktion af ekspressionsvektorer for E.coli ved fusion til E.coli B-galactosidasegenet.

Opfindelsen angår en nucleotidsekvens, der koder for et til kvægdiarré relateret polypeptid, altså et viralt polypeptid. Opfindelsen angår yderligere et rekombinant ekspressionssystem, der i en forligelig værtcelle 5 er i stand til at foretage produktion af et til kvægdiarré relateret polypeptid, og en rekombinant vektor og rekombinante værtceller. Derudover angår opfindelsen selve polypeptidet og en vaccine, der er effektiv mod kvægdiarrévirus samt fremgangsmåder til fremstilling af 10 henholdsvis polypeptidet og vaccinen. Endelig angår opfindelsen anvendelser af henholdsvis polypeptidet og vaccinen.

Sygelighed og dødelighed forårsaget af kvægdiarré-virus (BDV, "bovine diarrhea virus") blandt malke- 15 kvæg og okser er et overalt i verden uløst økonomisk problem. En subklinisk form karakteriseret ved høj sygelighed og lav dødelighed er endemisk og er forbundet med nedsat respiratorisk evne, neonatal diarré, ulcerationer i fordøjelseskanalen, immunodeficiens og, hos kalvebærende køer, abort og teratogenese. Sygdommen kan erken- 20 des hos kalve, men voksne bærere er vanskelige at identificere.

En akut form af sygdommen skyldes infektion af foetus inden for de første tre graviditetsmåneder. Forløbet af denne form af sygdommen er snigende. Kalvene 25 kan overleve den første infektion, men de, der overlever, bliver immunotolerante og udskiller levende vira. De kan ikke overleve en anden infektion. Da deres egenskab som bærere ikke kan påvises ved titrering af deres sera, er disse dyr ansvarlige for spredning af sygdommen 30 fra bestand til bestand.

BDV inficerer også svinebestande. Hos svin er det vigtigt at skelne mellem, om dyrene er inficeret med BDV eller svinekolera-virus, da svinekolera er en økonomisk

vigtig sygdom, mens kvægdiarré-infektion er af forbigående betydning og i de fleste tilfælde kan ignoreres. Svin inficeret med kolera skal slagtes, og da nuværende diagnosemetoder for svin ikke kan skelne mellem disse to  
5 infektionstyper, må svin, der i virkeligheden kun er inficeret med BDV, også tilintetgøres.

Nuværende påvisningsmidler for BDV-infektion hos kalve er ligeså mangelfulde, idet de er baseret på titrering for antistoffer i sera, hvilken titrering ikke  
10 kan påvise de immunotolerante kalve. Der ønskes således en diagnosemetode, der er i stand til både at påvise tilstedeværelse af virus'en hos nyfødte dyr med kroniske infektioner og at skelne mellem svinekolera-virusinfektion og BDV-infektion. Dette kunne opnås enten ved at anvende antistoffer med høj affinitet og specificitet  
15 for virus-partiklerne eller ved at anvende oligomere nucleinsyreprober, der har evnen til specifik hybridisering til de virale sekvenser.

Foruden behovet for forbedrede diagnostica er der for nærværende heller ingen effektiv vaccine, der er  
20 gunstig til at hindre spredning af sygdommen forårsaget af BDV. Det er ønskeligt, at en sådan vaccine skal tilvejebringe langvarig immunitet, ikke vil inficere foetus hos det inokulerede dyr og ikke har uønskede bivirkninger, såsom inducering af immunotolerance over for virus'en eller depression af immunsystemet. Disse karakteristika er vanskelige om ikke umulige at opnå hos en svækket eller dræbt virusvaccine. Sådanne vacciner udgør for størstedelens vedkommende teknikkens stadi  
25 (Saurat, P., et al., "La Maladie des Muqueuses" (1972), side 229-251. L'Expansion scientifique française, Paris).  
30 Nyligt har Fernelius, A.L., et al., (Am. J. Vet. Res. (1971) 32: 1963-1979) rapporteret om en vaccine frem-

stillet ud fra et højmolekylært opløseligt antigen vundet ved vægtfyldegradient-centrifugering fra BDV-virus dyrket i embryoniske kvægnyreceller.

De på området anvendte veje til påvisning af og beskyttelse mod viral kvæg-diarré har været i høj grad empiriske og har ikke gjort brug af forfinet kendskab til arten af den vektor, der forårsager sygdommen. Kvæg-diarré-virus'en er imidlertid, sammen med svinekolera-virus og tilgrænsende sygdomsvira, blevet klassificeret som en pestivirus, der tilhører familien Togaviridae (Porterfield, J.S., "The Togavirions, Biology, Structure, Replication", Schlesinger, W., Ed. (1980), Academic Press, side 17-24).

I analogi med andre togavira skulle disse vira indeholde et kapsidprotein og to eller tre membran-glycoproteiner (Horzinet, M.C., Non-arthropod-borne Togaviruses (1981), Academic Press, London), Epitoper, der er i stand til at fremkalde antistoffer forbundet med neutralisation og beskyttelse mod infektion, forventes at være indeholdt i membran-proteinerne (se f.eks. Bore, W., et al., J. Virol. (1984) 52:572-582). Pestivira er også karakteriseret ved opløselige antigener, der er omtrent 80 kD proteiner. Et 76 kD protein fra BDV har faktisk været anvendt som en forsøgsvaccine (Fernelius, A.L., et al., supra).

Opfindelsen tilvejebringer cDNA-kopier af den totale genomiske RNA-sekvens fra kvægdiarré-virus. Dette gør hele repertoire af peptider syntetiseret af virus'en tilgængelig og gør det muligt at fremstille proteiner, der indeholder epitoper, som er effektive og specifikke i henseende til at udvikle ønskede antistoffer og tilvejebringe celler, der er egnede for fremstilling af monoklonale antistoffer. Den primære struktur af genomet tilvejebringer også den nødvendige information

til at kunne konstruere oligomere sekvenser, der er anvendelige som diagnostiske prober.

Proteinprodukterne er således i stand til at tjene som vacciner til beskyttelse af dyr, der er genstand  
5 for infektion med denne virus, mod efterfølgende sygdom. Rådigheden over hele genomet giver mulighed for fremstilling af effektive proteiner, såsom større virionkomponenter og individuelle virion-underenheder, der ville være utilgængelige ved anvendelse af "nativ" produktionsteknik, dvs. fra viral infektion af vævdyrkede celler.  
10

I overensstemmelse hermed er den ovenfor definerede nucleotidsekvens ifølge opfindelsen ejendommelig ved, at den er en region af den genetiske sekvens for kvægdiarrévirus vist på figur 2 eller en kombination af regioner i denne sekvens. Hvis den omhandlede nucleotidsekvens er defineret ved, at den koder for et viralt polypeptid, er den ejendommelig ved, at dette i det væsentlige er identisk med det, der kodes for af den genomiske sekvens for kvægdiarrévirus vist på figur 2. Det rekombinante ekspressionssystem ifølge opfindelsen er ejendommeligt ved, at systemet omfatter en DNA sekvens som ovenfor defineret som en nucleotidsekvens, der operativt er forbundet til en med værten forligeligt kontrolsekvens. I en foretrukken udførelsesform findes opstrøms  
20 for nævnte DNA sekvens og i læseramme hermed en fusioneret nucleotidsekvens kodende for et værtprotein eller en del deraf. Den rekombinante vektor ifølge opfindelsen er ejendommelig ved, at den omfatter ovennævnte ekspressionssystem. De rekombinante værtceller ifølge opfindelsen  
30 er ejendommelige ved, at de er transformeret med ovennævnte vektor eller med en vektor omfattende den ovenfor definerede foretrukne udførelsesform for det opfinderske ekspressionssystem. Polypeptidet ifølge opfindelsen

sen er ejendommeligt ved, at det i det væsentlige er identisk med hele aminosyresekvensen som vist på figur 2 eller med en region eller kombinationer af regioner deraf. I en foretrukket udførelsesform er polypeptidet 5 ifølge opfindelsen yderligere fusioneret med et værtprotein eller en del deraf. Den ovenfor definerede vaccine mod kvægdiarrévirus er ejendommelig ved, at den omfatter polypeptidet ifølge opfindelsen samt farmaceutisk acceptable excipienser og i en foretrukket udførelsesform 10 yderligere en immunogen partikel som defineret i underkravene på vaccinen.

Fremgangsmåden ifølge opfindelsen til fremstilling af et polypeptid ifølge opfindelsen er ejendommelig ved, at man dyrker værtceller ifølge opfindelsen og udvinder 15 det rekombinante polypeptid. Endvidere er fremgangsmåden ifølge opfindelsen til fremstilling af en vaccine ifølge opfindelsen er ejendommelig ved, at man udøver ovennævnte fremgangsmåde og foretager en yderligere tilsætning af farmaceutisk acceptable excipienser.

20 Et opfinderisk polypeptid kan ifølge opfindelsen anvendes til fremstilling af prøvesæt til immunologisk påvisning af fusionsproteiner af kvægdiarrévirus. Og en opfinderisk nucleotidsekvens kan ifølge opfindelsen dels anvendes til konstruktion af som diagnostiske sonder 25 nyttige oligonucleotidsekvenser, og dels anvendes til fremstilling af en vaccine, der som bærer omfatter levende viral vektor, der tillader ekspresion af et ønsket kvægdiarrévirusantigen sammen med et protein fra bæreren i inficerede celler.

30 I det følgende beskrives opfindelsen med henvisning til tegningerne, på hvilke:

Fig. 1 viser kortet over overlappende segmenter af cDNA, der tilsammen udgør hele BDV-genom-sekvensen, og cDNA-fragmenter anvendt til at konstruere E. coli-ekspressionsvektorer, og

5 Fig. 2 viser den fuldstændige nucleotidsekvens for BDV-genomet. cDNA'et indeholder den identiske sekvens med den undtagelse naturligvis, at T vil erstatte U. Den reducerede aminosyresekvens, baseret på den åbne læseramme, og bekræftet ved ekspression af segmenter, er  
10 også vist.

### Udførelsesformer for opfindelsen:

#### A. Definitioner

Anvendt i nærværende beskrivelse skal en nucleotidsekvens "i det væsentlige identisk" med det eksemplificerede BDV-genom referere til en sekvens, der bevarer de afgørende egenskaber af det eksemplificerede polynucleotid. Et specifikt, men ikke-begrænsende eksempel på sådan substantiel ækvivalens ville være repræsenteret af en sekvens, der indkoder den identiske eller i det  
20 væsentlige identiske aminosyresekvens, men som, på grund af codon-degeneration, gør brug af andre specifikke codoner. Nucleotid-forandringer er i virkeligheden ofte ønskelige for at skabe eller udelade restriktionssteder, tilvejebringe oparbejdningssteder eller ændre aminosyre-  
25 sekvensen på måder, der ikke ugunstigt påvirker funktionalitet. "Nucleotidsekvens" refererer både til en ribonucleotid- og en deoxyribonucleotid-sekvens og omfatter den positive streng, som vist, og også den negative streng.

30 En DNA-sekvens "afledt af" nucleotidsekvensen, der omfatter BVD-genomet, refererer til en DNA-sekvens, der er omfattet af en region af den genomiske nucleotidsekvens eller en kombination af regioner af nævnte sekvens. Disse regioner er ikke nødvendigvis fysisk af-  
35 ledt af genets nucleotidsekvens, men refererer til poly-

nucleotider udviklet på en hvilket som helst måde, og som har den samme eller "i det væsentlige identiske" sekvens af baser som den, der foreligger i den eller de regioner, af hvilke(n) polynucleotidet er afledt. For eksempel omfatter typiske DNA-sekvenser "afledt af" BDV-genomet fragmenter, der indkoder specifikke epitoper, fragmenter, der indkoder dele af det virale polypeptid, sekvenser, der indkoder kapsid-proteinerne, sekvenser, der indkoder udeladte virioner, og sekvenser, der indkoder andre nyttige virale gener.

"Rekombinante værtceller", "værtceller", "celler", "cellelinier", "cellekulturer" og andre lignende udtryk, der angiver mikroorganismer eller højere eukaryotiske cellelinier dyrket som unicellulære enheder, er anvendt indbyrdes uvekslelige og refererer til celler, der kan, eller er blevet, anvendt som recipienter for rekombinant vektor eller andet transfer-DNA, og omfatter afkom af den oprindelige transficerede celle. Det skal forstås, at afkommet af en enkelt stamcelle ikke nødvendigvis er fuldstændigt identisk i morfologi eller i genomisk eller totalt DNA-komplement med den oprindelige stamcelle, hvilket kan skyldes tilfældig eller forsætlig mutation. Afkom af stamcellen, der har tilstrækkelig lighed med stamcellen til at blive karakteriseret ved den relevante egenskab, såsom tilstedeværelse af en nucleotidsekvens, der indkoder et ønsket peptid, er inkluderet i det afkom, der er tilsigtet med nævnte definition, og er omfattet af de ovennævnte udtryk.

"Styringssekvens" refererer til DNA-sekvenser, der er nødvendige for at bevirke ekspressionen af kodningssekvenser, til hvilke de er ligeret. Arten af sådanne styringssekvenser varierer i afhængighed af værtorganismen. I prokaryoter omfatter sådanne styringssekvenser i almindelighed en regulationsregionspromotor og ribosom-bindingssted og terminationssignaler, mens i

eukaryoter sådanne styringssekvenser i almindelighed omfatter promotorer, terminatorer og, i nogle tilfælde, transkriptionsfremmere. Angivelsen "styringssekvenser" skal som minimum omfatte alle komponenter, hvis tilstedeværelse er nødvendig for ekspression, og kan også omfatter yderligere komponenter, hvis tilstedeværelse er fordelagtig.

"Opererbart forbundet" refererer til en sidestilling, hvori de således beskrevne komponenter befinder sig i et indbyrdes forhold, der tillader dem at fungere på den for dem tilsigtede måde. En styringssekvens "opererbart forbundet" til en kodningssekvens er ligeret på en sådan måde, at ekspression af kodningssekvensen opnås under betingelser, der er kompatible med styringssekvenserne.

#### B. Generel beskrivelse

Centralt for den foreliggende opfindelse er tilvejebringelsen af en nucleotidsekvens indeholdende det totale genom fra kvægdiarré-virus. Tilgængeligheden af dette fuldstændige polynucleotid tillader konstruktion og fremstilling af oligomere prober til diagnose, af vacciner, der er effektive over for BDV, og af proteiner, der er anvendelige ved fremstilling af neutraliserende antistoffer. Sekvens-information, der er tilgængelig fra genomet, muliggør deducering af polypeptidets aminosyresekvens og udtænkning af lokationer for gunstige epitoper. Når først de ønskede sekvenser er valgt, kan endvidere passende fragmenter af genomet vindes og eksprimeres uafhængigt, hvorved der tilvejebringes ønskede polypeptider. Korte polypeptidfragmenter kan også syntetiseres kemisk og forbindes til bære-proteiner til brug som immunogener. Rekombinantmæssigt eksprimeredes polypeptider kan tilvejebringes under betingelser, der giver et gunstigt miljø for oparbejdning til for eksem-

pel konjugation med cellulære eller kunstige membraner, der således kunne bære epitop-stederne uden ulemperne ved at anvende en infektiøs virus. Pattedyr- og gærceller tilvejebringer egnede miljøer for sådan ekspression. 5 Endvidere kan epitoperne frembringes forbundet til et partikeldannende protein.

De ovennævnte fremstillede proteiner kan i sig selv anvendes som vacciner, eller de kan anvendes til at inducere immunokompetente B-celler i værter, hvilke B- 10 celler derefter kan anvendes til at frembringe hydridomer, der udskiller antistoffer, som er nyttige ved passiv immunoterapi og diagnose.

#### B.1. Nucleotidsekvens af BDV-genomet

15 BDV-genomsekvensen fandtes ud fra cDNA-kloner, der repræsenterer overlappende sektioner af det totale virale RNA-genom (Fig. 1). Det virale RNA blev isoleret fra virus dyrket på embryoniske kvægnyreceller. Det virale RNA blev fraktioneret på sucrosegradienter, og de 20 fraktioner, der indeholdt RNA af tilstrækkelig længde til at indeholde det intakte genom, blev samlet, ethanol-fældet og anvendt til at fremstille et cDNA-bibliotek. cDNA-insertioner blev screenet indledningsvis ved hjælp af et (+/-)-system. Positive hybridiseringer var 25 mod RNA isoleret fra virus efter lysis af inficerede celler, og negative hybridiseringer var mod RNA isoleret fra uinficerede celler. Én insertion med den rette (+/-)-respons anvendtes derefter som reference-klon til kortlægning af den resterende del af biblioteket. Ad- 30 skillige kolonier, der hybridiserede til den positive insertion, anvendtes til at opnå yderligere dele af det virale genom ud fra cDNA-biblioteket ved anvendelse af "walking"-teknik. Der vandtes ti cDNA-kloner, der repræsenterede overlappende dele af det virale genom, som 35 vist i Fig. 1, og disse blev underkastet restriktions-

kortlægning og sekvensbetemmelse. Den totale genom-sekvens blev udledt af disse ti cDNA-insertioner og er vist i Fig. 2.

Den illustrerede DNA-sekvens og dele deraf er direkte anvendelige som diagnostiske redskaber til påvisning af tilstedeværelse af BDV i inficerede dyr. De er særligt anvendelige til at skelne BDV-infektioner fra svinekolera-virus. Fremgangsmåder til anvendelse af DNA-hybridisering ved diagnosticering af sygdomme findes omtalt i US patentskrift nr. 4.358.535, Falkow. Som deri omtalt, kan biologiske prøver anvendes direkte til opnåelse af Southern blots ved anvendelse af egnede prober. Da BDV-genomet er forskelligt fra genomet af svinekolera-virus, kan specifikke dele af BDV-sekvensen anvendes til at påvise tilstedeværelse af tilsvarende komplementære sekvenser i biologiske prøver fra subjekter, der mistænkes for at have infektionen.

#### B.2. Fremstilling af virale polypeptidfragmenter i E.

##### 20 coli

Tilgængeligheden af den totale genom-sekvens muliggør konstruktion af ekspressionsvektorer, der indkoder formodede antigent aktive regioner af virion-proteinerne. Fragmenter, der indkoder de ønskede proteiner, vindes fra cDNA-klonerne ved anvendelse af konventionel restriktions-digerering og liggeres ind i en række vektorer indeholdende polylinker-steder i alle mulige læsrammer for at frembringe fusionsproteiner ved den C-terminale ende af  $\beta$ -galactosidase. Elleve dele af BVD-genomet blev eksprimeret som  $\beta$ -gal-fusioner i E. coli ved anvendelse af denne metode, som vist i Fig. 1. Disse dele vandtes ved restriktionskløvning og/eller ligation af de ti oprindelige kloner, eller de oprindelige klonedes sekvenser anvendtes direkte. De således tilvejebragte fusionsproteiner kan være immunogene.

### B.3. Fremstilling af antigene polypeptider og konjugation med bærer

Peptidregioner, der repræsenterer epitoper, kan syntetiseres ved anvendelse af kemiske metoder eller re-

5 kombinantmetoder og forsynes med for eksempel cystein-

grupper ved C-terminus, hvilket tilvejebringer midler til at forbinde peptiderne til neutrale bæreproteiner. En række teknikker til opnåelse af sådan binding kendes, inklusive dannelsen af disulfidbindinger ved anvend-

10 delse af sædvanlige reagenser, såsom N-succinimidyl-3-(2-pyridyldithio)propionat (SPDP) og succinimidyl-4-(N-maleimido-methyl)cyclohexan-1-carboxylat (SMCC) stammende fra Pierce Company, Rockford, Illinois. Disse reagenser skaber en disulfidbinding mellem dem selv og pep-

15 tid-cysteingrupper i det ene protein og en amidbinding gennem  $\epsilon$ -amino på en lysin, eller anden fri aminogruppe i det andet. Der kendes forskellige sådanne disulfid/amid-dannende midler. Se for eksempel Immun. Rev. (1982) 62:185. Andre bifunktionelle koblingsmidler danner en

20 thioether i stedet for en disulfid-binding. Mange af disse thioether-dannende midler er kommercielt tilgængelige og omfatter reaktive estre af 6-maleimidocaprønsyre, 2-bromeddikesyre, 2-iodeddikesyre, 4-(N-maleimido-methyl)cyclohexan-1-carboxylsyre og lignende. Carboxyl-

25 grupperne kan aktiveres ved at kombinere dem med succinimid eller 1-hydroxy-2-nitro-4-sulfonsyre-natriumsalt. Den foregående liste er ikke udtømmende, og modifikationer af de anførte forbindelser kan anvendes.

