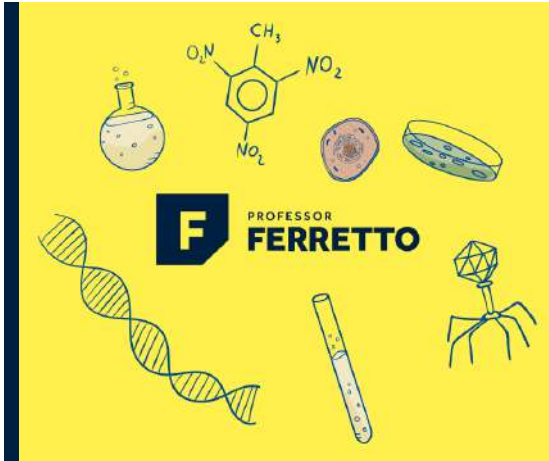


Biologia

PROFESSOR FLÁVIO LANDIM

ÁCIDOS NUCLEICOS PARTE II: TRANSCRIÇÃO E RNA



ASSUNTOS DA AULA.

Clique no assunto desejado e seja direcionado para o tema.

- [Genes e transcrição](#)
- [Descoberta do gene e conceito clássico de gene](#)
- [Unidade de transcrição gênica e transcrição](#)
- [DNA não codificante: "DNA lixo" ou Região Hipervariável de DNA](#)
- [Introns e Exons](#)
- [Splicing](#)
- [Diferenças entre as unidades de transcrição gênica em eucariontes e procariontes](#)
- [RNA](#)
- [Tipos de RNA](#)
- [DNA fita simples e RNA fita dupla em vírus](#)
- [Transcrição reversa em retrovírus](#)

Genes e transcrição

O DNA origina a partir de si o RNA, num processo denominado transcrição. O RNA, por sua vez, age ao nível dos ribossomos, estruturas que realizam a síntese proteica. Dessa maneira, o DNA desempenha mais dois importantes papéis:

- Controlando a síntese de RNA, ele controla a síntese proteica;

- Através da síntese proteica, há a produção de enzimas que controlarão todo o metabolismo celular; como o DNA controla a síntese de enzima, pode-se dizer que ele controla o metabolismo celular.

Em outras palavras, o DNA comanda e coordena todas as funções vitais, da reprodução ao metabolismo.

DESCOBERTA DO GENE E CONCEITO CLÁSSICO DE GENE

As bases da hereditariedade foram lançadas em 1866 pelo pesquisador austríaco Gregor Mendel, quando em seu celebre trabalho sobre ervilhas, ele afirmou que as características hereditárias eram determinadas por fatores segregantes, posteriormente conhecidos como fatores mendelianos, e daí como genes. (O termo gene só foi proposto pelo pesquisador dinamarquês Wilhem Ludwig Johannsen, em 1909.) De início, o gene mendeliano era abstrato, uma vez que não havia na época conhecimentos de Citologia e Bioquímica que pudessem embasar as ideias mendelianas de modo mais concreto. Juntamente com o fato de que Mendel vivia longe dos grandes centros científicos da segunda metade do século XIX, essa dificuldade de descrever concretamente o gene contribuiu para que suas pesquisas fossem completamente ignoradas pela comunidade científica daquela época.

Com os avanços da Biologia, sucessivas descobertas foram sendo feitas no sentido de levar a uma compreensão maior do significado dos genes. Em 1871, o médico suíço Friedrich Miescher isolou, pela primeira vez, a partir do núcleo da célula, uma substância por ele chamada de nucleína, e que mais tarde viria a ser conhecida como DNA. Em 1882, o biólogo alemão Walther Flemming descreveu a existência, no núcleo, de estruturas filamentosas denominadas cromossomos, e em 1883, o zoólogo belga Edouard Von Beneden descreveu seu comportamento segregante durante a meiose (ou seja, a separação dos cromossomos homólogos durante a divisão meiótica para a produção de gametas).

Em 1900, três pesquisadores trabalhando independentemente, o holandês Hugo De Vries, o belga Erick Tchesmack e o alemão Karl Correns, redescobriram os trabalhos de Mendel, e se iniciou a corrida pela identificação da localização celular dos genes e de sua natureza química.

Por volta de 1910, o biólogo norte-americano Thomas Morgan lança a Teoria Cromossômica da Herança, ao observar que o comportamento dos genes nas Leis de Mendel é exatamente igual ao comportamento dos cromossomos durante a meiose, sugerindo então que os genes estão dentro dos cromossomos.

Cromossomos eucariontes são estruturas constituídas por DNA e proteínas histônicas, e num primeiro momento, a maior parte da comunidade científica acreditou que os genes eram constituídos de histona, devido às proteínas serem formadas por 20 tipos de aminoácidos distintos, contra a simplicidade do DNA com seus 4 tipos de bases nitrogenadas. No entanto, em 1928, o microbiologista britânico Frederick Griffith, através de experimentos com bactérias da pneumonia, descobriu um princípio transformante que, incorporado por uma bactéria, dava a ela a capacidade de desenvolver novas características hereditárias. Esse princípio transformante era o DNA, e ele passou a ser visto pela ciência como o responsável pela determinação das características hereditárias, ou seja, os genes eram constituídos de DNA.

Mesmo antes da descoberta da natureza química dos genes, a ideia da maneira de funcionamento dos genes já havia sido proposta pelo médico inglês sir Archibald Garrod, que apresentou, em 1908, uma explicação para várias doenças caracterizadas pela incapacidade do corpo em realizar certas reações químicas. Ele sustentava que essa incapacidade era um defeito "inato", ou, em outras palavras, hereditário. Relacionou a falta de certas enzimas no homem com algumas anomalias do metabolismo, classificando-as como hereditárias, já que apareciam em proporções mendelianas nas famílias afetadas. Isso sugeria que os genes determinavam a produção de certas enzimas.

