



**I Escola Brasileira de Inteligência
Artificial e Bioinformática InBio
São Carlos**

**Módulo: Biologia
Molecular**

Responsável: Dr. Flávio Henrique Silva

Uma atividade de difusão do:



- Dezembro de 2001 -

Índice

Introdução	2
O Ácido Desoxirribonucléico – DNA.....	2
O Dogma Central da Biologia Molecular	5
Perpetuação da informação genética (replicação)	7
Decodificando o DNA (Transcrição e Tradução).....	9
Manipulação do DNA	11
Enzimas de restrição e vetores.....	11
Amplificação do DNA “in vitro”: A reação em cadeia da polimerase – PCR.....	13
Clonagem	15
de ácidos nucléicos	17
Seqüenciamento de DNA.....	18
Utilização da tecnologia do DNA recombinante	19
Produção de proteínas	19
Tipagem -DNA "fingerprint"	20
Aplicações na medicina	22
Terapia Gênica	23
Leitura Recomendada	24

Nota: Esta apostila foi elaborada por alunos do curso de Introdução à engenharia genética da UFSCar (ano de 1999), com supervisão, adaptações e edições do Prof. Flávio Henrique Silva.

Introdução

O Ácido Desoxirribonucléico - DNA

Foi intuitivamente que o homem começou a fazer uso da genética a seu favor. Já em 9000 A.C., mesmo sem compreender alguns conceitos, que seriam descobertos mais tarde, o homem selecionou as melhores sementes para o plantio e escolheu os animais mais vigorosos para a reprodução. Os primeiros filósofos da humanidade já falavam de alguns fenômenos genéticos (sem saber a suas causas) e com o desenvolver da sociedade pessoas de todas as áreas como médicos, matemáticos, físicos, padres e filósofos, também contribuíram com idéias para o entendimento da hereditariedade.

No século XVII, um monge Augustiniano chamado Gregor Mendel, deu o primeiro grande passo para desvendar a hereditariedade. Através da análise dos cruzamentos entre ervilhas, Mendel deduziu a presença de fatores hereditários que eram propagados de forma estável de geração a geração, sendo responsáveis pela formação de características individuais.

Com o avanço dos estudos de biologia celular, pôde-se determinar quais os principais componentes moleculares da célula. Durante muito tempo as proteínas carregaram o papel de propagadoras da informação hereditária. Em 1928, Frederick Griffith, um médico londrino, em experimentos com *Pneumococcus*, células bacterianas causadoras de pneumonia, descobriu o fenômeno da transformação. Neste experimento foi mostrado que *Pneumococcus* patogênicos, portadores de cápsulas, quando mortos por calor e colocados em contato com células de uma linhagem não encapsulada, e não patogênica, era capaz de transformar as células não patogênicas em células patogênicas encapsuladas e letais (**figura 1**).

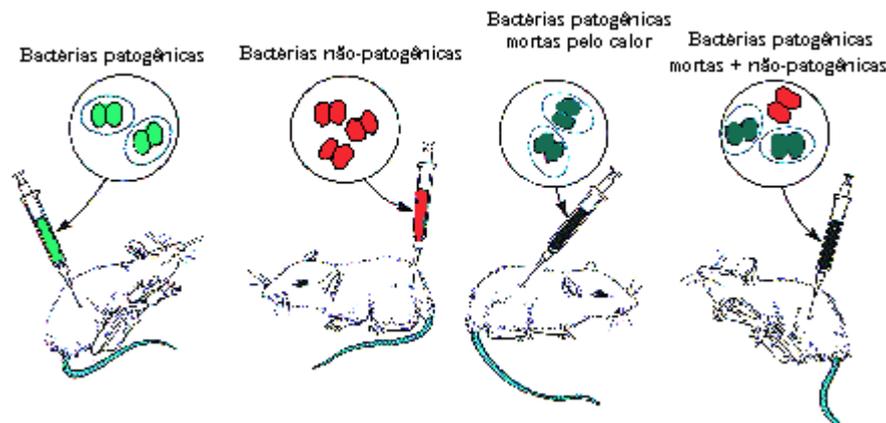


Figura 1- O experimento que levou Griffith a propor um "Princípio Transformante". Camundongos que recebiam a mistura de bactérias patogênicas mortas e não patogênicas vivas morriam por pneumonia e possuíam células encapsuladas (patogênicas em seu sangue).

Alguns anos mais tarde, em 1942, Avery e colaboradores, baseados no experimento de Griffith, separaram o extrato celular em várias frações (proteínas, DNA, RNA, lipídeos e carboidratos) e repetiram o experimento com os *Pneumococcus*. Apenas a fração contendo DNA recuperou a capacidade das bactérias formarem novamente a cápsula. A fração de DNA não perdia a sua capacidade transformante

quando tratada com proteases ou quando era aquecido, todavia, o mesmo não valia para o tratamento com DNase.

Outro experimento que ajudou a confirmar que o DNA era o material genético, foi o clássico experimento do liquidificador de Alfred Hershey e Martha Chase em 1952. Para esse experimento marcaram-se duas culturas de fagos T2, uma com fosfato radioativo (^{32}P) e outra com enxofre radioativo (^{35}S), sendo que o fosfato se incorporaria ao DNA e o enxofre às proteínas. De infecções paralelas de *E. coli* com fagos marcados com enxofre e fosfato radioativos foram tiradas amostras em diferentes tempos e estas, foram agitadas em um liquidificador de forma a separar as cápsulas dos fagos das bactérias. Após isso, as culturas foram centrifugadas, pois as cápsulas virais ficariam no sobrenadante e as células ficariam no precipitado. Foi constatado que grande parte do ^{32}P ficava na fração bacteriana, enquanto que a maior parte da fração protéica ficava no sobrenadante. Dessa forma, Hershey e Chase constataram que o que passava para dentro da bactéria, e que era responsável pela formação dos outros fagos, era o DNA e não as proteínas.

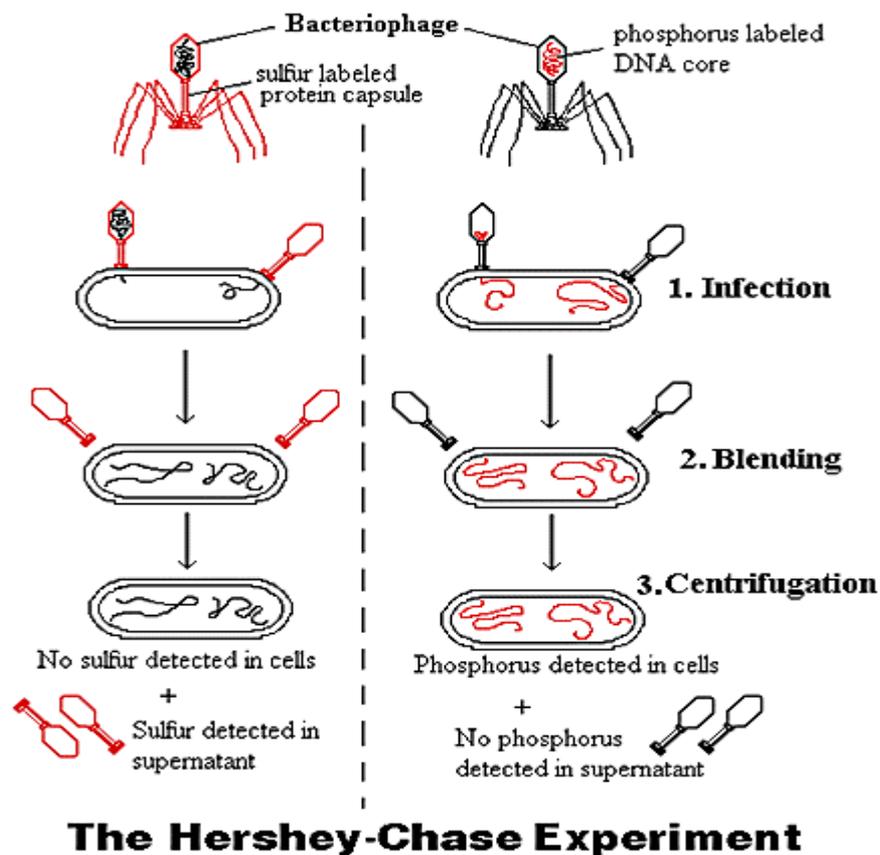


Figura 2- O experimento do liquidificador de Hershey e Chase para ajudar a confirmar que o DNA é o responsável pela hereditariedade.

Apesar da constatação de que o DNA desempenhava um papel imprescindível na hereditariedade, a comunidade científica ainda relutava um pouco para considerá-lo como o carregador da informação genética. Em 1953 James Watson e Francis Crick, baseados em vários trabalhos da época sobre o DNA (Como o trabalho de Chargaff, sobre a composição do DNA e as proporções das bases e os trabalhos de Wilkins com

difração de raios-X de moléculas de DNA), publicaram na revista científica *Nature* um trabalho denominado “Molecular Structure of Nucleic Acids- A Structure for Deoxyribose Nucleic Acid” (Estrutura molecular dos ácidos nucléicos- Uma Estrutura para o Ácido Desoxirribonucléico), que descrevia a estrutura do DNA.