Enhver bærer kan anvendes, der ikke i sig selv

30 inducerer dannelsen af antistoffer, der er skadelige for subjektet, såsom de forskellige serumalbuminer, tetanus-toxoider eller "keyhole limpet hemocyanin" (KLH).

Ved injektion i egnede subjekter resulterer konjugaterne i dannelsen af antisera, der indeholder immuno-

35 globuliner, som er specifikt reaktive over for ikke blot

disse konjugater, men også over for fusionsproteiner bærende de analoge dele af sekvensen, og over for hel BDV.

5 B.4. Fremstilling af pattedyr-cellemembraner indeholdende BDV-epitoper

Dele af cDNA-biblioteket omfattende BDV-genomet blev også ligeret ind i ekspressionsvektorer, der var kompatible med rekombinante pattedyr-værtceller, nedenfor illustreret ved en pattedyr/bakterietransportvektor 10 ("shuttle"-vektor") indeholdende en linkersekvens nedstrøms for den tidlige SV40-promotor, efterfulgt af den ligeledes af SV40 afledte polyA-sekvens. Alternative vektorer til denne særlige værtvektor, pSV7d, kunne naturligvis også anvendes. De pattedyr-kompatible vektorer 15 indeholdende kodningssekvenserne for de ønskede polypeptider transformeres derefter ind i egnede pattedyrceller til ekspression af sekvenserne og, i tilfælde af overflade-glycoproteiner, transport af det fremstillede protein til membranen. Cellerne bliver til slut udvun- 20 det og anvendt som hele celler ved fremstillingen af vacciner, eller membranerne opbrydes, og dele af membranerne anvendes tilsvarende, eller proteinerne renses og oparbejdes til vacciner.

25 B.5. Fremstilling af hybride partikel-immunogener indeholdende BDV-epitoper

Immunogeniteten af epitoperne af BDV kan også fremmes ved at fremstille dem i pattedyr- eller gær-systemer fusioneret med partikel-dannende proteiner, såsom 30 det der er associeret med hepatitis B-virus (HBV)-overfladeantigen (HBsAG). Konstruktioner, hvori en BDV-epitop er forbundet direkte til sekvenserne, der koder for partikeldannende protein, frembringer hybrider, der er immunogene med hensyn til BDV-epitopen såvel som med 35 hensyn til HBV-epitoper.

Hepatitis B-overfladeantigen har vist sig at blive dannet og samlet i *S. cerevisiae* (Valenzuela et al., Nature (1982) 298:344-350). Dannelsen af sådanne partikler har vist sig at forøge immunogeniteten af den monomere subenhed. Partiklerne kan også dannes ud fra konstruktioner, der indeholder præoverflade(pre-S)-regionen foruden det modne overfladeantigen. pre-S-Regionen indkoder en immunodominant HBV-epitop, og disse proteiner ekspri-  
5 meres i gær (Neurath et al., Science (1984) 224:392-  
10 394). Ekspresion af konstruktioner, der indkoder pre-  
B region fusioneret til partikel-dannende protein er be-  
skrevet i offentliggjort europæisk patentansøgning nr.  
EP-A-0 174 444. Ekspresion af kodningssekvenser for hy-  
bridpartikler indeholdende HBsAg og en heterolog epitop  
15 findes omtalt i U.S. patent nr. 4,722,840, idet indleve-  
ringsdatoen for den hertil svarende U.S. patentansøgning  
var 13. september 1984. Disse konstruktioner kan også  
eksprimeres i pattedyrceller, såsom kinesisk hamster-  
20 ovarieceller, ved anvendelse af en SV40-dihydrofolat-  
reduktasevektor (Michelle et al., Int. Symp. on Viral  
Hepatitis (1984)).

Endvidere kan dele af den sekvens per se, der ko-  
der for partikel-dannende protein, erstattes med codoner  
25 for en BDV-epitop. Ved denne udskiftning kan regioner,  
der ikke kræves for at formidle aggregationen af enhe-  
der under dannelse af immunogene partikler i gær eller  
pattedyr, udelades, hvorved yderligere antigene hepatis  
B-steder elimineres fra konkurrence med BDV-epito-  
30 pen.

#### B.6. Vaccinia-bærer

Store virusbærere med bredt vartområde har også  
været anvendt ved opbygning af vacciner ved integrering  
35 af de epitope regioner af det ønskede immunogen i det

virale bærer-genom. Vaccinia-virus har navnlig været anvendt til dette formål. For eksempel omtaler Smith, G.L. et al., Proc. Natl. Acad. Sci. (USA) (1983) 80: 7155-7159, integrationen af hemagglutinin-genet fra influenza-virus i vaccinia-genomet og anvendelse af den resulterende rekombinante virus som en vaccine. Tilsvarende klonede Panicali, D. et al., ibid (1982) 79:4927-4931, thymidinkinase-genet fra Herpes simplex-virus ind i vaccinia. Tilgængeligheden af BDV-genomet ifølge opfindelen giver tilsvarende muligheder. Rekombinationen foretages i almindelighed ved at co-inficere cellerne både med vaccinia-virus og med et chimerisk plasmid bærende den ønskede kodningssekvens under kontrol af de transkriptionelle regulationssignaler og RNA-startstedet fra det vaccinia-virusgen, der er nabostillet til en sekvens, der koder for translations-startsted/fremmed protein. Under infektion medfører ligheden mellem de flankerende DNA-sekvenser af de fremmede DNA-sekvenser og dem i vaccinia integration af den ønskede del af det chimeriske plasmid i vaccinia-genomet. Den resulterende rekombinante vaccinia kan udvindes fra de inficerede celler og anvendes ved dannelsen af en vaccine. Vaccinia-virus har et meget stort ( $120 \times 10^6$  dalton) genom og kan meget let dyrkes i kultur. Fremstillingen af store mængder billig immunogen vaccine er derfor let mulig.

#### B.7. Fremstilling af vacciner

Fremstilling af vacciner, der indeholder peptid-sekvenser som aktive bestanddele, er også velkendt teknik. Typisk fremstilles sådanne vacciner som injicerbare præparater, enten som flydende opløsninger eller suspensioner, men faste former egnede til opløsning eller suspendering i væske forud for injektion kan også fremstilles. Præparatet kan også være emulgeret, eller proteinet kan være indkapslet i liposomer. Den aktive immu-

nogene bestanddel er ofte blandet med excipienser, der er farmaceutisk acceptable og kompatible med den aktive bestanddel. Egnede excipienser er for eksempel vand, saltopløsning, dextrose, glycerol, ethanol eller lignende og kombinationer deraf. Vaccinen kan endvidere, om ønsket, indeholde mindre mængder af hjælpestoffer, såsom befugtningsmidler eller emulgeringsmidler, pH-puffermidler eller hjælpemidler, der fremmer vaccinenes effektivitet. Vaccinerne indgives konventionelt parenteralt, for eksempel ved injektion, enten subcutant eller intramuskulært. Yderligere kompositioner, der er egnede til andre anvendelsesmåder, omfatter suppositorier og, i nogle tilfælde, orale kompositioner. Til suppositorier kan traditionelle bindemidler og bærere omfatte for eksempel polyalkaliglycoler eller triglycerider. Sådanne suppositorier kan fremstilles ud fra blandinger indeholdende den aktive bestanddel inden for området 0,5% til 10%, fortrinsvis 1-2%. Orale kompositioner omfatter sådanne normalt anvendte excipienser som for eksempel farmaceutiske kvaliteter af manitol, lactose, stivelse, magnesiumstearat, natriumsaccharin, cellulose, magnesiumcarbonat og lignende. Disse kompositioner foreligger som opløsninger, suspensioner, tabletter, piller, kapsler, kompositioner med forhalet frigørelse eller pulvere og indeholder 10-95% aktiv bestanddel, fortrinsvis 25-70%.

Proteinerne kan oparbejdes til vaccinen som neutral form eller saltform. Farmaceutisk acceptable salte omfatter syreadditionssaltene (dannet med peptidets frie aminogrupper), og som er dannet med uorganiske syrer, såsom saltsyre eller phosphorsyre, eller sådanne organiske syrer som eddike-, oxal-, vin-, mandelsyre og lignende. Salte dannet med de frie carboxylgrupper kan også være afledt af uorganiske baser, såsom natrium-, kalium-, ammonium-, calcium- eller ferri-hydroxider, og

sådanne organiske baser som isopropylamin, trimethylamin, 2-ethylaminoethanol, histidin, procain og lignende.

Vaccinerne indgives på en måde, der er kompatibel med dosispræparatet, og i en sådan mængde, som vil være terapeutisk effektiv og immunogen. Den mængde, der skal indgives, afhænger af subjektet, der behandles, evnen hos subjektets immunsystem til at syntetisere antistoffer og den ønskede beskyttelsesgrad. Præcise mængder af aktiv bestanddel, der kræves til indgift, bestemmes af lægen og er speciel for hvert subjekt.

#### B.8. Fremstilling af kort over for BDV-epitoper

De immunogene proteiner eller immunokonjugaterne fremstillet som beskrevet ovenfor kan anvendes til at vinde lymfocytter fra perifert blod og miltceller i injektionsbehandlede pattedyr med henblik på at fremstille hybridomaer, der er i stand til at udskille monoklonale antistoffer rettet mod disse epitoper. De resulterende monoklonale antistoffer er særligt anvendelige ved diagnose, og de, der er neutraliserende, er anvendelige ved passiv immunoterapi.

#### C. Generelle metoder

Den generelle teknik, der anvendes ved ekstraktion af RNA fra virus'en, fremstilling og prøvning af cDNA-bibliotek, sekvensbestemmelse af kloner, konstruktion af ekspressionsvektorer, transformation af celler, og lignende er kendt teknik, og der findes laboratoriehåndbøger, der beskriver disse teknikker. Som generelle retningslinier er der imidlertid nedenfor omtalt nogle for nærværende tilgængelige kilder for sådanne metoder og for materialer til deres gennemførelse.

### C.1. Værter og ekspressions-styringssekvenser

Både prokaryotiske og eukaryotiske værtceller kan anvendes til ekspression af ønskede kodningssekvenser, når der anvendes passende styringssekvenser, der er kompatible med den designerede vært. Prokaryoter er mere anvendelige til kloning. Enten prokaryoter eller eukaryoter kan anvendes til ekspression. Blandt prokaryotiske værter anvendes oftest *E. coli*, mest for nemheds skyld. Ekspressions-styringssekvenser for prokaryoter omfatter 10 promotorer, eventuelt indeholdende operatordele, og ribosom-bindingssteder. Transfer-vektorer, der er kompatible med prokaryotiske værter, er sædvanligvis afledt af for eksempel pBR322, et plasmid indeholdende operoner, der medfører ampicillin- og tetracyclin-resistens, 15 og de forskellige pUC-vektorer, der også indeholder sekvenser, som medfører antibiotikum-resistens. De ovennævnte operoner kan anvendes som markører til opnåelse af gunstige transformanter ved udvælgelse. Almindeligt anvendte prokaryotiske styringssekvenser omfatter  $\beta$ -lactamase (penicillinase)- og lactose-promotorsystemer (Chang et al., *Nature* (1977) 198:1056), tryptophan (*trp*)-promotorsystemet (Goeddel et al., *Nucleic Acids Res.* (1980) 8:4057) og den  $\lambda$ -afledte  $P_L$ -promotor- og N-gen-ribosom-bindingssted (Shimatake et al., *Nature* (1981) 292:128). 20 De foregående systemer er særligt kompatible med *E. coli*. Om ønsket kan der, med tilsvarende styringssekvenser, anvendes andre prokaryotiske værter, såsom stammer af *Bacillus* eller *Pseudomonas*.

Eukaryotiske værter omfatter gær- og pattedyr- 30 cellekulturer. *Saccharomyces cerevisiae*, eller bagerigær, og *Saccharomyces carlsbergensis* er de mest anvendte gærværter, igen for nemheds skyld. Gær-kompatible vektorer bærer markører, der tillader udvælgelse af gunstige transformanter ved at bibringe auxotrope mutanter 35 prototrophi eller ved at bibringe vildtypestammer anti-

bioticum-resistens eller resistens mod tunge metaller. Gær-kompatible vektorer kan anvende 2 $\mu$ m-replikationsstarten (Broach, J., et al, Meth. Enz. (1983) 101:307), kombinationen af CEN3 og ARS1 eller andre midler til at sikre replikation, såsom sekvenser, der vil resultere i inkorporering af det passende fragment i værtcelle-genomet. Styringssekvenser for gærvektorer omfatter promotorer for syntesen af glycolytiske enzymer (Hess et al., J. Adv. Enzyme Reg. (1968), 7:149, Holland et al., Biochemistry (1978) 17:4900), og promotoren for 3 phosphoglyceratkinase (Hizeman et al., J. Biol. Chem. (1980) 255: 2073). Til gærekspresion kan terminatorer også være inkluderet, såsom dem, der er afledt af enolasegenet (Holland, M.J., J. Biol. Chem. (1981) 256:1385).

15 Særligt anvendelige styringssystemer inkluderer de heri specifikt beskrevne, som omfatter glycerinaldehyd-(3 phosphatdehydrogenase)(GAPDH)-promotoren eller den regulerbare alkoholdehydrogenase(ADH)-promotor, terminatorer ligeledes afledt af GAPDH, og, hvis der ønskes sekretion, ledersekvens fra gær- $\alpha$ -faktor. Disse systemer er beskrevet detaljeret i U.S. patent nr. 4,876,197 og U.S. patent nr. 4,870,008, idet indleveringsdatoerne for de hertil svarende U.S. patentansøgninger var henholdsvis 22. august 1983 og 12. august 1983.

25 Pattedyrccellelinier, der er tilgængelige som værter for ekspresion, omfatter mange udødeliggjorte cellelinier tilgængelige fra the American Type Culture Collection, omfattende HeLa-celler, kinesisk hamsterovarie(CHO)-celler, babyhamsternyre(BHK)-celler og en 30 række andre cellelinier. Egnede promotorer for pattedyrceller omfatter hovedsageligt virale promotorer, såsom den fra Simian virus 40 (SV40) (Fiers et al., Nature (1978) 273:113) eller andre virale promotorer, såsom Rous sarcoma-virus (RSV), adenovirus og kvæg-papilomavirus (BPV). Pattedyrceller kan også kræve terminator-

35

sekvenser. Vektorer, der er egnede til replikation i pattedyrceller, kan omfatte virale repliconer eller sekvenser, der sikrer integration af de passende sekvenser i vært-genomet.

5

### C.2. Transformationer

Den anvendte transformationsmetode afhænger af hvilken vært, der skal transformeres. Bakteriell transformation benytter i almindelighed behandling med calcium- eller rubidiumchlorid (Cohen, S.N., Proc. Natl. Acad. Sci. (USA) (1972) 69:2110, Maniatis et al., Molecular Cloning: A Laboratory Manual (1982), Cold Spring Harbor Press, side 254). Gærtransformationer kan gennemføres ved anvendelse af metoden ifølge Hinnen, A., et al., Proc. Natl. Proc. Acad. Sci. (USA) (1978) 75:1929-1933. Pattedyr-transformationer gennemføres ved anvendelse af calciumphosphatfældningsmetoden ifølge Graham og van der Eb, Virology (1978) 52:546, eller de forskellige modifikationer deraf.

20

### C.3. Vektor-konstruktion

Vektor-konstruktion benytter teknik, der nu er ganske godt klarlagt. Positions-specifik DNA-kløvning gennemføres ved behandling med egnet restriktionsenzym under betingelser, der generelt er specificeret af fabrikanterne af disse kommercielt tilgængelige enzymer (se f.eks. The New England Biolabs Product Catalog). I almindelighed bliver ca. 1 µg plasmid eller DNA-sekvens kløvet med 1 enhed enzym i ca. 20 µl pufferopløsning i en inkubationstid på ca. 1-2 timer ved ca. 37°C. Efter inkubation med restriktionsenzymet fjernes protein ved phenol/chloroform-ekstraktion, og DNA'et udvindes ved genudfældning med ethanol. De kløvede fragmenter kan adskilles ved anvendelse af polyacrylamid- eller agarosegelelektroforeseteknik ifølge de almene metoder beskrevet i Methods in Enzymology (1980) 65:499-560.

Kløvningsfragmenter med klæbeender kan gøres stumpendede ved anvendelse af *E. coli* DNA-polymerase I (Klenow) i nærværelse af de egnede deoxynucleotidtriphosphater (dNTP'er) under anvendelse af inkubations-  
5 betingelser, der er egnede for polymerasen. Polymerasen digererer fremspringende 3'-enkeltstreng, men udfylder fremspringende 5'-ender, alt efter de i blandingen tilstedeværende dNTP'er. Behandling med S1-nuclease kan også anvendes, da dette resulterer i hydrolyse af en-  
10 hver enkeltstrenget DNA-del.

Ligationer gennemføres ved anvendelse af standard-puffer- og -temperatur-betingelser under brug af T4 DNA-ligase og ATP. Ligationer med klæbeender kræver mindre ATP og mindre ligase end ligationer med stumpen-  
15 der. Når der som del af en ligationsblanding anvendes vektor-fragmenter, bliver vektor-fragmentet ofte behandlet med bakteriel alkalifosphatase (BAP) for at fjerne 5'-phosphatet og således hindre religation af vektoren. Alternativt kan restriktionsenzym-digerering af uønskede  
20 fragmenter anvendes til at hindre religation.

Ligationsblandinger transformeres ind i egnede kloningsværter, såsom *E. coli*, og gunstige transformanter udvælges for eksempel ved antibiotisk resistens og screenes for den korrekte konstruktion.

25

#### C.4. Konstruktion af ønskede DNA-sekvenser

Syntetiske oligonucleotider kan fremstilles ved anvendelse af et automatiseret oligonucleotid-syntetiseringsapparat, som beskrevet af Warner, B.D., et al.,  
30 DNA (1984) 3:401-411. Om ønsket kan disse syntetiske strenge kinaseres med henblik på mærkning med  $^{32}\text{P}$  ved anvendelse af et overskud af polynucleotidkinase i nærværelse af mærket ATP under standard-kinaseringsbetingelser.

35 DNA-sekvenser inklusive dem, der er isoleret fra genom- eller cDNA-biblioteker, kan modificeres ved posi-

tionsrettet mutagenese, som beskrevet af Zoller, M., et al., *Nucleic Acids Res.* (1982) 10:6487-6499. Kort angivet bliver det DNA, der skal modificeres, emballeret i phag som en enkeltstrenget sekvens og omdannet til et dobbeltstrenget DNA med DNA-polymerase under anvendelse som primer, af et syntetisk oligonucleotid, der er komplementært med den del af DNA'et, der skal modificeres, og har den ønskede modifikation inkluderet i dets egen sekvens. Det resulterende dobbeltstrengede DNA transformeres ind i en phag-støttende værtbakterie, og kulturer af den transformerede bakterie, der vil indeholde replikationer af hver streng af phagen, udplades i agar til opnåelse af plaquer. Teoretisk vil 50% af de nye plaquer indholde phag, der som enkeltstreng har den muterede form, mens 50% vil have den oprindelige sekvens. Replikater af plaquerne hybridiseres til kinaseret syntetisk probe ved temperaturer og under betingelser, der tillader hybridisering med den korrekte streng, men ikke med den umodificerede sekvens. De således identificerede, ønskede, modificerede sekvenser udvindes derefter og klones for at tjene som kilde for det ønske DNA.

