



(11) *Número de Publicação:* PT 981609 E

(51) *Classificação Internacional:* (Ed. 6)
C12N015/10 A C12Q001/68 B

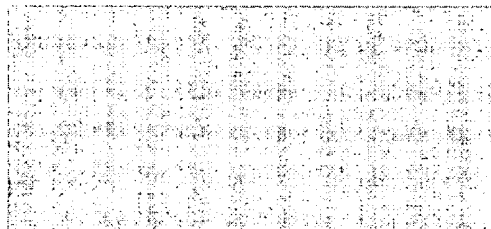
(12) *FASCÍCULO DE PATENTE DE INVENÇÃO*

(22) Data de depósito: 1998.05.13	(73) Titular(es): AZIGN BIOSCIENCE A/S VESTRE TEGLGADE 8 2450 COPENHAGEN SV DK
(30) Prioridade: 1997.05.13 DK 54797 1997.05.19 US 871030 1998.03.27 DK 43298	
(43) Data de publicação do pedido: 2000.03.01	(72) Inventor(es): WARTHOF, PETER, ROLF DK
(45) Data e BPI da concessão: 2001.09.19	(74) Mandatário(s): ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA RUA DAS FLORES 74 4/AND. 1294 LISBOA PT

(54) *Epígrafe:* MÉTODO PARA CLONAR ARNM E PARA EXIBIR TRANSCRITOS EXPRESSOS DIFERENCIALMENTE (DODET)

(57) *Resumo:*

MÉTODO PARA CLONAR ARNM E PARA EXIBIR TRANSCRITOS EXPRESSOS DIFERENCIALMENTE (DODET)





DESCRIÇÃO

"Método para clonar ARNm e para exhibir transcritos expressos diferencialmente (DODET)"

ANTECEDENTES DO INVENTO

O corpo humano é constituído principalmente por células especializadas que desempenham diferentes funções fisiológicas, organizadas em órgãos e tecidos. Todas as células humanas contêm ADN, disposto em séries de subunidades conhecidas como genes. Estima-se que existam aproximadamente 100 000 genes no genoma humano. Os genes são as impressões digitais das proteínas. As proteínas podem desempenhar uma grande variedade de funções biológicas, por exemplo, mensageiros, catalisadores e sensores. Estes compostos são responsáveis pela gestão da maioria das funções fisiológicas e bioquímicas, em humanos e em todos os outros organismos vivos. Nas últimas décadas, tem vindo a ser reconhecido que muitas doenças importantes têm uma base genética. Está agora bem estabelecido que os genes desempenham um papel importante no cancro, em doenças cardiovasculares, em desordens psiquiátricas, na obesidade e em doenças metabólicas. Recursos significativos estão a ser dirigidos à pesquisa genómica com base na noção de que as sequências de nucleótidos de um gene particular e o seu produto de proteína previsto conduzirão a uma compreensão da sua função em células ou tecidos saudáveis ou com funcionamento deficiente. Por sua vez, espera-se que esta compreensão conduza a abordagens terapêuticas e de diagnóstico, focalizadas em alvos moleculares associados ao gene e à proteína que este expressa. O primeiro passo do percurso para o desenvolvimento destas aplicações é identificar os genes especificamente envolvidos nas diferentes categorias de doenças. A aplicação deste conhecimento pode produzir novos e valiosos marcadores, identificando as regiões que produzem doenças importantes, para serem usados para benefícios de diagnóstico e terapêuticos.

Confrontados com a elevada complexidade do genoma humano, estão a ser usadas muitas abordagens para desenredar a relação entre a estrutura e a função do gene primário. Uma abordagem bem publicitada é concretizada no Projecto de Caracterização do Genoma Humano, onde a sequência de todos os genes individuais em todo o genoma humano está a ser persistentemente determinada. Presentemente, porém, pouca informação pode ser directamente recuperada sobre o funcionamento dos genes identificados e ainda menos acerca dos padrões de



expressão temporal e espacial do organismo em desenvolvimento ou maduro. Outras abordagens, tais como a sequenciação aleatória de ADNc, envolvem a determinação da sequência de todos os genes expressos num certo tecido, ou fase de desenvolvimento, de um organismo. Tal como várias outras estratégias, esta é morosa e propensa a numerosos problemas.

Ainda que a imensa quantidade de dados obtidos a partir de programas de sequenciação em grande escala seja um enorme benefício para a comunidade científica, um dos principais problemas encontrados com estas abordagens drásticas é a falta de informação específica que pode ser recuperada, sem significativamente mais trabalho, sobre a biologia de cada um dos genes individuais.

Várias outras abordagens têm sido empreendidas pelos biólogos moleculares para obter informação mais específica sobre o suporte genético de processos biológicos particulares. Estas abordagens assentam num conceito comum. Um gene, ou subconjunto de genes, é activado, iniciando o estado saudável, patológico ou de desenvolvimento, de um tipo de órgão ou célula.

Num grande número de sistemas experimentais, tem sido aplicado com sucesso o isolamento de genes, com base na sua expressão diferencial. A pesquisa diferencial e a hibridação subtractiva de bibliotecas de ADNc têm sido bem estabelecidas, cf. Zimmerman *et al.* (1980) e Davis *et al.* (1979). A pesquisa diferencial de bibliotecas funciona bem na prática para genes que são expressos em níveis elevados, mas os ARNm de pouca abundância são difíceis de isolar. A hibridação subtractiva proporciona uma pesquisa mais sensível, mas requer grandes quantidades de ARN. Mais recentemente, métodos de impressão digital de ARN (frequentemente designados por exibição diferencial ou DD/RT PCR) têm sido acrescentados a estas ferramentas, oferecendo novas e atractivas possibilidades para o isolamento de genes. Os métodos de impressão digital de ARN são baseados em PCR e, deste modo, não requerem grandes quantidades de ARN para as experiências. Para além disso, os métodos de impressão digital de ARN permitem a pesquisa de um grande número de amostras de ARN, para determinação de ARNm específicos, em simultâneo. Seria possível a investigação de uma grande gama de estados patogénicos em desenvolvimento e dos seus controlos. Até à data, dois métodos de impressão digital de ARN têm provado ser úteis para o isolamento de genes. Em 1992, Liang *et al.* publicaram um protocolo (Patente US 5 262 311), imediatamente após ter sido apresentado um protocolo de Welsh *et al.* (1992). Ambos os métodos começam com a síntese de ADNc a partir de ARN usando pelo menos um iniciador arbitrário para a iniciação da síntese das primeira e segunda cadeias.



Welsh *et al.* (1992) desenharam um protocolo no qual o mesmo oligómero 20-mero arbitrário é usado para a síntese da primeira e segunda cadeias. Usando iniciadores arbitrários, apenas um subconjunto dos ARNm será transcrito em ADNc. As amostras de ADNc são depois usadas para uma PCR padrão com os mesmos iniciadores. Um dos dNTP na mistura de PCR contém um marcador radioactivo (^{35}S ou ^{32}P) para visualização dos fragmentos de PCR por PAGE. Os métodos de Liang e Welsh baseiam-se em pelo menos um pequeno iniciador arbitrário para a selecção de ADNc específicos. Como consequência, as temperaturas de hibridação são baixas ($\sim 40^\circ\text{C}$), e todos os fragmentos de ADNc amplificados têm origem num certo grau de iniciação mal emparelhada. Mais tarde, vários outros grupos produziram refinamentos e optimizações conduzindo a uma abundância de artigos descrevendo a utilidade do método (Bauer e Warthoe *et al.* 1993; Warthoe *et al.* 1995; Liang e Warthoe *et al.* 1995; Rohde e Warthoe *et al.* 1996).

OBJECTIVO DO INVENTO

Um objecto do presente invento é proporcionar novos métodos e meios para investigação dos padrões de expressão em células, especialmente em células eucarióticas. Os resultados destas investigações podem ser usados no desenvolvimento de drogas, na identificação de genes, no diagnóstico de doenças etc., e deste modo estes métodos melhorados são altamente desejáveis.

SUMÁRIO DO INVENTO

No seu âmbito mais alargado, o invento refere-se a um método para a preparação de uma biblioteca subdividida de fragmentos de ADNc amplificados a partir de uma região de codificação de ARNm contido numa amostra, o método compreendendo os passos de

- a) submeter o ARNm derivado da amostra a transcrição inversa usando pelo menos um iniciador de ADNc, obtendo-se desse modo fragmentos de ADNc da primeira cadeia,
- b) sintetizar o ADNc da segunda cadeia, complementar com os fragmentos de ADNc da primeira cadeia, usando os fragmentos de ADN da primeira cadeia como molde, obtendo-se desse modo fragmentos de ADNc de cadeia dupla,
- c) digerir os fragmentos de ADNc de cadeia dupla com pelo menos uma endonuclease de restrição, obtendo-se assim fragmentos de ADNc clivados,

d) ligar pelo menos dois fragmentos de adaptador aos fragmentos de ADNc clivados obtidos no passo c), de modo a obter fragmentos de ADNc ligados, e

e) submeter os fragmentos de ADNc ligados obtidos no passo d) a um procedimento de amplificação molecular de modo a obter fragmentos de ADNc amplificados, no qual se usa, para um fragmento de adaptador usado no passo d), um conjunto de iniciadores de amplificação possuindo a fórmula geral



II

onde Com é uma sequência complementar com pelo menos a extremidade 5' de um fragmento de adaptador que está ligado à extremidade 3' de um fragmento de ADNc clivado, N é A, G, T ou C, e $n1$ é um inteiro ≥ 0 , e onde pelo menos um conjunto de iniciadores tem a fórmula geral II onde $n1 > 0$, sendo o referido pelo menos um conjunto capaz de iniciar a amplificação de qualquer sequência de nucleótidos ligada na sua extremidade 3' ao fragmento de adaptador complementar na sua extremidade 5' a Com.

Em comparação com a arte anterior, a vantagem global do invento é que a biblioteca resultante de fragmentos de ADNc contém sequências de ácido nucleico de todas as partes de ADNc que são produzidas no passo a). As técnicas da arte anterior que, *i.a.*, se baseiam numa iniciação de ADNc com poli-dT, têm tendência para apenas produzirem fragmentos derivados das longas regiões não traduzidas do ARNm. Adicionalmente, pela afinação das condições em cada passo, o método do presente invento resulta na reprodução altamente específica da informação das sequências que está presente no ARNm, mesmo em ARNm que apenas está presente em quantidades relativamente baixas. Adicionalmente, seleccionando a composição óptima da(s) endonuclease(s) é possível obter fragmentos de ADNc que são derivados, numa muito grande percentagem, do número total de genes transcritos em células relevantes.

O presente método permite a visualização pretendida de genes conhecidos usando combinações de iniciadores, correspondendo a sequências do gene de interesse. Isto tem a vantagem de todos os passos do procedimento e do sistema biológico poderem ser facilmente verificados. Também, análises de expressão muito específicas podem ser realizadas em genes relacionados com uma homologia muito elevada, o que não podia ser conseguido usando a tecnologia da hibridação.



Resumidamente, no método do invento, passos adicionais envolvem o isolamento de bandas de interesse a partir de um gel, a sua clonagem e sequenciação. A informação da sequência permite a reamplificação de bandas individuais, usando iniciadores com as extensões apropriadas de 3-4 nucleótidos. Quando realizadas sobre um gel, estas reacções mostrarão apenas uma ou apenas algumas bandas por pista, proporcionando uma determinação inequívoca da identidade da banda.

Uma vez que a presente tecnologia utiliza, para visualização, iniciadores marcados na sua extremidade, a tecnologia pode ser usada tanto com tecnologia padrão envolvendo radioactividade, como com iniciadores marcados por fluorescência, sem a necessidade de optimização adicional.

O invento refere-se também a métodos para detectar diferenças entre níveis de expressão em células que foram submetidas a condições diferentes e a métodos para o diagnóstico de doenças.

Apresenta-se a seguir uma breve discussão sobre os termos usados no presente pedido de patente:

"Uma biblioteca subdividida de fragmentos de ADNc amplificados" é, no presente contexto, uma biblioteca de fragmentos de ADNc amplificados que está dividida num certo número de amostras separadas, cada amostra sendo definida pelas sequências dos terminais dos fragmentos amplificados. Por exemplo, uma amostra pode conter fragmentos amplificados que são todos caracterizados por terem a sequência 5'-Com-AGC- numa das cadeias, enquanto que outra amostra contém fragmentos amplificados possuindo a sequência 5'-Com-AAT numa das cadeias. Para uma discussão do significado de "Com", cf. abaixo.

"Uma biblioteca normalizada" é uma biblioteca contendo uma representação substancialmente igual de cada ARNm, isto é, aproximadamente o mesmo número de cópias de cada ARNm.

"Transcrição Inversa" tem o seu significado usual na arte, isto é, síntese de ADN usando ARN como molde e efectuada por uma enzima possuindo actividade de transcriptase inversa.