Watson e Crick descreveram o DNA como uma dupla fita, enrolada em hélice ao redor de um eixo, sendo as fitas antiparalelas. O DNA possui uma estrutura periódica que se repete a cada 10 nucleotídeos. As bases nitrogenadas das duas fitas estão voltadas para o interior da hélice e pareiam de forma complementar entre si, na qual Adenina se liga a Timina e Guanina se liga a Citosina (**figura 3**).

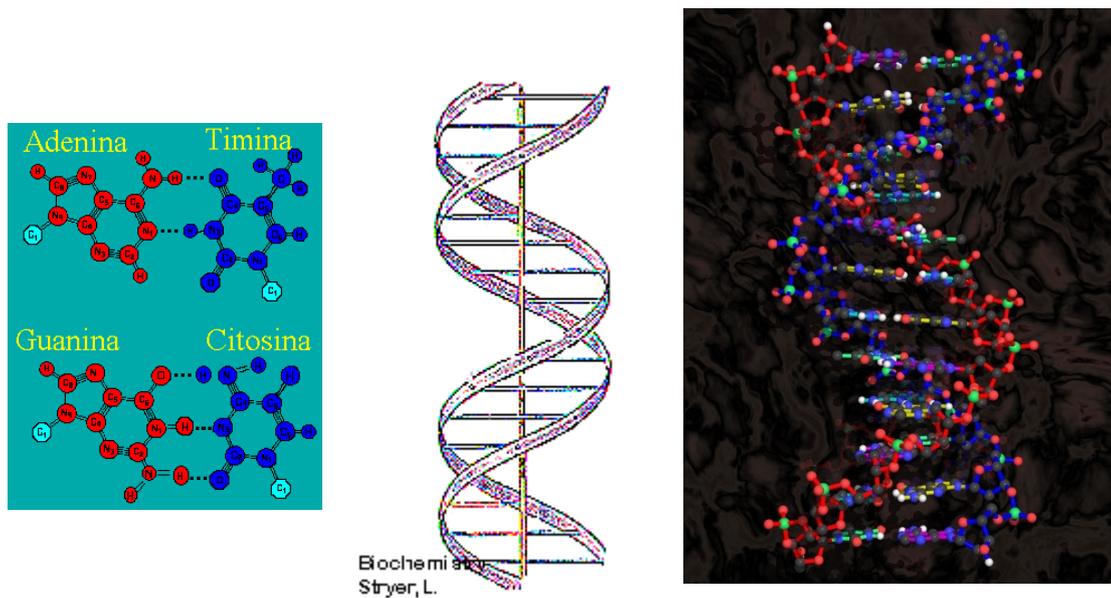


Figura 3 - A ligação entre os nucleotídeos, A com T e C com G, e duas representações da estrutura do DNA.

Somente depois do modelo de Watson e Crick o DNA foi considerado o material genético, pois sua própria estrutura já dava fortes indícios de como ocorreria a sua propagação (Replicação) e como era guardada a informação genética. No mesmo ano, Watson e Crick publicaram mais dois trabalhos falando sobre o assunto e em um deles, devido à obviedade do seu modelo, são discutidas as implicações moleculares do modelo na genética, comentando sobre as mutações e sobre a replicação. O reconhecimento final do trabalho de Watson e Crick veio em 1963 com o recebimento do Prêmio Nobel.



Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

Figura 4 – *Os descobridores da estrutura do DNA: Watson (à direita) e Crick (à esquerda)*

O Dogma Central da Biologia Molecular

O dogma central define o paradigma da biologia molecular, em que a informação é perpetuada através da replicação do DNA e é traduzida através de dois processos: A transcrição que converte a informação do DNA em uma forma mais acessível (uma fita de RNA complementar) e através da tradução que converte a informação contida no RNA em proteínas (**Figura 5**).

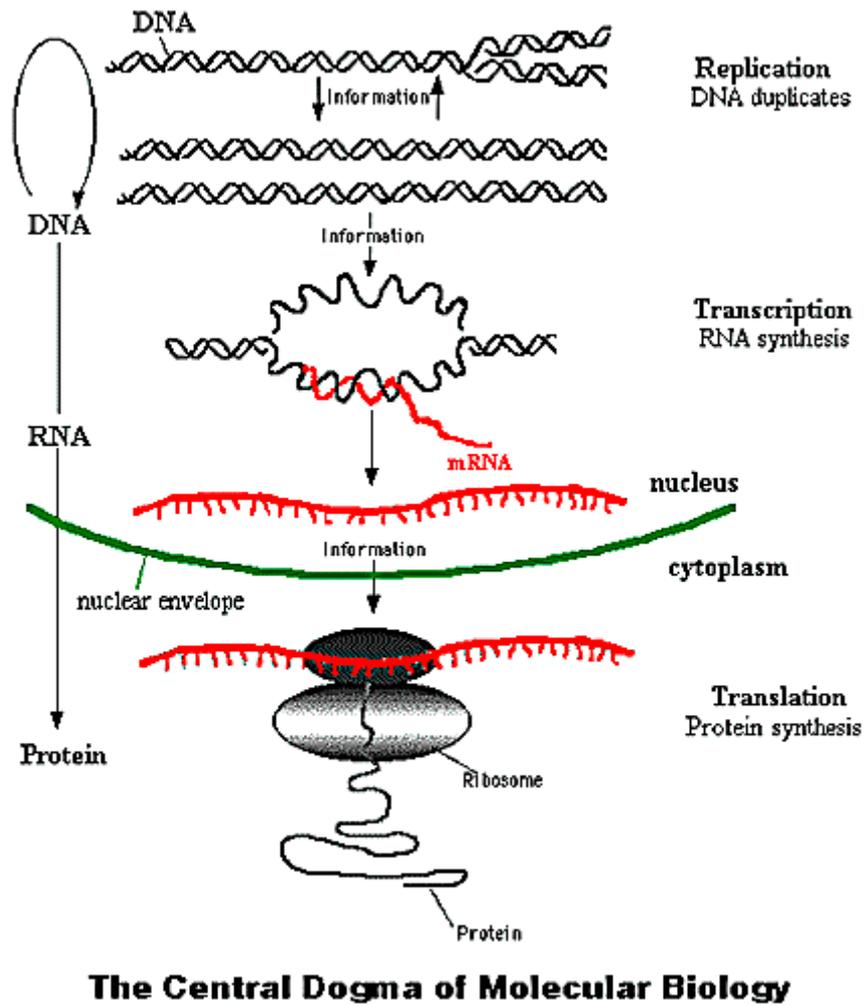


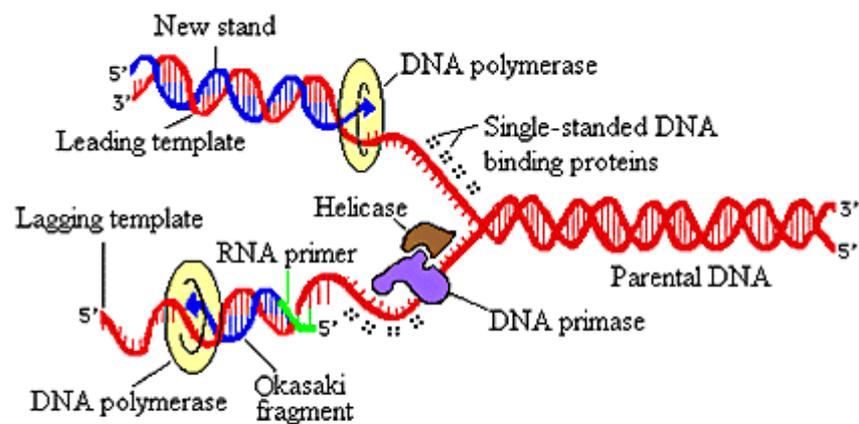
Figura 5. O Dogma Central da Biologia Molecular. A exceção é a replicação retroviral, na qual o RNA viral é molde para síntese do DNA do provírus.

Perpetuação da informação genética (replicação)

A replicação é o processo pelo qual uma molécula de DNA se duplica dando origem a duas moléculas idênticas a molécula inicial e envolve um conjunto de proteínas.

O primeiro passo para a replicação do DNA é a abertura das fitas, feita pela enzima **helicase**. Para manter as fitas desenroladas proteínas chamadas **SSBP** (*single strand binding protein*) se ligam nas fitas recém-desenroladas evitando que se associem de novo. Com o desenrolamento das fitas em um ponto, as regiões adjacentes sofrem um “super-enrolamento”, o que dificultaria a continuação do processo. As **topoisomerases** resolvem esse problema fazendo cortes em uma das fitas de DNA, que se desenrola, e religando-as, diminuindo a tensão provocada por esse “super-enrolamento”.

A síntese de novas fitas é feita pela enzima **DNA-polimerase**, sendo que esta enzima não pode sintetizar outra fita a partir de nucleotídeos livres. Dessa forma, a DNA polimerase precisa de um “**Primer**”, que é um pedaço de RNA sintetizado por uma RNA polimerase especial chamada **Primase** (Figura 6).



Collaboration of Proteins at the Replication Fork

Figura 6 – Esquema mostrando o papel das proteínas envolvidas na replicação.

Em bactérias existem três tipos de polimerases:

A **DNA polimerase I**, possui baixa capacidade de polimerização $5' \rightarrow 3'$ e é a única que possui atividade exonucleásica $5' \rightarrow 3'$ em DNA dupla fita.

A **DNA polimerase II**, possui uma capacidade de polimerização baixíssima e o seu papel na célula ainda não foi muito bem elucidado.