#### C.5. Hybridisering med probe

DNA-biblioteker probes ved anvendelse af metoden ifølge Grunstein og Hogness (*Proc. Natl. Acad. Sci. (USA)* (1975) 73:3961). Kort angivet bliver ved denne metode det DNA, der skal probes, immobiliseret på nitrocellulosefiltre, denatureret og præhybridiseret med en puffer indeholdende 0-50% formamid, 0,6M NaCl, 60mM natriumcitrat, 0,02% (vægt/vol.) af hvert af kvægserumalbumin, polyvinylpyrrolidon og Ficoll, 50mM natriumphosphat (pH 6,5), 1% glycin og 100 µg/ml denatureret bærer-DNA. Procentindholdet af formamid i pufferen samt tids- og temperaturbetingelserne ved præhybridiserings- og efterfølgende hybridiseringstrin afhænger af den øns-

kede stringens. Oligomere prober, der kræver lavere stringensbetingelser, anvendes i almindelighed med lave procentindhold af formamid, lavere temperaturer og længere hybridiseringstider. Prober indeholdende mere end 30 eller 40 nucleotider, såsom dem, der er afledt af cDNA eller genom-sekvenser, gør i almindelighed brug af højere temperaturer, f.eks. ca. 40-42°C, og et højt procentindhold, f.eks. 50% formamid. Efter præhybridisering bliver denne samme puffer, nu indeholdende den <sup>32</sup>P-kina- serede oligonucleotid-probe, tilsat til opnåelse af hybridisering. Radioautografi af de behandlede filtre viser lokationen af den hybridiserede probe, og de tilsvarende lokationer på replikafiltre, der ikke er blevet probet, kan derefter anvendes som kilde for det ønskede DNA.

#### C.6. Verifikation af konstruktion og sekvensbestemmelse

Til rutine-vektorkonstruktioner bliver ligationsblandinger transformeret ind i E. coli-stamme HB101 eller anden egnet vært, og gunstige transformanter udvælges ved antibiotisk resistens eller andre markører. Plasmider fra transformanterne fremstilles derefter ved metoden ifølge Clewell, D.B., et al., Proc. Natl. Acad. Sci. (USA) (1969) 62:1159, sædvanligvis efter chloramphenicol-forstærkning (Clewell, D.B., J. Bacteriol. (1972) 110:667. Det isolerede DNA isoleres og analyseres ved restriktionsanalyse eller sekvensbestemmes ved didoxymetoden ifølge Sanger, F., et al., Proc. Natl. Acad. Sci. (USA) (1977) 74:5463, som yderligere beskrevet af Messing et al., Nucleic Acids Res. (1981) 9:309, eller ved metoden ifølge Maxam et al., Methods in Enzymology (1980) 65:499.

#### D. Eksempler

De følgende eksempler skal illustrere, men ikke begrænse opfindelsen. De anførte fremgangsmåder, f.eks. under D.1 og D.2, kan, om ønsket, gentages, men dette er  
5 ikke nødvendigt, da der findes teknikker til konstruktion af de ønskede nucleotid-sekvenser baseret på den gennem opfindelsen tilvejebragte information. Ekspression er eksemplificeret i E. coli og i gær, men andre systemer er tilgængelige, som omtalt mere detaljeret un-  
10 der C.1. Yderligere epitoper afledt af genomstrukturen kan også fremstilles og anvendes til frembringelse af antistoffer, som omtalt nedenfor.

#### D.1. Fremstilling af cDNA

15

##### D.1.a. Fremstilling af BDV

Kvæg-embryonyre (BEKI)-celler blev dyrket i MEM (Earl's) indeholdende 0,85 g/l NaHCO<sub>3</sub> og 10% bestrålet  
20 foetal kalveserum. Den biologisk klonede Osloss-stamme af BDV blev 5 gange ledt gennem BEKI-celler ved en multiplicitet på 0,11. Cytipatiske effekter, bestående i klyngedannelse af celler efterfulgt af vacuolering og derefter celle-lysis, kunne let observeres fra den før-  
25 ste passage. Endelige titere ( $\sim 10^8$  pfu/ml) fandtes efter udvinding af virus ved frysning og optøning af inficerede celler.

Til virus-fremstillingen anvendtes 175 cm<sup>2</sup> formstofkolber af subconfluente BEKI-celler. Cellerne blev  
30 vasket 3 gange med infektionspuffer (MEM (Earl's) + 2,2 g/liter NaHCO<sub>3</sub>, pH 7,6) og blev derefter inficeret med 2 ml BDV i infektionspuffer ved en multiplicitet på 0,05 pfu/celle. Efter 1 time ved 35°C blev der tilsat 18 ml infektionspuffer, og cellerne blev inkuberet i 4-5  
35 dage ved 35°C, hvorefter cytopatisk effekt (vacuolering)

efterfulgt af celle-lysis) var højere end 80%. Ved en typisk produktion blev 150 kolber med celler inficeret. Mediet (ca. 3 liter) blev indsamlet og lagret ved 4°C. De resterende celler blev skrabet i 2 ml infektionspuffer/kolbe, underkastet 3 cykler med frysning og tønning, og den endelige suspension blev sat til infektionsmediet. Efter en centrifugering ved 10.000 g i 30 minutter blev supernatanten koncentreret 10 gange ved ultracentrifugering ved 120.000 g i 4 timer og 40 minutter ved 10 4°C.

Infektiøs virus havde en vægtfylde på 1,12 g/ml målt ved isopicnicbånddannelse i sucrose-vægtfyldegradient, og viste sig ved elektronmikroskopi som 45-55 nm sfæriske partikler. Virus-præparationerne blev neutraliseret med anti-BDV-antiserum fra kaniner injiceret med virus eller fra kvæg.

#### D.1.b Ekstraktion og rensning af viralt RNA

RNA blev isoleret fra virus-pellet'en ved CsCl/guanidiniumthiocyanat-metoden, som beskrevet af Chirgwin et al., Biochemistry (1979) 18:3294, og det rensede RNA blev lagret i 70% ethanol ved -20°C. Dette RNA-præparat indeholdt en stort mængde forurenende lavmolekylært cellulært RNA og intakt viralt RNA. Viralt RNA blev rensset yderligere ved sucrosevægtfyldegradient-centrifugering som følger:

En aliquot indeholdende en vurderet mængde på 5 µg BDV-RNA blev centrifugeret ved 10.000 g i 15 minutter ved 4°C. Pellet'en blev vasket med 80% ethanol, denatureret i 375 µl 99% DMSO (99%), 5 mM Tris-HCl (pH 7,5) og inkuberet i 5 minutter ved 37°C. Efter tilsætning af 1,125 ml 5 mM Tris HCl (pH 7,5), 1 mM EDTA, 1% Sarkosyl blev opløsningen opvarmet i 2 minutter ved 70°C og afkølet kraftigt på is. Denne opløsning blev fordelt på 5x15-30% sucrosegradienter i 5 mM Tris HCl (pH 7,5), 10

mM EDTA, 0,1M NaCl, 1% Sarkosyl (i sterile siliconiserede Beckman SW40-rør). En sjette gradient blev belastet med 3'-endemærket RNA som en markør (se nedenfor). Efter en centrifugering i 16 timer ved 19.000 rpm (20°C) blev 5 gradienterne fraktioneret (1 ml fraktioner). RNA'et fra hver fraktion af den gradient, der svarede til den, der indeholdt markør-mærket RNA, blev fældet med 2,5 voluminer ethanol i nærværelse af bærer-gær-RNA (10 µg) og underkastet formaldehyd-agarosegel-elektroforese. Lehrach 10 et al., Biochemistry (1977) 16:4743, for at bestemme hvilken fraktion, der indeholdt BDV-RNA-bånden. Fraktioner svarende til dem indeholdende BDV-RNA'et blev samlet fra de parallelle gradienter og fældet med 2,5 voluminer ethanol, vasket med 80% ethanol og lagret ved 15 -20°C i 70% ethanol.

Det rensede virale RNA blev mærket med  $^{32}\text{P}$ -pCp (3000 Ci/n mol) ifølge England et al., Meth. Enzymol. (1980) 65:65-74, og analyseret ved agarosegel-elektroforese i nærværelse af 2,2 M formaldehyd, som beskrevet 20 af Lehrach et al., (supra). Fluorografi gennemførtes med  $^3\text{H}$ -Enhancer (NEN) som anbefalet af fabrikanten.

Hovedparten af radioaktiviteten var forbundet med lavmolekylært RNA (mindre end 2 kb), men en lille mængde fandtes i et højmolekylært bånd på ca. 12,5 kb, identificeret som RNA ved hjælp af mærkningsegenskaber med 25 RNA-ligase, dets sensibilitet over for RNase og alkali og modstandsevne over for DNase og proteinase K. I overensstemmelse med andre rapporter vedrørende togavira fra flavivirus-gruppen båndt BDV-RNA'et ikke ti oligo-dT- 30 cellulose, hvilket viser enten fravær af en polyA-strækning ved 3'-enden, eller at polyA'et,, om tilstede, er meget kort. Kontrol-Sindbis-virus-RNA blev korrekt tilbageholdt af den samme søjle.

Disse egenskaber hos 12,5 kb båndet var identiske 35 med egenskaberne udvist af RNA ekstraheret fra BEKI-celler dyrket som følger:

BEKI-celler blev dyrket i 25 cm<sup>2</sup> formstofkolber, vasket 3 gange med infektionspuffer og inficeret ved multipliciteter på 50-100 pfu/celle med 1 ml BDV-opløsning. Efter 1 time ved 35°C blev der tilsat 4 ml infektionspuffer, og inkubationen blev fortsat. Efter 12, 15, 18, 21 og 36 timer (36 timer svarer til en komplet cyklus af BDV-replikation) blev det nyligt syntetiserede RNA mærket med <sup>3</sup>H-uridin (100 µCi/plade). Unificeret cellulært RNA indvundet efter 18 timers inkubation blev også analyseret. Efter 30 minutters mærkning blev det cellulære RNA ekstraheret under anvendelse af CsCl/guanidiniumthiocyanat-metoden ifølge Chirgwin et al., 1979 (supra). Pellet'en af RNA, vundet efter ultracentrifugering gennem en 5,7 M CsCl-pude, blev analyseret direkte ved formaldehyd-agarosegel-elektroforese, og gel blev tørret og fluorograferet. Ved alle de testede inkubationstider kunne et 12,5 kb bånd, der er fraværende i de uinficerede celler, påvises, hvilket bånd har de samme fysisk-kemiske egenskaber som vist af RNA'et ovenfor.

20

#### D.1.c. Fremstilling af cDNA

Det virale RNA isoleret fra virus'en under D.1.b. blev polyadenyleret ved metoden ifølge Sippel, Eur. J. Biochem. (1973), 37:31-40. Kort angivet blev den vurderede mængde på 0,7 µg af rensset BDV-RNA inkuberet i 5 ml 5 mM methykviksølvhydroxid i 10 minutter ved stuetemperatur og inkuberet i 6 minutter ved 37°C med 20 enheder polyA-polymerase (BRL) og 500 µCi <sup>3</sup>H-ATP (36 Ci/mmol, Amersham) i 50 µl 50 mM HCl (pH 7,5), 10 mM MgCl<sub>2</sub>, 2,5 mM MnCl<sub>2</sub>, 0,3 M NaCl, 1,5 mM 2-mercaptoethanol og indeholdende 2,5 µg RNase-frit BSA og 5 enheder human placentale ribonuclease-inhibitor (BRL). Efter phenol/chloroform-ekstraktion blev RNA'et rensset ved chromatografi på Sephadex G50 og fældet med 2,5 volumener ethanol. PolyA'et anvendtes til at fremstille prober og som template for cDNA-biblioteket.

Til fremstilling af prober blev 1  $\mu\text{g}$  polyA-RNA inkuberet i 10 minutter ved stuetemperatur i 5  $\mu\text{l}$  10 mM methykviksølvhydroxid og derefter 45 minutter ved 37°C med 40 enheder omvendt transkriptase i 100  $\mu\text{l}$  50 mM Tris HCl (pH 8,3), 10 mM  $\text{MgCl}_2$ , 1,5 mM 2-mercaptoethanol, 1 mM dATP, dGTP og dTTP, 10  $\mu\text{M}$  dCTP, 0,2 mg/ml actinomycin D, 5 enheder human placental ribonuclease-inhibitor, 500  $\mu\text{Ci}$  alpha- $^{32}\text{P}$ -dCTP (3000 Ci/mmol, Amersham) og 20  $\mu\text{g}$  oligonucleotider vundet ved partiel digereret med NDase I fra kalvethymus-DNA (vilkårlige primere). Efter 15 og 30 minutter blev der tilsat yderligere 10 enheder omvendt transkriptase. Efter phenol/chloroform-ekstraktion og Sephadex G50-søjlechromatografi blev RNA'et hydrolyseret med 0,1 M NaOH (1 time ved 65°C), hvorved vandtes enkeltstrengede cDNA-streng. Opløsningen blev neutraliseret med 0,1 M eddikesyre og sat direkte til hybridiserings-pufferen.

Til cDNA-biblioteket blev der anvendt to separate klonings-protokoller involverende dT (12-18)-primere eller vilkårlige (kalvethymus), DNA-afledte oligonucleotid-primere. RNA polyadenyleret in vitro som ovenfor beskrevet anvendtes. Omtrent 1  $\mu\text{g}$  polyadenyleret RNA blev inkuberet med 10 mM methykviksølvhydroxid i et 10  $\mu\text{l}$  volumen i 10 minutter ved stuetemperatur, og overskud af reagens blev titreret ved tilsætning af 1  $\mu\text{l}$  af 3M 2-mercaptoethanol-opløsning. Dette denaturerede polyA-RNA anvendtes straks i nærværelse af 50 mM Tris pH 8,0, 1 mM dATP, dGTP, dCTP og dTTP, 2,5  $\mu\text{g}/\text{ml}$  dT12-18-primere eller vilkårlige kalvethymus-oligonucleotid-primere, 10 mM  $\text{MgCl}_2$ , 10  $\mu\text{g}/\text{ml}$  actinomycin D, 100 enheder RNase-inhibitor (BRL) og 60 enheder omvendt transkriptase i et totalt volumen på 100  $\mu\text{l}$ .

Prøverne blev fortyndet til 400  $\mu\text{l}$  med en puffer indeholdende 10 mM Tris 7,0, 100 mM NaCl, 10 mM EDTA og 0,2% SDS, ekstraheret med phenol/chloroform, befriet

for dNTP'er ved Sephadex G50-chromatografi og fældet med ethanol.

Den fældede blanding af RNA- og cDNA-hybrider (10 µl) blev fortyndet i 50 ml S1-puffer (500 mM NaCl, 50 mM Na-acetat pH 4,5 og 1 mM ZnCl<sub>2</sub>) og digereret i 15 minutter ved stuetemperatur med 20 enheder S1-nuclease. Reaktionen blev stoppet ved fortynding af prøven til 500 ml med en puffer indeholdende 50 mM NaCl, 10 mM EDTA og 50 mM Tris pH 7,0, og digerering blev fortsat i 15 minutter ved stuetemperatur ved tilsætning af 20 µg/ml af RNAse A. Efter phenol- og chloroform-ekstraktion blev RNA:cDNA-hybriderne koncentreret ved ethanol-fældning og fraktioneret på en Sepharose CL4B-søjle fremstillet i en 1 ml formstof-pipette. Den udelukkede spids, indeholdende molekyler større end 80 basepar, blev samlet og ethanol-fældet, hvorved vandtes 50 ng hybrid for de dT-primede reaktioner og 200 ng hybrid for de med vilkårligt kalvethymus-fragment primede reaktioner.

Begge prøver blev "tail'et" for dC-rester under betingelser, der gav 15-25 rester pr. DNA- eller RNA-termini, og "annealet" til en dG-"tailet" pBR322-vektor lineariseret ved PstI-stedet (NEN) ved en vektor-koncentration på 0,1 µg/ml. De "annealede" plasmider blev transformeret ind i E. coli HB101 til Amp<sup>R</sup> for at opnå cDNA-biblioteket.

#### D.2. Screening af cDNA-biblioteket

Screening gjorde brug af en (+/-)-metode under anvendelse af mærkede dDNA'er fremstillet ud fra RNA isoleret fra uinficerede BEKI-celler (-probe) og fra RNA isoleret fra virus'en vundet efter fuldstændig lysis af cellerne (+probe). Kolonier af det E. coli-husede cDNA-bibliotek blev dyrket, lyseret på nitrocellulosefiltre (to replikaer) og probet. Hybridiseringspufferen anvendt til +probe indeholdt også et overskud af cellulært RNA

isoleret fra uinficerede BEKI-celler (10 mg/ml). De kolonier, der gav et klart signal med +probe og ingen respons med -proben, blev udvalgt. Ved denne metode blev der udvalgt 95 oligo dT-primede og 185 (vilkårlig primer)-primede kloner. Længden af insertionerne efter PstI-digerering varierede fra 400 til 4.000 basepar. Der vandtes intet virus-specifikt fuldlængde-cDNA.

En af klonerne, pDT28, med en 880 bp insertion udvalgtes til yderligere analyse. Dette fragment fra en PstI-digerering af plasmid DNA blev rensat ved acrylamidgel-elektroforese digereret med DdeI og MboI og derefter mærket med Klenow-fragmentet af DNA-polymerase I og de fire  $^{32}\text{P}$  dNTP'er, hvorved vandtes  $10^6$ - $10^4$  cpm/mg af insertion. Mærket insertion blev verificeret ved hybridisering til viralt RNA fraktioneret på en 0,9% agarosegel-elektroforese i nærværelse af formaldehyd (Smiley et al., Anal. Biochem. (1983) 131:365-372). Stringente hybridiseringsbetingelser anvendtes: præhybridiseringer og hybridiseringer foregik natten over ved  $42^\circ\text{C}$ , og 50% formamid anvendtes til hybridiseringer. Vask foregik ved  $65^\circ\text{C}$  først med  $2\times\text{SSC}$ , 0,1% SDS og derefter med  $0,2\times\text{SSC}$  og 0,1% SDS.

Ved den foregående verifikation anvendtes RNA fra uinficerede celler som negativ kontrol. Fravær af exogene virale sekvenser i genomet fra cellerne blev verificeret ved manglende evne hos cellulært DNA digereret med BamHI og EcoRI til at binde til pDT28 probe ved Southern blot-analyse. RNA'et fra inficerede celler efter 24 timers infektion ved en multiplicitet større end 1, og fra pellet'en af virus efter fuldstændig cellelysis anvendtes som positiver. Ingen hybridisering blev påvist med RNA'et fra de uinficerede celler, men insertionerne hybridiserede til et omtrent 13 kb-bånd af RNA'et isoleret fra de inficerede celler eller fra virus-pellet'en.

Plasmidet pDT28, der var påvist at indeholde en PstI-insertion, der binder til det virale RNA, anvendtes

til at probe cDNA-biblioteket for yderligere kloner, og den fuldstændige sekvens udvandedes ved "walking"-teknik. På denne måde vandtes otte yderligere plasmider, der spændte over hele 12,5 kb genomet af virusen. Positionerne af de overlappende insertioner er vist i Fig. 1. Som vist i Fig. 1, besidder pDT28-klonen en groft set central del af genomet. De 8 yderligere plasmider vundet fra cDNA-biblioteket på en måde analog med den ovenfor beskrevne, men ved anvendelse af den pågældende overlappende sekvens-holdige klon som probe, blev dyrket i *E. coli*, og plasmid-DNA'et blev isoleret. Insertionerne blev sekvensbestemt og påvist at indeholde overlappende dele. Resultaterne af denne sekvensbestemmelse er vist i Fig. 2, der tilvejebringer den totale genomiske RNA-sekvens konstateret ud fra insertionerne.

