

CÓDIGO GENÉTICO E SÍNTESE PROTEICA

Juliana Mara Stormovski de Andrade

As proteínas são as moléculas mais abundantes e funcionalmente diversas nos sistemas biológicos. Praticamente todos os processos vitais dependem desta classe de moléculas. Apresentam uma incrível diversidade de funções, embora todas compartilhem a característica estrutural comum de serem polímeros lineares de aminoácidos.

As proteínas são compostas de um ou mais polipeptídeos, que por sua vez são formados por aminoácidos. Embora mais de 300 aminoácidos diferentes tenham sido descritos na natureza, somente 20 são comumente encontrados como constituintes das proteínas dos mamíferos. Existem 20 aminoácidos diferentes para formar vários tipos de proteína, as quais diferem pela posição dos aminoácidos, como o alfabeto, que contém 23 letras formando milhares de palavras.

O DNA é uma molécula presente em todas as células nucleadas do organismo de seres vivos. No DNA nuclear e mitocondrial de cada célula estão contidos todos (e sempre os mesmos) os genes daquele organismo, genes esses responsáveis pela transmissão das características genéticas. O DNA nos fornece “receitas” de produção de proteínas.

A informação genética, armazenada nos cromossomos e transmitida às células filhas através da replicação do DNA, é expressa através da transcrição em RNA e, no caso, de RNAm, tradução subsequente em cadeias polipeptídicas. Este fluxo de informação do DNA ao RNA e à proteína é denominado de “dogma central” da biologia molecular, sendo descritivo de todos os organismos (com exceção de alguns vírus que usam o RNA como repositório de sua informação genética). O processo de tradução requer um código genético, através do qual a informação contida na sequência de ácidos nucleicos é expressa para produzir uma sequência específica de aminoácidos. A ligação molecular entre estes dois tipos relacionados de informação (o código de DNA dos genes e o código de aminoácidos das proteínas) é o ácido ribonucleico (RNA).

Os constituintes da molécula de DNA são denominados nucleotídeos. Um nucleotídeo é formado por três componentes básicos: uma base nitrogenada, um açúcar (desoxirribose) e um grupo fosfato. As bases nitrogenadas podem ser do tipo purinas (adenina A, e guanina G) ou pirimidinas (timina T, e citosina C). As bases purinas e pirimidinas, encontradas nos nucleotídeos podem ser sintetizadas de novo, ou ser obtidas através de rotas de salvamento que permitem a reutilização das bases pré-formadas resultantes do metabolismo normal da célula ou da dieta.

O CÓDIGO GENÉTICO

A informação genética é estocada no DNA por meio de um código (o código genético) no qual a sequência de bases adjacentes determina a sequência de aminoácidos no polipeptídeo codificado. O código genético, então é um dicionário que fornece a correspondência entre uma sequência de bases nucleotídicas e uma sequência de aminoácidos. Código genético=Receita genética (“livro de receitas de proteínas”). Em

teoria, são possíveis variações quase infinitas na disposição das bases ao longo de uma cadeia polinucleotídica. Uma vez que existem 20 aminoácidos diferentes e apenas quatro bases diferentes de RNA, uma única base não pode especificar cada aminoácido. Em qualquer posição existem quatro possibilidades (A, T, C, G). Assim, existem 4^n combinações possíveis em uma seqüência de n bases. Dessa forma, aminoácidos específicos não podem ser determinados por duplas de bases, porque são possíveis apenas 16 ou 4^2 pares diferentes. Entretanto, se trincas de bases forem traduzidas em aminoácidos, há 4^3 combinações possíveis de trincas, ou seja, 64 combinações podem ser obtidas, mais do que o suficiente para especificar cada aminoácido.

O código genético foi decifrado em 1953. Watson e Crick demonstraram, de modo elegante e apurado, que o código consiste em códons, cada um composto por uma trinca de bases nitrogenadas (tripletes). Dos 64 códons (RNAm) possíveis, três indicam o fim de um gene, e são conhecidos como códons finalizadores (ou sem sentido) porque designam o termino da tradução do mRNA neste ponto. São o UAA, o UGA e o UAG. Os outros 61 especificam aminoácidos. Como existem apenas 20 aminoácidos essenciais, isto significa que a maioria dos aminoácidos pode ser especificada por mais de um códon. Por exemplo, a leucina e a arginina são especificadas por seis códons. Apenas a metionina e o triptofano são cada um deles especificado por um único códon. O código genético é, portanto, dito “redundante” (ou degenerado). Embora um determinado aminoácido possa ser especificado por mais de um códon, cada códon só pode designar um aminoácido. Essa redundante descoberta é fundamental para, entre outras coisas, compreendermos que nem toda alteração no código genético leva a uma doença. Uma alteração de TTT para TTC, por exemplo, não deverá causar absolutamente nenhuma alteração no fenótipo de um indivíduo, porque ambos codificam o mesmo aminoácido. Porém há alterações na seqüência de ácidos nucleicos que podem resultar em um aminoácido inapropriado sendo inserido na cadeia polipeptídica, potencialmente causando uma doença ou mesmo a morte do organismo.