A **DNA polimerase III**, é a principal responsável pela síntese das fitas de DNA devido a sua alta capacidade de polimerização.

Após a síntese do primer de RNA, a DNA polimerase III pode começar a polimerização no sentido $5' \rightarrow 3'$. Além da polimerização, a DNA-polimerase III possui atividade exonucleásica $3' \rightarrow 5'$, essa atividade permite que logo após serem adicionados, os nucleotídeos sejam retirados e é conferido se o seu pareamento está correto (A com T e C com G). Esta atividade é chamada atividade Editorial. Como as fitas de DNA são antiparalelas, e a DNA polimerase caminha em um sentido no DNA, apenas uma das fitas será sintetizada continuamente, e tem o nome de “leading strand” (fita contínua) e a outra fita será sintetizada descontinuamente, chamada “lagging strand”(fita descontínua). A síntese da fita lagging é feita em pequenas partes. O *primer* de RNA é sintetizado e a fita lagging sofre um “loop” de forma que a polimerização ocorre na mesma direção que ocorre a polimerização da fita leading. Desta forma, a fita complementar à fita lagging é formada em pequenos fragmentos de aproximadamente 100 bases chamados **fragmentos de Okazaki**, que possuem esse nome devido ao seu descobridor, Reiji Okazaki.

Logo após a síntese das fitas pela DNA-polimerase III, entra em cena a DNA-polimerase I, que retira os primers de RNA (atividade exonucleásica $5' \rightarrow 3'$) e os substitui por nucleotídeos de DNA (desoxinucleotídeos). Após a substituição do primer, o primeiro nucleotídeo do fragmento de Okazaki não está ligado ao último nucleotídeo do fragmento anterior, então uma enzima chamada **ligase** catalisa essa ligação. Dessa forma, as duas fitas de DNA já estão terminadas e naturalmente se enrolam formando a dupla hélice.

A replicação em eucariotos segue o mesmo padrão, a maior das diferenças está nas polimerases que são cinco (α , δ , ϵ , β e γ) sendo uma exclusivamente mitocondrial (γ) e as outras quatro nucleares. As polimerases α , δ e ϵ se correlacionam funcionalmente com as polimerases I, III e II de procariotos respectivamente.

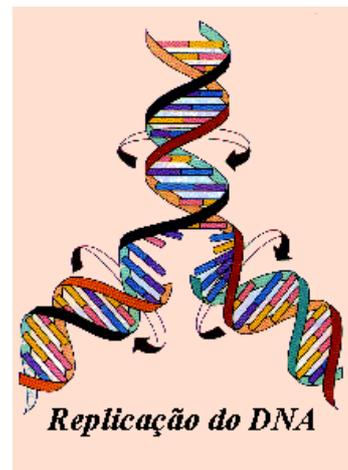
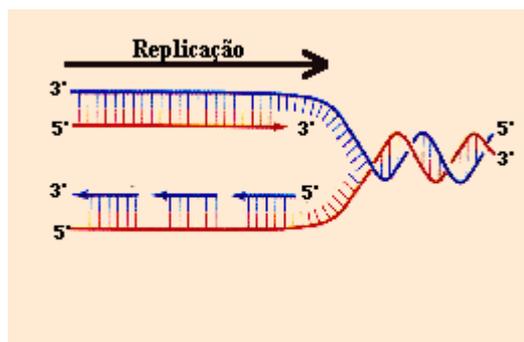


Figura 7 - Esquemas mostrando como ocorre a replicação contínua e descontínua e o produto final da replicação

Decodificando o DNA (Transcrição e Tradução):

A **Transcrição** é o processo de formação de uma fita de RNA complementar a uma região do DNA. Os RNAs formados durante a transcrição podem ser de três tipos: o **RNA_m** (mensageiro) que é aquele RNA que contém a seqüência que codifica uma proteína; o **RNA_t** (transportador) que carrega os aminoácidos até os ribossomos e possibilita a leitura da informação contida no RNA_m durante a tradução e o **RNA_r** (ribossômico) que faz parte da estrutura dos ribossomos.

A enzima responsável pela transcrição é a **RNA polimerase**, que nos procariotos é única, enquanto nos eucariotos elas são em três (RNA pol I, II e III).

Para iniciar a transcrição, a RNA polimerase deve reconhecer um local específico onde começará a síntese. Esse local chama-se promotor. O reconhecimento da RNA polimerase aos promotores se dá graças ao fator sigma, que liga-se à RNA polimerase fazendo com que estas tenham maior afinidade com as seqüências promotoras.

Os promotores contêm seqüências consenso localizadas antes do início da transcrição, a distâncias específicas. Os promotores procarióticos geralmente localizam-se na região -10 e -35 do início da transcrição e as seqüências consenso mais conhecidas são o TATA box na região -10 (TATAAT) e a seqüência TTGACA na região -35. Existem vários tipos de promotores e fatores sigma correspondentes e é essa variedade que permite que as funções celulares possam ser reguladas mantendo o equilíbrio das atividades celulares. Nos eucariotos, o processo de iniciação e regulação da transcrição é muito mais complexo, envolvendo um número e diversidade maior de seqüências promotoras e de fatores de transcrição (análogos ao fator sigma).

A RNA polimerase liga-se ao promotor e inicia a síntese da fita de RNA complementar a fita molde até parar em uma região chamada terminador, que também é uma seqüência de consenso.

Em procariotos, que não possuem envoltório nuclear, a transcrição ocorre no mesmo lugar onde ocorre a tradução, dessa forma, tão logo o RNA comece a ser formado, a tradução já se inicia. Por esse motivo diz-se que a transcrição e a tradução nos procariotos é acoplada. Nos eucariotos a transcrição ocorre no núcleo e a tradução ocorre no citoplasma. O RNA recém sintetizado nos eucariotos ainda precisa passar por várias modificações antes de estar pronto (retirada dos íntrons, adição de uma cauda de poli Adenina, adição de 7-Metil Guanosina na primeira base do RNA e outras).

A **Tradução** converte a informação na forma de trincas de nucleotídeos em aminoácidos, que darão origem a uma proteína. A transcrição ocorre em uma estrutura citoplasmática chamada ribossomo, formada por várias proteínas e RNA_r. O ribossomo é dividido em duas subunidades, uma subunidade menor (30S em bactérias e 40S em eucariotos) e uma subunidade maior (50S em bactérias e 60S em eucariotos).

Considerando o modelo bacteriano, o início da tradução se dá quando a subunidade 30S liga-se ao códon de iniciação AUG (com raras exceções, a tradução sempre começa no códon AUG, correspondente a uma metionina) e logo em seguida o met-tRNA e a subunidade 50S ligam-se ao complexo 30S-RNA_m, tudo isso com o auxílio de fatores de iniciação (IFs). O ribossomo reconhece a trinca correta da metionina, a iniciadora, através do pareamento de uma seqüência do RNA_r 16S da subunidade menor com uma seqüência no RNA_m que fica próxima ao início da tradução chamada seqüência de Shine-Delgarno.

O ribossomo possui dois sítios de entrada da RNA_t, o sítio P (peptídeo) e o sítio A (aminoácido). O primeiro RNA_t entra no sítio P e o segundo entra no A, algumas enzimas da subunidade 50S do ribossomo fazem a ligação do primeiro aminoácido com

o segundo, promovendo a ligação peptídica, dessa forma, no sítio A ficará o segundo RNAt com um dipeptídeo. Através de um processo chamado translocação o ribossomo se desloca um códon a frente de forma que o dipeptídeo-RNAt fica na sítio P. A partir daí, esse processo se repete, formando um tripeptídeo no sítio A que é translocado para o sítio P formando assim sucessivamente a proteína.

A tradução termina quando o ribossomo encontra um códon de terminação, que não codifica nenhum aminoácido, são eles: UGA, UAG e UAA. A estes códons se liga um fator para terminação da síntese (RF- Releasing Factor)

Após a tradução as proteínas ainda têm que passar por algumas modificações para que possam exercer adequadamente suas funções.