Orienteringen vist i Fig. 2 blev bestemt ved at subklone pDT28 ind i M13 i begge orienteringer, mærke den resulterende phag og anvende den mærkede phag som probe over for RNA, der var kendt som havende positiv polaritet. Dette foregik ved plet-hybridisering på nitrocellulosefiltre ved anvendelse af uinficeret celle-RNA, inficeret celle-RNA og viralt template-RNA. RNA'et fra inficeret celle og template-RNA'et skulle være af positiv polaritet. Derfor indeholder den M13-orientering, der hybridiserer til inficeret celle-RNA og viralt RNA, en streng af negativ art, og ud fra denne information kunne 5'- til 3'-sekvensen af insertioner fra pCT63 til pCT185 udledes.

Denne konklusion blev bekræftet ved analyse af sekvensen af pCT63, der viser dets evne til at danne den forventede hårnålestruktur ved 5'-enden, og ved fravær af yderligere kloner i cDNA-biblioteket med yderligere 5'-sekvenser foruden den fra pCT63.

D.3. Ekspresion af sekvenser, der indholder  $\beta$ Gal-BDV-fusioner i E. coli.

Tolv dele af BDV-genomet vandtes som følger:

(1) de totale cDNA-sekvenser per se, (2) produkter fra  
5 restriktionskløvning (med PstI eller BamHI eller begge)  
af de foregående cDNA'er, og (3) en ligeret sekvens vundet ved at ligere pCT185-cDNA'et med et fragment af et andet. (Se nedenstående tabel). Disse dele anvendtes til at indkode BDV-delene af fusions-proteinerne. Disse el-  
10 leve BDV-protein-indkodende sekvenser blev klonet ind i én af eller en blanding af pUR290, pUR291 og pUR292, der indeholder restriktionssteder, f.eks. BamHI- og PstI-steder i alle tre mulige læserammer med  $\beta$ -gal-codonerne, således at der blev indkodet fusionsproteiner ved den C-  
15 terminale del af  $\beta$ -galactosidaseproteinet (Ruther, U., et al., Embo. J. (1980) 2:1791-1794). Da alle tre mulige læserammer er tilvejebragt for de anvendte restriktionssteder, sikres den korrekte læseramme i mindst én af vektorerne for fusionsproteinet. Tabel I sammenfatter de  
20 fremstillede vektorer og BDV-sekvensen indeholdt i hver. Nucleotid-tal er som vist i Fig. 2.

Tabel I

	Navn	pUR-stam- produkt	BDV-insertion afledt af	BDV-nucleotider indeholdt i pUBVD-vektorer (tal som i Fig.2)
5	pUBVD1	pUR290	pCT63	1397-2607
	pUBVD2	pool	pCT36	2037-2574
	pUBVD4	pUR292	pCT183	2955-4560
	pUBVD5	pool	pDT28	5650-6450
10	pUBVD6	pUR290	pCT174	7225-10718
	pUBVD7	pool	pCT174	~9500-10811
	pUBVD8	pUR292	pCT65	10442-10811
	pUBVD9	pUR292	pDT65 + pCT185	10442-12470
	pUBVD10	pUR290	pCT185	11030-12457
15	pUBVD11	pUR290	pCT185	11405-12457
	pUBVD12	pUR291	pCT63	597-1397
	pUBVD13	pUR290	pDT28 + pDT17	~6000-~7800

Hver af de tolv cDNA-sekvenser blev blandet med  
 20 T4-ligase i nærværelse af PstI-digererede blandinger af  
 pUR290, 291 og 292 (eller af én af disse, hvis den kor-  
 rekte læseramme var udledt), og ligationsblandingen blev  
 transformeret ind i *E. coli* stamme D1210 (LacI<sup>-</sup>-mutant  
 af HB101) til Amp<sup>R</sup>. Gunstige transformanter blev bekræf-  
 25 tet ved hybridisering med mærket insertion, og isoleret  
 plasmid DNA blev analyseret ved restriktionsanalyse for  
 at bekræfte korrekt orientering. Ekspresion blev indu-  
 ceret i gunstige transformanter indeholdende korrekt  
 orienterede insertioner ved behandling med IPTG (1 mM)  
 30 på L-dyrkningsmedium indeholdende 40 µg/ml af ampicil-  
 lin. Tre timer efter induktion blev cellerne indvundet  
 og lyseret ved lydbehandling. Fusionsproteinerne blev  
 frembragt som inklusionslegemer, og inklusionslegemerne  
 blev indvundet ved metoden ifølge Klempnauer et al.,  
 35 Cell (1983) 33: 345-355, og lagret ved -20°C suspende-

ret i 10 mM Tris (pH 8,0), 1 mM EDTA. Omtrent 10-30 mg inklusionslegeme-proteiner vandtes pr. ml kultur.

#### D.4. Karakterisering af fusions-proteinerne

5 Fusions-proteinerne blev karakteriseret med hensyn til deres antigene egenskaber både i uopløselige og opløseliggjorte former.

Inklusionslegeme-proteiner opløseliggjort i 1% SDS eller 7 M urinstof efterfulgt af dialyse til en en-  
10 delig koncentration på 1 mg/ml er ikke-reaktive med sera fra inficerede kalve eller fra kaniner inficeret med rensset virus.

#### Fremstilling af antisera.

15 Både opløseliggjorte og ikke-opløseliggjorte inklusionslegemer blev injiceret i kaniner ved anvendelse af peri-lymfeknude-immuniseringer med 500 µg protein emulgeret med "Freund's complete adjuvant", idet der blev forstærket hver 4 uger (IM-injektion af 500 µg  
20 emulgeret i "adjuvant") og foretaget blodudtagning 10 dage efter forstærkning. Kontrol-antisera blev fremstillet fra inficerede kalve eller fra kaniner injiceret med rensset virus. Antisera'ne blev testet for immunoaktivitet ved ELISA og immunofluorescens, og ved Western  
25 blot og immunofældning.

Western blot og immunofældning gav komplementær information med hensyn til reaktivitet. Ved immunofældning omsættes den native proteinblanding med test-serum'et, og immunobundfaldet underkastes SDS-PAGE. Der-  
30 for fastsætter immunofældning immunoreaktivitet med det native protein.

Ved Western SDS blot-metoden gennemføres PAGE, før antisera'ne testes for fældning med proteinerne på gelen. Derfor fastsætter Western blot reaktivitet med  
35 denatureret protein.

Resultaterne fra disse processer er anført nedenfor.

#### Resultater

5 Kontrol-antisera'ne var immunoreaktive med hensyn til proteiner ekstraheret fra virus-pellet'en fremstillet på BEKI-celler og viste immunofældning med 76 kD proteinet, der formodedes at være den antigene hovedkomponent, såvel som mindre komponenter, der formodedes at  
10 være, ihvertfald delvis, virion-proteiner med molekylvægte på 36, 43, 47, 51 og 56 kD. Ingen immunofældning forekom, når kontrol-antisera'ne blev testet på Western blot. Kontrol-antisera over for infektion reagerer således med antigener i det native protein, men ikke efter  
15 denaturering.

#### Immunofældning og Western blot

De fleste af de antisera, der var dannet i respons til fusions-proteinerne, var negative både ved im-  
20 munofældningsprøven og, ligesom kontrol-antisera'ne, på Western blot.

Der var imidlertid undertagelser. Det antiserum, der var udviklet af fusionsprotein 7, immunofælder 36 kD-proteinet fra BEKI-dyrket virus og reagerer ved  
25 Western blot til 76 kD- og 51 kD-båndene. Antiserum fra fusion 5 immunofælder 3 størrelser af proteiner: 65, 98 og 105 kD, hvilket er størrelser, der ikke fældes af kontrol-antisera. Antiserum fra fusion 9 fælder et 58 kD-bånd, der heller ikke fældes af kontrol-antisera'ne.  
30 Betydningen af materialernes molekylvægt er ikke klar, da det ikke er klart, hvilket, om noget, af disse proteiner, der repræsenterer glycosylerede materialer med tilsvarende skift i molekylvægt.

ELISA

ELISA (gennemført ved metoden ifølge Bartlett et al., *Protides of the Biological Fluids*, H. Peeters, ed., Pergamon Press, Oxford, 1976, 24:767-770) gjorde brug af  
5 delvis rensede virus som antigen. Kun det antiserum, der var fremstillet mod fusionsprotein 7, var positivt ved en 1:40 titer. Serum fremstillet mod fusionsproteiner 5 og 11 havde titere på 1:4 og 1:8. Ikke-immune sera var negative.

10.

Immunofluorescens

Immunofluorescens gennemførtes ved anvendelse af mærkede levende eller fikserede inficerede celler. Det antiserum, der var fremstillet mod fusionsprotein 11,  
15 var let positivt med hensyn til immunreaktivitet med levende celler. På celler fikseret med methanol, acetone eller formaldehyd gav serum fremstillet fra fusionsprotein 7 den samme stærke respons som kontrol-antisera fra de inficerede dyr, mens antisera 5 og 3 var svagt positive  
20 over for proteiner, der var ekstraheret fra virus-pellet'en fremstillet på BEKI-celler.

## P A T E N T K R A V

1. Nucleotidsekvens, der koder for et til kvægdiarré relateret polypeptid, k e n d e t e g n e t ved, at den er en region af den genomiske sekvens for 5 kvægdiarrévirus vist på figur 2 eller en kombination af regioner i denne sekvens.

2. Nucleotidsekvens kodende for et viralt polypeptid, k e n d e t e g n e t ved, at dette i det væsentlige er identisk med det, der kodes for af den 10 genomiske sekvens for kvægdiarrévirus vist på figur 2.

3. Rekombinant ekspressionssystem, der i en forligelig værtcelle er i stand til at foretage produktion af et til kvægdiarrévirusrelateret polypeptid, k e n d e t e g n e t ved, at systemet omfatter 15 en DNA sekvens ifølge krav 1, der operativt er forbundet til en med værten forligelig kontrolsekvens.

4. System ifølge krav 3, k e n d e t e g n e t ved, at det opstrøms for den nævnte DNA sekvens og i 20 læseramme med denne yderligere inkluderer en fusioneret nucleotidsekvens kodende for et værtprotein eller en del deraf.

5. System ifølge krav 4, k e n d e t e g n e t ved, at den fusionerede DNA sekvens koder for en N-25 terminal del af  $\beta$ -galactosidase.

6. Rekombinant vektor, k e n d e t e g n e t ved, at den omfatter ekspressionssystemet ifølge krav 3.

7. Rekombinante værtceller, k e n d e t e g n e t ved, at de er transformeret med vektoren ifølge 30 krav 6 eller med en vektor omfattende systemet ifølge krav 4 eller 5.

8. Polypeptid, kendetegnet ved, at det i det væsentlige er identisk med hele aminosyresekvensen som vist på figur 2 eller med en region eller kombinationer af regioner deraf.

5 9. Polypeptid ifølge krav 8, kendetegnet ved, at det yderligere er fusioneret med et værtprotein eller en del deraf.

10. Vaccine, der er effektiv mod kvægdiarrévirus, kendetegnet ved, at den omfatter polypeptidet ifølge krav 8 samt farmaceutisk acceptable excipients.

11. Vaccine ifølge krav 10, kendetegnet ved, at den yderligere indeholder en immunogen partikel, hvilken partikel omfatter et polypeptid med en aminosyresekvens, der kan danne en partikel, når sekvensen frembringes i en eukaryotisk vært.

12. Vaccine ifølge krav 11, kendetegnet ved, at den partikeldannende aminosyresekvens er afledt af hepatits B virus.

20 13. Vaccine ifølge krav 12, kendetegnet ved, at den partikeldannende aminosyresekvens er afledt af HBsAg.

14. Fremgangsmåde til fremstilling af et polypeptid ifølge krav 8, kendetegnet ved, at man dyrker celler ifølge krav 7 og udvinder det rekombinante polypeptid.

15. Fremgangsmåde til fremstilling af en vaccine mod kvægdiarré, kendetegnet ved, at den omfatter udøvelse af fremgangsmåden ifølge krav 14 med en yderligere tilsætning af farmaceutisk acceptable excipients.

30

16. Anvendelse af et polypeptid ifølge et hvilket som helst af kravene 8 og 9 til fremstilling af prøvesæt til immunologisk påvisning af fusionsproteiner af kvægdiarrévirus.

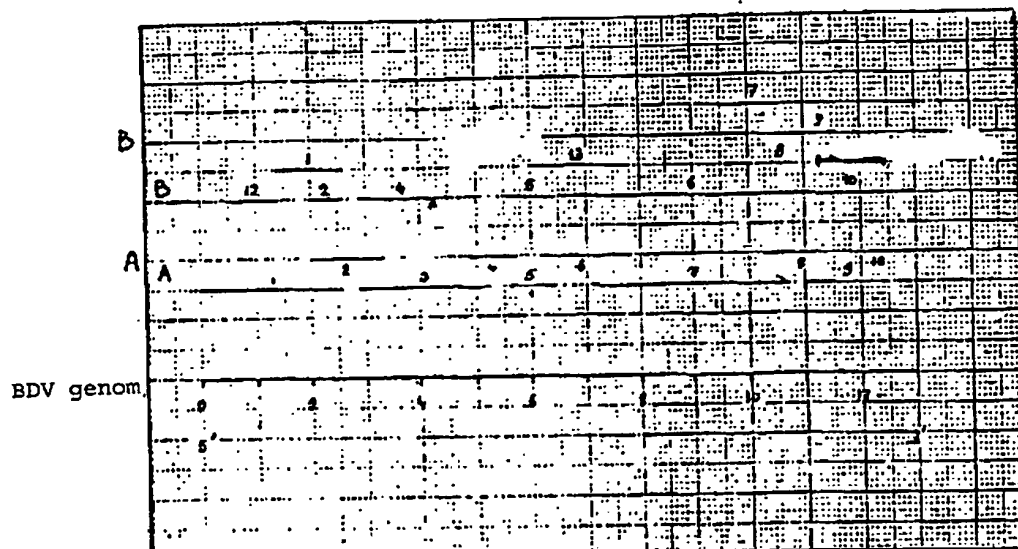
5           17. Anvendelse af en nucleotidsekvens ifølge krav 1 til konstruktion af som diagnostiske sonder nyttige oligonucleotidsekvenser.

18. Anvendelse af en nucleotidsekvens ifølge krav 1 til fremstilling af en vaccine, der som bærer  
10 omfatter levende viral vektor, der tillader ekspres-  
sion af et ønsket kvægdiarrévirusantigen sammen med  
et protein fra bæreren i inficerede celler.

15

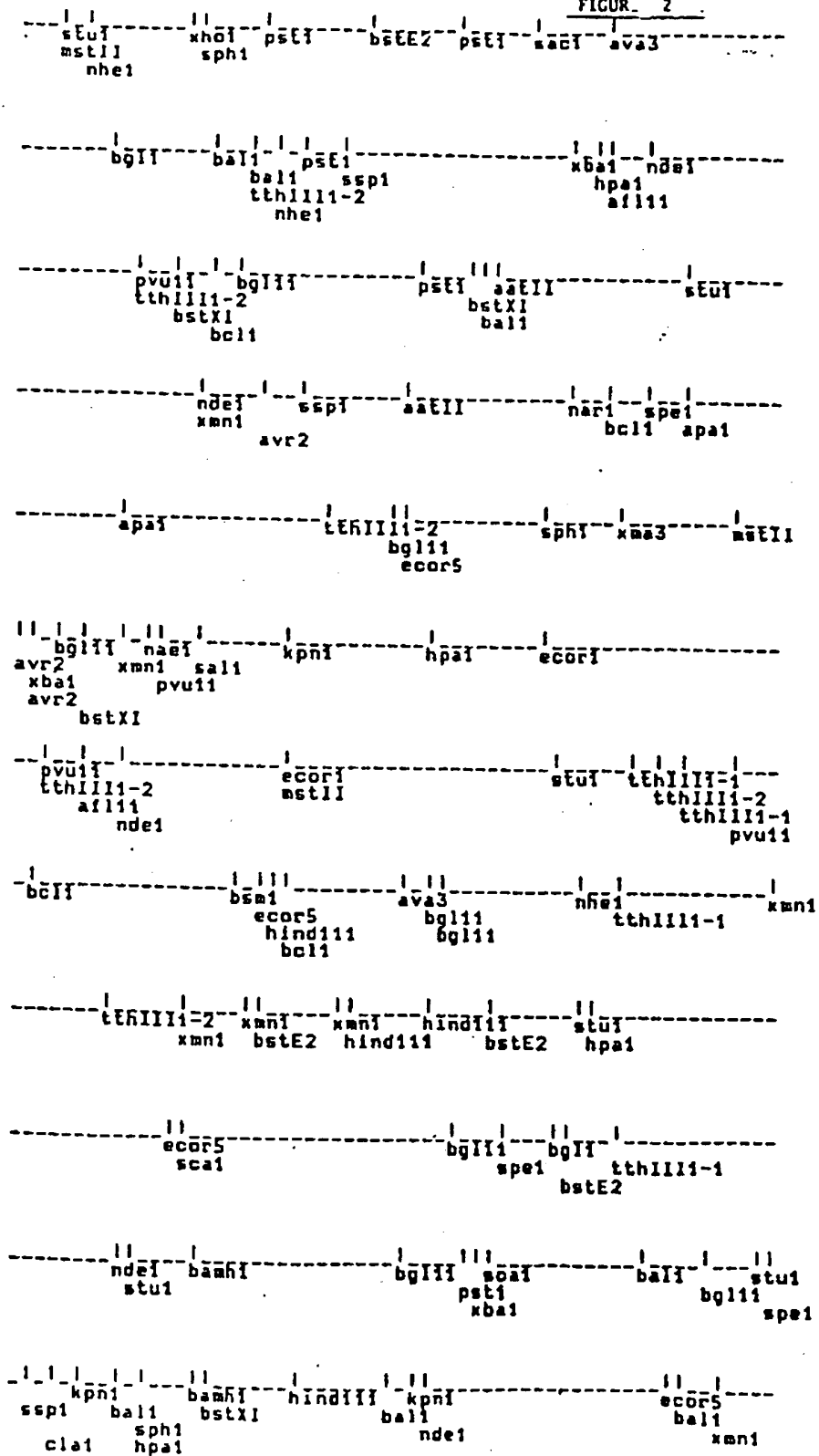
20

Figur 1



- A: Afbildning af 9 BDV cDNA kloner, der omspænder hele genomet. Klonerne blev fremstillet fra oligo dT primet cDNA (DT kloner) eller fra tilfældigt primet cDNA ved brug af kalvebrissel oligonucleotider (CT kloner). Navnene på klonerne er som følger: 1 - pCT63; 2 - pCT36; 3 - pCT183; 4 - pCT70; 5 - pDT28; 6 - pDT17; 7 - pCT174; 8 - pDT65; 9 - pCT185; 10 - pCT40.
- B: Anvendte cDNA fragmenter til konstruktion af ekspressionsvektorer for E.coli ved fusion til E.coli B-galactosidasegenet.