"Fragmento de adaptador" pretende designar uma sequência de ácido nucleico contendo uma sequência conhecida que pode ser usada como molde para um iniciador num procedimento de amplificação molecular subsequente tal como a PCR. O fragmento de adaptador é adicionalmente caracterizado pela sua capacidade de se integrar na extremidade de um fragmento de ADNc que foi previamente clivado com uma endonuclease de restrição no passo c). Na maior parte dos casos, a endonuclease de restrição deixa fragmentos possuindo "extremidades coesivas", com as quais o fragmento de adaptador hibridará facilmente e depois o fragmento de adaptador é ligado ao ADNc pela acção de uma ADN-ligase.

DESCRIÇÃO DETALHADA DO INVENTO

Descreve-se a seguir em detalhe o impacto de cada um dos passos, ver Figuras 1 e 7.

Passo a)

O objectivo do passo a) é produzir uma mistura de fragmentos de ADNc da primeira cadeia que é optimizada na sua composição para realizar os passos subsequentes. Aplicam-se um certo número de considerações:

Em primeiro lugar, para reduzir o "ruído de fundo", prefere-se que a hibridação do iniciador de ADNc ao ARNm no passo a) seja realizada sob condições de elevado rigor, assegurando-se assim que é introduzido um mínimo de erros de emparelhamento no ADNc em relação ao ARNm, isto é, a uma temperatura acima de 50°C.

Em segundo lugar, é desejável obter cópias de sequências derivadas de todas as partes de ARNm de modo a obter informação relativa à parte traduzida do ARNm. Os métodos da arte anterior para a transcrição inversa de material eucariótico têm muitas vezes utilizado poli-dT como iniciadores de ADNc. Contudo, esta estratégia tem a desvantagem de o material transcrito inversamente de um modo mais eficiente se situar na zona não traduzida dos genes de interesse. Deste modo, as únicas partes do ARNm que se tornam "visíveis" após, e.g., um procedimento de PCR, serão muito frequentemente derivadas de regiões não traduzidas do ARN. A razão para isto reside em dois aspectos. Em primeiro lugar, a abordagem de poli-dT tem a consequência de o ponto de iniciação da transcrição inversa se situar muito longe e.g. do codão de início em relação ao operão em questão. Em segundo lugar, o ARNm pode incluir estruturas (e.g. estruturas em "gancho" devido ao emparelhamento de bases intra-cadeias) que bloqueiam a transcrição inversa

e, por iniciar sempre a transcrição inversa num terminal de um gene, estas estruturas bloquearão estatisticamente a transcrição inversa de um certo número de regiões traduzidas.

No presente invento, prefere-se assegurar são produzidos no passo a) ADNc que são representativos de todo o gene, incluindo as regiões traduzidas.

Numa concretização preferida, a mistura reaccional com a amostra compreende um iniciador de ADNc, sendo o referido iniciador de ADNc suficientemente complementar com o ARN alvo presente na amostra para com ele hibridar e iniciar a síntese de uma molécula de ADNc de cadeia simples complementar ao referido ARN alvo e a mistura reaccional compreende um tampão apropriado que compreende todos os quatro trifosfatos de desoxirribonucleósido e um catião bivalente seleccionado de entre o grupo de Mg^{2+} e Mn^{2+} numa concentração entre 0,1 e 5 mM.

O objectivo do método do invento é obter uma subdivisão do ADNc produzido. Quando o ARNm é derivado de um sistema eucariótico, o pelo menos um iniciador de ADNc pode incluir uma cauda oligo ou poli dT na extremidade 5', com a fórmula geral $5'-dT_{n2}-V_{n3}-N_{n4}-3'$, na qual dT é desoxitimidilo, V é A, G ou C, N é A, G, C ou T, $n2$ é um inteiro ≥ 1 , $n3$ é 0 ou 1, se $n3$ for 0 então $n4$ é 0 e se $n3$ for 1 então $n4$ é um inteiro ≥ 0 . Claro que, quando $n3$ e $n4$ são ambos zero, então o iniciador é um iniciador de ADNc poli- ou oligo-dT vulgar. Contudo, quando $n3$ é 1, então o iniciador é de facto uma composição de iniciadores que será capaz de iniciar a transcrição inversa de qualquer ARNm possuindo uma cauda poli-A. Se a amostra original de ARN estiver subdividida, e cada sub-amostra for submetida a transcrição inversa que utiliza um dos iniciadores possíveis possuindo a fórmula anterior onde $n3$ é 1, então o resultado é um número de amostras de ADNc de cadeia simples que são diferentes umas das outras na extremidade 5'.

Por exemplo, quando $n3$ é 1, usam-se 3×4^{n4} grupos de iniciadores de ADNc, cada um dos grupos sendo distinto de qualquer outro em relação à estrutura $-V_{n3}-N_{n4}-$. Nesta concretização, a amostra de ARNm está convenientemente subdividida em 3×4^{n4} alíquotas que são submetidas, cada uma em separado, ao passo a) utilizando um dos 3×4^{n4} grupos de iniciadores de ADNc, obtendo-se desse modo uma subdivisão do ADNc da primeira cadeia em 3×4^{n4} amostras separadas. Normalmente, $n4$ será 0 ou 1, resultando em 3 ou 12 amostras respectivamente.

Quando o material de partida não é eucariótico ou quando não é intenção começar necessariamente a partir da parte do gene transcrito que está mais remota em relação ao codão de

início da tradução, o pelo menos um iniciador de ADNc não inclui uma cauda poli ou oligo dT na extremidade 5', ou, alternativamente, usam-se pelo menos dois iniciadores de ADNc dos quais pelo menos um inclui uma cauda poli ou oligo dT na extremidade 5' e dos quais pelo menos um segundo não inclui uma cauda poli ou oligo dT na extremidade 5'. Preferivelmente, o iniciador de ADNc que não inclui uma cauda poli ou oligo dT na extremidade 5' tem a estrutura seguinte:



na qual N é A, G, T ou C, x é um inteiro $1 \leq x \leq 20$. Claro que isto corresponde à iniciação de ADNc começando a partir de qualquer codão de paragem da tradução. Em relação às concretizações anteriores utilizando um iniciador com cauda poli- ou oligo-dT, isto é, preparando iniciadores possuindo todas as permutações possíveis representadas no grupo N_x , é possível compor os iniciadores de modo a corresponderem a qualquer sequência possível precedendo um codão de paragem, assegurando-se desse modo a iniciação de todas as sequências possuindo um codão de paragem na sua sequência.

Passo b)

Este passo é realizado por métodos bem conhecidos na arte. Contudo, prefere-se que o passo b) seja realizado sob condições que minimizem a formação de erros de emparelhamento entre os nucleótidos na primeira e na segunda cadeias de ADN. O procedimento de ADNc de cadeia dupla pode ser realizado de acordo com métodos padrão, conforme descrito em Sambrook *et al.* (1989). Contudo, uma vez que as polimerases padrão podem ter dificuldades na síntese de regiões contendo estruturas secundárias ou com um elevado teor de GC, a RNase H termostável (RNase H Termostável Hibridase, US 5 268 289) e a ADN-polimerase termostável rBst do *Bacillus stearothermophilus* auxiliam a ultrapassar algumas das limitações de que sofrem as polimerases padrão (polimerases de baixa temperatura).

Passo c)

Após a preparação e subdivisão ótima do ARNm obtido no passo b), cada uma das diferentes amostras de ADNc é digerida com pelo menos uma enzima de restrição para produzir fragmentos com um tamanho que possam ser separados usando um método de fraccionamento por tamanhos apropriado.

A escolha da enzima de restrição é baseada grandemente na frequência dos locais de clivagem numa dada amostra de ADNc. Demasiados locais de clivagem em cada fragmento de

ADNc resultarão em fragmentos muito pequenos, e *vice versa*. Optimamente, a pelo menos uma enzima deverá clivar cada ADNc para dar fragmentos com o tamanho desejado. Estatisticamente, não é possível clivar cada ADNc mas, por outro lado, uma muito grande percentagem pode ser clivada seleccionando uma enzima ou combinação de enzimas adequada. Prefere-se que o método do invento utilize pelo menos uma enzima de restrição seleccionada de modo a assegurar que pelo menos 60% dos ADNc são clivados, mas percentagens mais elevadas tais como pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, ou mesmo pelo menos 85%, são mais preferidas.

O invento deverá usar enzimas de restrição que deixem extremidades protuberantes (extremidades coesivas) nos terminais do ADN após a digestão no passo c), uma vez que isto facilita grandemente a introdução dos fragmentos de adaptador no passo d).

Como será evidente do anteriormente exposto, é importante a frequência com a qual a endonuclease de restrição faz a clivagem. A pelo menos uma enzima de restrição é preferivelmente seleccionada de modo a clivar cada ADNc completo numa média de cerca de 3 fragmentos. Deverá ser entendido que alguns dos ADNc obtidos a partir dos passos anteriores não serão de todo cortados (ainda que isto tenha uma incidência rara quando a(s) enzima(s) de restrição é (são) cuidadosamente escolhida(s)) enquanto que outros serão cortados com uma frequência elevada. Tem sido salientado que a utilização de um cortante raro de 4 bases como a pelo menos uma endonuclease de restrição (tal com os cortantes de 4 bases *Acil*, *Alul*, *Bfal*, *BstUI*, *Csp6I*, *Dnpl*, *DnplI*, *HaeIII*, *HhaI*, *HinP1I*, *HpaII*, *Mbol*, *MnII*, *MseI*, *MspI*, *NlaIII*, *RsaI*, *Sau3AI*, *TaII*, *TaqI* e *Tsp509I*), assegura o desempenho óptimo do método do invento. Pela utilização de um tal cortante raro de 4 bases, a utilização de apenas 1 enzima de restrição no passo c) é suficiente e resulta numa produção superior.

Alternativamente, pode-se usar uma combinação de endonucleases de restrição, na qual um equilíbrio de e.g., cortantes de 6 bases e cortantes de 4 bases assegura uma distribuição razoável de dimensões de fragmento. Por exemplo, a utilização de uma primeira enzima de restrição (e.g., um cortante de 6 bases), que cliva estatisticamente pelo menos 20% do ADNc completo, derivado da amostra de ARNm, em dois subfragmentos, e de uma segunda enzima de restrição (e.g., um cortante de 4 bases), que cliva estatisticamente pelo menos 50% dos referidos subfragmentos em 3 subfragmentos adicionais, resultará também numa série de fragmentos adequados para posterior fraccionamento por tamanhos.



Passo d)

A(s) mistura(s) obtida(s) no passo c) é(são) depois submetida(s) a uma reacção na qual se adicionam fragmentos de adaptador a ambas as extremidades dos fragmentos de ADNc de cadeia dupla obtidos. Como acima mencionado, esta parte do procedimento é grandemente facilitada por os fragmentos de ADNc clivados possuírem extremidades protuberantes "coesivas", porque se podem preparar facilmente fragmentos de adaptador pré-desenhados que se ajustam a estas extremidades protuberantes.

Os fragmentos de adaptador (ou âncoras) são adicionados aos fragmentos clivados de modo a obter "ordem no caos" no passo subsequente. Adicionando sequências conhecidas aos terminais dos fragmentos clivados, são criados alvos para iniciadores de amplificação específicos, que podem ser desenhados especificamente com o objectivo de amplificar as sequências que estão de acordo com os fragmentos de adaptador. O material assim obtido (molde primário) pode ser pré-amplificado, usando iniciadores complementares das sequências de adaptador ligadas, dando origem a um molde secundário. A pré-amplificação do molde primário permite a produção de quantidades virtualmente ilimitadas de molde a partir de uma preparação de ARN, evitando a necessidade de isolamentos repetidos.

A sequência de adaptador é assim seleccionada de modo a servir como ponto de partida para a polimerização de ADN numa reacção, e.g., PCR. As sequências de adaptador são construídas de modo a que os locais específicos de endonucleases não sejam regenerados após a ligação do referido adaptador.

Pelo menos um fragmento de terminação é também ligado à extremidade 3' de cadeias simples de fragmentos de ADNc clivados, introduzindo o referido pelo menos um fragmento de terminação um bloqueio contra a polimerização de ADN no sentido 5' → 3', iniciando-se a partir do pelo menos um fragmento de terminação, e o referido pelo menos um fragmento de terminação sendo incapaz de hibridar com qualquer iniciador dos pelo menos dois conjuntos de iniciadores no passo f) durante o procedimento de amplificação molecular.

O aspecto anterior é um procedimento muito importante quando combinado com a utilização de detecção efectuada por iniciadores marcados no passo de amplificação, no qual apenas um membro do par de iniciadores está marcado enquanto que o outro é desenhado de modo a dividir os produtos amplificados de acordo com a sua composição de bases adjacente ao fragmento de adaptador. Um aspecto importante é que um fragmento de ADNc de cadeia simples

que foi provido de um fragmento de terminação não será amplificado, porque nenhum dos iniciadores será capaz de hibridar com os produtos de um primeiro ciclo de polimerização, no qual um tal fragmento era o molde, ver Figura 7. Em segundo lugar, a abordagem abre a possibilidade de remover o "ruído" de fundo numa fase de detecção subsequente.