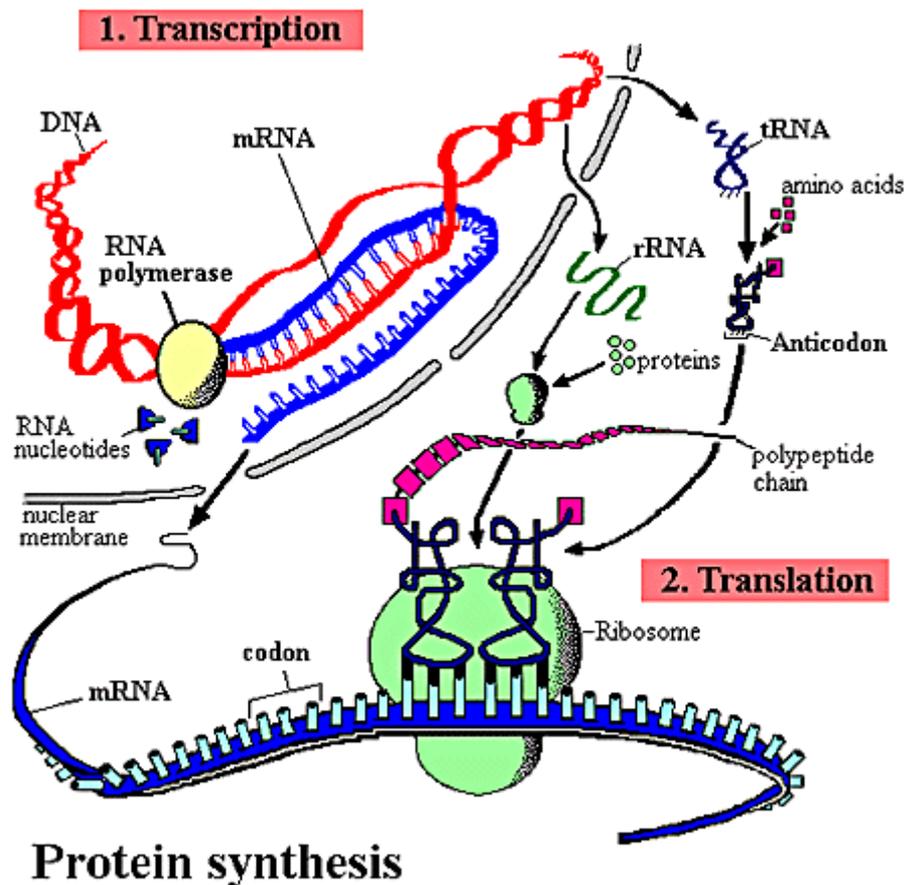


Figura 8 - Esquema mostrando a transcrição e a tradução em eucariotos.

Manipulação do DNA

Enzimas de restrição e vetores.

Quando se trata de biologia molecular, as enzimas de restrição e os vetores sempre estarão presentes, pois são usados em quase todos (se não em todos) os trabalhos dessa área.

As enzimas de restrição são enzimas que cortam o DNA em seqüências específicas. As enzimas de restrição fazem parte de um sistema celular de modificação e restrição que tem como função proteger a célula contra DNAs “estranhos”. O sistema funciona da seguinte forma: Enzimas específicas metilam algumas bases em algumas seqüências específicas do DNA da célula e dentro dessa célula existem enzimas de restrição que reconhecem essas seqüências e clivam o DNA, a não ser que esteja metilado. Dessa forma, o DNA celular está a salvo das enzimas de restrição e qualquer DNA estranho que entre na célula e possua a seqüência de reconhecimento das enzimas de restrição será clivado.

Depois da descoberta desse sistema várias enzimas de restrição de vários organismos foram identificadas e hoje centenas são comercialmente avaliáveis.

A seguir uma tabela com alguns exemplos de enzimas de restrição, sua origem e seus sítio de clivagem: (os espaços entre as bases mostram os pontos de corte).

Tabela 1. Exemplos de enzimas de restrição e organismos de origem

Denominação	Origem	Seqüência de reconhecimento
<i>Eco</i> RI	<i>Escherichia coli</i> RY 13	5' G AATTC 3' 3' CTTAA G 5'
<i>Hind</i> III	<i>Haemophilus influenzae</i> Rd	5' A AGCTT 3' 3' TTCGA A 5'
<i>Kpn</i> I	<i>Klebsiela pneumoniae</i>	5' GGTAC C 3' 3' C CATGG 5'
<i>Xho</i> I	<i>Xantomonas hilicola</i>	5' C TCGAG 3' 3' GAGCT C 5'

Vetores são o meio pelo qual pode-se artificialmente ou naturalmente, carrear informações para dentro de um organismo. Os vetores mais comuns são os plasmídeos, os bacteriófagos, os cosmídeos e os YACs (Yeast Artificial Chromosome - cromossomo artificial de levedura).

Os plasmídeos são sem dúvida, os vetores mais usados. Os plasmídeos são DNAs circulares, que existem naturalmente em bactérias todavia, plasmídeos foram manipulados pelo homem com várias finalidades. Os plasmídeos geralmente possuem um gene de resistência a um antibiótico, uma região de clonagem com sítios de enzimas de restrição, para inserir DNAs de qualquer origem e uma origem de replicação para que esse plasmídeo possa se replicar dentro da célula. Os plasmídeos podem conter promotores de transcrição fortes para expressar proteínas dentro de outros organismos, sendo chamados assim de plasmídeos de expressão (**Figura 9** - Esse assunto será tratado com mais detalhes em outra parte dessa apostila).

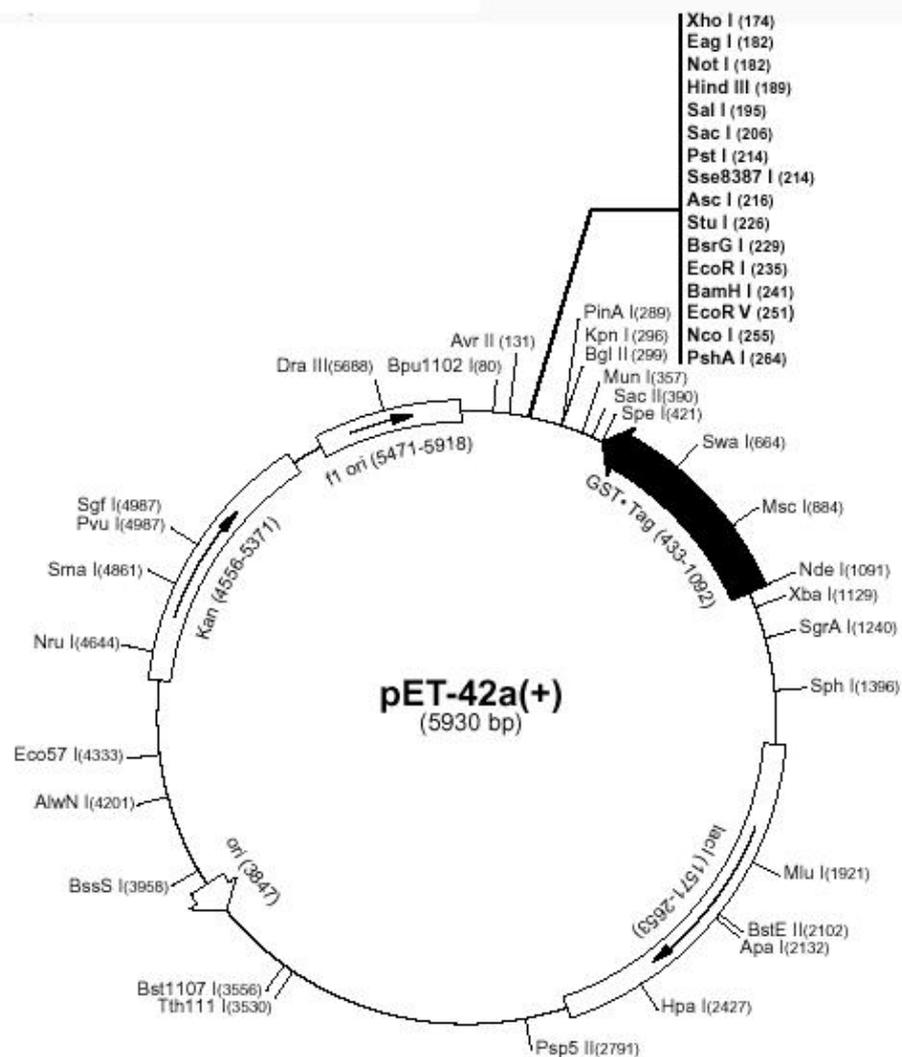


Figura 9. Exemplo de um plasmídeo contendo resistência ao antibiótico Kanamicina. O plasmídeo em questão é utilizado para expressão de proteínas heterólogas em bactérias. A proteína é expressa sob controle de um promotor viral (fago T7) e em fusão com a Glutathione S-Transferase, o que facilita processos posteriores de purificação.

Os bacteriófagos são vírus que infectam bactérias. Para serem utilizados como vetores do seu material genético são retiradas determinadas seqüências que são substituídas pela seqüência de interesse. A molécula recombinante é introduzida em fagos “vazios” que são então utilizados para infectar bactérias e propagar o DNA de interesse.

Os cosmídeos são muito semelhantes a plasmídeos, possuindo porém uma seqüência de fago, que possibilita seu posterior “empacotamento” em cápsulas virais. Diferente de plasmídeos, eles comportam insertos de tamanhos maiores, acima de 20 Kb.

Os YACs são, como o nome já diz, cromossomos artificiais de leveduras e geralmente são usados quando se quer ter um grande pedaço de DNA exógeno dentro de uma levedura. Possuem seqüências centroméricas e teloméricas para que funcione perfeitamente como um cromossomo.

Amplificação do DNA “in vitro”: A reação em cadeia da polimerase - PCR

O avanço das técnicas de Biologia Molecular tem revolucionado os conhecimentos sobre o genoma, incluindo o da espécie humana, quanto a sua estrutura, expressão e função. Entretanto, um fator limitante sempre foi a pequena quantidade de material disponível para o estudo.