FIGUR 2



Figur. 2

$\frac{1}{\text{pVII}}$ ----- $\frac{1}{\text{aE11}}$   
 nco1

- 1 ATGTATACGAGAATTTCCTAACCTCGTATACATATTGGGCATTCTAAAAATAAATTAGGCC  
 CATATGCTCTTAAACGGATTGGAGCATATGTATAACCCGTAAGATTTTATTTAATCC5G  
 58 stu1, 61 ms111,
- 63 TAAGGGACAAATCCTCCTTAGCGAAGGCCGAAAAGAG6CTAGCCATGCCCTTAGTAGGAC  
 ATTCCTGTTTAGGAGGAATCGCTTCCGGCTTTTCTCCGATC66TACGGGAATCATCCTG  
 100 nhe1,
- 123 TAGCAAAAACAAGGAGGGTAGCAACAGTGGTGAGTTCGTTGGATGGCTGAAGCCCTGAGTA  
 ATCGTTTTGTTCCCTCCCATCGTTGTCAACCACTCAAGCAACCTACC6ACTTCGG6ACTCAT
- 183 CAGGGTAGTCGTACAGTGGTTCGACGCTTCTGTGTGACAAAGCCTCGAGGTGCCACGTGGACG  
 GTCCCATCAGCAGTCACCAAGCTGCGAAGCACACTGTTCCGAGCTCCACGGTGCACCTGC  
 223 xho1,
- 243 AGGGCATGCCACAGCACATCTAACCTGAGCGGGGGTCTTCAGGTGAAA6CGGTTTAA  
 TCCCGTACGGGTGTCTGTAGAAATTGGACTCGCCCCAAGCAA6TCCACTTTCGCCAAATT  
 246 sph1,
- 303 CCAACCGCTACGAATACAGCCTGATAGGGT6CTGCAGAG6CCCACTGTATTGCTACTAAA  
 GGTGGCGATGCTTATGTCCGACTATCCCAC6ACGTCTCCGGGT6ACATAAC6ATGATT  
 334 pst1,
- 363 <sup>MetGluLeuIleThrAsnGluLeuLeuTyrLysThrTyr</sup>  
 AATCTCTGCTGTACATGGCACATGGAGTTGATTACAAATGAACTTTTATACAAAACATAC  
 TTAGAGACGACATGTACC6GTACCTCAACTAATGTTTACTTGAAAATATGTTTTGTATG
- 423 <sup>LysGlnLysProAlaGlyValGluGluProValTyrAsnGlnAlaGlyAspProLeuPhe</sup>  
 AAACAAAACCCGCTGGAGTGGAGGAACCAAGTATATAACCAAGCAGGTGACCCCTTGT  
 TTTGTTTTGGGCGACCTCACCTCCTGGTTCATATATTG6TTCCTCCACTGGGAACAAA  
 468 bstE2,
- 483 <sup>GlyGluArgGlyValValHisProGlnAlaThrLeuLysLeuProHisLysArgGlyGlu</sup>  
 GGCGAGAGAGGAGTGGTTTCATCC6CAGGCCGACGCTAAAACCTGCCACATAAAAAGAGGGGAG  
 CCGCTCTCTCCTACCAAGTAGGCGTCC6CTGCGATTTTGACGGTGTATTTTCTCCCCTC
- 543 <sup>ArgGluValProThrAsnLeuAlaSerLeuProLysArgGlyAspCysArgSerGlyAsn</sup>  
 C6CGAAGTACCTACTAATCTGGCGTCTCTGCCAAAAGAGGGTACTGCA6GTCCGGGTAAAC  
 GCGCTTCATGGATGATTAGACCGCAGAGACGGTTTTTCTCCACTGACGTCCAGCCCATG  
 587 pst1,
- 603 <sup>SerLysGlyProValSerGlyIleTyrLeuLysProGlyProLeuPheTyrGlnAspTyr</sup>  
 AGCAAGGGACCCGIGAGTGGAAATCTACCTGAAACCGGG5CC6TTATTCTACCAGGATTAC  
 TCGTTCCTTGGGCACTCACCTTAGATGGACTTTGGCCCC6CAATAAGATGGTCTTAATG
- 663 <sup>LysGlyProValTyrHisArgAlaProLeuGluPhePheGlnGluAlaSerMetCysGlu</sup>  
 AAAGGACCCGTCATCATAGAGCTCCATTGGAGTCTTTTCAGGAAG6CCTCTATGTGTAG  
 TTTCTGGGAGATAGTATCTCGAGGTAACCTCAA6AAAGTCTTTC6GAGATACACACTC  
 682 sac1,
- 723 <sup>ThrThrArgArgIleGlyArgValThrGlySerAspGlyLysLeuTyrHisIleTyrVal</sup>  
 ACAACTAGAGGATTGGGAGAGTAACCTGGTAGTGATGGTAAATTGTACCACATTTATGTG  
 TGTGATCTTCTAACCCCTCTCATTGACCATCACTACCATTTAACATGGTGTAAATACAC
- 783 <sup>CysIleAspGlyCysIleIleValLysSerAlaThrLysTyrHisGlnLysValLeuLys</sup>  
 TGCATAGATGGATGCATAATAGTTAAGAG6CGCCACAAAATATCATCAAAA6GTACTCAA  
 ACGTATCTACCTACGTATTATCAATTCTCG6GGTGTTTTATAGTAGTTTTCCATGAGTT  
 794 ava3,
- 843 <sup>TrpValHisAsnLysLeuAsnCysProLeuTrpValSerSerCysSerAspThrLysAla</sup>  
 TGGGTCCACAACAACTAAATTC6CCTCTATGGGTTTCAAGCTGCTCCGACACAAAAGCA  
 ACCAGGTGTGTTGATTTAACGGGAGATACCCAAAGTTC6AC6A66CTGTGTTTTCGT

Figur 2

903 GluGlyAlaThrArgLysLysGlnGlnLysProAspArgLeuGluLysGlyArgMetLys  
GAAGGGGGCGACAAGAAAGAAGCAACAAAACCAGATAGGCTGGAAAAGGGGAGGATGAAG  
CTTCCCCTGCTTCTTTCTTCTGTTGTTTTGGTCTATCCGACCTTTTCCCCCTCTACTC

963 IleThrProLysGluSerGluLysAspSerLysThrLysProProAspAlaThrIleVal  
ATAACTCCTAAAGAGTGGGAGAAAGATAGTAAGACCAAAACCGCCAGATGCTACGATAGTG  
TATTGAGGATTTCTCAGCCTCTTTCTATCATTCTGGTTTGGCGGTCTACGATGCTATCAC

1023 ValAspGlyValLysTyrGlnValLysLysGlyLysIleLysSerLysAsnThrGln  
GTAGATGGTGTCAAATATCAGGTAAAGAAAAAGGAAAATCAAGAGTAAGAATACCCAG  
CATCTACCACAGTTTATAGTCCATTTCTTTTTCCCTTTTAGTCTCTATTCTATGGGT

1083 AspGlyLeuTyrHisAsnLysAsnLysProGlnGluSerArgLysLysLeuGluLysAla  
GACGGTTTGTACCACAACAAAATAAACCTCAAGAGTCAAGCAAGAACTAGAGAAAAGCC  
CTGCCAACATGGTGTGTTTTATTGGAGTCTCAGTGCCTTCTTGATCTCTTTCCG

1140 bgl1,

1143 LeuLeuAlaIrpAlaValIleAlaLeuValLeuPheGlnValAlaValGlyGluAsnIle  
CTGTTGGCATGGCAGTAATAGCTTGGTTTTGTTTCAAGTCCAGTGGGAGAGAACATA  
GACAACCGTACCCGTCATTATCGGAACCAAAACAAAGTTCAAGCTCACCTCTCTGTAT

1203 ThrGlnTrpAsnLeuGlnAspAsnGlyThrGluGlyIleGlnArgAlaMetPheGlnArg  
ACACAATGGAACCTTACAGACAATGGGACGGAAAGGAATACAACGGCCATGTTCCAAAGA  
TGTGTTACCTTGAATGTTCTGTTACCCTGCTTCTTATGTTGCCCGGTACAAGTTTCT

1263 GlyValAsnArgSerLeuHisGlyIleTrpProGluLysIleCysThrGlyValProSer  
GGCGTAAATAGAAGTCTGCATGGGATCTGGCCAGAGAAAATCTGTACAGGTGTCCCCTCC  
CCGCATTTATCTTCAGACGTACCCTAGACCGGTCTCTTTAGACATGTCCACAGGGGAGG

1290 ball,

1323 HisLeuAlaThrAspThrGluLeuLysAlaIleHisGlyMetMetAspAlaSerGluLys  
CACTTGGCCACTGATACAGAACTGAAGGCAATTGATGGTATGATGGATGCTAGCGAGAG  
GTGAACCGGTGACTATGTCTTGACTTCCGTTAAGTACCATACTACTACGATCGCTCTTC

1327 ball, 1333 tth111r, 1371 nhw1,

1383 ThrAsnTyrThrCysCysArgLeuGlnArgHisGluTrpAsnLysHisGlyTrpCysAsn  
ACAAATTACACATGCTGAGGCTCCAACGTCATGAGTGGAAACAAGCATGGTGGTGAAT  
TGTTAATGTGTACGACGTCCGAGGTTGCGGTACTCACCTGTTCTACCAACCACGTTA

1397 pst1,

1443 TrpTyrAsnIleGluProTrpIleValLeuMetAsnLysThrGlnAlaAsnLeuAlaGlu  
TGGTACAATATTGAACCTTGGATTGTTCTCATGAATAAAACCAAGCCAACCTTGTCTGAG  
ACCATGTTAACTTGAACCTAACAAAGTACTTATTTTGGGTTCCGTTGGAAACGACTC

1449 asp1,

1503 GlyGlnProProArgGluCysAlaValThrCysArgTyrAspArgAspSerAspLeuAsn  
GGTCAGCCACCAAGGAGTGTGCCGTTACATGCCGATGACCCGAGATAGTGAACCTAAAT  
CCAGTGGTGGTTCCCTCACACGGCAATGTACGGCCATACTGGCTCTACTACTGGATTTA

1563 ValValThrGlnAlaArgAsnSerProThrProLeuThrGlyCysLysLysGlyLysAsn  
ETAGTAACACAAGCTAGGAACAGCCCCACACCATGACAGGCTGCAAGAAAGGCAAGAAC  
CATCATTGTGTTGATCCTTGTGCGGGTGTGTAACCTGTCCGACGTTCTTTCCGTTCTG

1623 PheSerPheAlaGlyValLeuValGlnGlyProCysAsnPheGluIleAlaValSerAsp  
TTCCTTTGCAAGTGTGTTGGTACAAGGGCCCTTGCACCTTGAATAGCTGTAAAGTGA  
AAGAGGAAACGTCACACAACCATGTTCCCGAAACGTTGAACTTTATCGACATTCACTA

1683 ValLeuPheArgGluHisAspCysThrSerValIleGlnGlyThrAlaHisTyrLeuVal  
GTGCTGTTTAGAGAGCAGCAATGCAAGTGTGATTCAAGGCACGGCTCACTATCTGGTA  
CAGGACAAATCTCTGCTAACGTGTTCACACTAAGTTCCTGCGGAGTGTATAGACCAT

1743 AspGlyMetThrAsnSerLeuGluSerAlaArgGlnGlyThrAlaLysLeuThrThrTrp  
GACGGGATGACCAATTCCTAGAAAAGTCCAGGCAAGGGACCGCAAATTAACCTACTTGG  
CTGCCCTACTGGTTAAGAGATCTTTACGGTCCGTTCCCTGGCCTTCAATTGATGAACC

1760 xba1, 1790 hpa1,

1803 LeuGlyArgGlnLeuLysLysLeuGlyLysLysLeuGluAsnLysSerLysThrTrpPhe  
TTGGGTAGGCAGCTTAAGAAACTAGGGAAGAACTGGAAAACAAGAGTAAGACATGTTTT  
AACCCATCCGTGGAATCTTTGATCCCTTGTGACCTTTTGTCTCATTCTGTACAAA

1815 a111,

1863 GlyAlaTyrAlaAlaSerProTyrCysGluValGluArgArgLeuGlyTyrIleTrpTyr  
GGGGCATATGCAGCCTCTCCCTACTGCSAGGTAGAACGGAGGCTTGGTTACATCTGTAT  
CCCCGTATACGTGGAGAGGGATGACGCTCCATCTTGCTCCGAACCAATGTAGACCATA

Figur 2

1923 ThrLysAsnCysThrProAlaCysLeuProLysAsnThrLysIleValGlyProGlyArg  
 ACAAAGAATTGCACCCCTGCCTGTITACCAAAAAATACAAAGATCGTGGCCCCGGTAGG  
 TGTTCCTAACGTTGGGGACGGCAAAATGGTTTTTATGTTTCTAGCAACCCGGGCCATCC

1983 PheAspThrAsnAlaGluAspGlyLysIleLeuHisGluMetGlyGlyHisLeuSerGlu  
 TTCGACACCAATGCGGAGGATGGTAAAATACTGCATGAGATGGGGGGCCACTTGTACAGAG  
 AAGCTGTGGTTACGCCTCCTACCATTTTATGACGTA CTCTACCCCCGGTGAACAGTCTC

2043 ValLeuLeuLeuSerValValValLeuSerAspPheAlaProGluThrAlaSerValVal  
 GTGCTACTACTCTCAGTGGTAGTGGCTTCCGATTTCCGCTCCAGAGACAGCCAGTGTGGTA  
 CACGATGATGAGAGTCACCATCACGAAAGGCTAAAGCGAGGCTCTGTCTGGTACACCAT

2103 TyrLeuIleLeuHisPheSerIleProGlnGlyHisThrAspIleHisAspCysAspLys  
 TATTTAAATTTCTACATTTCTCCATCCACAAGGACACACTGACATACATGACTGTGATAAA  
 ATAAATTAAGATGTAAGAGGTAGGGTGTCTCTGTGTGACTGTATGTACTGACACTATTT

2163 AsnGlnLeuAsnLeuThrValGlyLeuThrThrAlaGluValIleProGlySerValTrp  
 AACCAACTAAACCTCACCGTAGGACTCACAAACAGCTGAAGTAATACCTGGGTTCAGTTTGG  
 TTGGTTGATTTGGAGTGGCATCCTGAGTGTTCGACTTCATTATGGACCCAGTCAAACC

2194 pvu11, 2209 tth1111,

2223 AsnLeuGlyLysTyrValCysIleArgProAspTrpTrpProTyrGluThrAlaThrPhe  
 AATTTGGGCAAATATGTTTTGTATAAGACCAGATTGCTGGCCCTTATGAGACAGCCACGTTT  
 TAAACCCGTTTATACAAACATATTCTGGTCTAACCCACCGGAATACTCTGTCTGGTGAAG

2250 bst11,

2283 LeuValPheGluGluValGlyGlnValIleArgIleValLeuArgAlaLeuArgAspLeu  
 CTAGTGTGGTGAAGAGGTGGGTCAAGTGATCAGGATAGTCTTGGGGCTTTAAGAGATCTA  
 GATCACAAACTTCTCCACCAGTTCCTACTAGTCTTATCAGAACTCCGAAATTCCTCTAGAT

2308 bcl1, 2336 bgl11,

2343 ThrArgIleTrpThrAlaAlaThrThrThrAlaPheLeuValCysLeuValLysValVal  
 ACGCGCATTTGGACCGCTGCTACGACTACTGCATTCCTGGTATGTCTGGTGAAGGTGGTG  
 TGCGCGTAAACCTGGCGACGATGCTGATGACGTAAGGACCATAACAGACCCTTCCACCAC

2403 ArgGlyGlnValLeuGlnGlyIleLeuTrpLeuIleLeuIleThrGlyAlaGlnGlyLeu  
 AGAGGCCAAGTGTGCAAGGCATACTGTGTTGATACTCATAACAGGGGCGACAGGGCTC  
 TCTCCGGTTCACAACGTTCCGTATGACACCAACTATGAGTATTGTCCCCGTGTTCCCGAG

2463 ProValCysLysProGlyPheTyrTyrAlaIleAlaLysAsnAsnGluIleGlyProLeu  
 CCAGTTTGCAAACCCTGGCTTTTACTACGECATAGCCAAAAATAATGAGATCGGCCCTCTT  
 GGTCAAACGTTTGGGCGGAAAAATGATGCGGTATCGGTTTTTATTACTCTAGCCGGGAGAA

2523 GlyAlaThrGlyLeuThrThrGlnTrpTyrGluTyrSerAspGlyMetArgLeuGlnAsp  
 GGGGCTACGGGCTCACCACTCAGTGTATGAATACTCGGATGGGATGCGGCTG6CAGGAC  
 CCCCAGTCCCCGGAGTGGTGGTACCATACTTATGAGCCTACCCTACGCCGACGCTCTG

2574 pst1,

2583 ThrGlyValValValTrpCysLysGlyGlyGluIleLysTyrLeuIleThrCysGluArg  
 ACGGGAGTTGTAGTGTGGTGTAAAGGTGGAGAGATCAAATATCTAATTACATGTGAGAGG  
 TSCCCTCAACATCACACCACATTTCCACCTCTCTAGTTTATAGATTAATGTACACTCTCC

2643 GluAlaArgTyrLeuAlaIleLeuHisThrArgAlaLeuProThrSerValValPheGlu  
 GAAGCCAGGTATCTGGCCATTTCTACACAGGAGGCCCTGCCGACGCTCTGTAGTATTGAA  
 CTTCCGGTCCATAGACCCGGTAAGATGTGTCTCTCGGGACGGCTGACACATCATAAACTT

2647 bst11, 2656 ball, 2684 sat11,

2703 LysIleIleAspGlyLysGluGlnGluAspValValGluMetAspAspAsnPheGluLeu  
 AAAATCATAGATGAAAAGAACAAGAGGACGTAAGTGGAAATGGATGATAACTTTGAACTC  
 TTTTAGTATCTACTTTTCTTGTCTCTGTCATCACCTTACCTACTATTGAAACTTGAG

2763 GlyLeuCysProCysAspAlaLysProLeuValArgGlyLysPheAsnThrThrLeuLeu  
 GGTCCTTTCCCGTGTGATGCTAAACCTTGGTAAGGGGAAAATTAATAACAACACTTCTG  
 CCAGAAAACGGGCACACTACGATTTGGGAACCATTCCCTTTTAAATTATGTTGTGAAGAC

2823 AsnGlyProAlaPheGlnMetValCysProIleGlyTrpThrGlyThrValSerLeuCys  
 AATGGGCCAGCCTTCCAGATGGTTTGCCTATAGGATGGACAGGAACCTGTAGTCTGTGT  
 TTACCCGGTGGGAAGGTCTACCAACGGGATATCTACTGTCTTGTACTGACACTCAEACACA

2883 HisTrpSerAsnLysAspThrLeuAlaMetThrValValArgThrTyrLysArgHisArg  
 CACTGGTCCAATAAGGATACGTTAGCCATGACCGTTGTACGAACATACAAGAAGGACAGG  
 GTGACCAGGTTATTCCTATGCAATCGGTACTGGCAACATGCTTGTATGTTCTCCGTGTC

2940 stu1,

2943 ProPheProPheArgGlnGlyCysIleThrGlnLysValIleGlyGlyAspLeuTyrAsp  
 CCTTTCCCTTTAGGCAAGGCTGCATTACCCAGAAAGTCACTGGGGGAGACCCTACCGAT  
 GGAAGGGGAAAATCCGTTCCGACGTAATGGGCTTTCAGTAGCCCCCTCTGGAGATGCTG

Figur 2

1003 TGTGCTTGGGAGGGAAC TGGACTTGTGTACCGGGGACATACTACGATATGTAGATGGG  
ACACGGAACCTCCCTTGACCTGAACACATGGCCCCCTGTATGATGCTATAACATCTACCC

3063 ProValGluSerCysLysTrpCysGlyTyrLysPheHisLysSerGluGlyLeuProHis  
CCTGTGAGTCTTGCAAGTGGTGTGGTTACAAGTTTCATAAAAGTGAGGGTCTGCCACAC  
GGACAGCTCAGAACGTTCAACACCAATGTTCAAAGTATTTCACTCCCAGACGGTGTG

3123 PheProIleGlyLysCysLysLeuLysAsnGluSerGlyTyrArgGlnValAspGluThr  
TTCCCAATTGGTAAAGTGAAGCTGAAGAATGAAAGTGGCTACAGACAAGTAGATGAGACC  
AAGGGTTAACCGTTACCGTTCGACTTCTTACTTTCACCGATGTCTGTTTACTTACTCTGG

3183 SerCysAsnArgAspGlyValAlaIleValProThrGlySerValLysCysLysIleGly  
TCTTGCAACAGAGACGGTGTGGCTATAGTACCAACTGGTTCGGTGAAATGCAAGATAGGG  
AGAACGTTGTCTCTGCCACACCGATATCATGTTGACCAAGCCACTTTACGTTCTATCCC

3243 AspThrValValGlnValIleAlaMetAspAspLysLeuGlyProMetProCysArgPro  
GACACAGTGGTGAAGTCATAGCAATGGATGATAAGCTAGGGCCTATGCCCTGCAGACCA  
CTGTGTCACCAAGTTCAGTATCGTTACCTACTATTGATCCCGGATACGGAACGTCGGT