Na década de 1940, uma série de experiências realizadas pelos geneticistas norte-americanos George Beadle e Edward Tatum, com fungos do gênero *Neurospora*, confirmaram a predição de Garrod de que os genes controlavam a síntese de enzimas. Beadle e Tatum irradiavam fungos *Neurospora* com raios X, de modo a induzir mutações, e observaram que cada mutação levava o fungo mutante à impossibilidade de produzir certa enzima, confirmando a ideia de que os genes controlam a síntese de enzimas, o que é descrito pela conhecida "**Teoria 1 Gene, 1 Enzima**". Com a confirmação de que nem todas as proteínas eram

enzimas, essa Teoria foi modificada para “**1 Gene, 1 Proteína**”, e com o reconhecimento de proteínas de estrutura quaternária, formadas por mais de uma cadeia peptídica, para “**1 Gene, 1 Peptídio**”. Assim, como a hemoglobina é constituída por dois tipos de cadeia, alfa e beta, ela é uma proteína codificada por dois genes, um para cada tipo de cadeia peptídica (localizados, inclusive, em cromossomos distintos).

A definição clássica de gene afirma que **gene é um segmento de DNA com informação para a síntese de um polipeptídio ou de um RNA**. Assim, existem genes que codificam a produção de moléculas de RNA além do RNA mensageiro que é traduzido em peptídios, como o RNA ribossômico e o RNA transportador, os quais não são traduzidos em peptídios.

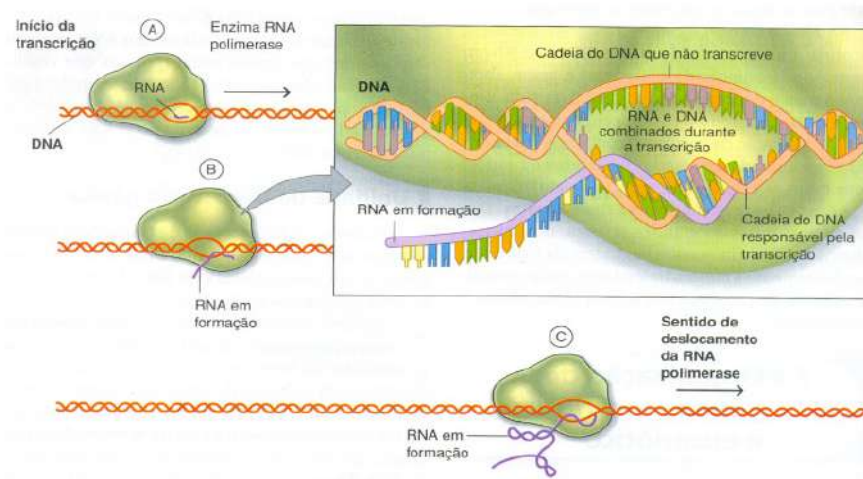
Tome nota:

UNIDADE DE TRANSCRIÇÃO GÊNICA E TRANSCRIÇÃO

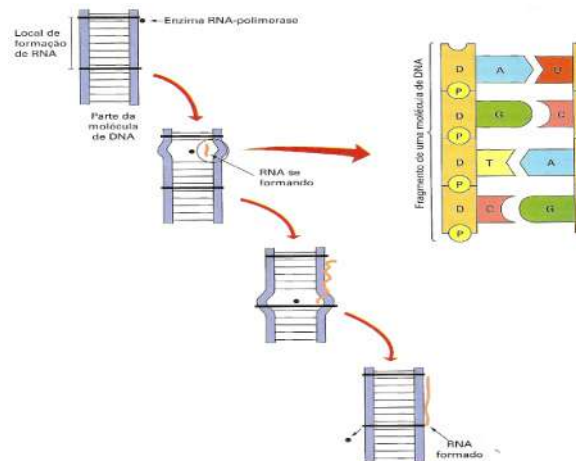
Uma consequência da melhor compreensão a respeito dos ácidos nucleicos foi o desenvolvimento do conceito de unidade de transcrição gênica. A **unidade de transcrição gênica** pode ser descrita como um segmento de DNA que é transcrito de forma contínua em uma molécula de RNA. Esse segmento de DNA apresenta uma sequência inicial de nucleotídeos denominada **região promotora (ou promotor)**, que corresponde ao sítio de ligação da enzima RNA polimerase, responsável pelo processo de transcrição, e uma sequência terminal de nucleotídeos denominada **sequência de término de transcrição**, que determina o desligamento da enzima RNA polimerase da molécula de DNA e o consequente término da transcrição.

O processo de transcrição de um gene se inicia com o encaixe da enzima RNA polimerase na região promotora do DNA. A partir daí, a RNA polimerase vai deslizando ao longo do DNA, promovendo a quebra das pontes de hidrogênio entre os nucleotídeos que ligam as duas fitas, e com isso, a abertura da dupla hélice e separação das duas fitas ao nível da unidade de transcrição. Ocorre então o pareamento de ribonucleotídeos (nucleotídeos de RNA), por pontes de hidrogênio, à fita molde do DNA equivalente ao gene, obedecendo à regra de pareamento entre guanina e citosina, citosina e guanina, timina e adenina, e adenina e uracila (uma vez que não há timina na molécula de RNA). É importante lembrar que, apesar de o DNA ser formado por duas fitas, o gene corresponde a apenas uma delas, sendo a fita complementar ao gene desprovida de função em termos de transcrição, servindo apenas para manter a estrutura em dupla hélice e consequentemente a estabilidade do DNA. A fita de DNA que corresponde ao gene é identificada pela região promotora onde a RNA polimerase se liga, não havendo promotor na fita complementar que não será transcrita.