Uma característica significativa do código genético é ser “universal”, ou seja, virtualmente todos os organismos vivos usam os mesmos códigos de DNA para especificar aminoácidos. Uma exceção conhecida a esta regra é a das mitocôndrias, as quais têm suas próprias moléculas de DNA extranuclear. Vários códons do DNA mitocondrial codificam aminoácidos diferentes dos códons do DNA nuclear.

O código genético é extremamente conservado. Os mesmos tripletes correspondem aos mesmos aminoácidos, seja em seres humanos, seja em bactérias. A correspondência entre códons específicos e aminoácidos é conhecida como código genético.

- . Em síntese pode-se dizer que o Código Genético tem as seguintes características:
- **Especificidade**
- **Universalidade**
- **Redundância**

A informação genética está contida no DNA dos cromossomos dentro do núcleo celular, mas a síntese de proteínas, durante a qual a informação codificada no DNA é usada, ocorre no citoplasma.

Devido à compartimentalização das células eucarióticas, a transferência de informação do núcleo para o citoplasma é um processo muito complexo. Enquanto o DNA se forma e se replica no núcleo da célula, ocorre no citoplasma a síntese de proteínas. A

informação contida no DNA deve, portanto, ser transportada para o citoplasma, e assim usada para ditar a composição das proteínas. Isto envolve dois processos, transcrição e tradução. Resumidamente, o código do DNA é transcrito para o RNA mensageiro, que então deixa o núcleo para ser traduzido em proteínas.

RNA

Existem três tipos principais de RNA que participam do processo da síntese protéica: RNA ribossômico (RNAr), RNA de transferência (RNAt) e RNA mensageiro (RNAm). Assim como o DNA, esses três tipos de RNA são moléculas poliméricas não ramificadas, compostas de mononucleotídeos unidos por ligações fosfodiéster. Entretanto, eles diferem do DNA em vários aspectos; por exemplo, são consideravelmente menores, e contêm ribose em vez de desoxirribose e uracil em vez de timina. Os três principais tipos de RNA diferem um do outro em termos de tamanho, função e modificações estruturais especiais.

RNA ribossômico: é encontrado em associação com uma série de proteínas diferentes, como componente dos ribossomos, as estruturas complexas que servem como sítios para a síntese de proteínas. No citosol eucariótico, existem quatro espécies de RNAr de tamanhos diferentes (28S, 18S, 5,8S e 5S). ["S" é a unidade Svedberg, relacionada ao peso molecular do composto]. Juntos constituem até 80% do RNA da célula.

RNA de transferência: é a menor das três principais moléculas de RNA (4S), tem entre 74 e 95 resíduos de nucleotídeos de tamanho, e forma de trevo. Existe no mínimo um tipo específico de molécula de RNAt para cada um dos 20 aminoácidos comumente encontrados nas proteínas. Juntos, eles constituem cerca de 15% do RNA da célula. As moléculas de RNAt contêm bases incomuns que possuem extenso pareamento de bases intracadeia. Cada RNAt serve como um "adaptador", que transporta seu aminoácido específico ao sítio de síntese de proteínas. Lá, ele reconhece o termo do código genético que especifica a adição de seu aminoácido à cadeia peptídica em formação.

RNA mensageiro: compreende somente cerca de 5% do RNA da célula e é o tipo mais heterogêneo de RNA em termos de tamanho. O RNAm transporta a informação genética do DNA ao citosol, onde é usado como molde para a síntese de proteínas. As características estruturais especiais do RNAm eucariótico incluem uma longa seqüência de nucleotídeos adenina (uma "cauda poli-A") no 3'-terminal da cadeia de RNA, mais uma "cabeça" no 5'-terminal (cap 5').