3301 nde1,

3303 TyrGluIleIleProSerGluGlyProValGluLysThrAlaCysThrPheAsnTyrThr  
TATGAAATCATTCCCAGTGAAGGGCCGGTAGAAAAGACGGCATGTACCTTCAACTACACA  
ATACTTTAGTAAGGGTCACTCCCCGGCCATCTTTTCTGCCGTACATGGAAGTTGATGTGT

3306 xmn1,

3363 LysThrLeuLysAsnLysTyrTyrGluProArgAspAsnTyrPheGlnGlnTyrMetLeu  
AAAACATTAAAGAACAAGTATTATGAGCCTAGGGATAATTTTCCAACAATACATGTTA  
TTTTGTAATTTCTTGTTCATAACTCGGATCCCTATTAATAAAGGTTGTTATGTACAA

3390 avr2,

3423 LysGlyGluTyrGlnTyrTrpPheAspLeuGluIleThrAspHisHisArgAspTyrPhe  
AAAGGGGAGTACCAATATTGGTTTGACTAGAGATCACTGACCACCACCGGGATTACTTC  
TTTCCCCTCATGGTTATAACCAAACTGGATCTCTAGTGACTGGTGGTGGCCCTAATGAAG

3436 esp1,

3483 AlaGluSerLeuLeuValIleValValAlaLeuLeuGlyGlyArgTyrValLeuTrpLeu  
GCTGAGTCCCTACTGGTGATAGTGGTTGCACTCCTGGGCGGTAGGTACGTGCTCTGGTTA  
CGACTCAGGGATGACCACTATACCAACGCTGAGGACCCGCCATCCATGCACGAGACCAAT

3543 LeuValThrTyrMetIleLeuSerGluGlnMetThrSerGlyArgProValTrpAlaGly  
CTGGTTACATATGATCCTATCAGAACAAATGACCTCGGGACGTCCAETATGGGCAAGT  
GACCAATGTATATACTAGGATAGTCTTGTCTTACTGGAGCCCTGCAGGTATACCCGTCCA

3583 aat11,

3603 GluIleValMetMetGlyAsnLeuLeuThrHisAspSerIleGluValValThrTyrPhe  
GAAATAGTGATGATGGGCAACCTGTAACACATGACAGTATTGAAGTGGTGACTTATTTT  
CTTTATCACTACTACCGTTGGACGATTGTGTACTGTATAACTTCAACCTGAATAAAG

3663 LeuLeuLeuTyrLeuLeuLeuArgGluGluAsnIleLysLysTrpValIleLeuIleTyr  
TTACTACTATACTACTACTAAGAGAGGAAAACATCAAAAAATGGGTTATACTTATATAC  
AATGATGATATGGATGATGATTCTCTCTTTTGTAGTTTTTACCAATATGAATATATG

3723 HisIleIleValMetHisProLeuLysSerValThrValIleLeuLeuMetValGlyGly  
CACATCATAGTAATGCACCCACTAAAATCAGTGACGGTGATACTGCTAATGGTTGGAGGG  
GTGTAGTATCATTACGTGGGTGATTTTAGTCACTGCCACTATGACGATTACCAACCTCCC

3783 MetAlaArgAlaGluProGlyAlaGlnSerPheLeuGluGlnValAspLeuSerPheSer  
ATGGCAAGGGCAGAACCGGTCGCCAGAGCTTCTTAGAGCAGGTGGACCTGAGTTTTTCA  
TACCGTTCGCTCTGGTCCGCGGGTCTCGAAGGATCTCGTCCACCTGGACTCAAAAAGT

3801 nar1,

3843 MetIleThrLeuIleValValGlyLeuValIleAlaArgArgAspProThrValValPro  
ATGATCAGCTCATTGTAGTAGGTCGGTCAATGCCAGGGCGGALCCCCACTGTGGTGCCA  
TACTAGTCCGAGTAACATCATCCAGACCAGTAACGGTCCGCGCTGGGGTGACACCACGGT

3844 bcl1, 3902 spe1,

3903 LeuValThrIleValAlaAlaLeuArgValThrGlyLeuGlyPheGlyProGlyValAsp  
CTAGTCACAATAGTTGCAGCACTAGGGTAACGGGACTAGGCTTTGGGCCCGAGTGGAT  
GATCAGTGTATCAACGTCGTGACTCCCATTTGCCCTGATCCGAAACCCGGGCTCACCTA

3948 apa1,

Figur 2

763 ValAlaMetAlaValLeuThrLeuThrLeuLeuMetIleSerTyrValThrAspTyrPhe  
GTAGCTATGGCAGTCCCTAACCTTGACCTACTGATGATTAGTTATGTGACAGACTACTTC  
CATCGATACCGTCAGGATTGGAAGTGGGATGACTACTAATCAATACACTGTCTGATGAAG

823 ArgTyrLysArgTrpLeuGlnCysIleLeuSerLeuIleAlaGlyValPheLeuIleArg  
AGGTACAAAAGGTGGCTACAATGTATCCTCAGCTTAATAGCCGGGGTTTTCTTATACGA  
TCCATGTTTTCCACCGATGTTACATAGGAGTCGAATTATCGGCCCCAAAAGGAATATGCT

083 SerLeuLysHisLeuGlyGluIleGluThrProGluLeuThrIleProAsnTrpArgPro  
AGCCTTAAACATCTGGGCGAGATTGAGACCCCTGAGCTGACCATACCGAAGTGGAGGCCA  
TCGGAATTTGTAGACCCGCTCTAACTCTGGGACTCGACTGGTATGGCTT6ACCTCCGGT

143 LeuThrPheIleLeuLeuTyrLeuThrSerAlaThrValValThrArgTrpLysValAsp  
CTAACCTTCATACTATTGTACCTGACTTCAGCAACAGTTGTACACAGTGGAAAGTTGAC  
GATTGGAGTATGATAACATGGACTGAAGTCGTTGTCAACAGTGTGCTACCTTCAACTG

203 IleAlaGlyIleLeuLeuGlnGlyProGlnSerPheCysOP  
ATAGCTGGCATATTACTGCAAGGGCCCCAATCCTTCTGCTGATTGCCACCTATGGGCT  
TATCGACCGTATAATGACGTTCCCGGGGTTAGGAAGACGACTAACGGTGGATACCCGA  
4224 apa1,

261 GACTTCTGACCCCTTGATTTGATCCTGCCACCCACGAATTAGTCAAGTTGACTACCTG  
CTGAAGGACTEGGAACATAACTAGGACGGGTGGGTGCTTAATCAGTCAACAT6ATGGAC

321 AAGACCGTCAAGACTSATGTGGAAAAGAGTTGECTAGGGGGGGTGGACTACAAGACAATT  
TTCTGGCAGTTCTGACTACACCTTTTCTCAACCGATCCCCCACCTGATGTTCTGTTAA

381 MetAspGluSerGlyGluGlyValTyrLeuPheProSerLysGln  
GGCTCTATTTATGATATGGATGAAAGTGGAGAGGGCGTGTACCTTTTCCCATCCAAACAG  
CCGAGATAAAATACTATACCTACTTTCACCTCTCCCGCACATGGAAAAGGGTAGGTTGTCT

4441 AsnGlyLysLysAsnValSerIleLeuLeuProLeuIleArgAlaThrLeuIleSerCys  
AATGGCAAGAAAAATGTAGCATACTCTTGGCCCTCATTAGAGCTACACTAATAAGCTGT  
TTACCGTTCTTTTACAGTCGTATGAGAACGGGGAGTAATCTCGATGTGATTATTCGACA  
4497 tth1111,

4501 IleSerSerLysTrpGlnMetValTyrMetAlaTyrLeuThrLeuAspPheMetTyrTyr  
ATCAGCAGCAAATGGCAGATGGTGTACATGGCTTACCTAACCTTGGACTTTATGACTAC  
TAGTCGTGCTTTACCGTCTACCACATGTACCGAATGGATTGGAACCTGAAATACATGATG

4561 IleHisArgLysValIleGluGluIleSerGlyGlyThrAsnValIleSerArgValIle  
ATAGCAGAAAAGTTATAGAAGAGATCTCA6GGGGCACCAATGTGATATCTA6GGT6ATA  
TATGTCTTTCCAATATCTTCTAGAGTCCCCCGTGGTTACACTATAGATCCCACATAT

4583 bq111, 4605 ecor5,

4621 AlaAlaLeuIleGluLeuAsnTrpSerMetGluGluGluGluSerLysGlyLeuLysLys  
GCAGCACTCATAGAGCTAAACTGGTCTATGGAAGAAGAAGAAAGCAAGGGLTAAAGAAAG  
CGTCGTGAGTATCTCGATTTGACCAGATACCTTCTTCTTCTTCTTCTTCCCGAATTTCTTC

4681 PhePheIleLeuSerGlyArgLeuLysAlaLeuIleIleLysHisLysValArgAsnGln  
TTTTTTATACTATCTGGGAGGTTGAAGGCCCTTATAATAAAGCATAAG6TTAGGAACCA6  
AAAAAATATGATAGACCCCTCCAACCTCCGGGAATATTATTTCTGATTCCAATCCTTGGTC

4741 ThrValAlaSerTrpTyrGlyGluGluGluValTyrGlyMetProLysValValThrIle  
ACCGTAGCAAAGCTGGTATGGGGAGGAAGAAGTCTAC66LATGCCAAAAGTAGTGACCATA  
TGGCATCGTTC6ACCATACCCCTCCTTCTTCTCAGATGCCGTAC6GTTTTCTACTG6TAT

4778 sph1,

4801 IleArgAlaCysSerLeuAsnLysAsnLysHisCysIleIleCysThrValCysGluAla  
ATAAGGGCTTGCTCACTAAACAAGAACAAGCATTGCATAATATGCACAGTATGTGAGGCT  
TATTC6CGAACGAGTGATTTGTTCTT6TTCGTAACGTATTATACGT6TCAACACTCCGA

4861 LysLysTrpLysGlyGlyAsnCysProLysCysGlyArgHisGlyLysProIleThrCys  
AAGAAGTGGAAAGGGTGGCAACTGCCCTAAATGCGGGCCGACGGGAAGCCCTCACTTGT  
TTCTTCACTTCCCACCGTTGACGGGATTTACGCGGGCG6TGCCTTCC6GTA6TGAACA  
4893 xaa3,

4921 GlyMetThrLeuAlaAspPheGluGluArgHisTyrLysArgIlePheIleArgGluGly  
GG6ATGACTCTAGCGGATTTTGAAGAGAGCCACTACAAGAGAATTTTATAAGAGAGGGT  
CCCTACTGAGATCGCCTAAACTTCTCTCGTGATGTTCTCTTAAAGTATTCTCTCCCA

4981 ThrPheGluGlyProPheArgGlnGluHisSerGlyPheValGlnTyrThrAlaArgGly  
ACATTGAAAGGACCTTCCAGCCAGGAACATAGCG6TTTTGTACAATACACCECTAG6GGA  
TGTAA6CTTCTGGGAAGTCCGTCTTGTATCGCCCAAACATGTTATGTG6CGATCCCT

Figur 2

341 GlnLeuPheLeuArgAsnLeuProIleLeuAlaThrLysValLysMetLeuMetValGly  
 CAATTGTTCCCTGAGGAATTAACCCATATTGGCAACCAAGTAAAAATGCTTATGGTAGGC  
 GTTAAACAAGGACTCCTTAAATGGGTATAACCGTTGGTTTCATTTTTACGAATACCATCCG  
 5049 mst11,

101 AsnLeuGlyValGluIleGlyAspLeuGluHisLeuGlyTrpIleLeuLysMetGlnIle  
 AACCTAGGGGTAGAAATCGGTGATCTAGAACACCTAGGATGGATCTTAAAAATGCAGATC  
 TTGGATCCCATCTTTAGCCACTAGATCTTGTGGATCCTACCTAGAATTTTTACGTC TAG  
 5103 avr2, 5124 xba1, 5133 avr2, 5156 bgl11,

161 PheValLysThrLeuThrGlyLysThrIleThrLeuGluValGluProSerAspThrIle  
 TTCGTGAAAACCCCTGACCGCAAGACCATCACCTGGAGGTGGAGCCCAAGTGCACCATC  
 AAGCACTTTTGGGACTGGCCGTTCTGGTAGTGGACCTCCACCTCGGGTCACTGTGGTAG  
 5186 bst11,

221 GluAsnValLysAlaLysIleGlnAspLysGluGlyIleProProAspGlnGlnArgLeu  
 GAGAACGTGAAGGCCAAGATCCAGGATAA6GAAGGCTTCCCCCTGACCAAGCAGAGGCTC  
 CTCTTGCACTTCCGGTCTAGGTCCTATTCTTCCGTAAGGGGACTGGTCTGCTCCGAG  
 5251 xmn1,

281 IlePheAlaGlyLysGlnLeuGluAspGlyArgSerLeuSerAspTyrAsnIleGlnLys  
 ATCTTTGCCGGCAAGCAGCTGGAAGATGGCCGCTCTCTTTCTGATTACAACATCCAGAAA  
 TAGAAACGGCCGTTCTCGACCTTCTACC66CGAGAGAAAGACTAAT6TT6TA6GTCTTT  
 5287 nae1, 5296 pvu11,

341 GluSerThrLeuHisLeuValLeuArgLeuArgGlySerGlyProAlaValCysLysLys  
 GAGTCGACCCTGCACCTGGTCTCGTCTGAGGGGTAGTGGGCTGCCGTGTGCAAAAAG  
 CTCAGCT66GACGTGGACCAAGGCGAGACTCCCATCACCCGACGGCACACGTTTTTC  
 5343 sal1,

3401 IleThrGluHisGluLysCysHisValAsnIleLeuAspLysLeuThrAlaPhePheGly  
 ATTACTGAGCATGAGAAATGCCATGTCAACATACTAGACAAAATTGACCGCATTTTTCCGGG  
 TAATGACTCGTACTCTTACGGTACAGTTGTATGATCTGTTAACTGGCGTAAAAAGCCC

5461 ValMetProArgGlyThrThrProArgAlaProValLysIleProThrAlaLeuLeuLys  
 GTTATGCCAAGAGGTACCACACCAAGGGCTCCGGTGAAGATTCCAACCGCATTTGCTAAAA  
 CAATACGGTTCTCATGGTGTGGTTCCCGAGGCCACTTCTAAGGTTGGCGTAACGATTT  
 5473 kpn1,

5521 ValArgArgGlyLeuGluThrGlyTrpAlaTyrThrHisGlnGlyGlyIleSerSerVal  
 GTGAGGAGGGGACTGGAAACCGGATGGGCCTACACACACCAAGGCGGATAAGCTCAGTA  
 CACTCCTCCCTGACCTTTGGCCTACCCGGATGTGTGTGGTTCCGCCATTTCGAGTCAT

5581 AspHisValThrAlaGlyLysAspLeuLeuValCysAspSerMetGlyArgThrArgVal  
 GTCATGTCACCGCAGGCAAGACCTACTGGTTTGTGATAGTATGGGTAGGACAAGAGTG  
 CTGGTACACTGGCGTCCGTTTCTGGATGACCAAACTATCATACCCATCCTGTTCTCAC

5641 ValCysGlnSerAsnAsnLysLeuThrAspGluThrGluTyrGlyValLysThrAspSer  
 GTTTGCCAAAGTAACAACAAGTTAACTGATGAGACAGAATATGGTGTCAAGACGGACTCC  
 CAAACGGTTTCATTGTTGTTCAATTGACTACTCTGTCTTATACCACAGTTCCTGAGG

5661 hpa1,

5701 GlyCysProAspGlyAlaArgCysTyrValLeuAsnProGluAlaValAsnIleSerGly  
 GGTATGTCAGATGGTGGCAGGTGCTACTGATTAATCCAGAGGCAGTAAATATATAC66G  
 CCTACAGTCTACCACGGTCCACGATGCATAATTA66TCTCCGTCATTTATATAGTCCC

5761 SerLysGlyAlaAlaValHisLeuGlnLysThrGlyGlyGluPheThrCysValThrAla  
 TCCAAGGGAGCTGCTGTACACCTCCAAAAACAGGTGGGGAATTCACATGTGTTACTGCA  
 A66TCCCTCGACGACATGTGGAGGTTTTTGTCCACCCCTTANGTGTACACAATGACGT  
 5800 ecor1,

5821 SerGlyThrProAlaPhePheAspLeuLysAsnLeuLysGlyTrpSerGlyLeuProIle  
 TCGGGAACCTCCAGCCTTCTTTGACCTGAAAAATTT6AAGGGATGGTCAGGTCTACCCATA  
 AGCCCTTGAGGTCGGAAGAAACTGGACTTTTTAACTTCCCTACCA6TCCAGAT666TAT

5881 PheGluAlaSerSerGlyArgValValGlyArgValLysValGlyLysAsnGluGluSer  
 TTTGAGGCTTCTAGTGGCAGGGTGGTCCGCGAGGTTAAAGTAGGAAAGAAAT6AGGAATCC  
 AAACCTCGAAGATCACCGTCCACCAAGCCGTCATTTTCAATTTTCACTCTTACTCCTTAGG

5941 LysProThrLysLeuMetSerGlyIleGlnThrValSerLysSerThrAlaAspLeuThr  
 A66CCCAAAAATTAATGAGTGGTATCCAAACCGTCTCAAAAAGCACAGCCGATTTAACA  
 TTCGGGTGTTTTAATTACTCACCATAGGTTGGCAGAGTTTTTCGTGTGGGCTAAATTTG

Figur 2

6001 GluMetValLysLysIleThrSerMetAsnArgGlyAspPheLysGlnIleThrLeuAla  
GAGATGGTCAAGAAGATAACCAGCATGAACAGGGGAGACTTTAAGCAGATAACCCCTTGCA  
CTCTACCAGTCTTCTATTGGTCTGACTTGTCCCTCTGAAATTCGTCTATTGGGAACT

6061 ThrGlyAlaGlyLysThrThrGluLeuProLysAlaValIleGluGluIleGlyArgHis  
ACAGGGGCAGGAAAACTACAGAAGTCCCAAAGGCAGTGTATAGAGGAGATAGGAAGACAC  
TGTCCCCGTCTTTTGTATGTCCTGAGGGTTCCCTCACTATCTCTCTATCCTTCTGTG

6121 LysArgValLeuValLeulleProLeuArgAlaAlaAlaGluSerValTyrGlnTyrMet  
AAGCGGGTGTAGTGTATTACCATTGAGAGCAGCAGCTGAGTCAGTCTATCAATACATG  
TTCGCCCCAGATCAGGAATATGGTAACTCTGTCCTCACTCAGTCAGATAGTTATGTAC  
6155 pvu11, 6158 tth1111,

6181 ArgLeuLysHisProSerIleSerPheAsnLeuArgIleGlyAspMetLysGluGlyAsp  
AGATTGAAACATCCAGTATCTCTTCAACTTAAAGATAGGGGACATGAAAGAAAGGGGAC  
TCTAACTTTGTAGGGTCATAGAGGAAGTTGAATTCCTATCCCTGTACTTTCTCCCTG

6210 sfl11,

6241 MetAlaThrGlyIleThrTyrAlaSerTyrGlyTyrPheCysGlnMetProGlnProLys  
ATGGCAACTGGGATCACCTACGCCTCATATGGATATTTTCCAAATGCCGAGCCGAAG  
TACCCTTGACCCTAGTGGATGCGGATATACCTATAAAAACGGTTTACGGCTCGGCTTC

6266 nde1,

6301 LeuArgAlaAlaMetValGluTyrSerTyrIlePheLeuAspGluTyrHisCysAlaThr  
CTCAGGGCCGCAATGGTAGAGTATTCAACATATTTCTGGATGAGTACTACTGTCTACT  
GAGTCCCGGCTTACCATCTATAAGTATGTATAAAGACCTACTCATAGTGCACAGATGA