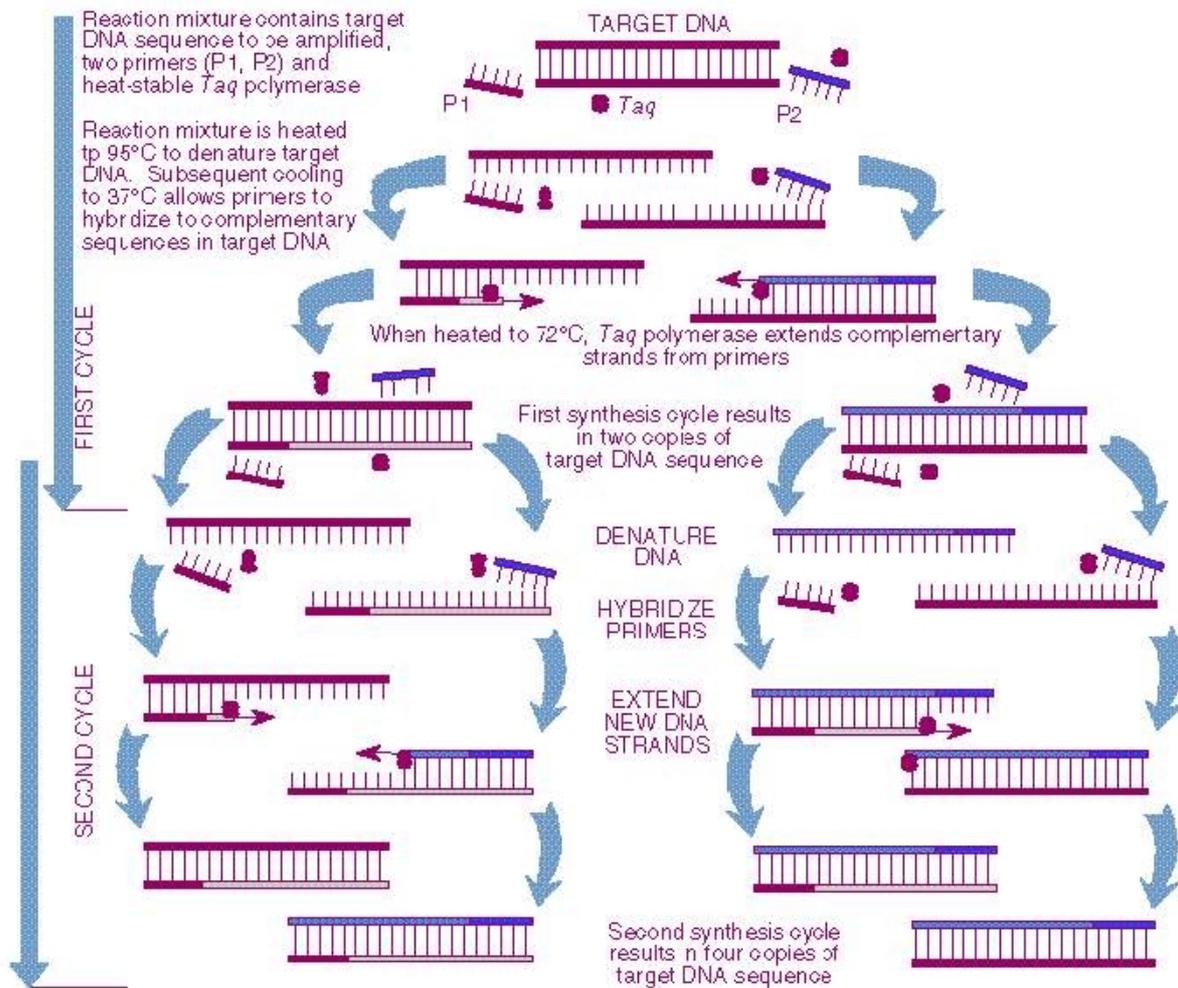
O método para análise de DNA proposto por Saiki e cols. em 1985, o da Reação em Cadeia da Polimerase (PCR), permitiu um grande avanço neste sentido. Este método tem como princípios básicos a de oligonucleotídeos (*primers*) a regiões específicas da molécula de DNA e a ação da DNA polimerase. A partir daí tornou-se possível amplificar um número ilimitado de seqüências alvo, as quais podem ser estudadas por métodos convencionais de análise do DNA.

A técnica de PCR tem por objetivo amplificar uma seqüência alvo específica por meio de uma enzima termoestável, *in vitro*. Neste método utiliza-se além desta enzima, os quatro nucleotídeos e um par de *primers* (aproximadamente 20pb) que flanqueiam a região a ser amplificada. Estes primers são utilizados para direcionar a síntese do DNA, em ciclos repetidos.

Em cada ciclo, as fitas servem de molde para a geração de novas fitas, como ilustra a **figura 10**.

ORNL-DWG 91M-17475

DNA Amplification Using Polymerase Chain Reaction



Source: *DNA Science*, see Fig. 13.

Figura 10. Representação esquemática da amplificação utilizando PCR.

Cada ciclo se inicia com a desnaturação da dupla-hélice do DNA, elevado à altas temperaturas (91-95° C), por aproximadamente 1 minuto. Esta etapa é seguida da dos *primers* ao DNA molde, a temperaturas que variam de 53-65° C por 1 a 2 minutos, posteriormente ocorre a alongação de cadeia por ação da polimerase, em geral entre temperaturas de 65-72° C durante 2 a 5 minutos. O número de ciclos varia de acordo com o objetivo e condições utilizadas.

Até 1987, a técnica da PCR era limitada principalmente devido à instabilidade térmica da enzima, cuja atividade diminuía com o calor, se fazendo necessário sua reposição ao término de cada ciclo. Este problema foi solucionado com a utilização de uma outra polimerase descoberta em *Thermus aquaticus*, a Taq DNA polimerase, que se apresenta extremamente estável em altas temperaturas, dispensando novas adições a cada ciclo.

Clonagem

Uma das contribuições mais valiosas da tecnologia genética para a pesquisa biológica consiste na possibilidade de clonagem de seqüências de DNA. Para isto, um fragmento de DNA de qualquer origem é incorporado em um plasmídeo pequeno, com o qual o hospedeiro (em geral a bactéria *Escherichia coli*) será transformado. Este plasmídeo, incluindo o fragmento de DNA nele incorporado, se reproduz no hospedeiro, originando um certo número de cópias de si e conseqüentemente do DNA estranho.

O uso de enzimas de restrição é um dos passos essenciais desta técnica; com elas o preparado de DNA em questão é cortado em fragmentos precisamente definidos. Se a enzima de restrição usada for do tipo que corta as duas fitas do dúplice assimetricamente, os fragmentos por ela originados possuirão extremidades em fita única, com seqüências complementares idênticas. Em temperaturas fisiológicas, os cortes realizados pela enzima de restrição nas fitas de DNA, mesmo que deslocadas algumas bases, causam uma quebra irreparável no dúplice: as poucas pontes de hidrogênio restantes não são o bastante para manter os fragmentos juntos. Em temperaturas mais baixas, quando o movimento térmico das moléculas é menor, algumas poucas pontes de hidrogênio já são suficientes para manter a união da estrutura. A conseqüência prática disso é que os fragmentos, só por resfriamento se “colam” uns aos outros devido a renaturação das curtas extremidades em fita única (resultado da clivagem por enzimas de restrição). Por intermédio da enzima ligase, a continuidade dessas fitas pode ser restaurada juntando dessa forma duas fitas de DNA de origens diferentes. Quando junta-se um pedaço qualquer de DNA a um plasmídeo, mantendo a possibilidade de se propagar indefinidamente este DNA, dá-se o nome de “clone”.

Nem sempre a eficiência com que o DNA a ser clonado é incorporado ao plasmídeo é satisfatória. Por isso desenvolveram-se diferentes sistemas seletivos para as situações desejadas. Por exemplo, pode-se escolher como veículo de clonagem um plasmídeo contendo um fator de resistência a um antibiótico, de modo que, em meio contendo o antibiótico em questão, somente os hospedeiros que aceitaram o plasmídeo com o fator de resistência podem formar colônias. O plasmídeo com o fator de resistência para ser interessante neste caso tem que estar carregando o DNA a ser clonado. Para se distinguir entre veículos com ou sem o DNA exógeno, lança-se mão de um outro fator no plasmídeo localizado no ponto de ação da enzima de restrição. Se a estrutura anelar do plasmídeo for restaurada e o gene indicador voltar a sua função normal, então o veículo está sem o DNA estranho; porém se o fragmento de DNA tiver sido inserido nesse ponto, o gene indicador perde sua função. Para este segundo

indicador pode-se utilizar outra resistência a antibióticos: as colônias crescidas no primeiro meio seletivo podem ser transferidas para outro meio, pela técnica de carimbo, testando assim sua resistência ao segundo antibiótico. Hoje vários outros métodos foram desenvolvidos para se confirmar a presença de um fragmento inserido em um plasmídeo, todavia, a maioria se baseia nesse princípio de possuir um gene indicador no sítio de restrição onde será feita a clonagem. Recentemente têm-se utilizado genes “suicidas”, os quais, se não forem interrompidos por um DNA clonado, produzirão proteínas tóxicas para a célula. Portanto, apenas serão viáveis colônias portando plasmídeos recombinantes.

A primeira clonagem foi feita no ano de 1973 por Stanley Cohen e Herbert Boyer, e inicialmente as clonagens eram feitas usando-se como inserto fragmentos de DNA resultantes de clivagem de DNAs totais, dessa forma, o rendimento das clonagens era muito baixo. A técnica de PCR veio como um grande aliado para a clonagem, pois com ela, pode-se clonar ampliações, aumentando assim o rendimento (devido ao aumento no número de insertos) e a confiabilidade do experimento.