TRADUÇÃO

Um grande número de componentes são requeridos para a síntese da cadeia polipeptídica. Estes incluem todos os aminoácidos que são encontrados no produto final, o RNAm a ser traduzido, os RNAt, ribossomos funcionais, fontes de energia e fatores proteicos necessários à iniciação, alongação e terminação da cadeia polipeptídica.

A tradução é o processo pelo qual o mRNA fornece um molde para a síntese de um polipeptídeo. O mRNA é transportado do núcleo para o citoplasma, onde a seqüência de RNA é codificada, ou traduzida, para determinar a seqüência de aminoácidos na proteína que está sendo sintetizada. O mRNA não pode, entretanto, se ligar diretamente a aminoácidos. O RNA transportador (tRNA) fornece a ligação molecular entre a seqüência

de bases do mRNA e a seqüência de bases da proteína. A ligação do anticódon do RNAt ao códon do RNAm segue as regras da ligação complementar e antiparalela - isto é, o códon do RNAm é lido de 5' para 3' por um anticódon pareado em orientação invertida (3'-5').

No citoplasma, o mRNA é traduzido em proteína pela ação de uma variedade de moléculas de tRNA, cada uma específica para um determinado aminoácido. O RNAt tem a função de transferir os aminoácidos corretos para suas posições ao longo do molde de mRNA, para que sejam adicionados à cadeia polipeptídica crescente.

O tRNA tem um sitio em sua ponta 3' para a ligação de um aminoácido por uma ligação covalente. Na ponta oposta do trevo há uma seqüência de três nucleotídeos chamada de anticódon. Esta seqüência faz um pareamento complementar de bases com um códon apropriado no mRNA. O mRNA, portanto, especifica a seqüência de aminoácidos agindo por meio do tRNA.

O local citoplasmático da síntese de proteínas é o ribossomo, que consiste em proteínas enzimáticas e RNA ribossomal (mais abundante). A função do rRNA é auxiliar a ligação do mRNA e do tRNA ao ribossomo. O ribossomo primeiramente liga-se a um sítio de iniciação na seqüência do mRNA. Este sítio consiste de um códon específico, AUG, que especifica o aminoácido metionina. O ribossomo então liga o tRNA a sua superfície, de modo que possa haver pareamento entre o tRNA e o mRNA. O ribossomo move-se ao longo da seqüência de mRNA, códon por códon, no sentido usual 5' para 3'. O ribossomo então desliza ao longo do mRNA a cada três bases, alinhando o códon seguinte para o reconhecimento por outro tRNA com o próximo aminoácido. À medida que cada códon é processado, um aminoácido é traduzido pela interação entre mRNA e tRNA. A ligação entre o códon e o anticódon coloca o aminoácido apropriado na posição seguinte no ribossomo para formação de uma ligação peptídica com a ponta carboxila e a cadeia polipeptídica crescente.

Neste processo, o ribossomo fornece uma enzima que catalisa a formação de ligações peptídicas covalentes entre aminoácidos adjacentes, resultando em um polipeptídeo crescente. Quando o ribossomo chega a um códon finalizador na seqüência de mRNA, terminam a tradução e a formação de polipeptídeo. O terminal amino (NH₂) do polipeptídeo correspondente à ponta 5' do filamento de mRNA, e o terminal carboxila (COOH) corresponde à ponta 3'. Quando a síntese esta completa, o mRNA, o ribossomo e o polipeptídeo se separam um do outro. O polipeptídeo é então liberado para o citoplasma.

A tradução de um mRNA processado é sempre iniciada em um códon AUG, que especifica metionina. A metionina é, portanto, o primeiro aminoácido codificado (amino-terminal) de cada cadeia polipeptídica, embora em geral seja removido antes que a síntese da proteína esteja completa. O códon para metionina (iniciador AUG) estabelece a matriz de leitura do mRNA.

Todo esse especializado e refinado processo pode ser dividido em três etapas básicas:

Iniciação:

A iniciação da síntese de proteínas envolve a reunião de componentes do sistema de tradução antes que ocorra a formação da ligação peptídica. Estes componentes incluem as duas subunidades ribossômicas, o RNAm a ser traduzido, o aminoacil-RNAt para metionina, especificado pelo códon iniciador AUG na mensagem, o GTP (o qual fornece energia ao processo) e fatores de iniciação que facilitam a montagem deste complexo de iniciação.