Normalmente, o pelo menos um fragmento de terminação compreende ou é uma sequência de nucleótidos modificada quimicamente, tal como, por exemplo, uma sequência de nucleótidos que compreende um didesoxinucleótido na extremidade 3'; esta técnica de terminação é bem conhecida, e.g., da técnica de sequenciação de terminação de cadeias de acordo com Sanger. Sob circunstâncias normais, de acordo com o invento, o didesoxinucleótido deverá estar ligado de um modo covalente à cadeia de nucleótidos de modo a evitar a perda do didesoxinucleótido durante os ciclos de amplificação subsequentes. É conseguida uma estabilização superior se o didesoxinucleótido estiver fosforilado.

Conforme anteriormente mencionado, a ligação dos fragmentos de adaptador e/ou de terminação aos fragmentos de ADNc clivado no passo d) é convenientemente conseguida por hibridação dos fragmentos de adaptador a extremidades coesivas do ADNc resultante da clivagem no passo c) e submetendo o produto à acção de uma enzima possuindo actividade de ADN-ligase. Pode-se usar qualquer ADN-ligase adequada conhecida na arte.

Passos e) e f)

O passo e) do método do invento resulta no ordenamento final dos fragmentos de ADNc modificados do passo d).

O iniciadores possuindo a estrutura com a fórmula I (passo f) são desenhados de modo a amplificarem selectivamente fragmentos de ADNc de cadeia dupla sintetizados obtidos no passo b) ou subconjuntos predefinidos dos fragmentos adaptados obtidos no passo d). Podem ser vislumbradas várias formas de conseguir isto, mas a estratégia principal consiste em iniciar a amplificação numa série de reacções em separado onde a sequência de nucleótidos de um iniciador numa reacção assegura que os produtos amplificados dessa reacção são diferentes dos obtidos a partir de qualquer uma das outras reacções e que todas as reacções resultam na amplificação de todos os fragmentos obtidos no passo d), respectivamente.

Ainda que o pelo menos um conjunto de iniciadores de amplificação com a fórmula I tenha um n_1 que é ≥ 0 , prefere-se que $n_1=1$, $n_1=2$, $n_1=3$ ou $n_1=4$ num dos iniciadores, porque o número

de fragmentos de iniciador, a ser usado nas reacções de modo a cobrir todos os trechos de nucleótidos possíveis adjacentes ao fragmento Con ou de adaptador, é facilmente manejável. Por exemplo, se $n_1=5$, seria necessário usar $4^5=1024$ iniciadores diferentes de modo a obter a amplificação de todas as sequências de nucleótidos possíveis adjacentes ao fragmento de adaptador relevante, e uma vez que a concretização preferida do invento requer que cada um destes iniciadores seja usado numa reacção em separado, o trabalho envolvido seria problemático.

Prefere-se também que em um dos iniciadores $n=0$, e é especialmente preferido que este iniciador seja marcado, de modo a facilitar a determinação dos fragmentos amplificados. Assim, nas concretizações mais preferidas do invento, os fragmentos de ADNc adaptados são amplificados num certo número de reacções em separado nas quais se usa um iniciador marcado (que é normalmente idêntico em todas as reacções) e pelo menos um iniciador não marcado que é um membro do conjunto de iniciadores acima descrito onde $n > 1$. Prefere-se que este conjunto de iniciadores de amplificação com a fórmula I na qual $n_1 > 0$ compreenda todas as combinações e permutações possíveis de A, G, T e C no grupo N_{n_1} , uma vez que isto assegurará que todos os fragmentos de ADNc possíveis podem ser amplificados pelo conjunto.

Deste modo, os fragmentos de ADNc ligados são subdivididos num certo número de amostras antes da amplificação molecular no passo f), e cada amostra é submetida à amplificação usando um subconjunto do conjunto de iniciadores de amplificação e, na concretização mais preferida, os fragmentos de ADNc ligados do passo d) são subdivididos em 4^{n_1} amostras que são, cada uma, submetidas em separado ao passo f), onde se usa um iniciador de amplificação marcado conforme acima descrito ($n_1=0$) e um iniciador do conjunto de iniciadores de amplificação conforme acima definido ($n > 0$), sendo o referido um iniciador distinto de qualquer um dos iniciadores usado para a amplificação de quaisquer outras amostras. Utilizando esta abordagem, os fragmentos de ADNc originalmente transcritos de modo inverso e clivados são subdivididos em 4^{n_1} amostras que podem ser submetidas a passos adicionais.

Passos Adicionais e Aplicações

O material obtido a partir da série de reacções acima descrita pode agora ser utilizado de várias formas. Normalmente, é realizado um passo adicional de separação dos fragmentos amplificados obtidos a partir do procedimento de amplificação molecular. Este resulta numa mistura de fragmentos amplificados que são separados e.g., por separação de tamanhos, pela mobilidade



em electroforese em gel ou por qualquer método cromatográfico adequado. Além disso, é normalmente realizado um passo de identificação (e.g., por visualização destes fragmentos separados) para "efeitos de arquivo"; a mistura separada de fragmentos será normalmente comparada com algum tipo de referência que pode ser material obtido do mesmo ou de outro tipo de célula.

A visualização dos fragmentos separados pode ser conseguida, conforme acima mencionado, pelo facto de um dos iniciadores na reacção de amplificação estar marcado, mas estão obviamente disponíveis outros métodos. Por exemplo, uma sonda especificamente marcada que, e.g., se liga a uma das sequências de adaptador, permitirá visualizar os fragmentos, mas também nucleótidos marcados, que foram incorporados nos fragmentos durante o procedimento de amplificação (e.g., uma PCR), constituirá obviamente meios adequados para a detecção (e.g., por incorporação de alfa dNTP radioactivo ou fluorescente no fragmento de ADNc durante a PCR, onde N = A, C, T, U ou G).

Contudo, prefere-se que a visualização de Fragmentos Derivados de ARN (RDF, "RNA Derived Fragments") específicos seja conseguida usando iniciadores que estão marcados radioactivamente ou por fluorescência e que são homólogos dos adaptadores. As temperaturas de hibridação comparativamente elevadas (valor limite de 65°C a 56°C) que são preferivelmente usadas asseguram que os eventos de polimerização terão origem predominantemente a partir de uma iniciação perfeita de sequências de adaptador e de bases selectivas adjacentes. As intensidades de banda são grandemente uma função da concentração inicial do molde enquanto que as intensidades de banda dos métodos de Exibição Diferencial ("Differential Display") originais dependem da qualidade do emparelhamento entre o molde individual e o iniciador. A visualização de ARNm raros usando os métodos do presente invento será menos dificultada pela representação excessiva de sinal de ARNm altamente abundantes. Tal como no caso da iniciação arbitrária, a amplificação com erros de emparelhamento e os RDF abundantes ficam fora de competição na amplificação de fragmentos raros de pares de bases perfeitos. As nossas experiências sugerem que uma quantidade tão pequena quanto 100 moléculas pode ser detectada em ensaios de rotina num dado molde. Isto corresponde a menos do que 1 transcrito por célula no tecido original.

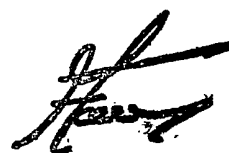
Uma outra parte do invento refere-se à utilização do método do invento para comparar os níveis de expressão em diferentes células. Uma forma de o conseguir consiste em determinar a alteração da expressão, por comparação com a expressão numa célula de referência ou num grupo de células de referência, de um produto de expressão numa célula ou grupo de células que



foi submetido a um primeiro conjunto de condições que influenciam o padrão de expressão da referida célula ou grupo de células, sendo a referida célula ou grupo de células de referência submetidos a um segundo conjunto de condições, compreendendo o método proporcionar uma amostra contendo ARN da célula ou grupo de células e submeter a amostra ao método do invento para subdivisão, obtendo-se desse modo dados descrevendo os fragmentos de ADNc amplificados derivados da amostra, proporcionar dados de referência descrevendo fragmentos de ADNc amplificados derivados de uma amostra de referência contendo ARN da célula de referência ou grupo de células de referência, sendo os dados de referência obtidos submetendo previamente a amostra de referência ao método do invento, subsequentemente efectuar uma comparação entre os dados e os dados de referência para identificar os fragmentos de ADNc que são expressos em diferentes níveis nos dois conjuntos de dados, e depois utilizar os fragmentos de ADNc expressos diferencialmente para determinar quais os produtos de expressão que são sujeitos a uma alteração no nível de expressão. Por outras palavras, o método do invento é realizado duas vezes com base em duas amostras de ARN diferentes derivadas de células submetidas a condições diferentes.

Normalmente, os dados e os dados de referência são seleccionados de entre o grupo consistindo em pesos moleculares aparentes dos fragmentos de ADN amplificados, M dos fragmentos de ADNc amplificados, quantidade absoluta dos fragmentos de ADN amplificados, e quantidades relativas dos fragmentos de ADN amplificados. Os dados de referência podem ainda ser extraídos a partir de uma base de dados contendo os dados de referência definidos anteriormente e opcionalmente informação adicional referente à origem genética de cada fragmento de ADNc amplificado a partir da referência.

Relacionado com o aspecto anterior, o invento permite também o diagnóstico de uma doença que é caracterizada por um desvio (aumento ou redução) no nível de expressão do pelo menos um produto de expressão em pelo menos um tipo de célula, compreendendo o método proporcionar uma amostra contendo ARN derivado de pelo menos um tipo de célula, submeter a amostra ao método do invento obtendo desse modo dados descrevendo os fragmentos de ADNc amplificados derivados da amostra, proporcionar dados de referência descrevendo fragmentos de ADNc amplificados derivados de uma amostra de referência contendo ARN derivada do mesmo tipo de células de um sujeito não sofrendo da doença, sendo os dados de referência obtidos submetendo previamente a amostra de referência ao método de acordo com o invento, e subsequentemente efectuar uma comparação entre os dados e os dados de referência em relação aos fragmentos de ADNc que se sabe estarem relacionados com a doença, e avaliar se existe ou



não uma diferença significativa entre os dados e os dados de referência de modo a estabelecer se o nível de expressão do produto de expressão apresenta ou não desvios.

Tal como para a concretização anterior, também aqui os dados e os dados de referência são seleccionados de entre o grupo consistindo em pesos moleculares aparentes dos fragmentos de ADN amplificados, M_r dos fragmentos de ADN amplificados, quantidade absoluta dos fragmentos de ADN amplificados e quantidades relativas dos fragmentos de ADN amplificados, e também aqui os dados de referência podem ser extraídos a partir de uma base de dados contendo os dados de referência definidos anteriormente e opcionalmente informação adicional referente à origem genética de cada fragmento de ADNc amplificado a partir da referência.

As misturas de fragmentos amplificados obtidos do passo f) do método do invento podem também ser usados para preparação de uma superfície (microplaca, "chip") revestida com fragmentos de ADNc. Isto pode ser realizado

- submetendo uma amostra contendo ARN ao método de subdivisão do invento incluindo os passos de separação, e
- transferindo os fragmentos de ADNc amplificados separados para uma superfície de microplaca adaptada para se ligar de modo estável aos fragmentos de ADNc amplificados separados mantendo o seu padrão de distribuição espacial relativa.

Alternativamente, uma tal microplaca pode ser preparada

- submetendo uma amostra contendo ARN ao método do invento sem realizar a separação, e depois
- separando, por electroforese, os fragmentos de ADNc amplificados assim obtidos sobre uma superfície particular adaptada para se ligar de modo estável aos fragmentos de ADNc amplificados separados mantendo o seu padrão de distribuição relativa após electroforese. Nesta concretização, a electroforese é preferivelmente na forma de micro-electroforese.

A transferência para uma superfície é realizada preferivelmente por uma técnica de transferência electroforética, e/ou por técnicas bem conhecidas de acoplamento foto-activado de química orgânica ou inorgânica.

O invento refere-se também a uma superfície que pode ser obtida pelo método acima mencionado para a sua preparação. Estas superfícies são consideradas novas e inventivas, uma vez que as "microplacas de ADN" conhecidas se baseiam na introdução específica de um arranjo de fragmentos de ácido nucleico de estrutura conhecida, enquanto que o presente método proporciona um "semi-arranjo" contendo fragmentos de ADNc caracterizando uma "situação" específica para um tipo de células específico, de acordo com a Figura 8.

Uma tal superfície pode, *i.a.*, ser usada para a pesquisa de genes dentro de uma família de genes. É proporcionada a "microplaca com o arranjo" e depois deixa-se hibridar uma sonda marcada (que é representativa de uma família de genes) com a microplaca sob condições pouco rigorosas, *i.e.*, nas condições descritas nas páginas 94-106 de "Nucleic acid hybridisation. A practical approach", editado por H.D. Hames & S.J. Higgins, IRL Press. Um número de fragmentos acoplados à microplaca hibridarão com a sonda e estes fragmentos podem ser subsequentemente identificados, isolados e sequenciados/caracterizados de modo a determinar se são ou não representativos da mesma família de genes.