O pareamento de novos nucleotídeos de RNA ao gene forma, de início, uma molécula híbrida, com uma fita de DNA molde e uma fita de RNA ligadas por pontes de hidrogênio. Ao chegar à sequência de término, a enzima RNA polimerase se solta do DNA e a transcrição termina, sendo o RNA liberado e as duas fitas de DNA voltando a se ligar por pontes de hidrogênio para retomarem a estrutura em dupla hélice. Devido à ação precisa da RNA polimerase, o RNA transcrito apresentará bases nitrogenadas rigorosamente complementar à da cadeia de DNA que serviu de modelo.



Representação da ação da RNA polimerase na transcrição gênica. No detalhe, a correspondência entre as seqüências de bases nitrogenadas do RNA transcrito e do DNA modelo.



Visão simplificada da transcrição.

DNA NÃO CODIFICANTE: “DNA LIXO” OU REGIÃO HIPERVARIÁVEL DE DNA

Em eucariontes, entre duas unidades de transcrição consecutivas, existem seqüências de **DNA não codificante**, ou seja, **seqüências de bases nitrogenadas cujas informações não são traduzidas em proteínas**. Devido à sua aparente falta de função, esses trechos de DNA foram descritos, de início, como “**DNA lixo**”. Esse termo tem caído em desuso devido a descobertas que esclarecem algumas possíveis utilidades para essas regiões, o que será descrito logo a seguir. Entretanto, o termo **região hipervariável** continua em uso pelo fato de que a seqüência de bases nitrogenadas nessas regiões variam enormemente de um indivíduo para outro, sendo obrigatoriamente distintas entre indivíduos distintos, com exceção de clones (como gêmeos univitelinos). Por isso, esses trechos são usados na comparação realizada nos exames de DNA, uma vez que suas seqüências exclusivas de bases nitrogenadas permitem a diferenciação genética entre dois indivíduos (novamente, com exceção de clones como gêmeos univitelinos).

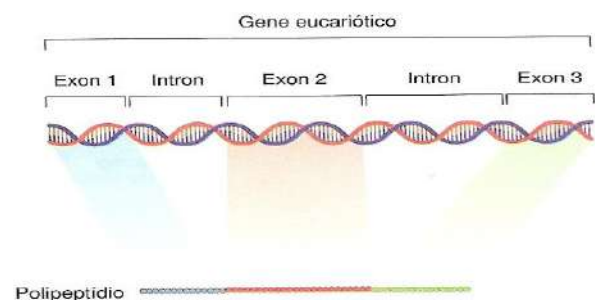
Genes determinam a sequência de aminoácidos de uma proteína, de modo que, proteínas idênticas numa espécie são codificadas por genes iguais, com a mesma sequência de nucleotídeos. Assim, o gene para a síntese de insulina em humanos, por exemplo, é igual em todos os indivíduos. Se o gene para insulina sofrer mutações, passará a produzir uma molécula de insulina com algum aminoácido alterado, ou seja, com sua estrutura primária alterada, e conseqüentemente todos os demais níveis de estrutura proteica alterados (uma vez que dependem da estrutura primária). Essa insulina alterada não deverá ser funcional, levando o portador dessa mutação à morte. Como genes codificam proteínas, a maioria das mutações nos genes leva à formação de proteínas defeituosas e possível morte, de modo que os genes são altamente conservados, ou seja, genes para uma mesma característica são idênticos mesmo entre indivíduos distintos. Já o “DNA lixo” não codifica proteínas, de modo que, caso sofra mutações, essas não acarretarão prejuízo algum ao indivíduo. Por isso, o “DNA lixo” pode acumular mutações ao longo das gerações, o que explica ser hipervariável.

A origem da região hipervariável de DNA está provavelmente em dois eventos. Parte dela deve ser proveniente do material genético de vírus que se incorporaram ao genoma celular ao realizar ciclo lisogênico e não passar ao ciclo lítico. Assim, esse material genético viral foi sendo replicado ao longo da evolução da célula e acumulando mutações que o descaracterizou. Outra parte da região hipervariável deve corresponder a genes fósseis ou pseudogenes, ou seja, genes que não têm função atualmente, mas que tinham função em espécies ancestrais. O sequenciamento do genoma humano (veja Projeto Genoma Humano, mais à frente) e do genoma do chimpanzé, nosso parente mais próximo dentre as espécies atuais, mostram genes humanos semelhantes a trechos de “DNA lixo” de chimpanzés, e genes de chimpanzés semelhantes a trechos de “DNA lixo” humanos. Explica-se: como a espécie humana e o chimpanzé guardam um ancestral comum, alguns genes desse ancestral se modificaram em DNA não codificante em chimpanzés, mas permaneceram como genes em humanos, e vice-versa.

Estima-se que algo em torno de 98,5% do DNA humano corresponda a DNA não codificante. Ou seja, apenas cerca de 1,5% do DNA humano corresponde aos genes. Tão estranho quanto o fato de que a maior parte do DNA não codifica genes, é o fato de que a maior parte do DNA não codificante se encontra não entre genes (unidades de transcrição) consecutivos, mas dentro dos genes, na forma de introns.