6361 ProGluGlnLeuAlaValIleGlyLysIleHisArgPheSerGluSerIleArgValVal  
CCTGAGCAGTGGCTGTATAGGAAAAATTCACAGATTTTCTGAAAGCATAAAGGGTGGT  
GGACTCGTCAACCGACAGTATCTTTTAAAGTGTCTAAAAGACTTTCGTATCCCAAA

6421 AlaMetThrAlaThrProAlaGlySerValThrThrThrGlyGlnLysHisProIleGlu  
GCTATGACCGCCACCCAGCAGGGTCACTAACTACAACAGGGCAAAAACACCCAAATAGAA  
CGATACTGGGGTGGGGTCTGCCAGTCATTGATGTTGTCCCTTTTGTGGGTTATCTT

6481 GluPheIleAlaProGluValMetLysGlyGluAspLeuGlySerGlnPheLeuAspIle  
GAATTGACCTCCTGAGGTGATGAAAGGGGAAAGACTTGGAAAGCCAGTTCCTTGACATA  
CTTAAGTATCGAGGACTCCACTACTTCCCTTCTGAAACCTTCGTCAGGAAGCTGTAT

6481 ecor1, 6493 mst11,

6541 AlaGlyLeuLysIleProValGluGluMetLysGlyAsnMetLeuValPheValProThr  
GCGGGGCTAAAAATCCGGTTGAGGAGATGAAGGGTAACATGCTGGTCTTCCACCCACA  
CGCCCCGATTTTTAGGGCCAACCTCTACTTCCCATTTGTACGACCAGAAGCATGGGTGT

6601 ArgAsnMetAlaValAspValAlaLysLysLeuLysAlaLysGlyTyrAsnSerGlyTyr  
AGAAACATGGCAGTTGATGTAGCCAAAGAACTAAAAGCCAAAGGGCTACAACCTCAGGGTAT  
TCTTTGTACCCTCAACTACATCGGTTCTTGTATTTTGGGTTCCCGATGTTGAGTCCCAT

6661 TyrTyrSerGlyGluAspProAlaAsnLeuArgValValThrSerGlnSerProTyrVal  
TACTACAGTGGGGAAGALCCGCTAACTTGAGGTTGGTAACATCACAGTCCCATACGCTC  
ATGATGTCACCCCTTCTGGCCGATTGAACTCCACCATTTGTAGTGTCAAGGGTATGCG

6721 ValValAlaThrAsnAlaIleGluSerGlyValThrLeuProAspLeuAspThrValVal  
GTAGTAGCCCAATGCCATTGAGTCAGGGGTAACGCTGCCAGATTTAGATACAGTTGTT  
CATCATGGGTGGTACGGTAACTCAGTCCCATTTGCGACGGTCTAAATCTATGTCAACAA

6781 AspThrGlyLeuLysCysGluLysArgValArgValSerSerLysIleProPheIleVal  
GACACAGGTCGAAGTGTGAAAAGAGGGGTGAGGGTGTATCAAAAATACCTTTCATAGTA  
CTGTGTCCAGACTTCACACTTTTCTCCCACTCCCAAGTAGTTTTTATGAAAAGTATCAT

6841 ThrGlyLeuLysArgMetAlaValThrValGlyGluGlnAlaGlnArgArgGlyArgVal  
ACAAGCCTTAAAAGAAATGGGTGTCACTGTGGGCGAACAGGCTCAGCCAAAGAGGCAAGGTA  
TGTCCGGAATTTTCTACCGACAGTGACACCCGCTTGTCCGATCGCTTCTCCCTCCCAT

6843 stu1,

6901 GlyArgValLysProGlyArgTyrTyrArgSerGlnGluThrAlaThrGlySerLysAsp  
GGTAGAGTGAAGCCCGGTAGGTACTATAGAAGCCAGGAAACAGCCGACCCGGTCAAGGAC  
CCATCTCACTTCCGGCCATCCATGATATCTTCCGTCTTTGTCTGCTGCCCCAGTTTCTG

6945 tth1111,

6961 TyrHisTyrAspLeuLeuGlnAlaHisArgTyrGlyIleGluAspGlyIleAsnValThr  
TACCACATGACCTGTTACAGGCACACAGGTATGGGATAGAGATGGAATCAACGTGACA  
ATGGTGATACTGGACAATGTCCGTGTGTCCATACCCCTATCTTCTACCTTAGTTCAGTGT

6973 tth1111, 7017 tth1111,

Figur 2

921 LysSerPheArgGluMetAsnTyrAspTrpSerLeuTyrGluGluAspSerLeuLeulle  
 AASTCCTTTAGGGAATGAATTACGATTGGAGCCTGTACGAGGAGGACAGCTTGTCTGATA  
 TTCAGGAAATCCCTTACTTAATGCTAACCTCGGACATGCTCCTCTGTGAAACGACTAT

9084 Thr61nLeuGluIleLeuAsnAsnLeuLeuIleSerGluAspLeuProAlaAlaValLys  
 ACCCAGCTGGAGATACTGAACAATCTACTCATCTCTGAAGACTACCAGCAGCAATAAA  
 TGGGTCGACCTCTATGACTTGTAGATGAGTAGAGACTTCTGGATGGTCGTCATTTT  
 7084 pvu11,

7141 AsnIleMetAlaArgThrAspHisProGluProIleGlnLeuAlaTyrAsnSerTyrGlu  
 AACATCATGCAAGGACTGATCACCCAGAACCAATCCAGCTTGCATACAACAGTTATGAG  
 TTGTAGTACCGTTCCTGACTAGTGGGCTTGGTTAGGTCGAACGTATGTTGTCAATACTC  
 7158 bcl1,

9201 ValGlnValProValLeuPheProLysIleArgAsnGlyGluValThrAspThrTyrGlu  
 GTCCAGGTCCCTGTACTGTTCCAAAAAAGGAATGGGGAGGTTACAGATACTTACGAG  
 CAGGTCAGGGACATGACAAAGGTTTTTATTCCCTTACCCCTCCAATGCTATGAAATGCTC

9261 AsnTyrSerPheLeuAsnAlaArgLysLeuGlyGluAspValProValTyrIleTyrAla  
 AACATCATGCAAGGACTGATCACCCAGAACCAATCCAGCTTGCATACAACAGTTATGAG  
 TTGTAGTACCGTTCCTGACTAGTGGGCTTGGTTAGGTCGAACGTATGTTGTCAATACTC

9321 ThrGluAspGluAspLeuAlaValAspLeuLeuGlyLeuAspTrpProAspProGlyAsn  
 ACCGAAGATGAAGACTGAGCAGTAGACTTCTAGGCTTGGACTGGCCGACCCAGGGAAC  
 TGGCTTCTACTTCTGGACCTCATCTGGAGATCGAACCTGACCGGGCTGGGTCCTTTC

9381 GlnGlnValValGluThrGlyLysAlaLeuLysGlnValValGlyLeuSerSerAlaGlu  
 CAGCAAGTAGTGGAGACTGGGAAAGCACTGAAGCAAGTGGTAGGACTGTCCTCTGCTGAG  
 GTCGTTCTACCTCTGACCCCTTCTGACTTCTGTTCAACATCTGACAGGAGACGACTC  
 7440 bsm1,

9441 AsnAlaLeuLeulleAlaLeuPheGlyTyrValGlyTyrGlnAlaLeuSerLysArgHis  
 AATGCCCTGCTCATAGCCCTGTTGGGTATGTAGGATATCAAGCTTTGTCAAAAAGACAC  
 TTACGGGACGAGTATCGGGACAAACCCATACATCTATAGTTGAAACAGTTTTTCTGTG  
 7475 ecor5, 7481 hind111,

9501 ValProMetIleThrAspIleTyrThrIleGluAspGlnArgLeuGluAspThrThrHis  
 CCCCCAATGATCAGACATATACACCATAGAAATCAAAAGACTAGAGGACACAAACCCAC  
 CAGGGTTACTAGTGTCTGTATATGTGTATCTTCTAGTTTCTGATCTCCTGTGTTGGGTG  
 7508 bcl1,

7561 LeuGlnTyrAlaProAsnAlaIleArgThrGluGlyLysGluThrGluLeuLysGluLeu  
 CTCGAATATGCACCTAATGCTATAAGAACTGAGGGGAAGGAGACTGAACTAAAGGAATTA  
 GAGGTTATACGTGGATTACGATATCTTGACTCCCTTCTCTGACTTGATTTCTTTAAT

7621 AlaValGlyAspMetAspArgIleMetGluSerIleSerAspTyrAlaSerGlyGlyLeu  
 GCAGTGGGTGACATGGACAGAATCATGGAATCCATCTCAGATTATGCATCAGGAGGTTG  
 CGTCAACCCACTGACTCTGTAGTACCTTAGGTAGAGTCTAATACGTAGTCTCTCCCAAC  
 7664 ava3,

7681 ThrPheIleArgSerGlnAlaGluLysValArgSerAlaProAlaPheLysGluAsnVal  
 ACATTCATAAGATCTCAGGCAGAGAAAGTAAGATCTGCCCTGCAATCAAAAGAAACGTE  
 TGTAAATATCTAGAGTCCGCTCTTTTCATTCTAGACGGGGACGTAAGTTCTTTTGCAC  
 7690 bgl11, 7711 bgl11,

7741 GluAlaAlaLysGlyTyrValGlnLysPheIleAspAlaLeuIleGluAsnLysGluThr  
 GAAGCTGCAAAAGGTTACGTCCTCAAAAGTTTATTGATGCTCTTATTGAAACAAAGAAAC  
 CTTGACGTTTTCCATGCAAGGTTTTCAAAATACTACGAGAATAACTTTTGTCTTTTGG

7801 IleIleArgTyrGlyLeuTrpGlyThrHisThrAlaLeuTyrLysSerIleAlaAlaArg  
 ATAATCAGATATGGCTTATGGGGAACACACACGGCCTTTACAAGATATTGCCGCAAGA  
 TATTAGTCTATACCGAATACCCCTTGTGTGTGCCGTGAAATGTTCTATAACGGCGTTCT

7861 LeuGlyHisGluThrAlaPheAlaThrLeuValIleLysTrpLeuAlaPheGlyGlyGlu  
 CTGGGGCATGAAACAGCATTTGCTACGCTAAGTATAAGTGGCTAGCCTTCCGGGGGTGAG  
 GACCCCGTACTTGTGCTAAACEATCCGATCACTATTTACCGATCGGAAAGCCCCCACTC  
 7902 nhe1,

7921 ProValSerAspHisValArgGlnAlaThrValAspLeuValValTyrTyrValMetAsn  
 CCGGTGTAGATCATGTGAGACAGGCGACCTTGAACCTGCTTATTATGATGATGAAAC  
 GGCCACAGTCTAGTACACTCTGTCCGCTGGCAACTGGACCAGCAATAATACACTACTG  
 7954 hha1114

Figur 2

7981 LysProSerPheProGlyAspSerGluThrGlnGlnGluGlyArgArgPheValAlaSer  
AAACCCCTCTTCCCAGGGGATCCGAAACCCAGCAGGAGGGGAGGCCGATTTCGTTGCCAGC  
TTTGGGAGAAAGGGTCCCCTAAGGCTTTGGGTCGTCTCCCTCCGCTAAGCAACGGTCG

8041 TTATTCAICTCCGCTCTGGCAACCTACACATACAAGACTTGGAACTACCACAACCTCTCC  
AATAAGTAGAGGCGAGACCGTTGGATGTGTATGTTCTGAACCTTGATGGTGTGGAGAGG

8101 LysValValGluProAlaLeuAlaTyrLeuProTyrAlaThrSerAlaLeuLysMetPhe  
AAGGTAGTAGAACCCAGCTTTGGCATACTCCCTACGCTACCAATACAAAACCTTACCTC  
TTCCATCATCTTGGTCGAAACCGTATGGAGGGGATGCGATGGTCACGTGACTTTTACAAG

8151 xmn1,

8161 ThrProThrArgLeuGluSerGluValIleLeuSerThrThrIleTyrLysThrTyrLeu  
ACCCCAACTAGACTGGAGAGCGAGTTATACTTAGCACTACAATATACAAAACCTTACCTC  
TGGGGTTGATCTGACCTCTCGCTCCAATATGAATCGTGATGTTATATGTTTGAATGGAG

8221 SerIleArgLysGlyLysSerAspGlyLeuLeuGlyThrGlyIleSerAlaAlaMetGlu  
TCAATAAGGAAGGGGAAAAGTGTGGACTCTTGGGTACAGGGAATAGTGCAGCAATGGAA  
AGTTATTCCTTCCCCTTTTCACTACCTGAGAACCCATGTCCCTAATCACGCCGTTACCTT

8281 IleLeuSerGlnAsnProValSerValGlyIleSerValMetLeuGlyValGlyAlaIle  
ATTCTGTCACAGAACCCTGATCGGTAGGCATATCTGTTATGCTGGGGTGGGGGCAATT  
TAAGACAGTGTCTTGGGCCATAGCCATCCGTATAGACAATACGCCCCACCCCGGTAA

8284 lth1111,

8341 AlaAlaHisAsnAlaIleGluSerSerGluGlnLysArgThrLeuLeuMetLysValPhe  
GCECTCACAAATGCCATTGAGTCTAGCGAACAAAAAGGACCCTGTTGATGAAAGTGTTC  
CGGCGAGTGTACGGTAACCTCAGATCGCTTGTTTTTCTGGGACAACCTACTTTCACAAG

8391 xmn1,

8401 ValLysAsnPheTrpSerGlnAlaAlaThrAspGluLeuValLysGluAsnProGluLys  
GTAATAAACTTCTGGAGCCAGGCAGCAACAGATGAATGGTGAAGGAAAATCCAGAAAA  
CATTTTTGAAGACCTCGGTCCGTGCTACTTAACCACTTCTTTTAGGTCCTTTT

8461 IleIleMetAlaLeuPheGluAlaValGlnThrIleGlyAsnProLeuArgLeuIleTyr  
ATAATAATGGCCCTATTTGAAGCAGTTCAGACAATGGTAACCTCTGAGGCTTATATAT  
TATTATTACCGGATAAACTTCGTCAGTCTGTTAACCACTTGGGAGACTCCGAATATATA

8479 xmn1, 8497 bstE2,

8521 HisLeuTyrGlyValTyrTyrLysGlyTrpGluAlaLysGluLeuSerGluArgThrAla  
CACCTGTATGGAGTTTACTACAAAGGCTGGGAAGCAAAAGAACTATCCGAGAGGACAGCA  
GTGGACATACCTCAAATGATGTTCCGACCCTTCGTTTTCTGATAGGCTCTCTGTCGT

8581 GlyArgAsnLeuPheThrLeuIleMetPheGluAlaPheGluLeuLeuGlyMetAspSer  
GGCAGGAACCTGTTCACTTTGATAATGTTTCAAGCTTTCGAACCTGTTAGGGAATGGACTCT  
CCGTCTTGGACAAGTGAACCTATTACAAGCTTCGAAAGCTTGACAATCCCTACCTGAGA

8586 xmn1, 8612 hind111,

8641 GluGlyLysIleArgAsnLeuSerGlyAsnTyrIleLeuAspLeuIleTyrSerLeuHis  
GAAGGGAAAGATAAGGAACCTGTCTGGAATATATCTTGGATTTGATCTATAGTTTACAT  
CTTCCCTTCTATTCTTGGACAGACCTTAAATATAGAACCCTAAACTAGATATCAAATGTA

8701 LysGlnIleAsnArgSerLeuLysLysValValLeuGlyTrpAlaProAlaProPheSer  
AAACGATAAACAGAAAGCTTGAAGAAAGTGGTCTTGGGTTGGCTCCCGCACCTTTAGT  
TTTGTCTATTTGCTTCGAACTCTTCCACAGGACCCCAACCCGAGGGCTGGAAAAATCA

8715 hind111,

8761 CysAspTrpThrProSerAspGluArgIleArgLeuProThrAspAsnTyrLeuArgVal  
TGTGACTGGACTCCTAAGTATGAGAGAATTAGGTTACCCACAGACAACCTATCTAAGAGTG  
ACACTGACCTGAGGATCACTACTCTTAATCCAATGGGTGCTGTTGATAAATCTCAC

8792 bstE2,

8821 GluThrLysCysProCysGlyTyrGluMetLysAlaLeuArgAsnValSerGlySerLeu  
GAGACTAAGTCCCATGTGGTATGAGATGAAAGCACTAAGCAACGTTAGTGGCAGTCTT  
CTCTGATTACCGGTAACCAATACTCTACTTTCGTGATTCCTTGAATCACCGTCAGAA

8881 ThrIleValGluGluLysGlyProPheLeuLysArgAsnArgProGlyArgGlyProVal  
ACTATAGTGAAGAGAAAGGGCTTTTCTCTGTAGGAACAAGCCCTGGTAGAGGGCAGTT  
TGATATCACCTTCTTTCCCGGAAAAGAGACATCCTTGTCCGGACCATCTCCCGGTCAA

8920 stu1, 8938 hpa1,

AsnTyrArgValThrLysTyrIyrAspAspAsnLeuAlaGluIleLysProValArgArg  
AACTATAGAGTTACAAAATACTATGATGACAACCTCGCAGAGATAAAGCCAGTTCGAAGA  
TTGATATCTCAATGTTTTATGATACTACTGTTGGAGCGTCTCTATTTCCGGTCAAGCTCT

Figur 2

9001 LeuGluGlyLeuValGluHisTyrTyrLysGlyValThrAlaArgIleAspTyrGlyLys  
 CTAGAAGGACTCGTGGACACTATTACAAGGTGTACAGCAAGGATAGATTATGGTAAG  
 GATCTTCCTGAGCACCTCGTATAATGTTCCACAGTGTCTTCTATCTAATACCGTTC

9061 GlyLysMetLeuLeuAlaThrAspLysTrpGluValGluHisGlyIleValThrArgLeu  
 GGAAAAATGCTGTTAGCCACTGATAAATGGAGGTGGAGCACGATATCGTAACTAGGTTG  
 CCTTTTACGACAATCGGTGACTATTTACCCTCCACTCGTGCCATAGCATTGATCCAAC

9121 AlaLysLysTyrThrGlyValGlyPheLysGlyAlaTyrLeuGlyAspGluProAsnHis  
 GCGAAGAAGTACACTGGTGTGGGTTCAAGGAGCATAACCTGGGTGACGAGCCCAACCC  
 CGCTTCTTCATGTGACCACAACCCAAGTTCCTTCGTATGGACCCACTGCTCGGGTGGTG

9181 ArgAspLeuValGluArgAspCysAlaThrIleThrLysAsnThrValGlnPheLeuLys  
 CGTGACCTAGTGGAAAGAGACTGTGCAACCATAAACAAAAATACAGTTCAGTTTTTGA  
 GCACTGGATCACCTTCTCTGACACGTTGGTATTGGTTTTTATGTCAAETCAAAAACCTT

9241 MetLysLysGlyCysAlaPheThrTyrAspLeuSerLeuSerAsnLeuThrArgLeuIle  
 ATGAAGAAGGCTGTGTCATTACCTATGACTTGTCCCTGTCCAATTTGACCAAGTTAAT  
 TACTTCTTCCGACACGTAAATGGATACTGAACAGGGACAGGTTAAACTGTTCAATTA

9301 GluLeuValHisLysAsnAsnLeuGluGluLysAspIleProAlaAlaThrLeuThrThr  
 GAATTGGTGACAAAAATAACCTTGAAAGAAAGACATAACAGCCGACATTAACAACA  
 CTAAACCAGTGTTTTTATTGGAACCTCTCTTCTGTATGGTCCGGCGGTGAATTTGT

9361 CysLeuAlaTyrThrPheValAsnGluAspIleGlyThrIleLysProValLeuGlyGlu  
 TCGCTAGCTTACACATTTGTGAATGAAGATATCGGGACTATAAAACCACTACTGGGGAG  
 ACGGATCGAATGTGTAACACTTACTTCTATAGCCCTGATATTTGGTTCATGACCCCTC