Outra utilização destes "semi-arranjos" é na determinação da diferença no padrão de expressão entre uma primeira célula ou tipo de células e uma segunda célula ou tipo de células, compreendendo o método proporcionar amostras de ARN ou ADNc marcados da primeira e segunda células ou tipos de células, e subsequentemente pôr em contacto cada uma destas amostras com uma superfície de microplaca como anteriormente se descreveu, e subsequentemente detectar a quantidade e a distribuição de ARN ou de ADNc marcados, ligados, de cada amostra.

Em todas as circunstâncias, a superfície de microplaca com o ADNc a ela ligado poder ser produzida, *e.g.*, pelos métodos descritos na EP-0 654 061.

Ainda uma outra parte do invento refere-se a um método para pesquisar as interações entre uma proteína pré-seleccionada e um fragmento de polipéptido, o método compreendendo a preparação de uma biblioteca subdividida de fragmentos de ADNc amplificados resultantes do passo e), opcionalmente, a adaptação dos terminais dos membros da biblioteca de modo a facilitar a inserção num vector, a inserção dos fragmentos em vectores, a transformação de uma população de células hospedeiras adequado com os vectores, a cultura das células hospedeiras sob condições que permitam a expressão de fragmentos de ADNc correctamente inseridos pela célula

hospedeira e subsequentemente o ensaio dos fragmentos de polipéptido codificados pelos fragmentos de ADNc inseridos para determinar a interacção com a proteína pré-seleccionada.

Uma forma conveniente de conseguir isto é por via de uma técnica de dois híbridos, na qual as células hospedeiras são células eucarióticas (tais como células fúngicas, especialmente, células de levedura) que são acopladas ou transfectadas com material de ácido nucleico codificando a proteína pré-seleccionada, o acoplamento/transfecção bem sucedidos da(s) célula(s) resultando numa célula ou células nas quais a interacção entre a proteína pré-seleccionada e um fragmento de polipéptido dá origem a um sinal detectável.

Estes métodos têm atraído recentemente grande atenção, *i.a.*, como consequência da revelação em Fromont-Racine *et al.*, Nature Genetics 16, 277-282 (1997), que é aqui incorporada como referência.

Um sistema conveniente para proporcionar o sinal detectável é pela utilização da Proteína Verde Fluorescente ("Green Fluorescent Protein"), revelada na EP-A-0 569 170, no qual se usam como repórter as alterações no espectro fluorescente devidas a interacções.

DESCRIÇÃO DAS CONCRETIZAÇÕES PREFERIDAS

Em primeiro lugar, descrevem-se resumidamente os desenhos.

Fig. 1

A base da Exibição de Transcritos Expressos Diferencialmente ("Display Of Differentially Expressed Transcripts").

Fig. 2

Desenho da âncora e do iniciador de PCR

Fig. 3

Um autorradiograma de um gel de DODET usando a preparação celular descrita no Exemplo 1; células PC12 de feocromocitoma de rato foram estimuladas com Factor de Crescimento de Nervo (NGF, "Nerve Growth Factor") e com Factor de Crescimento de Epiderme (EGF, "Epidermal Growth Factor").

87 220

EP 0 981 609/PT

18

Pistas 1 - 24, transcrição inversa usando o iniciador poli T ancorado 5'-T₂₅AA-3'

Pistas 25 - 48, transcrição inversa usando o iniciador poli T ancorado 5'-T₂₅GC-3'

As pistas 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45 representam as células PC12 não tratadas.

As pistas 2, 6, 10, 14, 18, 22, 26, 30, 34, 38, 42 e 46 representam as células PC12 tratadas com o factor NGF durante 60 minutos.

As pistas 3, 7, 11, 15, 19, 23, 27, 31, 35, 39, 43 e 47 representam as células PC12 tratadas com o factor NGF durante 90 minutos.

As pistas 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44 e 48 representam as células PC12 tratadas com o factor EGF durante 90 minutos.

Pistas 1 - 48, usando os pares seguintes para as amplificações pré-PCR:

Iniciador de pré-amplificação *TaqI*: 5' -CAGCATGAGTCCTGACCGA

Iniciador de pré-amplificação *BclI*: 5' -CTCGTAGACTGCGTACCGATCA

Para a segunda amplificação por PCR, usaram-se os seguintes pares de iniciadores:

Pistas 1 - 4

5' -CATGAGTCCTGACCGAA

5' -GACTGCGTACCGATCAA (marcação na extremidade 5')

Pistas 5 - 8

5' -CATGAGTCCTGACCGAA

5' -GACTGCGTACCGATCAC (marcação na extremidade 5')

Pistas 9 - 12

5' -CATGAGTCCTGACCGAA

5' -GACTGCGTACCGATCAG (marcação na extremidade 5')

87 220

EP 0 981 609/PT

19

Pistas 13 - 16

5' -CATGAGTCCTGACCGAA

5' -GACTGCGTACCGATCAT

(marcação na extremidade 5')

Pistas 17 - 20

5' -CATGAGTCCTGACCGAC

5' -GACTGCGTACCGATCAA

(marcação na extremidade 5')

Pistas 21 - 24

5' -CATGAGTCCTGACCGAC

5' -GACTGCGTACCGATCAC

(marcação na extremidade 5')

Pistas 25 - 48

Combinação de iniciadores repetida das pistas 1 - 24

Fig. 4a

Transferência de Northern da sequência RDF01 a partir do ARN celular total.

a) células PC12 não tratadas, b) tratamento com NGF durante 60 minutos, c) tratamento com NGF durante 90 minutos, e d) tratamento com EGF durante 90 minutos.

Fig. 4b

Controlo da carga; os extractos de ARN foram submetidos a electroforese sobre um gel de agarose a 1,2% contendo brometo de etídio e usados como controlo para determinar a concentração relativa de ARN em cada pista; a), b), c), e d) os mesmos que na Figura 4a.

Fig. 4c

Transferência de Northern da sequência RDF02 a partir do ARN celular total.

a) células PC12 não tratadas, b) tratamento com NGF durante 60 minutos, c) tratamento com NGF durante 90 minutos, e d) tratamento com EGF durante 90 minutos.

Fig. 4d

Controlo da carga; os extractos de ARN foram submetidos a electroforese sobre um gel de agarose a 1,2% contendo brometo de etidio e usados como controlo para determinar a quantidade relativa de ARN em cada pista; a), b), c), e d) os mesmos que na Figura 4c.

Fig. 5

Pesquisa de genes modulados por um factor de crescimento.

Pista 1	Marcador de tamanhos em bp (150 bp, 200 bp, 250 bp)
Pistas 2-6	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAA
Pistas 7-11	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAC
Pistas 12-16	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAG
Pistas 17-21	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAT
Pistas 22-26	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GCA

Na pista 11 observa-se uma regulação negativa ("*downregulation*") após 6 dias de tratamento, enquanto que na pista 16 se observa uma regulação positiva ("*upregulation*") após 6 dias de tratamento. Ambas as modulações são devidas ao factor de crescimento, uma vez que a regulação apenas é observada quando o factor de crescimento activo está presente.

Fig. 6

Pesquisa de genes envolvidos na resistência bacteriana.

Pista 1	Marcador de tamanhos em bp (150 bp, 200 bp, 250 bp e 300 bp)
Pistas 2-5	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAA
Pistas 6-9	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAC
Pistas 10-13	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAG
Pistas 14-17	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GAT
Pistas 18-21	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GCA
Pistas 22-25	Iniciador de amplificação 5'—Com- N_{n1} —3' onde N_{n1} é GCC

Nas pistas 8-9 observa-se uma regulação negativa, nas pistas 20-21 observa-se uma regulação positiva. Ambas as modulação genéticas são de potenciais genes envolvidos na resistência à bacteriamicina, Inosina.

Fig. 7

Princípio da tecnologia usada nos Exemplos 4 e 5.

Após a síntese de ADNc-ds, o ADN é digerido com uma endonuclease de 4 pares de bases e ligam-se âncoras às extremidades do ADNc-ds. Usando iniciadores de desenho especial, obtêm-se os perfis de expressão por amplificação dos ARNm em diferentes janelas (sub-frações) de expressão. O número de janelas de expressão depende da complexidade da amostra, *i.e.*, 64 janelas de expressão num eucariota.

Fig. 8

Princípios de uma superfície de ADN (microplaca de ADN) para a identificação de genes.

Após separação por tamanhos dos fragmentos de ADN, os fragmentos de ADN são transferidos para uma membrana de *nylon* usando um princípio electroforético. A membrana é hibridada com uma sonda de ADN complexa gerada usando o princípio do invento. Alternativamente, a membrana pode ser hibridada com um único gene para identificar novos membros de uma família de genes particular. Na coordenada x, a membrana está dividida em 64 janelas de expressão e na coordenada y está dividida em tamanho de pares de bases (de 50 pares de bases a 1200 pares de bases) de acordo com o princípio descrito na Figura 7.

Fig. 9

Princípio de geração de 64 amostras de ADNc da extremidade 3'

Passo 1

Produção de ADNc de cadeia simples usando o oligonucleótido $5'-con_1-T_nV$, onde con_1 é um oligonucleótido com entre 1-100 nucleótidos, n está entre 5-40 e V é uma mistura de A, C e G.

Passo 2

As sínteses de ADNc de cadeia dupla são produzidas usando $5'-con_2N_x$, onde con_2 é um oligonucleótido com entre 1-100 nucleótidos, x está entre 1-10 e N é uma mistura de A, C, T e G. Na síntese, o ADNc-ds é sintetizado pela enzima de Klenow com o oligonucleótido acima descrito.



Passo 3

Pré-amplificação de ADNc de cadeia dupla para amplificar o ADNc de cadeia dupla; o ADNc é amplificado por PCR usando uma combinação de iniciadores con₁ e con₂.

Passo 4

O ADNc pré-amplificado é adicionalmente amplificado e separado em 64 amostras usando uma combinação de um con₁ marcado e de 64 iniciadores con₂NNN num procedimento de amplificação por PCR, onde NNN são combinados de 64 maneiras diferentes usando os nucleótidos A, T, G e C.

Passo 5

Cada uma das 64 amostras é separada usando o princípio da electroforese Page.

EXEMPLOS

De modo a verificar a funcionalidade do invento, descrevem-se a seguir exemplos nos quais se utilizou um sistema celular eucariótico em desenvolvimento, o feocromocitoma PC12.

O Factor de Crescimento de Nervo (NGF) induz a paragem do crescimento e a excrescência de neurónios no sistema de células PC12 *in vitro*. Outros factores de crescimento, tais como o Factor de Crescimento de Epiderme (EGF), suportam a sobrevivência e estimulam o crescimento. Os genes induzidos pelo NGF incluem os genes precoces imediatos, que codificam os factores de transcrição, tais como o c-fos e o c-myc. Crê-se que os produtos dos genes precoces imediatos estão envolvidos na regulação da expressão dos genes associados com o fenótipo neuronal, por exemplo, neurofilamentos, periferina, GAP43 e transina.

De modo a identificar novos genes precoces envolvidos na diferenciação e proliferação neuronal, usou-se, para identificar estes genes, o método DODET seguinte.

Nos exemplos seguintes, demonstra-se como o método do invento pode ser aplicado de um modo eficiente a estes sistemas celulares.

EXEMPLO 1

Cultivaram-se (*in vitro*) células PC12 de feocromocitoma de rato na presença e na ausência de Factor de Crescimento de Nervo (NGF) e de Factor de Crescimento de Epiderme (EGF) sob condições de crescimento descritas algures (Saltiel *et al.* 1996).

O ARN total foi isolado usando o método padrão de passo único de Chomczynski e Sacchi de acordo com Sambrook *et al.* 1989.

A concentração de ARN total foi determinada espectrofotometricamente e depois ajustada a 0,2 µg/µl. Este ARN foi usado directamente na análise de Northern.

Para o DODET, transcreveram-se inversamente 4 × 0,5 µg de ARN total em amostras separadas usando o iniciador 5'–T₂₅AA–3'. O mesmo procedimento foi realizado usando os iniciadores poli-dT ancorados 5'–T₂₅GC–3', obtendo-se um total de 2 × 4 × 0,5 µg de ARN.

Síntese da primeira cadeia

20,0 µl	quantidade de ARN total entre 0,3 e 1,0 µg de ARN
3,0 µl	5'–T ₂₅ AA–3' Conc. 100 ng/µl ou 5'–T ₂₅ GC–3' Conc. 100 ng/µl
5,0 µl	tampão de ADNc 10 × (tampão B da Epicentre Technologies # R19250)
2,0 µl	DNTP (25 mM da Pharmacia Biotech)
1,0 µl	SuperScript II RT (200 U/µl) (Gibco BRL # 18064-014)
5,0 µl	Retrotherm RT (1 U/µl) (Epicentre Technologies # R19250)
14,0 µl	H ₂ O

Para obter uma especificidade elevada, a reacção de ADNc foi incubada a 50°C durante 30 minutos seguindo-se 1 hora de incubação a 70°C.