INTRONS E EXONS

Em 1978, o geneticista norte-americano Walter Gilbert sugeriu o termo introns (do inglês *intra-genic region*, região intragênica) para designar as regiões de DNA não codificante dentro dos genes (“DNA lixo” dentro dos genes), e o termo exons (do inglês *expressed region*, região expressa) para designar as regiões de DNA codificante dentro dos genes (regiões que são traduzidas em sequências de aminoácidos nas proteínas). Costuma-se, então, descrever os genes eucariontes como interrompidos, uma vez que ocorrem introns dentro dos genes.



Uma comparação clássica que ajuda a esclarecer o conceito de introns e exons dentro dos genes, diz respeito à comparação entre o gene e uma receita de bolo (o bolo, no caso, corresponde à proteína codificada pelo gene). Ao invés de a receita ser escrita de modo contínuo, ela apresenta, em determinados pontos do texto, uma série de interrupções correspondendo a informações sem sentido (letras, palavras ou frases sem significado). Para fazer o bolo, as informações sem sentido (introns) devem ser ignoradas pelo executor da receita (ribossomo, que sintetiza a proteína), que deve tomar o cuidado de utilizar apenas as informações pertinentes (exons).

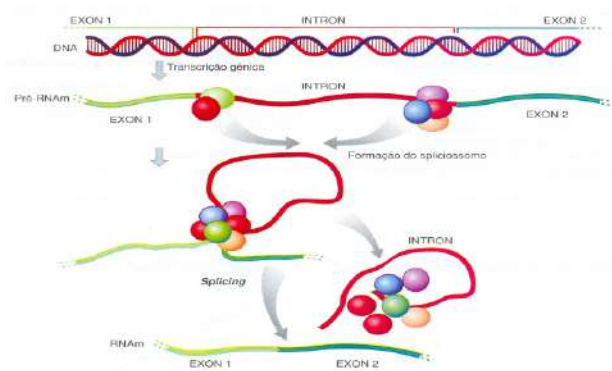
SPLICING

A enzima RNA polimerase, ao percorrer uma unidade de transcrição eucariótica, transcreve tanto os exons como os introns do gene, formando uma primeira molécula de RNA transcrito conhecida como **pré-RNA mensageiro** (ou RNA heterogêneo, devido à presença de introns). Se esse pré-RNA mensageiro for traduzido, o ribossomo organizará uma molécula com uma sequência de aminoácidos que resultará em uma proteína não funcional, uma vez que os introns apresentam informações sem sentido, que não guardam relação com a estrutura da proteína a ser produzida. Assim, os introns devem ser removidos do pré-RNA mensageiro.

A remoção dos introns do pré-RNA mensageiro ocorre através de um mecanismo conhecido como **splicing**, que ocorre ainda dentro do núcleo, e pode ser comparado ao processo de edição de um texto ou vídeo, com o corte dos trechos indesejados (introns) e a união dos trechos de interesse (exons). O termo *splicing*, inclusive, vem do inglês e poderia ser traduzido por "corte e emenda", ou seja, corte e remoção dos introns e emenda dos exons. A nova molécula de RNA gerada após o *splicing*, e contendo apenas exons, é chamada agora de RNA mensageiro.

O splicing, processo de corte e emenda do RNA pré-mensageiro, é realizado por um sistema denominado de **spliciossomo**, constituído de proteínas e um tipo especial de RNA conhecido como snRNA (do inglês *small nuclear RNA*, pequenos RNA nucleares). As partículas componentes do spliciossomo têm a habilidade de reconhecer as extremidades de um intron, se ligando então a elas. Em seguida, essas partículas spliciossômicas se unem, aproximando as extremi-

dades do intron e cortando a molécula de RNA pré-mensageiro nos limites entre o intron e os dois exons adjacentes, os quais são imediatamente unidos entre si. Como mencionado anteriormente, apenas com a remoção dos introns e a emenda dos exons do pré-RNA mensageiro, é que o RNA, agora na forma de RNA mensageiro, será traduzido. Para isso, o RNA mensageiro sai do núcleo para o citoplasma, onde se liga aos ribossomos para que ocorra a tradução da informação genética, resultando na síntese proteica.



Representação do mecanismo de splicing.

Tome nota:

Na maioria dos genes, a quantidade de introns equivale a mais da metade do gene e, conseqüentemente, do pré-RNA mensageiro. Para alguns genes, no entanto, os introns correspondem a bem mais do que a metade do gene e do pré-RNA mensageiro. Num exemplo bem conhecido, o gene que codifica a proteína distrofina, presente nos músculos, tem 2 milhões de nucleotídeos, de modo que seu pré-RNA mensageiro também apresenta 2 milhões de nucleotídeos. O *splicing*, então, remove 78 introns do pré-RNA mensageiro, equivalendo a mais de 1,9 milhão de nucleotídeos, resultando num RNA com apenas 14 mil nucleotídeos.

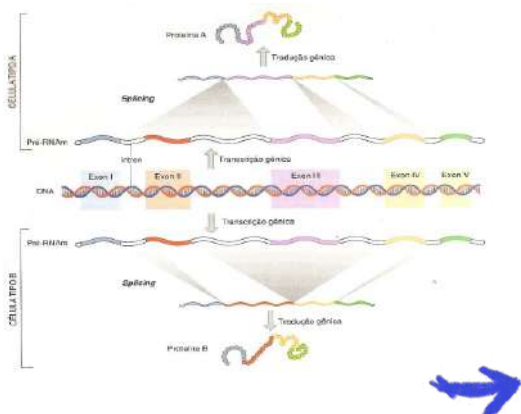
À primeira vista, todo esse trabalho envolvendo a transcrição de um pré-RNA mensageiro enorme, e posterior *splicing* para a produção de um RNA mensageiro bem menor, parece corresponder a uma enorme perda de tempo e de energia para a célula. Qual seria a vantagem que a presença de introns e do mecanismo de *splicing* traz para a célula? Algumas descobertas recentes apontam uma possível explicação de como um sistema tão complexo poderia resultar em benefícios para a célula.