9388 scor5, 9408 sca1,

9421 ArgValIleAlaAspProValValAspIleAsnLeuGlnProGluValGlnValAspThr  
 ABAGTGATAGCCGACCCAGTGGTAGACATTAACCTTACAACCAAGTGCAGGTGCAACA  
 TCTCACTATCGGCTGGGTCAACATCTGTAATGAATGTTGGTCTTCAGTCCACCTATGT

9481 SerGluValGlyIleThrLeuValGlyArgAlaAlaLeuMetThrThrGlyIleThrPro  
 TCAGAGGTTGGGATCACTCTGTTGGAGAGCAGCCTTGTGACAACAGGTATTACACCC  
 AGTCTCCAACCTTAGTGAGACCAACCTTCTGCTCGGAACACTGTTGTCCATAATGTGG

9541 ValValGluLysThrGluProAsnAlaAspGlySerProSerSerIleLysIleGlyLeu  
 GTGGTTGAAAAACAGAGCCTAATGCCGATGELAGTCCAAGCTCTATAAAGATTGGACTG  
 CACCAACTTTTTTGTCTCGGATTACGGCTACCGTCAGGTCGAGATATTTCTAACCTGAC

9601 AspGluGlyCysTyrProGlyProArgProGlnAspHisThrLeuAlaAspGluIleHis  
 GACGAAGGATGTTACCCAGGGCTAGACCECAAGACACACTTTAGCTGACGAATACAT  
 CTGCTTCTACAATGGGTCCCGGATCTGGCGTCTGGTGTGAAATCGACTGCTTTATGTA

9661 SerArgAspGluArgProPheValLeuValLeuGlySerArgSerSerMetSerAsnArg  
 TCTAEGGATGAAAGGCCCTTTGTTTTGGTCTTGGGTTCAAGAAGTTCATGTCAAATAGA  
 AGATCCCTACTTCCGGAAACAAAACCAAGACCAAGTCTTCAAGGTACAGTTTATCT

9721 AlaLysThrAlaArgAsnIleAsnCysThrGlnLysArgProGlnGluIleArgAspLeu  
 GCAAAAACCTGTAGAAACATCAACTGTACACAGAAAAGACCCCAAGAAATTAGAGATCTG  
 CGTTTTGACGATCTTGTAGTTGACATGTGTCTTTTCTGGGGTCTTTAATCTTAGAC

9774 bgl11,

9781 MetAlaGlnGlyArgMetLeuValValAlaLeuArgSerPheAsnProGluLeuSerGlu  
 ATGGCACAAAGGCGTATGCTAGTAGTGGCTTAAAGAAAGTTCAATCCTGAGTTGCTGAA  
 TACCGTGTCCCGCATACGATCATCAACCAAAATCTTCAAAGTTAGGACTCAACAGACT

9840 spa1,

9841 LeuValAspPheLysGlyThrPheLeuAspArgValAlaLeuGluAlaLeuSerLeuGly  
 CTAGTTGATTTCAAGGGGACTTTCTTGGATAGGGTTGCCCTTGAAGCCCTTAGCCCTGGG  
 GATCAACTAAAGTCCCTGAAAGAACCTATCCCAACGGAACCTTCCGGAAATCGGACCC

9900 bgl1,

9901 ProGlyArgProLysGlnValThrThrAlaThrValLysGluLeuLeuGluGlnGluGlu  
 CCGGGAAGGCCCAAGCAGGTAACCAAGCCACAGTTAAGGAGTTGCTAGAGCAAGAGAA  
 GGCCCTTCCGGGTTCTCATTGTTGTCTGGTGTCAATTCTCAACGATCTGTTCTCCT

9918 betE2,

9961 GlnValGluIleProAsnTrpPheGlyAlaAspAspProValPheLeuGluValAlaLeu  
 CAACTCGAGATCCCAACTGTTTGGTGGGATGACCCAGTCTTCTTGGAAAGTAGCTCTG  
 GTTCAGCTCTAGGGGTTGACCAAGCCACGCTACTGGGTGAGAAGAACCTTATCGAGAC

9994 tth111,

(Figur 2)

10021 LysGlyAspLysIyrHisLeuValGlyAspValAspLysValLysAspGlnAlaLysGly  
 AAGGGTGACAAATACCACTTAGTAGGTGATGTAGATAAGTAAAAGATCAAGCAAAGGGA  
 TTCCTACTGTTTATGGTGAATCATCCACTACATCTATTTTCATTTCTAGTTCGTTTCCCT

10051 LeuGlyAlaThrAspGlnThrArgIleValLysGluValGlyAlaArgThrTyrThrMet  
 CTAGGGGCCACGGACCAAAGTAAAGTAAAGAAAGTAGGTGCGAGAACCTACACAAATG  
 GATCCCCGGTGCTGGTTTGTATCTTATCATTTTCTTCATCCACGCTCTTGGATGTGTTAC

10141 LysLeuSerSerTrpPheLeuGlnAlaSerSerLysGlnMetSerLeuThrProLeuPhe  
 AAGCTGTCTAGTIGGTTTCTTCAAGCATCAAGTAAACAGATGAGCTTGACCCCTTTGTTC  
 TTCGACAGATCAACCAAAGAAGTTCGTAGTTCATTTGTCTACTCGAAGCTGGGAAACAAG

10201 GluGluLeuLeuLeuArgCysProProLysMetLysAsnAsnLysGlyHisIleGlySer  
 GAGGAAGCTGTTCCTTCCCTCCCAAGATGAAGAACAATAAAGGGCATATCGGATCA  
 CTCCTTGACAACGAAGCAACGGGAGGGTCTACTTCTTGTATTTCCTGATAGCCTAGT

10261 AlaTyrGlnLeuAlaGlnGlyAsnTrpGluProLeuAspCysGlyValHisLeuGlyThr  
 GCCTACCAACTAGCTCAGGCAACTGGGAACCCCTCGATTGTGGAGTACACCTGGGCAAC  
 CGGATGGTGTGATCGAGTCCCGTTGACCCCTTGGGGAGCTAACACCTCATGTGGACCCGTTGG

10321 IleProAlaArgArgValLysIleHisProTyrGluAlaTyrLeuLysLeuLysAspLeu  
 ATAGCTGCCAGGAGGGTAAAGATCCACCCATATGAAGGCTATCTGAAACTGAAGGATTA  
 TATGGACGGTCTCCCATTTCTAGGTGGGTATACTCCGGATAGACTTTGACTTCTTAAT

10349 nde1, 10355 stu1,

10381 LeuGluGluGluGluArgLysProGluGlyArgAspThrValIleArgGluHisAsnLys  
 TTAGAAGAAGAGAGAGGAAGCCAGAGGGTAGAGATACAGTGATAAGAGAACAATAACAAG  
 AATCTTCTTCTCTCTCTCTCCGCTCCCATCTCTATGTCACTATTCTCTTGTATTGTTT

10441 TrpIleLeuLysLysValArgProProArgLysProGlnTyrLysGluAsnProGlnPro  
 TGGATCCTCAAAAAGTGAGGCCACCAAGGAACCTCAATCAAAAAGAAAATCCTCAACCC  
 ACCTAGGAGTTTTTCACTCCGGTGGTCTTGGAGTTATGTTTCTTTTAGGAGTTGGG

10442 bamh1,

10501 TrpLysAlaIleArgAlaThrArgLeuGluLysGlyIleLysGluThrSerIleIleThr  
 TGGAAAGCTATCAGAGCAACTAGACTAGAGAAGGGCATAAAAGAAAATCTATAATAACC  
 ACCTTTCGATAGTCTCGTTGATCTGATCTCTCCCGTATTTTCTTTGTAGATATTATTGG

10561 LysLeuAlaSerIleLeuThrGlyAlaGlyIleArgLeuGluLysLeuProValValArg  
 AAATTTGGCCTCCATACTAACAGGTGCAGGAATAAGGCTGGAAAATTTGCCAGTCTTAGA  
 TTTAACCGGAGGTATGATTTCCAGCTCTTATTCCGACCTTTTAAACCTCAGCAATCT

10621 AlaGlnThrAspHisLysSerPheHisGluAlaIleArgAspLysIleAspLysAsnGlu  
 GCCCAAAGTACATAAAAGTTTCCATGAGGCAATCAGAGATAAGATAGACAAGAACGAA  
 CGGGTTTACTGATTTTCAAAGGTACTCCGTTAGTCTCTATTCTATCTGTTCTTGCTT

10681 AsnGlnGlnSerProGlyLeuHisAspLysLeuLeuGluIlePheHisThrIleAlaGln  
 AATCAGCAGAGCCAGGATTACATGATAAATTTGTTAGAGATCTTTCACACAATAGCCCAA  
 TTAGTCGTCTCGGGTCTAATGTACTATTTAAACAATCTAGAAAAGTGTGTTATCGGGTT

10718 bgl11,

10741 ProSerLeuLysHisThrTyrGlyGluValThrTrpGluGlnLeuGluAlaGlyIleAsn  
 CCCAGCCTAAAGCACACTTACGGTGAAGTGAAGTGGGAACAGCTTGAGGCAGGGAATCAAC  
 GGGTCCGATTTCTGTGAATGCCGCTTCACTGCACCCTTGTGCAACTCCGTCCTGATTTG

10801 ArgLysGlyAlaAlaGlyPheLeuGluLysLysAsnLeuGlyGluValLeuAspSerGlu  
 AGAAAAGGGGCTGCAGGCTTTCTAGAAAAGAAAGAAATCTTGGAGAAAGTACTGGACTCAGAG  
 TCTTTTCCCGACGTCGAAAAGATCTTTTCTTCTTAGAACCTCTTCATGACCTGAGTCTC

10811 pst1, 10821 xba1, 10845 sca1,

10861 LysHisLeuValAspGlnLeuIleArgAspLeuLysThrGlyArgLysIleArgTyrTyr  
 AAGCACCTGGTGGACCAACTAATCAGAGACTGAAAACAGGACGGAAGATAAGATATTAT  
 TTCGTGGACCACTGGTTGATTAGTCTCTGGACTTTTGTCTTCTCTATTCTATAATA

10921 GluThrAlaIleProLysAsnGluLysArgAspValSerAspAspTrpGlnAlaGlyAsp  
 GAGACAGCAATACCTAAGAACGAGAGAGGGATGTCAGTGACGATTGGCAAAGCAGGGGAC  
 CTCTGTCTGTTATGGATTCTTGTCTTCTCCCTACAGTCACTGCTAACCGTTCTCCCTG

10981 IleValAspGluLysLysProArgValIleGlnTyrProGluAlaLysThrArgLeuAla  
 ATAGTTGATGAAAAGAAACCAAGAGTGTATCAATACCCTGAAGCTAAGCAAGACTGGCC  
 TATCAACTACTTTTCTTGGTTCTCACTAAGTTATGGGACTTCTGATTCTGTTCTGACCGG

11036 bal1,

11041 IleThrLysValMetTyrAsnTrpValLysGlnGlnProValValIleProGlyTyrGlu  
 ATCACTAAAGTTATGTACAACTGGGTGAAGCAGCAGCCTGTTGTGATCCCAAGGATGAA  
 TAGTGATTTCAATACATGTTGACCCACTTCGTCTGCGACAACTAGGGTCCCATACTT

Figur 2

11101 GlyLysThrProLeuPheLysIlePheAsnLysValArgLysGluTrpAspLeuPheAsn  
 GGAAGACCCCAATTATTCAAGATCTTTAAACAAGGTAAGAAAGGAATGGGACCTGTTCAAT  
 CCCTTCGGGGTAATAAGTICTAGAAATTGTTCCATTCTTCCCTTACCCTGGACAAGTTA  
 11120 bgl11,

11161 GluProValAlaValSerPheAspThrLysAlaTrpAspThrGlnValThrSerArgAsp  
 GAGCCAGTAGCTGTGAGTTTGGATACTAAGGCCYGGGACACCAAGTCACTAGTAGGGAT  
 CTCGGTCATCGACACTCAAACACTATGATTCCGGACCCCTGTGGGTTCACTGATCATCCCTA  
 11189 stu1, 11209 spa1,

11221 LeuArgLeuIleGlyGluIleGlnLysTyrTyrTyrArgLysGluTrpHisLysPheIle  
 CTACGGCTTATTGGTGAATTCAAAAATATTACTACAGGAAGGAGTGGCACAAATTCATC  
 GATGCCGAATAACCACTTAAGTTTTTATAATGATGTCTTCCCTCACCGTGTTTAAGTAG  
 11246 ssp1, 11278 cla1,

11281 AspThrIleThrAspHisMetValGluValProValIleThrAlaAspGlyGluValTyr  
 GATACCATCACCGACCACATGGTGGAGGTACCAGTCATAACAGCAGATGGTGAAGTATAC  
 CTATGGTAGTGGCTGGTGTACCACCTCCATGGTCAGTATTGTCTGTACCACTTCATATG  
 11307 kpn1,

11341 IleArgAsnGlyGlnArgGlySerGlyGlnProAspThrSerAlaGlyAsnSerMetLeu  
 ATAAGAAATGGACAAAGGGGTAGTGGCCAGCCAGACACAAGCGCAGGTAAACAGCATGCTA  
 TATTCTTTACCTGTTTCCCATCACCGGTGGTCTGTGTTCGCGTCCATTGTCTGACGAT  
 11364 ball, 11393 sph1,

11401 AsnValLeuThrMetMetTyrAlaPheCysGluSerThrGlyValProTyrLysSerPhe  
 AATGTGTTAAACAATGATGTATGCCCTTCTGTGAAAGTACGGGGTTCATATAAGAGTTTT  
 TTACACAATTGTTACTACATACGGAAAGACACTTTCATGCCCCCAAGGTATATTCTCAAAA  
 11406 hpa1,

11461 AsnArgValAlaAroIleHisValCysGlyAspAspGlyPheLeuIleThrGluArgGly  
 AATAGAGTGGCAAGGATCCATGTCTGTGGGGATGACGGCTTCCGATAACAGAGAGGGGG  
 TTATCTCAACGTTCTAGGTACAGACACCCCTACTGCCGAAGGACTATTGTCTCTCCCC  
 11474 bamh1, 11478 bstXI,

11521 LeuGlyThrLysIleCysGlnGlnArgAspAlaAsnPheCysMetArgArgAlaSerSer  
 CTGGGCACATAAATTTGCCAACAAAGGGATGCAAACTTCTGCATGAGGCGGGCAAGCTCA  
 GACCCGTGATTTTAAACGGTGTGTTCCCTACGTTTGAAGACGTACTCCGCCCTTCGAGT

11581 LysAsnAsnArgArgGlyLysAsnGluSerLeuProIleGlyLeuArgHisArgValLeu  
 AAAAATAACAGAAGGGGAAAGAATGAAAGCTTGCCATATAGGTTTGAAGCATAGAGTTTTG  
 TTTTATTGTCTTCCCTTTCTTACTTTCGAACGGATATCCAAACTCCGTATCTCAAAA  
 11607 hind111,

11641 LeuProHisThrSerProArgLysCysLeuIleIleProAlaAlaThrTrpProValGly  
 CTCCACACACCAGTCCCGTAAGTCTCTGATAATACCAGCAGCTACATGGCCGGTAGGC  
 GAGGCTGTGTGTGAGGGCATTACAGACTATTATGGTGGTGGTGTACCGGCCATCCG

11701 ThrAlaIleIleLeuSerLysMetAlaAsnLysIleGlyLeuSerGlyGluArgGlyThr  
 ACTGCCATTATATTCAAAGATGGCCAACAAGATTEGATTAAGTGGAGAGAGGCTACC  
 TGACGGTAATATAATAGTTTCTACCGGTTGTTTCTAACCTAATTCACCTCTCTCTCATGG
 11723 ball, 11755 kpn1,

11761 ThrAlaTyrGluLysAlaValAlaPheSerPheLeuLeuMetTyrSerTrpAsnProLeu  
 ACGGCATATGAAAAGGCAGTGGCTTTCAAGTTTCTGTTGATGTACTCCTGGAATCCACTT  
 TGCCGTACTTTTTCCGTACCCGAAAGTCAAAGAACAACACTACATGAGGACCTTAGGTGAA  
 11765 nde1,

11821 ValArgArgIleCysLeuLeuValLeuSerGlnHisProGluThrAlaProSerThrGln  
 GTAAGGAGGATTTGCTCCTGGTCTTTTACAGCATCCAGAAACAGCTCCATCAACCCAG  
 CATTCTCTAAACAAGGACCAAGAAAGTGTCTGAGGCTTTTGTGAGGTAAGTGGGTC

11881 ThrSerTyrTyrTyrLysGlyAspProIleGlyAlaTyrLysAspValIleGlyLysAsn  
 ACCTCTTACTATTATAAAGGAGACCCCAATAGGGCTATAAAGATGTTATAAGGAAAT  
 TGGAGAAATGATAATATTTCTCTGGTTATCCCGGATATTTCTACAATATCTTTTTTA

11941 LeuSerGluLeuLysArgThrGlyPheGluLysLeuAlaAsnLeuAsnLeuSerLeuSer  
 CTGAGTGAACATAAAGGACGGGTTTTGAAAAATTGGCTAATCTAAATCTAAGCCTGTCC  
 GACTCACTGATTTTTCTGCCAAAACTTTTAACCATAGATTAGATTTCGGACAGG

Figur 2

12001 ThrLeuGlyIleTrpSerLysHisThrSerLysArgIleIleGlnAspCysValThrIle  
 AACTAGGAATCTGGTCCAAACATACAAGTAAACGAATAATCCAGGACTGTGTAACCATC  
 TGTGATCCTTAGACCAGGTTTGTATGTTTCATTGCTTATTAGGTCCTGACACATTGGTAG

12061 GlyLysGluAspGlyAsnTrpLeuValAsnAlaAspArgLeuIleSerSerLysThrGly  
 GGGAAAGAGGACGGCAATTGECTGGTAAATGCCGACAGGCTGATATEAAGCAAAACTGGC  
 CCCTTTCTCCTGCCGTYAACCGACCATTTACGGCTGTCCGACTATAGTTCGTTTGGACC

12102 ecor5, 12117 ball,

12121 HisLeuTyrIleProAspLysGlyTyrThrLeuGlnGlyLysHisTyrGluGlnLeuGln  
 CATCTGTACATACCTGACAAAGGTTATACATTACAGGGAAAACACTATGAACAACCTTCAA  
 GTAGACATGTATGGACTGTTTCCAATATGTAATGTCCCTTTTGTGATACTTGTGAASTT

12169 xmn1,

12181 LeuGlnAlaArgThrSerProIleMetGlyValGlyThrGluArgTyrLysLeuGlyPro  
 TTGCAGGCAAGAAGTACGCCAATCATGGAGTATGGACAGAGAGATATAAACTAGGTCCT  
 AACGTCGGTTCCTGATCGGTTAGTACCCTCATCCCTGTCCTCTATATTGATCCAGGA

12241 IleValAsnLeuLeuLeuArgArgLeuLysValLeuLeuMetAlaAlaValGlyAlaSer  
 ATAGTAAACTTGTCTGCTGAGGAGGTTGAAAGTCTGCTTATGGCAGCTGTCCGTTCCAGC  
 TATCATTTGAACGACGACTCCTCCAACCTTTCAGGACGAATACCCTGACAGCCACGGTCC

12284 pvu11,

SerOP  
 12301 AGTTGAAATAAATGTATATATTGTACATAAATCTGTATTTGTATATATTATATAAACT  
 TCAACTTTATTTACATATATAACATGTATTTAGACATAAACATATATAATATATATTGA

12361 TAGTTGAGATTAGTATGTATATATAGTTATCTACCTCAAGTAAACACTACACTCAATGCA  
 ATCAACTCTAATCATCACTATATATCAATAGATGGAGTTTCATTTGTGATGTGAGTTACGT

12421 CACAGCACTTGTAGCTGTATGAGGAAACACCCGACGTCATGTTGGACTAGGSAAGACCC  
 GTGTGCTGAAATCGACATACTCCCTTGTGGGCTGCAGGTACCAACCTGATCCCTTCTGGC

12452 sat11, 12457 nco1,

12481 TTAACAGCCCCA  
 AATTGTCGGGGT