Síntese da segunda cadeia

À reacção da primeira cadeia, adicionar os componentes seguintes:

15,0 μ l	tampão de ADNc 10 $\times$
3,0 μ l	RNase Termostável Hibridase (1 U/ μ l) (Epicentre Technologies #R39050)
1,0 μ l	ADN-polimerase termostável rBst (1 U/ μ l) (Epicentre Technologies #BH1100)
81,0 μ l	H ₂ O

Incubar a 65°C durante 1 hora.

O ADNc de cadeia dupla resultante foi extraído com fenol e precipitado e de novo suspenso em 20 μ l de H₂O. Metade deste volume foi verificado sobre gel; se fosse observada uma mancha entre 100 bp e 3000 bp, o resto do ADNc era usado para a produção do molde de DODET. Os ADNc resultantes foram digeridos com 10 U de cada uma das enzimas de restrição termostáveis *TaqI* e *BclI* a 50°C durante 2 horas. A esta mistura, adicionaram-se adaptadores de DODET e ligaram-se às extremidades dos fragmentos de restrição com ADN-ligase de *T₄* (1 U), resultando o molde primário. Realizaram-se 8-15 ciclos de pré-amplificação não radioactiva, usando iniciadores complementares com os adaptadores de DODET, numa pequena alíquota (1/10 do volume) do molde primário (desnaturação a 94°C, 30 s; hibridação a 56°C, 30 s; polimerização a 72°C, 1 min.). Os produtos da amplificação (designados por molde secundário) foram também verificados sobre um gel de agarose a 1,5%. Como era de esperar, as dimensões dos fragmentos estavam predominantemente entre 100 bp e 1000 bp. Todas as reacções de amplificação foram realizadas num aparelho de ciclos térmicos "PE-9600 thermocycler" usando ADN-polimerase *Taq*, ambos da Perkin Elmer Corp. (Norwalk, CT, E.U.A.). O molde final foi depois diluído 10 vezes com H₂O.

Apresentam-se a seguir os adaptadores ligados aos fragmentos de restrição, a pré-amplificação e a PCR activa:

Adaptador *TaqI*

5' -CAGCATGAGTCCTGAC
TACTCAGGACTGGC-5'

Iniciador de pré-amplificação *TaqI*:

5' -CAGCATGAGTCCTGACCGA

Iniciador de amplificação *TaqI*:

5' -CATGAGTCCTGACCGAN
(N = A ou C ou G ou T)

Adaptador *Bcl*I: 5' -CTCGTAGACTGCGTACC
CTGACGCATGGCTAG-5'

Iniciador de pré-amplificação *Bcl*I: 5' -CTCGTAGACTGCGTACCGATCA

Iniciador de amplificação *Bcl*I: 5' -GACTGCGTACCGATCAN
(N = A ou C ou G ou T)

Para a PCR, estavam disponíveis todas as combinações diferentes de uma extensão (acima representada por N), resultando num total de 4² combinações de iniciador. Todos os oligonucleótidos foram obtidos da DNA Technology (Aarhus, Dinamarca).

A marcação radioactiva do iniciador *Bcl*I foi realizada usando 1 U de polinucleótido-quinase de T₄. O tratamento por ciclos térmicos foi realizado essencialmente como acima se descreveu mas com 35 ciclos e incluindo um limite de 11 ciclos (a temperatura de hibridação foi reduzida de 65°C até 56°C em passos de 0,7°C durante 11 ciclos e subsequentemente mantida a 56°C durante 23 ciclos). As amostras foram depois fervidas, após a adição de corante e de formamida a 50%, e separadas num gel do tipo de sequenciação de poliacrilamida a 5% (GIBCO BRL Life Technologies Inc., Gaithersburg, MD, E.U.A.). Todos os geles foram corridos em condições padrão, de modo que o marcador de 70 bp estivesse a 3 cm do fundo do gel, proporcionando uma boa resolução entre 70-800 bp. Os geles foram depois secos directamente sobre papel Whatman 3M num secador de placas de gel. Os fragmentos de ADN marcados foram visualizados por autorradiografia. Os geles e as películas foram marcados posicionalmente antes do desenvolvimento. As extensões selectivas de 1 base foram seleccionadas empiricamente para dar aproximadamente 50 fragmentos marcados radioactivamente por pista.

As bandas, identificadas no autorradiograma como interessantes, foram alinhadas com marcações sobre a película e o gel foi desidratado, e aquelas foram excisadas. Os fragmentos excisados foram monitorizados quanto à sua actividade. Os fragmentos de gel foram isolados usando GENE CLEAN (BIO101, Califórnia, E.U.A.). O ADN foi depois recuperado de acordo com as recomendações do fabricante. Os fragmentos de ADN puderam então ser novamente amplificados usando as mesmas condições de PCR e os mesmos iniciadores que foram usados na PCR inicial; contudo, 15 ciclos produziram geralmente produto suficiente para clonagem. A clonagem foi

conseguida usando produto de PCR não purificado e o vector display-p123T (Display Systems Biotech, E.U.A.). Usaram-se as condições recomendadas pelo fabricante.

EXEMPLO 2

A Figura 3 mostra um gel DODET típico produzido por amplificação de um molde derivado do tratamento de células PC12 com NGF ou com EGF.

Transcreveu-se inversamente o ARN total com os iniciadores ancorados poli-dT T₂₅GC-3' e 5'-T₂₅AA-3', e após a ligação da âncora realizou-se a pré-amplificação com pares de iniciadores de pré-amplificação *BclI* e *TaqI*.

Mostram-se 6 de 16 combinações possíveis de iniciadores, usando 1 base selectiva, em cada local de enzima de restrição (Figura 3). Os produtos visíveis maiores (Figura 3) têm aproximadamente 1000 bp de tamanho e a extremidade inferior do gel corresponde a aproximadamente 100 bp. Nesta janela de tamanhos podem ser contadas uma média de 50 bandas para cada combinação de iniciadores. Na Figura 3 podem ser detectados vários padrões de expressão.

Devido principalmente às condições rigorosas possíveis no DODET, a resolução do padrão de bandas é elevada, enquanto que o nível de fundo permanece em níveis aceitáveis (Figura 3). Adicionalmente, as modificações bastante radicais na intensidade de bandas individuais durante o período de tratamento não parecem afectar os padrões de outras bandas na mesma pista.

Deste modo, é possível concluir que a PCR permanece proporcionalmente independente da concentração dos substratos individuais na reacção.

A utilização de uma combinação optimizada de protocolos padrão acima descritos para isolamento, reamplificação e clonagem de RDF individuais, permitiu a identificação de um número de transcritos associados a eventos de diferenciação e proliferação.

Isolaram-se quatro RDF para análise posterior, Figura 3, bandas a, b, c e d.

A análise das sequências revelou que RDF a = RDF b e RDF c = RDF d, conforme ilustrado na Figura 3.



Em todos os casos, puderam ser recuperadas sequências de terminal apropriado com as extensões de 1 base selectiva correctas usadas na PCR, demonstrando o rigor e a fidelidade do sistema (não se mostram resultados).

EXEMPLO 3

Durante a pesquisa dos sistemas celulares PC12 tratados com NGF e EGF com combinações de iniciadores diferentes, isolaram-se dois RDF (designados por RDF01 e RDF02) exibindo uma expressão diferencial durante o tratamento com NGF (RDF01 = RDF a = RDF b e RDF02 = RDF c = RDF d, na Figura 3).

Após reamplificação, subclonagem e sequenciação de ADN, a análise de ADN posterior revelou, no sistema celular PC12, dois RDF desconhecidos regulados positivamente, após 60 minutos de tratamento com NGF ou após 90 minutos de tratamento com EGF. As sequência de nucleótidos de ambos os RDF01 e RDF02 mostram menos do que 10% de homologia com qualquer gene existente nas bases de dados GeneBank ou EMBL.

A expressão de RDF01 e de RDF02 foi adicionalmente analisada usando transferência de Northern (Figura 4a - 4d). Puderam aqui ser claramente detectados transcritos após tratamento de 60 minutos com NGF ou tratamento de 90 minutos com EGF das células PC12, confirmando os resultados obtidos usando o método DODET, conforme se ilustra na Figura 3.

Experiências para clonar todo o comprimento de RDF01 e de RDF02, e a caracterização biológica do seu envolvimento na diferenciação e proliferação do sistema celular PC12 estão actualmente sob investigação.

EXEMPLO 4

Pesquisa de genes modulados por um factor de crescimento

Tratou-se uma linha de células humana com um factor de crescimento e isolou-se o ARN em vários instantes de tempo como a seguir se indica:

- 1 Células sem qualquer tratamento (pistas 2, 7, 12, 17, 22)



- 2 Células tratadas com agente adjuvante, 1 dia (pistas 3, 8, 13, 18, 23)
- 3 Células tratadas com agente adjuvante e com factor de crescimento, 1 dia (pistas 4, 9, 14, 19, 24)
- 4 Células tratadas com agente adjuvante, 6 dias (pistas 5, 10, 15, 20, 25)
- 5 Células tratadas com agente adjuvante e com factor de crescimento, 6 dias (pistas 6, 11, 16, 21, 26)

O ARN humano foi isolado e a análise para identificação de genes foi realizada essencialmente como se descreve na legenda da Fig. 3. Na Fig. 5 mostram-se 5 de 64 iniciadores de amplificação, cada um cobrindo uma certa porção da amostra de ARNm na linha de células humana. A análise da expressão foi realizada num ALFexpress, um analisador de fragmentos automático da Pharmacia Biotech, usando um marcador Cy5.

Na pista 11, observa-se uma regulação negativa após 6 dias de tratamento. Na pista 16 observa-se uma regulação positiva após 6 dias de tratamento. Ambas as modulações são devidas ao factor de crescimento, uma vez que a regulação apenas é observada com o factor de crescimento activo presente.

EXEMPLO 5

Pesquisa de genes envolvidos na resistência bacteriana a antibióticos

Tratou-se uma estirpe de *Listeria monocytogenes* com a bacteriomicina, Inosina. Investigou-se adicionalmente o ARN de uma estirpe resistente à Inosina.

- 1 Clone bacteriano 1 sem qualquer tratamento (pistas 2, 6, 10, 14, 18)
- 2 Clone bacteriano 2 sem qualquer tratamento (pistas 3, 7, 11, 15, 19)
- 3 Clone bacteriano 3 resistente à Inosina (pistas 4, 8, 12, 16, 20)
- 4 Clone bacteriano 4 resistente à Inosina (pistas 5, 9, 13, 17, 21)

O ARN bacteriano foi isolado por técnicas padrão e a análise para a identificação de genes foi realizada de acordo com o Exemplo 4 e Fig. 5, com a excepção de se ter usado um iniciador 5'-NNNNNNYYA para a síntese da primeira cadeia.

87 220

EP 0 981 609/PT

29

Na Figura 6 mostram-se 6 de 64 iniciadores de amplificação, cada um cobrindo uma certa porção da amostra de ARNm no sistema de células procarióticas. A análise da expressão foi realizada num ALFexpress, um analisador de fragmentos automático da Pharmacia Biotech, usando um marcador Cy5.

Nas pistas 8 e 9 observa-se uma regulação negativa e nas pistas 20 e 21 observa-se uma regulação positiva. Ambas as modulações de genes são potenciais genes envolvidos na resistência a Inosina.

Lisboa, 14. DEZ. 2001

Por AZIGN BIOSCIENCE A/S

- O AGENTE OFICIAL -

○ ADJUNTO



Eng.º ANTÓNIO JOÃO
DA CUNHA FERREIRA
Ag. Of. Pr. Ind.
Rua das Flores, 74 - 4.º
1200-195 LISBOA

REIVINDICAÇÕES

1. Método para a preparação de uma biblioteca subdividida normalizada de fragmentos de ADNc amplificados da região de codificação de ARNm contido numa amostra, compreendendo o método os passos de

- a) submeter o ARNm derivado da amostra a transcrição inversa usando pelo menos um iniciador de ADNc, obtendo-se desse modo fragmentos de ADNc da primeira cadeia,
- b) sintetizar o ADNc da segunda cadeia complementar com os fragmentos de ADNc da primeira cadeia usando os fragmentos de ADN da primeira cadeia como molde, obtendo-se desse modo fragmentos de ADNc de cadeia dupla,
- c) digerir os fragmentos de ADNc de cadeia dupla com uma endonuclease de restrição, deixando a referida endonuclease de restrição extremidades protuberantes (extremidades coesivas) de dimensão similar nos terminais do ADN após digestão, obtendo-se desse modo fragmentos de ADNc clivados,
- d) adicionar pelo menos dois fragmentos de adaptador contendo sequências conhecidas aos fragmentos de ADNc clivados obtidos no passo c), sendo os referidos pelo menos dois fragmentos de adaptador capazes de se ligar especificamente às extremidades coesivas do ADNc de cadeia dupla produzido no passo c), sendo o primeiro fragmento de adaptador capaz de hibridar com o iniciador possuindo a fórmula I no passo e), sendo o segundo fragmento de adaptador um fragmento de terminação que introduz um bloqueio contra a polimerização de ADN no sentido 5'→3' iniciando-se a partir do referido fragmento de terminação e sendo o referido fragmento de terminação incapaz de hibridar com qualquer iniciador dos pelo menos dois conjuntos de iniciadores no passo e) durante o procedimento de amplificação molecular, sendo os pelo menos dois fragmentos de adaptador ligados aos fragmentos de ADNc clivados obtidos no passo c), de modo a obter fragmentos de ADNc ligados,
- e) subdividir os fragmentos de ADNc ligados obtidos no passo d) em 4^n amostras onde $1 \leq n \leq 4$, e



2/8

f) submeter cada amostra de fragmentos de ADNc ligados obtidos no passo e) a um procedimento de amplificação molecular de modo a obter fragmentos de ADNc amplificados, no qual se usa, para um fragmento de adaptador usado no passo d), um conjunto de iniciadores de amplificação possuindo a fórmula geral



I

onde Com é uma sequência complementar pelo menos à extremidade 5' de um fragmento de adaptador que está ligada à extremidade 3' de um fragmento de ADNc clivado, N é A, G, T ou C, possuindo o primeiro iniciador a fórmula geral I onde $n1 = 0$, e possuindo o segundo iniciador a fórmula geral I onde $1 \leq n1 \leq 4$, sendo o referido segundo iniciador capaz de iniciar a amplificação de qualquer sequência de nucleótidos ligada na sua extremidade 3' ao fragmento de adaptador complementar na sua extremidade 5' a Com.