SPLICING ALTERNATIVO

A espécie humana apresenta entre 100 e 120 mil tipos distintos de proteínas. De acordo com o modelo clássico que descreve o gene, a Teoria “1 Gene, 1 Peptídio”, deveria haver número idêntico de genes, ou mesmo maior, se for lembrado que algumas proteínas, como a hemoglobina, apresentam mais de uma cadeia peptídica codificadas por genes diferentes. Assim, até o início dos anos 2000, acreditava-se que a espécie humana apresentava em seu genoma entre 100 e 120 mil genes.

Em 2003, foi concluído o Projeto Genoma Humano, o qual de-terminou a sequência de bases nitrogenadas de todo o DNA humano, incluindo os 24 tipos de cromossomos humanos (22 autossomos e os cromossomos sexuais X e Y). Dos resultados advindos do sequenciamento do genoma humano, um dos que trouxe maior surpresa foi o de que o número de genes na espécie humana está estimado em apenas 25 mil genes. Esse dado colocou em xeque a Teoria “1 Gene, 1 Peptídio”: se cada gene codifica um peptídio, com 25 mil genes no genoma, deveria haver apenas 25 mil proteínas diferentes no corpo humano, e não 100 mil. Como os 25 mil genes poderiam então codificar 100 mil proteínas distintas?

Os estudos realizados para responder a esse questionamento levaram a uma descoberta surpreendente. Um mesmo gene é sempre transcrito em um mesmo pré-RNA mensageiro, mas cada célula pode realizar o *splicing* de modo distinto, removendo não apenas os introns, como também alguns exons, resultando em moléculas de RNA mensageiro distintas, num processo denominado de *splicing* alternativo. Assim, graças ao **splicing alternativo**, um mesmo gene pode ser traduzido em várias proteínas distintas, dependendo da maneira que a célula faz o *splicing*, ou seja, de qual exon ou quais exons são removidos do pré-RNA mensageiro junto aos introns. Observe no modelo abaixo:



Representação do mecanismo de *splicing* alternativo. Observe que nas células A e B o gene é transcrito em um mesmo pré-RNA mensageiro, incluindo os cinco exons (I, II, III, IV e V). No entanto, o *splicing* ocorreu de modo diferente, uma vez que na célula A, o exon II foi removido junto com os introns, estando esse trecho de informação ausente no RNA mensageiro e, portanto, não codificando aminoácidos na proteína. De modo semelhante, na célula B, o exon III foi removido junto com os introns, sendo o RNA mensageiro resultante do *splicing* diferente, por não conter este outro trecho de informação, e a proteína resultante diferente por não conter esta outra sequência de aminoácidos.

A descoberta do *splicing* alternativo, além de explicar como o pequeno número de genes humanos consegue produzir uma grande quantidade de tipos de proteínas, explica também a diferença de complexidade dos organismos vivos. Por exemplo, o primeiro organismo pluricelular a ter seu genoma sequenciado foi o verme nematelminto *Caenorhabditis elegans*, um animal microscópico extremamente simples, mas com cerca de 21 mil genes, contra os cerca de 25 mil genes humanos. A pequena diferença no número de genes não explicaria, a princípio, a enorme diferença de complexidade entre o *C. elegans* e a espécie humana. Entretanto, a habilidade das células humanas em fazer *splicing* alternativo, possibilita a produção de um número de tipos de proteína muito maior, o que justifica, então, a diferença de complexidade. Por exemplo, estima-se que mais de 60% dos genes humanos apresentam *splicing* alternativo, o que explica porque o número de tipos de proteínas humanas é tão superior ao número de genes.

Em algumas espécies, um número menor de genes pode codificar um número maior de proteínas devido a mecanismos de *splicing* alternativo, e o grau de complexidade dos organismos está relacionado, consequentemente, não ao número de genes, mas à habilidade de utilizar um mesmo gene na síntese de várias proteínas.

O *splicing* alternativo trouxe, entretanto, um problema em termos de definição do gene. O conceito clássico de gene, de acordo com a Teoria "1 Gene, 1 Peptídio", se mostrou inadequado com a possibilidade de um mesmo gene codificar várias proteínas. Assim, pode-se afirmar que um conceito atual de **gene** está ainda em construção, mas uma maneira moderna de descrever o gene é se referir a ele como **um trecho de DNA que pode ser transcrito numa molécula de RNA, seja ele um pré-RNA mensageiro, um RNA ribossômico ou um RNA transportador**. No caso do pré-RNA mensageiro, ele pode ser editado em RNA mensageiro pelo *splicing* e então traduzido em um peptídio.

LEITURA COMPLEMENTAR

Um conceito em apuros

Quinze anos atrás, o historiador e filósofo da Biologia Richard Burian observou: "Há um fato estabelecido sobre a estrutura do DNA, mas não há nenhum fato estabelecido sobre o que é o gene". Nesse ínterim, as coisas só pioraram. Como resultado do progresso espetacular na identificação, mapeamento e sequenciamento de determinados genes, aprendemos uma quantidade imensa de fatos sobre a estrutura e a função do material genético, e muito do que aprendemos fica fora da moldura do nosso quadro original. As complicações criadas pelos novos dados são imensas. Tomadas em seu conjunto, elas ameaçam colocar o próprio conceito de "o gene" - quer como uma unidade de estrutura, quer como unidade de função, em flagrante desmantelamento.