Várias técnicas utilizam essa ferramenta, entre elas pode-se citar o seqüenciamento de DNA e a produção de proteínas em laboratório, que serão explicadas em outros capítulos dessa apostila.

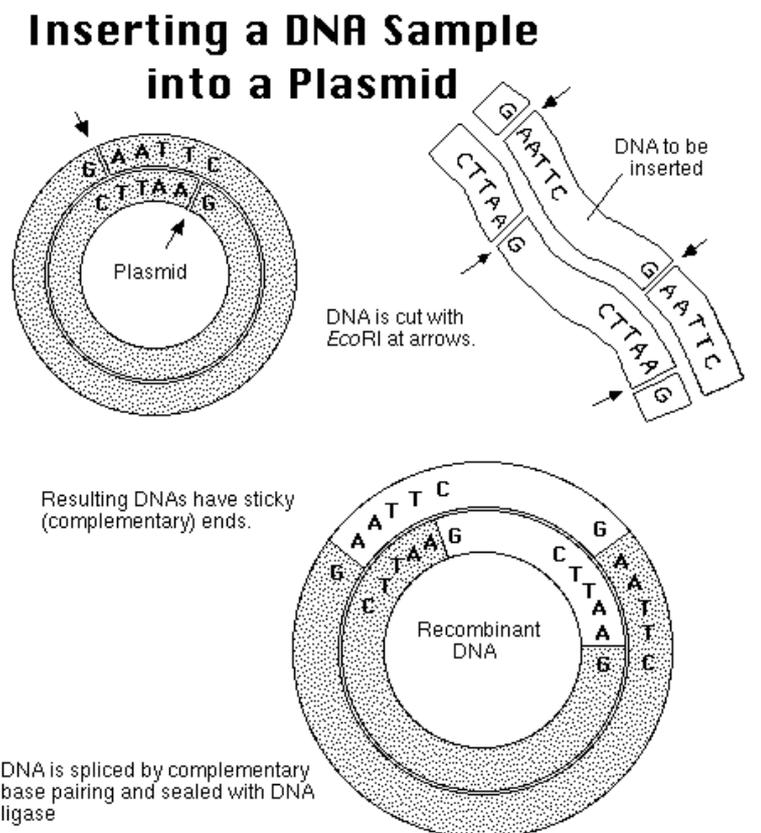


Figura 11 – Esquema mostrando a clonagem de um fragmento de DNA em um plasmídeo usando-se a enzima *Eco RI*.

Hibridação de ácidos nucleicos

DNAs com seqüências de bases específicas podem ser identificados por uma técnica desenvolvida por Edwin Southern conhecida como “Southern Blotting”. Essa técnica baseia-se em fazer uma eletroforese em gel, e transferir o DNA, na forma de fita simples, do gel de eletroforese para uma membrana de nitrocelulose ou nylon (que possui afinidade por DNA), que será uma réplica do gel. Essa transferência é feita em meio alcalino para desnaturar o DNA, deixando-o assim em fita simples. Essa membrana, que contém o DNA desnaturado, é então colocada em contato com fragmentos de DNA, complementares a seqüência alvo, marcados com fósforo radioativo ou fosfatase alcalina. Esses fragmentos marcados, chamados sondas, hibridarão nas suas regiões complementares e quando a membrana for lavada, apenas as sondas hibridadas permanecerão na membrana. Essa membrana é então colocada em contato com um filme fotográfico, que apresentará marcas na altura das bandas que contiverem a seqüência alvo, devido as sondas que emitem radiação (no caso das marcadas com fósforo radioativo) ou algum outro comprimento de onda capaz de marcar o filme fotográfico (outros tipos de marcação de sondas).

Seqüências de RNAs podem também ser detectadas por uma técnica, que usa o mesmo princípio chamada “Northern Blotting”; assim como proteínas que podem ser detectadas usando-se seus anticorpos como sonda na técnica chamada “Western Blotting” ou “Imunoblotting”.

A Hibridação é uma técnica muito utilizada em laboratórios de biologia molecular. Suas aplicações vão desde tipagem de indivíduos (teste de paternidade, identificação de criminosos) até a descoberta de novos genes envolvidos nos mais diversos processos celulares.

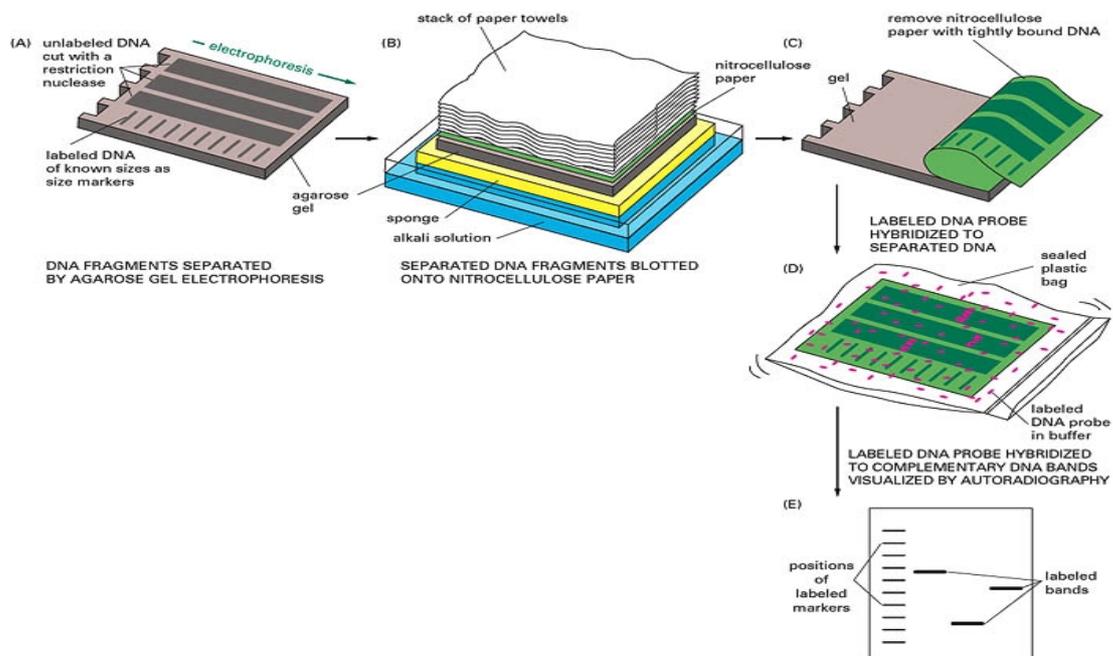


Figura 12 – Esquema mostrando o procedimento da técnica de Hibridação.

Seqüenciamento de DNA

Antes de 1975, pensar em tentar seqüenciar cromossomos era praticamente inconcebível. Na melhor das hipóteses, era uma tarefa trabalhosa requerendo muitos anos de trabalho. No final de 1976, o cromossomo inteiro do fago X174, de 5.387 nucleotídeos tinha sido seqüenciado. Hoje é possível obter as seqüências de nucleotídeos de genes de cromossomos eucarióticos inteiros, e estes dados são armazenados em bancos de dados em computadores para referência posterior.

Foi desenvolvido nesta época, quase que simultaneamente, duas técnicas que promoveram uma revolução na "arte" de seqüenciar DNA. O desenvolvimento de outras técnicas, como a descoberta de enzimas de restrição (uso na preparação de segmentos específicos de cromossomos); o aperfeiçoamento da eletroforese em gel até o ponto em que fragmentos de DNA que diferem em comprimento por um único nucleotídeo pudessem ser resolvidos, e o desenvolvimento de técnicas de clonagem de genes (preparação de grandes quantidades de um determinado gene ou seqüência de DNA de interesse) viabilizaram o seqüenciamento.

Em 1975, F. Sanger e colaboradores desenvolveram um método enzimático com terminadores de cadeia específicos, para gerar quatro populações de fragmentos que terminam em As, Gs, Cs e Ts, respectivamente.

O método de Sanger utiliza de terminadores de cadeia chamados de 2'3'-didesoxinucleotídeos trifosfatos. As DNA polimerases necessitam de um OH- 3' livre na fita de DNA iniciadora. Se um 2' 3' didesoxinucleotídeo é adicionado à extremidade de uma cadeia, ele irá bloquear a extensão subsequente desta cadeia, uma vez que os 2' 3' -dideoxinucleotídeos possuem um grupo OH- 3' trocado por um H.

O seqüenciamento é feito da seguinte maneira:

Quatro reações são feitas em paralelo, cada uma delas contendo um dos quatro terminadores de cadeia 2', 3'-didesoxi: ddGTP, ddATP, ddCTP e ddTTP. As quatro reações contêm: uma fita molde (simples fita), a seqüência de nucleotídeos a ser determinada, uma fita iniciadora com uma hidroxila 3' livre (normalmente obtida por síntese química), os quatro precursores de DNA: dGTP, dATP, dTTP e dCTP, um deles pelo menos radioativo (^{32}P -dCTP), e o "fragmento Klenow" da DNA polimerase I de *E. Coli*. O "fragmento Klenow" representa 2/3 DNA polimerásica I de *E. Coli* produzido por clivagem com enzima proteolítica tripsina. O fragmento possui as atividades polimerásica $5' \rightarrow 3'$ e exonucleásica $3' \rightarrow 5'$ da DNA polimerase I, mas não possui atividade exonucleásica $5' \rightarrow 3'$. Esta atividade exonucleásica $5' \rightarrow 3'$ é crítica para a técnica, pois se esta atividade estiver presente, ela irá reduzir a fita iniciadora a partir da extremidade 5'.