2. Método de acordo com a reivindicação 1, no qual a endonuclease de restrição do passo c) é um cortante raro de 4 bases, sendo o referido cortante raro de 4 bases seleccionado de modo a clivar cada ADN completo numa média de cerca de 3 fragmentos.

3. Método de acordo com a reivindicação 1 ou 2, no qual o cortante raro de 4 bases é seleccionado de entre o grupo consistindo em *Acil*, *Alul*, *Bfal*, *BstUI*, *Csp6I*, *Dnpl*, *Dnpll*, *HaeIII*, *HhaI*, *HinP1I*, *HpaII*, *Mbol*, *MnlI*, *MseI*, *MspI*, *NlaIII*, *RsaI*, *Sau3AI*, *TaiI*, *TaqI* e *Tsp509I*.

4. Método de acordo com qualquer uma das reivindicações 1-3, no qual o fragmento de terminação no passo d) compreende ou é uma sequência de nucleótidos modificada quimicamente, sendo a sequência de nucleótidos modificada um didesoxinucleótido ligado de modo covalente à extremidade 3' da cadeia de nucleótidos.

5. Método de acordo com qualquer uma das reivindicações 1-4, no qual o conjunto de iniciadores de amplificação com a fórmula I na qual $1 \leq n1 \leq 4$ compreende todas as combinações e permutações possíveis de A, G, T e C no grupo N_{n1} .

6. Método de acordo com qualquer uma das reivindicações 1-5, no qual o conjunto de iniciadores de amplificação com $n1=0$ na fórmula I está marcado.



7. Método de acordo com qualquer uma das reivindicações 1-6, no qual o conjunto de iniciadores de amplificação com a fórmula I onde $1 \leq n \leq 4$ tem $n=3$.

8. Método de acordo com qualquer uma das reivindicações anteriores, que compreende o passo adicional de separação dos fragmentos amplificados obtidos a partir do procedimento de amplificação molecular do passo f)

9. Método de acordo com a reivindicação 8, no qual a separação é realizada por electroforese em gel ou por cromatografia.

10. Método de acordo com a reivindicação 8 ou 9, que compreende adicionalmente o passo de identificação de fragmentos amplificados separados.

11. Método de acordo com a reivindicação 10, no qual a identificação é realizada por visualização.

12. Método de acordo com a reivindicação 11, no qual se visualizam nucleótidos marcados, sendo os nucleótidos marcados parte de uma sonda ou dos fragmentos amplificados.

13. Método de acordo com a reivindicação 12, no qual os nucleótidos marcados são os nucleótidos marcados sendo parte dos iniciadores marcados conforme definidos na reivindicação 6.

14. Método de acordo com a reivindicação 13, no qual a visualização é realizada por incorporação de alfa dNTP radioactivo ou fluorescente no fragmento de ADNc durante a PCR, onde $N = A, C, T, U$ ou G .

15. Método de acordo com qualquer uma das reivindicações anteriores, no qual o passo b) é realizado sob condições que minimizam a formação de erros de emparelhamento entre nucleótidos na primeira e segunda cadeias de ADNc.

16. Método de acordo com qualquer uma das reivindicações anteriores no qual o ARNm é de origem eucariótica, *Archae* ou procariótica.



17. Método para determinação da presença de um produto de expressão numa célula ou grupo de células, compreendendo o método proporcionar uma amostra contendo ARN da célula ou grupo de células e submeter a amostra ao método de acordo com qualquer uma das reivindicações 1-16 e depois efectuar uma comparação dos fragmentos de ADNc amplificados assim identificados, com dados de saída de uma base de dados, compreendendo os referidos dados de saída de uma base de dados uma lista gerada por computador de pesos moleculares de fragmentos de ADN de restrição de sequências conhecidas, sendo a referida lista preparada por

- entrada e armazenagem de dados de sequências de ADN numa base de dados como sequências de ADN virtuais,
- subsequentemente, simulação da clivagem das sequências de ADN virtuais com pelo menos uma nuclease de restrição e armazenamento dos produtos de clivagem simulada resultantes como fragmentos de ADN clivados virtualmente,
- simulação da ligação aos fragmentos clivados virtualmente dos pelo menos dois fragmentos de adaptador e armazenamento dos resultados como fragmentos de ADN ligados virtualmente,
- para cada combinação individual de iniciadores usada no passo f), agrupamento dos fragmentos de ADN ligados virtualmente susceptíveis de amplificação pela referida combinação de iniciadores no mesmo grupo,
- determinação, em cada grupo, do peso molecular absoluto e/ou relativo de cada fragmento de ADN ligado virtualmente, e
- obtenção de dados de saída do conteúdo de cada grupo na forma de uma lista compreendendo os pesos moleculares absolutos e/ou relativos dos fragmentos ligados virtualmente no grupo.

18. Método de acordo com a reivindicação 17, no qual os dados de entrada de sequências de ADN estão ligados a dados referentes à origem genética dos dados de sequências de ADN e opcionalmente a dados referentes a características funcionais relacionadas com a origem genética.

19. Método de acordo com a reivindicação 17, no qual a indicação de dados de saída compreende adicionalmente informação acerca da origem genética do fragmento de ADN ligado



virtualmente e opcionalmente informação acerca de características funcionais associadas à origem genética.

20. Método de acordo com qualquer uma das reivindicações 17-19, no qual a comparação é efectuada por entrada de dados dos fragmentos de ADNc amplificados identificados num formato que permite a comparação automática com os dados de saída da base de dados, ou, alternativamente, por obtenção de dados de saída da base de dados num formato que permite a comparação directa entre os fragmentos de ADNc amplificados separados e os dados de saída da base de dados.

21. Método para determinar a alteração na expressão, em comparação com a expressão numa célula de referência ou num grupo de células de referência, de um produto de expressão numa célula ou grupo de células que foi submetido a um primeiro conjunto de condições que influenciam o padrão de expressão da referida célula ou grupo de células, sendo a referida célula ou grupo de células de referência submetidos a um segundo conjunto de condições, compreendendo o método proporcionar uma amostra contendo ARN da célula ou grupo de células e submeter a amostra ao método de acordo com qualquer uma das reivindicações 10-14, obtendo-se desse modo dados descrevendo os fragmentos de ADNc amplificados derivados da amostra, proporcionar dados de referência descrevendo fragmentos de ADNc amplificados derivados de uma amostra de referência contendo ARN da célula de referência ou grupo de células de referência, sendo os dados de referência obtidos submetendo previamente a amostra de referência ao método de acordo com qualquer uma das reivindicações 10-14,

subsequentemente efectuar uma comparação entre os dados e os dados de referência para identificar os fragmentos de ADNc que são expressos em diferentes níveis nos dois conjuntos de dados, e depois utilizar os fragmentos de ADNc expressos diferencialmente para determinar quais os produtos de expressão que são sujeitos a uma alteração no nível de expressão.

22. Método de acordo com a reivindicação 21, no qual os dados e os dados de referência são seleccionados de entre o grupo consistindo em pesos moleculares aparentes dos fragmentos de ADN amplificados, M_r dos fragmentos de ADN amplificados, quantidade absoluta dos fragmentos de ADN amplificados, e quantidades relativas dos fragmentos de ADN amplificados.

23. Método de acordo com a reivindicação 21, no qual os dados de referência são extraídos a partir de uma base de dados contendo os dados de referência definidos na



reivindicação 22 e opcionalmente informação adicional referente à origem genética de cada fragmento de ADNc amplificado a partir da referência.

24. Método para diagnóstico de uma doença num sujeito, sendo a referida doença caracterizada por um desvio no nível de expressão de pelo menos um produto de expressão em pelo menos um tipo de célula, compreendendo o método proporcionar uma amostra contendo ARN derivada do pelo menos um tipo de célula, submeter a amostra ao método de acordo com qualquer uma das reivindicações 10-14 obtendo-se desse modo dados descrevendo os fragmentos de ADNc amplificados derivados da amostra, proporcionar dados de referência descrevendo fragmentos de ADNc amplificados derivados de uma amostra de referência contendo ARN derivada do mesmo tipo de células de um sujeito não sofrendo da doença, sendo os dados de referência obtidos submetendo previamente a amostra de referência ao método de acordo com qualquer uma das reivindicações 10-14, e

subsequentemente, efectuar uma comparação entre os dados e os dados de referência em relação aos fragmentos de ADNc que se sabe estarem relacionados com a doença, e avaliar se existe ou não uma diferença significativa entre os dados e os dados de referência de modo a estabelecer se o nível de expressão do produto de expressão apresenta ou não desvios.

25. Método de acordo com a reivindicação 24, no qual os dados e os dados de referência são seleccionados de entre o grupo consistindo em pesos moleculares aparentes dos fragmentos de ADN amplificados, M_r dos fragmentos de ADN amplificados, quantidade absoluta dos fragmentos de ADN amplificados e quantidades relativas dos fragmentos de ADN amplificados.

26. Método de acordo com a reivindicação 24, no qual os dados de referência são extraídos de uma base de dados contendo os dados de referência definidos na reivindicação 25 e opcionalmente informação adicional referente à origem genética de cada fragmento de ADNc amplificado a partir da referência.

27. Método para preparação de uma superfície (microplaca) revestida com fragmentos de ADNc, compreendendo o método,

- submeter uma amostra contendo ARN ao método de acordo com qualquer uma das reivindicações 8-14, e



- transferir os fragmentos de ADNc amplificados separados para uma superfície de microplaca adaptada para se ligar de modo estável aos fragmentos de ADNc amplificados separados mantendo o seu padrão de distribuição relativa espacial.

28. Método para preparação de uma superfície (microplaca) revestida com fragmentos de ADNc, compreendendo o método

- submeter uma amostra contendo ARN ao método de acordo com qualquer uma das reivindicações 1-7,
- separar, por electroforese, os fragmentos de ADNc amplificados assim obtidos sobre uma superfície particular adaptada para se ligar de modo estável aos fragmentos de ADNc amplificados separados mantendo o seu padrão de distribuição relativa após electroforese.

29. Método de acordo com a reivindicação 28, no qual a electroforese é na forma de micro-electroforese.

30. Método de acordo com qualquer uma das reivindicações 27-29, no qual a transferência é realizada por uma técnica de transferência electroforética

31. Método de acordo com qualquer uma das reivindicações 27-29, no qual a transferência é realizada por técnicas de química orgânica ou inorgânica foto-activadas.

32. Superfície possuindo ADNc a ela ligado de um modo estável, podendo a referida superfície ser obtida pelo método de acordo com qualquer uma das reivindicações 27-31.

33. Método para a pesquisa de genes dentro de uma família de genes, compreendendo o método proporcionar uma superfície de acordo com a reivindicação 32, na qual o ADNc ligado de um modo estável à superfície é hibridado sob condições pouco rigorosas com um ácido nucleico marcado de um modo detectável que é representativo de uma família de genes e subsequentemente analisar os fragmentos da microplaca na qual ocorreu a hibridação de modo a determinar se estes fragmentos estão ou não relacionados com a mesma família de genes.

34. Método para determinação da diferença no padrão de expressão entre uma primeira célula ou tipo de células e uma segunda célula ou tipo de células, compreendendo o método proporcionar amostras de ARN ou ADNc marcados da primeira e segunda células ou tipos de

células, e subsequentemente pôr em contacto cada uma destas amostras com uma superfície de acordo com a reivindicação 32, e subsequentemente detectar a quantidade e a distribuição de ARN ou de ADNc marcados ligados de cada amostra.