Técnicas e dados da análise do sequenciamento levaram à identificação não apenas de genes interrompidos, mas também de genes repetidos, genes superpostos, DNA críptico, transcrição reversa, genes nidados e promotores múltiplos (permitindo que a transcrição se inicie em locais alternativos e de acordo com critérios variáveis). Todas essas variações confundem de forma desmedida a tarefa de definir o gene como uma unidade estrutural.

Similarmente, a descoberta de um elaborado processo de edição ao qual o transcrito primário RNA pré-mensageiro é submetido, de mecanismos regulatórios operando ao nível da síntese de proteínas, e outros ainda operando no nível da função de proteínas, confunde nossos esforços para dar ao gene uma definição funcional precisa. Como Peter Portin observa: "Nosso conhecimento da estrutura e função do material genético extrapolou a terminologia tradicionalmente usada para descrevê-lo. Pode-se argumentar que o antigo termo gene, essencial em uma etapa inicial da análise, não seja mais útil". William Gelbart, trabalhando na vanguarda da genética molecular, concorda ao sugerir que o gene pode ser "um conceito

cuja época passou". "Ao contrário dos cromossomos", escreve Gelbart, "os genes não são objetos físicos mas meramente conceitos que adquiriram muita bagagem histórica nas últimas décadas." Claro que o conceito de gene desempenhou um papel crucial em nos trazer ao conhecimento atual dos fenômenos biológicos, mas hoje, ele sugere, "podemos bem ter chegado ao ponto em que o uso do termo 'gene' pode de fato ser um obstáculo ao nosso entendimento".

Existe mais que um pouco de ironia no atual estado das coisas, pois nunca na história do gene o termo teve tanta proeminência, tanto na imprensa científica quanto na leiga. Diariamente somos informados da identificação de novos genes "causadores de doenças", e a lista das doenças "genéticas" correspondentes cresce a cada dia. Da mesma forma, dizem-nos que muito da conduta humana, antes imaginada como voluntária, ou culturalmente induzida, é um produto de nossos genes. Certamente fez-se um progresso espantoso na compreensão da importância de mutações genéticas na incidência de muitas doenças (incluindo alguns distúrbios da conduta). Um certo número de condições mórbidas foi definitivamente associado a mutações em genes específicos. Os casos mais simples e claros são os distúrbios em genes únicos (Tay-Sachs, doença de Huntington, fibrose cística, talassemia e fenilcetonúria [PKU], entre outras). Tais exemplos, no entanto, continuam raros, e mesmo nesses casos bem demarcados, muito ainda precisa ser compreendido sobre o processo que liga o gene defeituoso ao desencadeamento da doença.

Em condições que sabidamente envolvem a participação de muitos genes (tais como certos tipos de doenças cardíacas, acidentes vasculares cerebrais, psicoses, diabetes), as limitações do entendimento atual são ainda mais notáveis. O efeito final é que, muito embora tenhamos nos tornado extraordinariamente eficientes em identificar riscos genéticos, a possibilidade de benefícios médicos significativos – benefícios que há apenas uma década esperava-se que surgissem logo depois, nos calcanhares das novas técnicas de diagnóstico – recua para um futuro cada vez mais distante. Como escreve D. J. Weatherall, diretor do Instituto de Medicina Molecular da Universidade de Oxford: "Transferir genes para um novo ambiente e seduzi-los a fazer seu trabalho, com todos os mecanismos regulatórios sofisticados que estão envolvidos, tem-se mostrado, até agora, uma tarefa difícil demais para os geneticistas moleculares". Parte da dificuldade, é claro, está em compreender o que é que os genes fazem.

Em outras palavras, por trás da chamada disparidade terapêutica entre diagnóstico genético e benefícios médicos, está a complexidade da dinâmica regulatória, que agora coloca o próprio conceito de gene em risco. Na realidade, "ironia" pode ser uma palavra fraca demais para descrever as incongruências de nossa situação atual. Pois o fato básico é que, no mesmo momento em que o discurso sobre os genes passou a dominar tão poderosamente nosso discurso biológico, as façanhas das novas técnicas analíticas na biologia molecular e todo o peso das descobertas que elas propiciaram, trouxeram o conceito do gene à beira do colapso. O que é um gene hoje? Quando ouvimos os modos pelos quais o termo é hoje usado pelos biólogos em atividade, descobrimos que o gene se tornou muitas coisas – não mais uma única entidade, mas uma palavra de grande plasticidade, definida somente pelo contexto experimental específico no qual é utilizada.

Extraído de Amabis & Martho, Biologia das Populações

Tome nota:

DIFERENÇAS ENTRE AS UNIDADES DE TRANSCRIÇÃO GÊNICA EM EUKARIOTES E PROCARIOTES

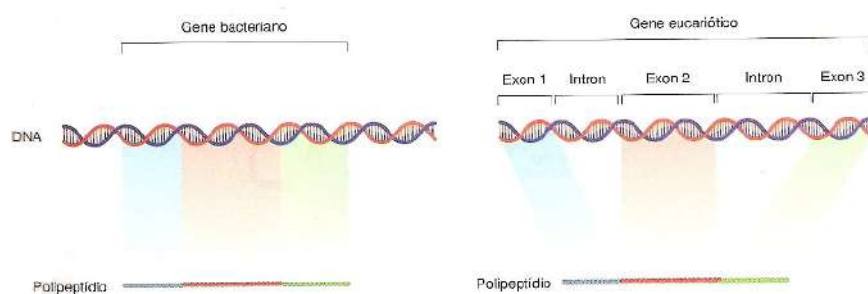
Em **células eucarióticas**, as unidades de transcrição gênica são intercaladas por trechos de “DNA lixo” não codificante, o qual não é transcrito nem traduzido. Cada unidade de transcrição gênica contém um único gene, sendo transcrito em uma única molécula de RNA. Assim, como o RNA corresponde ao transcrito de um único gene (ou cistron), esse **RNA** é dito **monocistrônico**.