O importante nesta técnica de seqüenciamento é usar a relação correta do didesoxirribonucleosídeo trifosfato normal (por exemplo, dGTP na reação 1) e de didesoxirribonucleosídeo trifosfato terminador (por exemplo ddGTP na reação 1), de modo a obter uma população de cadeias nascentes terminando em todas as posições de nucleotídeos possíveis (por exemplo, em todos os Cs de fita-molde na reação 1). A relação de 100 desoxirribonucleosídeos trifosfatos para 1 didesoxirribonucleosídeo trifosfato normalmente dá a freqüência de terminação desejada em cada sítio potencial de terminação.

Os produtos das quatro reações são então separados por eletroforese em gel de poliacrilamida e as posições dos produtos de reação radioativos (cadeias nascentes de DNA) são determinadas por autorradiografia. Uma vez que as cadeias menores migram distâncias maiores no gel, a seqüência de nucleotídeos definida pelas cadeias nascentes

é dada pela leitura na direção $5' \rightarrow 3'$, a partir da parte de baixo (anodo) até o topo (catodo) do gel.

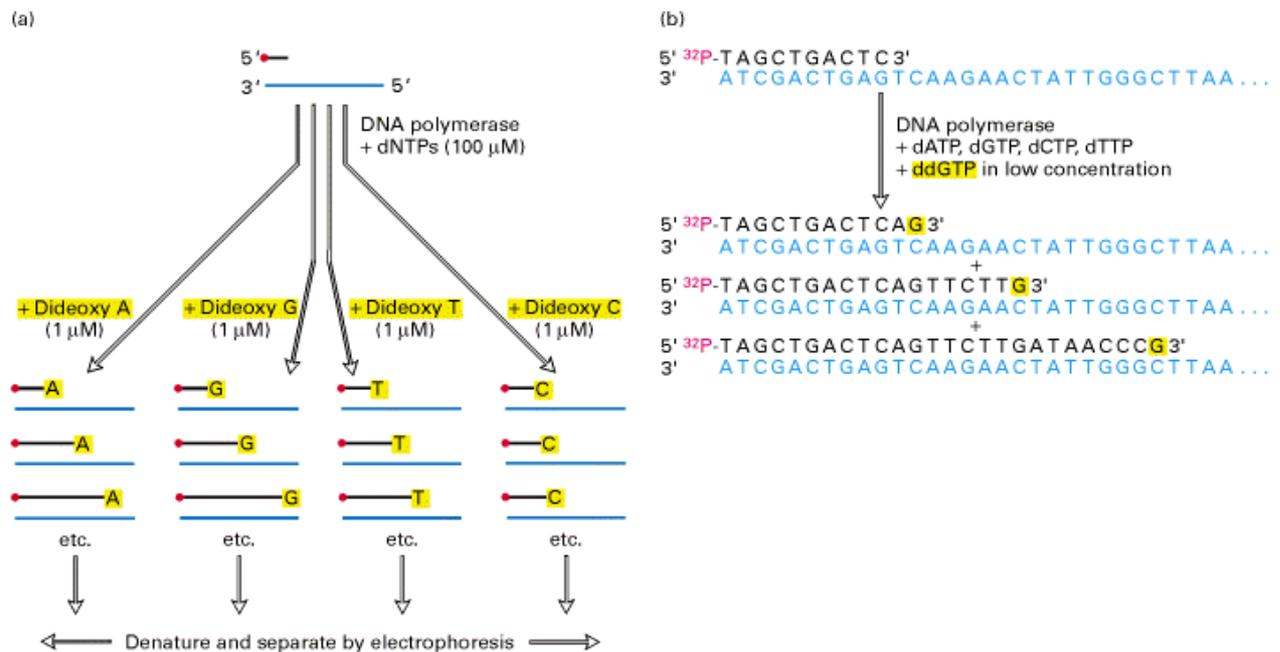


Figura 13 - Esquema mostrando como é feito o seqüenciamento de DNA pelo método dos desoxirribonucleotídeos desenvolvido por Sanger

Atualmente não são utilizados nucleotídeos radioativos. Os didesoxinucleotídeos são marcados com substâncias fluorescentes que emitem em comprimentos de onda diferentes. Desta forma, os fragmentos que incorporam estes nucleotídeos podem ser identificados após excitação com laser e identificação de cada fluorescência. Isso possibilita que a reação seja feita em um único tubo. Além disso, são utilizadas enzimas termoestáveis em substituição à Klenow e à T7 DNA polimerase. A reação, que não é uma amplificação, apenas extensão, é realizada de forma cíclica. Isso faz com que possam ser obtidas seqüências a partir de uma quantidade menor de DNA e também auxilia no seqüenciamento de regiões que possuem estruturas que, em temperaturas mais baixas, impedem a ação das polimerases. A utilização de temperaturas mais altas (70°C) evita este problema. Atualmente são utilizados seqüenciadores automáticos que permitem a realização de seqüências em larga escala.

Utilização da tecnologia do DNA recombinante

Produção de proteínas

A produção de proteínas em laboratório é uma técnica que vem sendo cada vez mais utilizada. Essa produção é feita através de uma técnica chamada *expressão heteróloga*, que consiste em produzir em um organismo proteínas originárias de outros (Ex: produção de hormônios humanos em leveduras).

Para modificar esses organismos, normalmente utiliza-se de plasmídeos como vetores. Esses vetores devem possuir algumas características desejáveis: um promotor de transcrição

forte, específico para o organismo que produzirá a proteína, que funcionará somente na presença de alguma substância indutora (IPTG, Metanol, Celulose, Arabinose, etc); um gene de seleção, que confere uma característica diferencial entre as células com e sem o vetor (crescimento em meio mínimo, resistência a antibióticos, etc) e um sítio para clonagem.

Em um sistema bastante utilizado, a fase aberta de leitura do gene que codifica a proteína a ser produzida é clonada no vetor, após o promotor forte e uma seqüência de 18 (dezoito) nucleotídeos que codificam 6 (seis) histidinas, que servirão para futura purificação em coluna de afinidade. Esse vetor recombinante servirá para transformar o organismo que produzirá a proteína. Esse vetor pode ficar na célula de duas formas: como uma estrutura circular de replicação autônoma (plasmídeo), ou pode integrar-se no genoma do organismo através de uma recombinação homóloga.

As células transformadas são crescidas até ficarem bem vigorosas e então adiciona-se a substância indutora que promoverá a transcrição e tradução do gene de interesse em grandes quantidades.

A proteína de interesse estará então em grande quantidade na célula, todavia, para tê-la na forma pura precisa-se separá-la das outras proteínas celulares. Foi pensando nisso que as 6 (seis) histidinas foram adicionadas à proteína, pois essa “cauda de histidina” possui afinidade por níquel. Desta forma, após separação de todas as proteínas celulares por uma coluna com uma resina complexada com níquel, as proteínas que possuírem a cauda de histidina (somente a de interesse) ficarão grudadas na resina e as outras passarão direto. Para retirar a sua proteína da coluna de níquel, é só passar um tampão que reduza a afinidade das histidinas ao níquel, geralmente contendo um competidor.

A técnica de expressão pode apresentar muitas variações, desde o promotor usado até o organismo produtor. Cada um desses fatores confere vantagens e desvantagens em relação a outro. Como exemplo pode-se comparar a expressão em fungos e em bactérias: As bactérias possuem crescimento muito mais rápido, todavia geralmente apresentam problemas ao expressar proteínas muito grandes (formação de corpos de inclusão), além disso, não realizam modificações pós-traducionais, enquanto que os fungos, apesar de crescerem mais lentamente possuem mecanismos que possibilitam que a proteína fique na sua forma ativa e com modificações pós-traducionais. São desvantagens destes últimos a difícil manipulação e a presença de proteases.

Em última escala, a transformação pode ser feita em organismos inteiros, criando os organismos transgênicos. Há toda uma discussão ética por trás dos transgênicos, todavia, os transgênicos podem em muitos casos ser a solução para vários problemas mundiais como a fome e a falta de espaço.

A produção de proteínas em laboratório é um grande avanço para a ciência, graças a ela, facilitou-se muito a determinação da estrutura de proteínas, que podem estar envolvidas em doenças, possibilitando a fabricação de drogas para bloquear ou controlar a sua ação ou as proteínas purificadas podem ser usadas diretamente no tratamento de doenças como diabetes (insulina), nanismo (Hormônio de Crescimento) entre outras.

Tipagem -DNA "fingerprint"

DNA "fingerprint" é um termo utilizado para referir-se ao teste de identificação de indivíduos através do exame de DNA. Este termo impressão digital nada tem a ver com o recurso datiloscópico tradicional. Mas foi adotado porque não há, também, duas estruturas de DNA iguais entre as pessoas. Trata-se de um método revolucionário para a identificação de indivíduos, a partir do qual é possível estabelecer, com precisão absoluta, vínculos genéticos para efeito de investigação de paternidade, casos de crianças trocadas na maternidade ou

seqüestradas, construção de árvores genealógicas para estabelecimento de direito a heranças e esclarecimentos de crimes.