35. Método para pesquisar as interações entre uma proteína pré-seleccionada e um fragmento de polipéptido, compreendendo o método a preparação de uma biblioteca subdividida de fragmentos de ADNc amplificados de acordo com o método de qualquer uma das reivindicações 1-14, opcionalmente a adaptação dos terminais dos membros da biblioteca de modo a facilitar a inserção num vector, a inserção dos fragmentos em vectores, a transformação de uma população de células hospedeiras adequadas com os vectores, a cultura das células hospedeiras sob condições que permitam a expressão de fragmentos de ADNc correctamente inseridos pela célula hospedeira, e subsequentemente o ensaio dos fragmentos polipeptídicos codificados pelos fragmentos de ADNc inseridos para determinar a interacção com a proteína pré-seleccionada.

36. Método de acordo com a reivindicação 35, no qual o ensaio dos fragmentos polipeptídicos é realizado por uma técnica de dois híbridos, na qual as células hospedeiras são células eucarióticas que são acopladas ou transfectadas com material de ácido nucleico codificando a proteína pré-seleccionada, o acoplamento/transfecção bem sucedidos da(s) célula(s) resultando numa célula ou células nas quais a interacção entre a proteína pré-seleccionada e um fragmento polipeptídico dá origem a um sinal detectável.

37. Método de acordo com a reivindicação 36, no qual se usa no acoplamento/transfecção, um fungo, tal como uma célula de levedura.

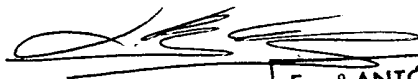
38. Método de acordo com a reivindicação 36 ou 37, no qual o sinal detectável é proporcionado pela Proteína Verde Fluorescente.

Lisboa, 14. DEZ. 2001

Por AZIGN BIOSCIENCE A/S

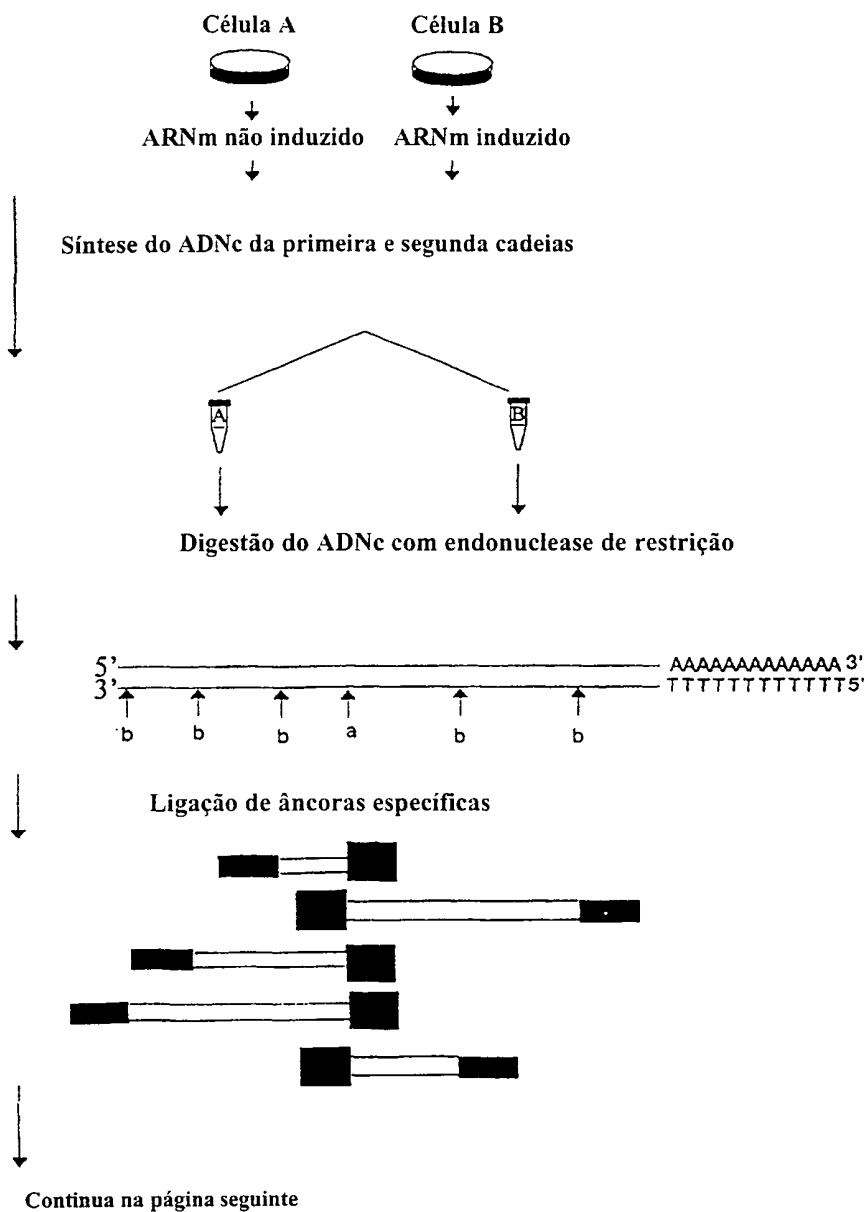
- O AGENTE OFICIAL -

ADJUNTO



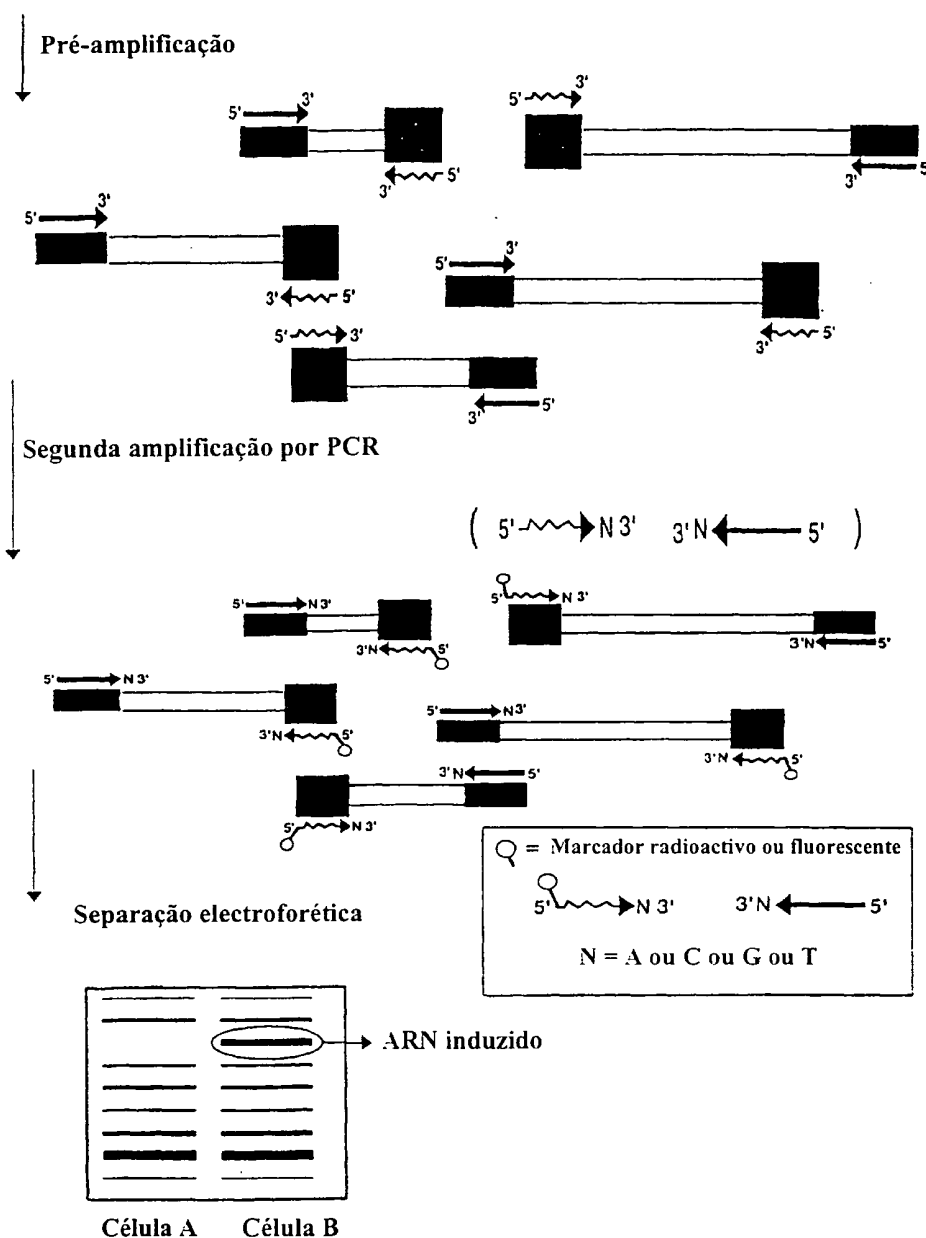
Eng.º ANTÓNIO JOÃO
DA CUNHA FERREIRA
Ag. Of. Pr. Ind.
Rua das Flores, 74 - 4.º
1200-195 LISBOA

Fig. 1 Base do DODET

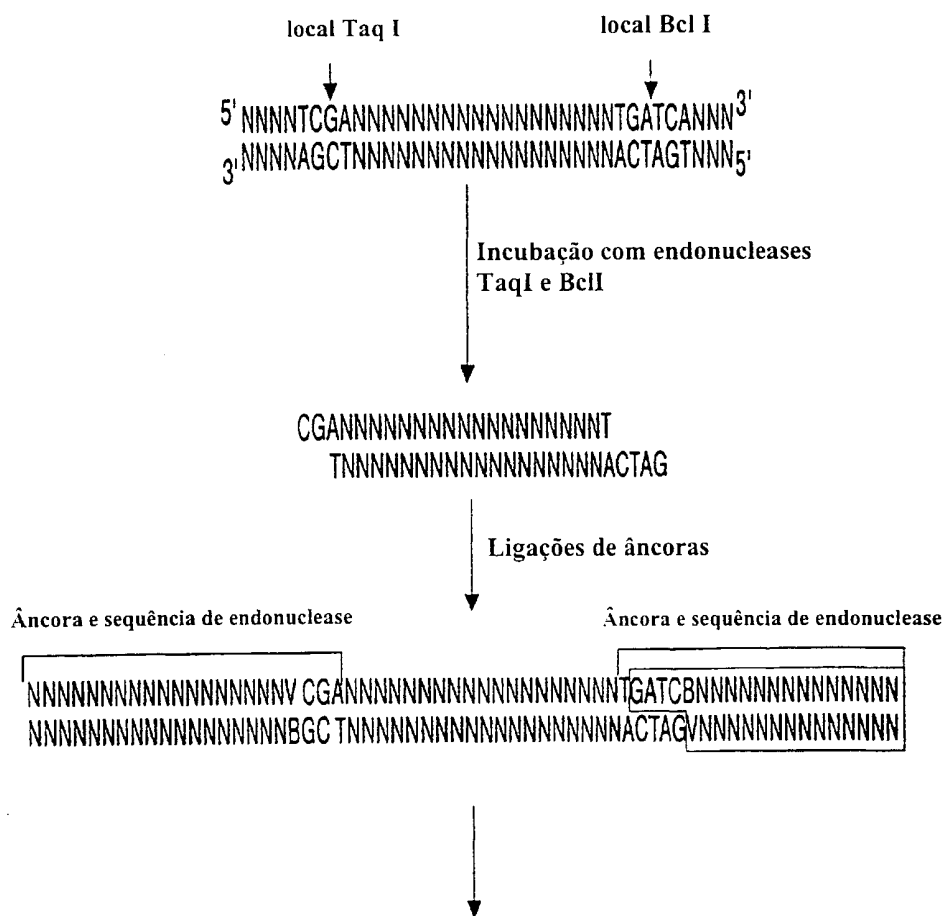


Handwritten signature

● **Fig. 1** Base do DODET (continuação)



● **Fig. 2** Desenho da âncora e do iniciador de PCR



Continua na página seguinte



5' NNNNNNNNNNNNNNNNNNNVCGANNNNNNNNNNNNNNNNNNNTGATCBNNNNNNNNNNNNNNNNNN 3'
3' NNNNNNNNNNNNNNNNNNNBGCNNNNNNNNNNNNNNNNNNNACTAGVNNNNNNNNNNNNNNNNNN 5'

Segunda amplificação por PCR

5'NNNNNNNNNNNNNNNNNNVCGAT 3'
5'NNNNNNNNNNNNNNNNNNVCGAG 3'
5'NNNNNNNNNNNNNNNNNNVCGAC 3'
5'NNNNNNNNNNNNNNNNNNVCGAA 3'

[illegible]