Além disso, as unidades de transcrição gênica eucariotes apresentam trechos de “DNA lixo” não codificante dentro do gene, ou seja, **introns**, os quais são transcritos no pré-RNA mensageiro e removidos no **splicing**, estando ausentes no RNA mensageiro traduzido em peptídios.

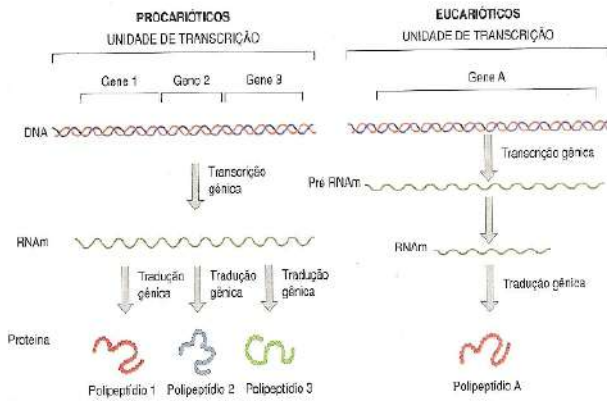
Em **células procarióticas**, ou seja, bacterianas, as unidades de transcrição gênica não são intercaladas por trechos de “DNA lixo”, sendo contínuas. Cada unidade de transcrição gênica contém vários genes justapostos, de modo que o RNA transcrito corresponde à informação de vários genes, sendo esse RNA dito policistrônico. O RNA mensageiro policistrônico é então traduzido em vários peptídios.

Dá-se o nome de **operon** ao conjunto de genes localizados na unidade de transcrição gênica procariótica e transcritos num mesmo RNA mensageiro policistrônico. O operon inclui a região promotora, que corresponde ao sítio de ligação da RNA polimerase, e a sequência de término, que corresponde ao sítio de desligamento da RNA polimerase.

O exemplo mais conhecido de operon corresponde ao operon da lactose em bactérias *Escherichia coli*, em que a unidade de transcrição gênica contém três genes que codificam enzimas relacionadas ao metabolismo da lactose, a beta-galactosidase, a galactosídeo-permease e a acetilase. Os três genes são transcritos em uma única molécula de RNAm policistrônico, e os ribossomos traduzem esse RNAm simultaneamente em três trechos para sintetizar os três peptídios.



Representação da organização de um gene não-interrompido de uma bactéria, sem introns, e de um gene interrompido de organismo eucariótico, com introns e exons.



À esquerda, representação de uma unidade de transcrição gênica em procariontes, com genes colineares transcritos em um RNA mensageiro policistrônico e traduzidos em três peptídios distintos. À direita, representação de uma unidade de transcrição gênica em eucariontes, com um único gene transcrito em um RNA mensageiro monocistrônico, o pré-RNA mensageiro, que passa pelo splicing e gera o RNA mensageiro definitivo, que é traduzido em um único peptídio.

É importante observar que, em procariontes, a sequência de aminoácidos de um peptídio corresponde exatamente à sequência de bases do segmento de DNA que foi transcrito para o RNAm. Por isso, pode-se dizer que, em bactérias, há colinearidade entre as cadeias polipeptídicas e os segmentos de DNA que as codificam.

RESUMO

Procariontes	Eucariontes
Há vários genes justapostos, sem trechos de DNA não codificante entre eles, formando unidades denominados operons	Sempre há trechos de DNA não codificante separando os genes
Apresentam RNAm policistrônicos , ou seja, vários genes são transcritos numa única molécula de RNAm, que então codifica vários peptídios simultaneamente	Apresentam RNAm monocistrônicos , ou seja, cada RNAm é produto da transcrição de um único gene
Não apresentam introns em seus genes, de modo que não há necessidade de ocorrer splicing	Apresentam introns em seus genes, de modo que há necessidade de ocorrer splicing

Ocorre DNA não codificante em procariontes apenas em alguns trechos entre operons, e em pequena quantidade. Já em eucariontes, o DNA não codificante entre genes e dentro dos genes (introns), somados perfazem cerca de 98,5% do material genético.

Alguns poucos organismos eucariontes apresentam unidades de transcrição gênica com genes justapostos e RNA mensageiros policistrônicos, como ocorre com alguns vermes nematódeos como o *Caenorhabditis elegans*.

LEITURA COMPLEMENTAR

Operon

Num organismo pluricelular, todas as células somáticas são formadas por meio da divisão celular mitótica. Isso quer dizer que todas as células contêm todos os cromossomos e, portanto, todos os genes. É evidente que nem todos os genes funcionam em todas as células: células da pele e do intestino têm necessidades distintas; diversas proteínas são fabricadas nelas, e por ação de genes diferentes. Em ocasiões posteriores, trataremos de como e porque ocorre a diferenciação celular.

Em unicelulares, como as bactérias, há processos de regulação de quais genes atuam num dado momento, e esse mecanismo é conhecido dos cientistas em modelos ainda hipotéticos, mas como forte evidência de serem os mecanismos corretos. Em 1961, os franceses Jacob e Monod propuseram o mecanismo que acreditamos ser o correto.