O cientista inglês, Allec Jeffreys, na Universidade de Leicester, descobriu que certas regiões do DNA humanos podiam ser examinadas através de ferramentas de engenharia genética.

Até o final da década de oitenta a Ciência Médica dispunha dos testes clássicos, envolvendo os antígenos eritrocitários e o exame do HLA (Human Leukocyte Antigen). O advento do exame de DNA proporcionou um avanço significativo em certas eventualidades, principalmente quando as probabilidades de paternidade, calculadas através de testes convencionais, eram baixas ou de magnitude insuficiente para que a paternidade biológica fosse claramente estabelecida. Deve-se ter em mente que o exame de DNA exclui, praticamente, 100% dos falsos pais biológicos, e possibilita o cálculo da probabilidade de paternidade sempre em valores acima de 99,9%. O exame de DNA já se encontra padronizado em seus procedimentos técnicos (American Association of Blood Banks 1990) e se encontra validado pela justiça e plenamente aceito em vários países. De posse desta ferramenta de alta tecnologia e poder de resolução sem precedentes é, então, possível a elucidação de casos complexos de investigação de paternidade, inclusive aqueles em que o suposto pai encontra-se falecido ou não se encontra disponível para o teste.

Em humanos, apenas 5% do genoma total contém informação de aproximadamente 50.000 a 100.000 genes que governam funções celulares. Os restantes 95% de DNA, não codificam proteínas e é onde se encontram as regiões repetitivas, que são seqüências que se repetem várias vezes. Existem dois tipos de repetições no genoma, uma quando as seqüências repetidas ocorrem uma ao lado da outra (em "tandem") e a outra é quando as repetições ocorrem de forma aleatória no genoma. Várias classes de seqüências repetitivas de DNA têm sido escritas e caracterizadas em várias espécies de animais e vegetais. Estas diferem no número e na composição dos nucleotídeos. A diferença entre dois indivíduos pode ser determinada através do polimorfismo dessas regiões repetitivas do DNA, que diferem de tamanho (número de repetições) dentro de uma população.

A caracterização pode ser feita de duas formas: por PCR ou por hibridação. Por PCR utiliza-se primers que flanqueiam as regiões repetitivas, amplificando-as. Dependendo do tamanho da região repetitiva, os fragmentos poderão ser maiores ou menores e poderão ser diferenciados através de eletroforese em gel. A limitação dessa técnica por PCR está no tamanho da região repetitiva que não pode ser muito grande, dessa forma, o número de alelos geralmente é menor. A caracterização por hibridação consiste em cortar o DNA com enzimas de restrição, correr em gel de agarose, transferi-lo para uma membrana e usar como sondas a seqüência cerne da repetição. A Hibridação permite identificar regiões repetitivas maiores (algumas vezes de dezenas de milhares de bases), que possuem um maior polimorfismo e são mais confiáveis.

O resultado de cada uma dessas regiões e dado por um padrão de duas faixas (bandas) representando os dois alelos daquele loco (já que o ser humano é diploide). Com a análise de várias regiões repetitivas, pode-se então obter um padrão (quase) individual.

Nos casos de criminalística, o DNA do criminoso obtido na cena do crime é comparado com o DNA dos suspeitos, podendo-se dessa forma, identificar com grande confiabilidade se algum dos suspeitos é o criminoso. A extração de DNA pode ser feita de sangue, fios de cabelo (bulbo), pêlos, manchas de sangue ou esperma, ossos, pedaços de pele e até de insetos hematófagos (carrapatos e ácaros).

Nos casos de paternidade, o padrão do filho é colocado em comparação com o do suposto pai e com o da mãe, um dos alelos (bandas) do filho vem da mãe, o outro, vem do pai verdadeiro. Analisando-se várias regiões pode-se então incluir o pai com grande grau de certeza (aprox. 99,99%) ou excluí-lo da paternidade do filho.

O diagrama abaixo mostra um caso de identificação de paternidade.

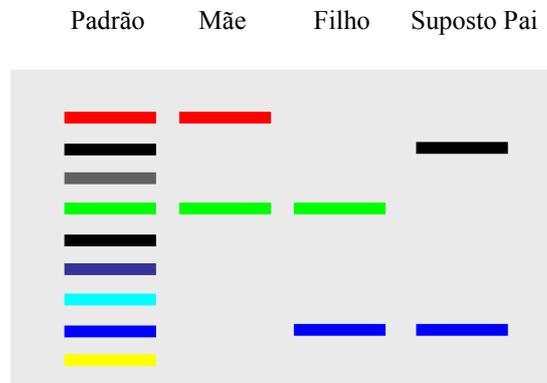


Figura 14- Esquema mostrando resultado de um teste de paternidade para um loco. Um alelo do filho vem da mãe (verde) e o outro (azul) tem que vir do pai biológico. Esse é um caso de inclusão no qual o suposto pai é o pai da criança (ao menos considerando esse loco).

Aplicações na medicina

O primeiro trabalho documentado sobre as aplicações dos produtos da PCR data de 1985 e relata a mutação que causa a anemia falciforme.

Nos dias de hoje, a biologia molecular tem contribuído muito com a medicina, principalmente no que diz respeito a diagnósticos de doenças. Distrofia muscular do tipo Duchene e do tipo Becker, fibrose cística do pâncreas e AIDS são apenas alguns exemplos de doenças que podem ser diagnosticadas por técnicas de biologia molecular.

O PCR é a técnica mais usada para diagnósticos de doenças. Por PCR, pode-se também detectar mutações em oncogenes (p53, K-ras e BRCA 1 e 2); detectar a causa de infecções através de marcadores moleculares específicos, canalizando o tratamento de forma eficiente; verificar tendência genéticas a enfarto através da tipagem do gene ACE (enzima conversora de angiotensina) e fazer vários diagnósticos pré-natais.

Como exemplo de um diagnóstico feito por PCR, pode-se citar o diagnóstico da Distrofia muscular de Duchene (DMD). Essa doença caracteriza-se por uma degeneração progressiva e irreversível dos músculos esqueléticos. É determinada por um gene do cromossomo X, e o padrão de herança é recessivo ligado ao X, afetando apenas indivíduos do sexo masculino.

O gene da DMD foi clonado em 1987 e possui 2,5 milhões de pares de bases, de onde cerca de 99% são íntrons.

A identificação da deleção em pacientes com suspeita de DMD/DMB é extremamente importante para o prognóstico clínico e o aconselhamento genético.

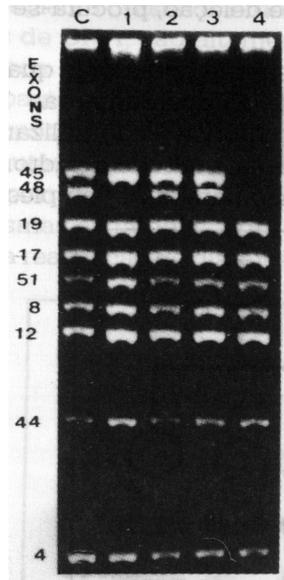


Figura 15 – Eletroforese em gel de agarose mostrando 4 resultados de diagnósticos de DMD, “C” representa o controle (indivíduo sem deleções); os indivíduos 2 e 3 não apresentam deleção para nenhum dos éxons testados; o indivíduo 1 apresenta deleção do éxon 48 e o indivíduo 4 apresenta deleção nos éxons 45 e 48 (Figura retirada do Capítulo 7 do livro de Ácidos Nucléicos, capítulo escrito pelas Dras Maria Rita Passos Bueno e Mayana Zatz. O livro foi organizado pelo Dr. Francisco J. S. Lara)

Terapia Gênica

Além do diagnóstico a biologia molecular proporciona também procedimentos para correção de doenças genéticas. Desta forma, genes defeituosos podem ser substituídos por genes normais utilizando vetores específicos (retrovírus, adenovírus, etc...) para carrear estes genes. Isso já foi feito com sucesso para corrigir a Síndrome da Imunodeficiência Congênita, causada por mutações no gene que codifica a enzima Adenosina Deaminase (ADA) e também para a Fibrose Cística, causada por mutações na CFTR (Cystic Fibrosis Transmembrane Regulator). Embora alguns sucessos já tenham sido obtidos, faz-se necessária ainda muita pesquisa, principalmente relativa ao desenvolvimento de vetores seguros para transporte dos genes de interesse. Além disso, este tipo de terapia ainda é muito cara, tornando-a impraticável nos moldes atuais.

Leitura Recomendada

Francisco J.S. Lara - **de Ácidos Nucléicos** (1995) - Sociedade Brasileira de Genética

Lodish et al., **Molecular Cell Biology** (1999) – W.H. Freeman and Company, New York, 4th Ed

Watson et al., **Recombinant DNA** (1992) – Ed. Scientific American Books, New York, 2nd Ed

Strachan and Read , **Human Molecular Genetics** (2000) - BIOS Scientific Publishers Ltd, 2nd Ed

Ninfa and Ballou, **Fundamental Laboratory Approaches for Biochemistry and Biotechnology** (1998) - Fitzgerald Science Press, Inc., Bethesda, Maryland, USA

Lubert Stryer, **Bioquímica** (1996) – Guanabara Koogan S.A., RJ, Brasil, 4^a Ed.

Benjamin Lewin, **Genes VII** (2000) – Oxford University Press, Inc., New York

Voet and Voet, **Biochemistry** (1997) – John Wiley & Sons, New York, 2nd Ed.