Alongamento:

A alongação envolve a adição de aminoácidos à extremidade carboxila da cadeia polipeptídica em formação. Durante a alongação, os ribossomos movem-se do 5'-terminal ao 3'- terminal do RNAm que está sendo traduzido. Há vários fatores de alongamento envolvidos com esse processo. A formação das ligações peptídicas é catalizada pela peptidiltransferase. Após formar-se a ligação peptídica, o ribossomo avança três nucleotídeos em direção ao 3'- terminal do RNAm. Isto causa a liberação do RNAt descarregado

Terminação:

A terminação ocorre quando um dos três códons de encerramento movem-se ao sítio. O fator de liberação faz a proteína recém sintetizada ser liberada do complexo ribossômico e causa a dissociação entre o ribossomo e o RNAm. O polipeptídeo recém sintetizado pode sofrer modificações subseqüentes; as subunidades ribossômicas, o RNAm, o RNAt e fatores proteicos podem ser reciclados e usados para sintetizar outro polipeptídeo.

Modificações Pós-Traducionais:

Antes que um polipeptídeo recém-sintetizado possa começar a sua existência como proteína funcional, freqüentemente sofre outro processamento chamado de modificação pós-traducional. Estas modificações podem ter uma variedade de formas, incluindo a clivagem em unidades polipeptídicas menores, ou uma combinação com outros polipeptídeos para formar uma proteína maior. Outras modificações possíveis incluem a adição de cadeias laterais de carboidratos ao polipeptídeo. Estas modificações são necessárias, por exemplo, para produzir o dobramento apropriado da proteína final, ou para estabilizar sua estrutura. A cadeia polipeptídica que é o produto primário da tradução é dobrada e associada em uma estrutura tridimensional específica que é determinada pela própria seqüência de aminoácidos. Duas ou mais cadeias polipeptídicas, produtos do mesmo gene ou de genes diferentes, podem se combinar para formar um único complexo protéico final. Os produtos protéicos também podem ser modificados quimicamente por, por exemplo, adição de fosfato ou carboidratos em sítios específicos. Outras modificações podem envolver a clivagem da proteína, seja para remover seqüências amino-terminais específicas após terem direcionado uma proteína para o seu local correto dentro da célula, ou para dividir a molécula em cadeias polipeptídicas menores.

Nem todos os genes estão “funcionando” (se expressando) em todas as células. Alguns genes que codificam para polipeptídeos essenciais a todos os tecidos se expressam em todas as células do organismo, porém alguns genes, cujos polipeptídeos codificados têm uma função muito específica, só se expressam em determinado tecido, ou em determinados momentos da vida ou do ciclo circadiano.

O fato de todo o conteúdo de genes, o genoma de um organismo, estar presente em absolutamente todas as células, a despeito de estar se expressando ali ou não, permite que o resultado de uma análise do DNA seja o mesmo quando realizada em qualquer célula do organismo. Por exemplo, para verificarmos se um paciente tem uma alteração hereditária no gene da desidrogenase alcoólica, gene esse que se expressa preferencialmente nos hepatócitos, não é preciso realizar uma biópsia de fígado, para estudar a atividade enzimática. Basta analisarmos o gene que codifica a desidrogenase alcoólica presente em uma célula do sangue, ou da pele, ou da saliva, ou da urina, pois a constituição do DNA do gene da desidrogenase alcoólica dessas células será a mesma do DNA do hepatócito.

Mitocôndria

A mitocôndria contém moléculas circulares de DNA (DNAm). Assim como ela tem genes específicos, ela tem um código genético diferente. O DNAm é herdado matematicamente, porque a mitocôndria da célula espermática não entra no ovo fertilizado (embora estudos recentes estejam evidenciando a presença de mitocôndrias nos espermatozoides). O DNAm possui uma elevada velocidade de mutação (cerca de 10 vezes maior que o DNA), e acredita-se que possam contribuir para uma série de doenças humanas.

BIBLIOGRAFIA CONSULTADA

1. MULLER RF: Emery's – Elements of Medical Genetics, 11ed. 2001.
2. JORDE, CAREY, BAMSHAD: Genética Médica, 2ed. 1999.
3. Thompson e Thompson – Genética Médica, 6ed. 2002.
4. WILSON GN: Clinical Genetics A Short Course. 2000.
5. CARAKUSHANSKY G: Doenças Genéticas em Pediatria. 2001.
6. ALBERTS B; et all: Molecular Biology of The Cell, 4ed. 2002.