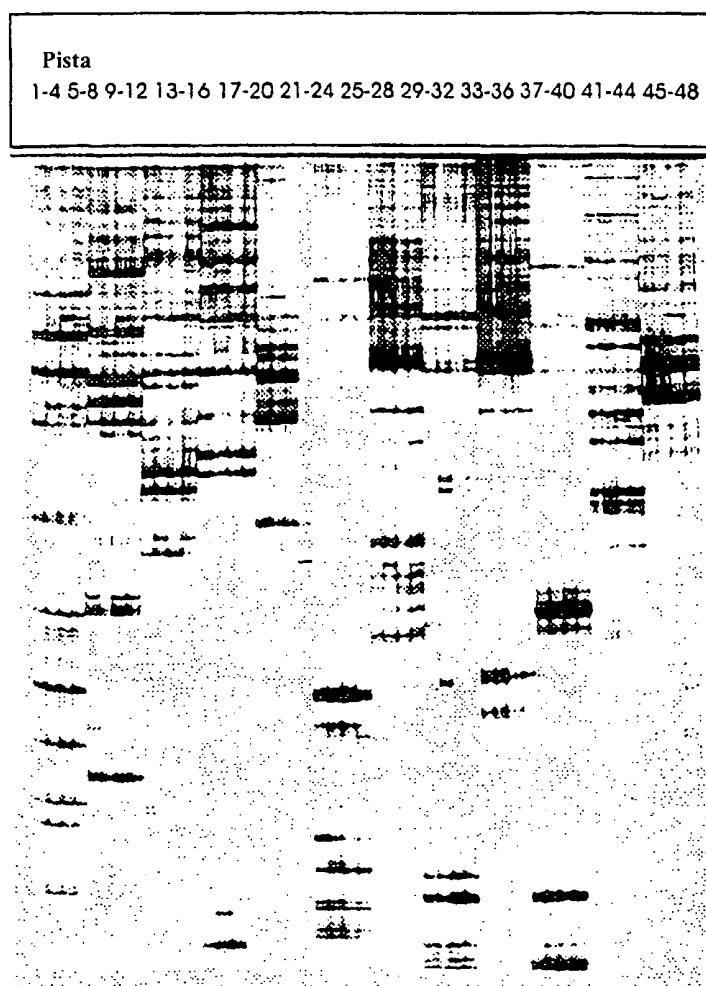
3' AACTAGVNNNNNNNNNNNNNNNNNN—5'
3' CACTAGVNNNNNNNNNNNNNNNNNN—5'
3' GACTAGVNNNNNNNNNNNNNNNNNN—5'
3' TACTAGVNNNNNNNNNNNNNNNNNN—5'

-O = Marcador
radioactivo ou
fluorescente

Separação electroforética

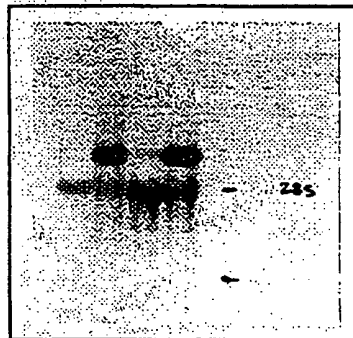


Fig. 3



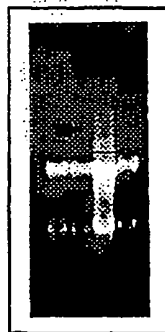
[Handwritten signature]

Fig. 4a



a b c d

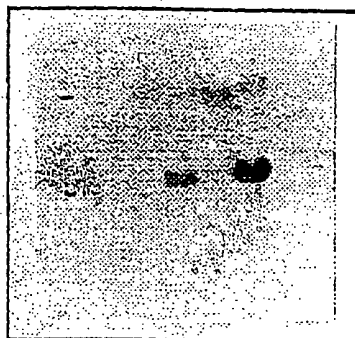
Fig. 4b



a b c d

[Handwritten signature]

Fig. 4c



a b c d

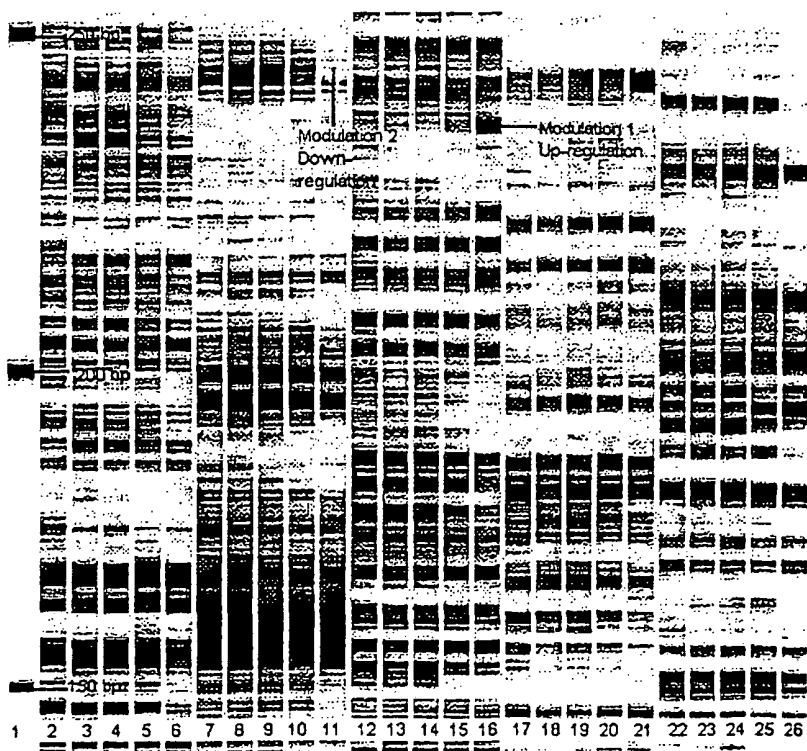
Fig. 4d



a b c d

[Handwritten signature]

Fig. 5



[Handwritten signature]

Fig. 6

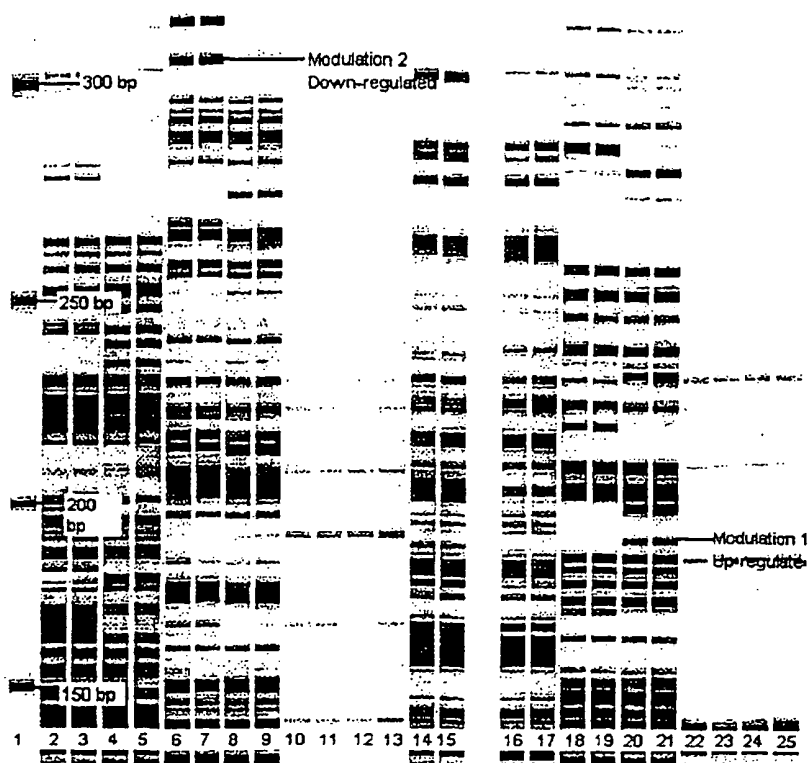


Fig. 7

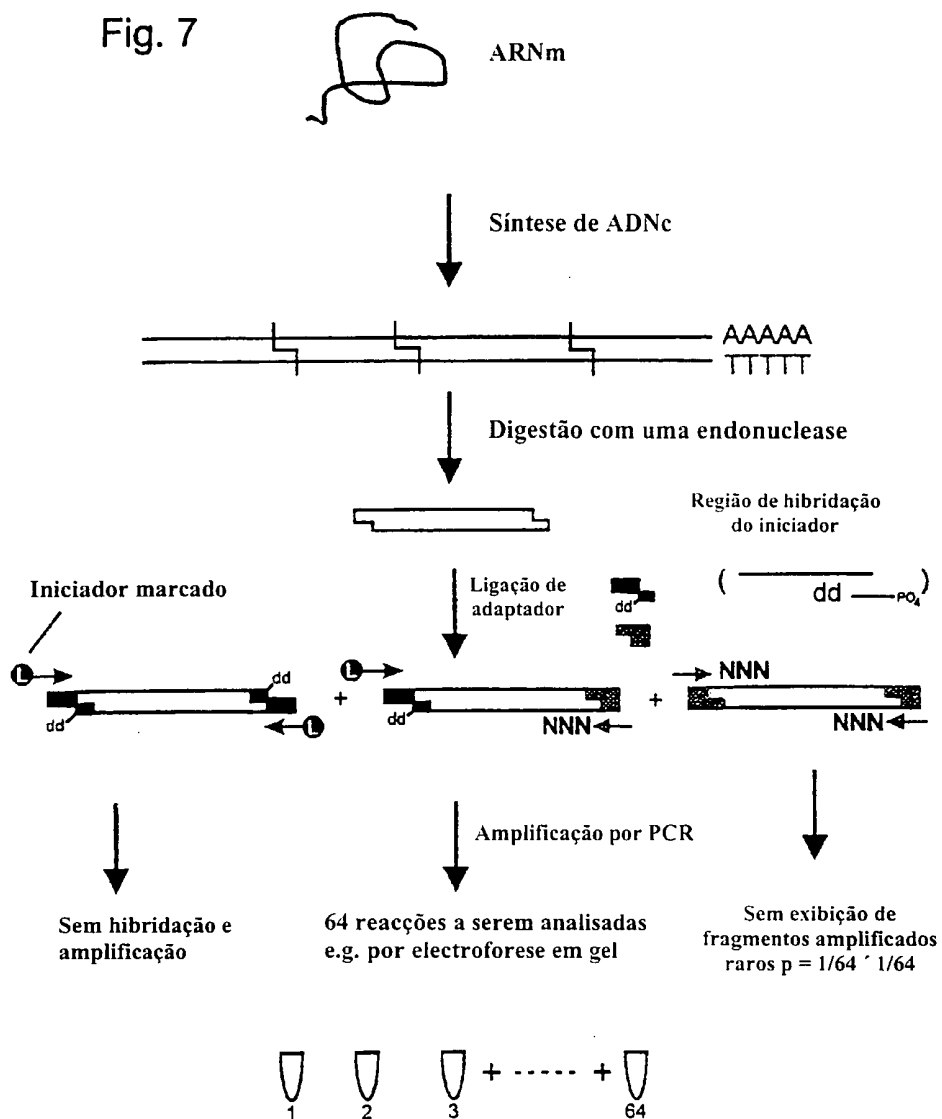


Fig. 8

Princípio do arranjo de exibição

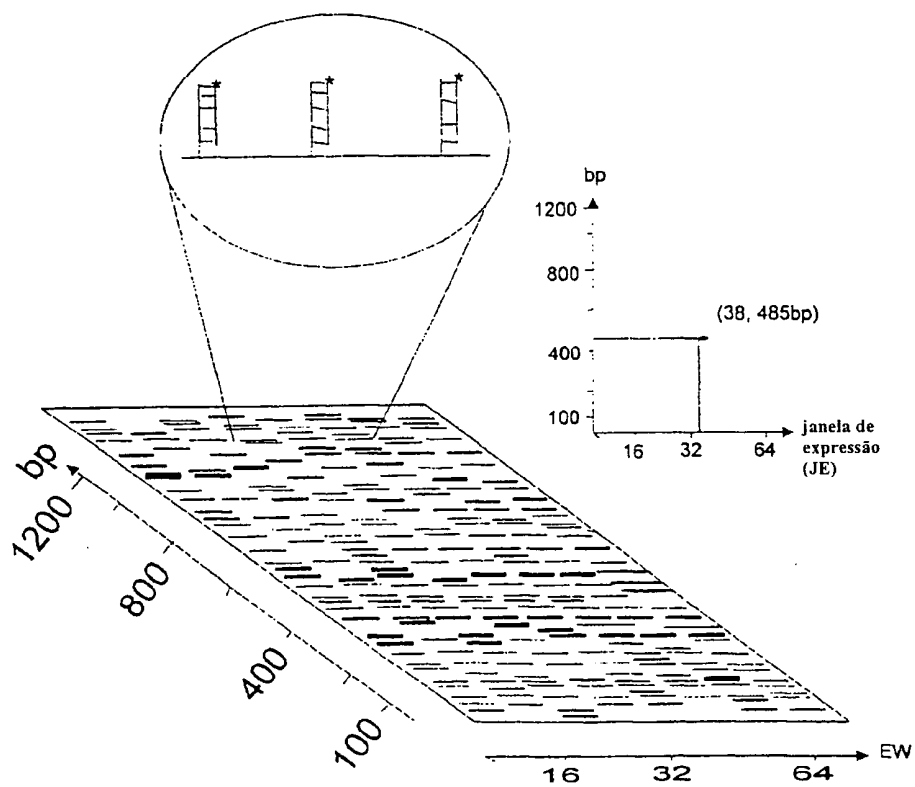


Fig. 9