Em 1950, esses pesquisadores descobriram que a bactéria *Escherichia coli* não sintetiza β -galactosidase (enzima que hidrolisa a lactose) na ausência de lactose. Porém, quando esse açúcar era fornecido à célula, logo se iniciava a produção da enzima. Aparentemente, o gene para a produção da enzima é "ligado" pela presença da lactose e "desligado" pela ausência da mesma. A lactose, dessa forma, parece induzir a produção da enzima.

A partir desse dado, eles perceberam que provavelmente existe algum mecanismo de feedback negativo intracelular: os genes deviam responder à estimulação "ambiental" de substâncias presentes no citoplasma.

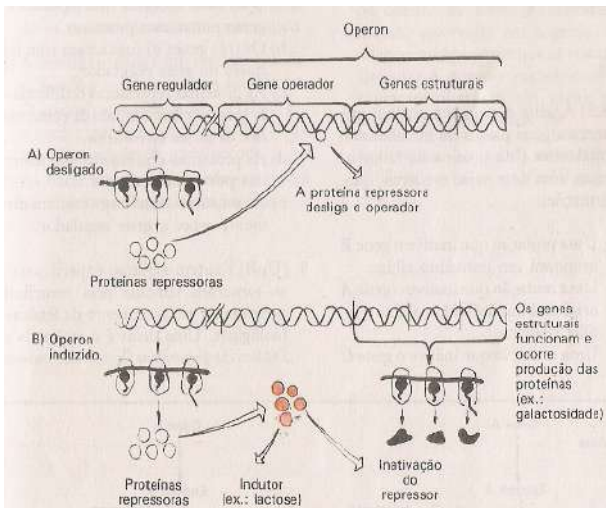
Jacob e Monod elaboraram então um modelo de regulação que, resumidamente, consiste no seguinte: há um grupo de genes alinhados ao longo do mesmo pedaço de DNA; esse grupo é chamado de **operon**. No operon da lactose, existe uma série de três **genes estruturais**, que contêm informação para a síntese de três proteínas: a β -galactosidase e mais duas outras enzimas envolvidas na absorção e utilização da lactose. Os três genes estruturais são precedidos por dois outros genes, o **gene promotor**, primeiro gene do operon, no qual se liga a enzima RNA polimerase para transcrever os genes estruturais, e um **gene operador**, entre o gene promotor e os estruturais. O gene operador se liga a substâncias chamadas **repressoras**, que impedem que a RNA polimerase se ligue ao gene promotor e transcreva os genes estruturais.

O repressor, que impede a transcrição dos genes estruturais, é produzido por outro gene, chamado gene regulador, que não faz parte do operon.

Na ausência de lactose, o regulador produz o repressor, não ocorrendo transcrição. Porém, quando há lactose na célula, esta provavelmente se liga ao repressor, modificando a forma de sua molécula. O repressor modificado fica inativo e permite a passagem da RNA polimerase, pois não consegue mais se combinar com o gene operador. Isso permite a síntese da β -galactosidase e das demais proteínas dos genes reguladores. (Interessante é que os três genes são transcritos num único RNAm, dito policistrônico, uma vez que codificou vários cístrons, formando as três proteínas pregadas, só sendo separadas após a saída das três do ribossomo.

Esse modelo do operon explica o papel da lactose como "indutor" da produção de enzimas com ela relacionadas. É óbvio que depois de digerida a lactose, o repressor torna a bloquear a passagem da RNA polimerase e a transcrição dos genes estruturais.

Deve ficar claro que, embora extremamente engenhosa, essa explicação não deixa de ser um modelo hipotético. Esse modelo tem como base observações em células bacterianas. Há pouca evidência que em células eucarióticas o mecanismo seja o mesmo, uma vez que há muito mais genes, influenciados por outras substâncias, como hormônios, e ainda há vários genes inativos de acordo como o mecanismo de diferenciação.



Operon da lactose.

RNA

A estrutura primária do **RNA** ou **ARN** é semelhante à do DNA, exceto pela substituição da desoxirribose pela ribose e da timina pela uracila, como foi anteriormente discutido. A estrutura espacial difere, de modo que o RNA é formado por apenas uma cadeia de polinucleotídeos.

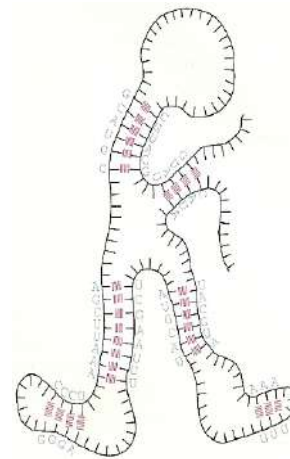
Apesar da molécula de RNA possuir só uma tira, o RNA não é uma estrutura linear simples e lisa. As moléculas de RNA possuem extensas regiões de complementação, nas quais as pontes de hidrogênio entre os pares AU e GC são formadas unindo diferentes porções da mesma molécula. Como resultado disso, a molécula dobra-se sobre si mesma, formando estruturas denominadas alças. Nestes casos, a molécula de RNA adota uma estrutura helicoidal comparável à do DNA nas áreas em que há pareamento de bases. A estrutura compacta das moléculas de RNA dobradas sobre si mesmas tem importantes consequências biológicas, como na estrutura em "folha de trevo" do RNA transportador.

TIPOS DE RNA

O RNA desempenha um papel imprescindível no processo de síntese proteica. Para isso, ele assume três diferentes formas: **RNA ribossômico**, **RNA mensageiro** e **RNA transportador (ou transferidor)**.

RNA RIBOSSÔMICO (RNAr)

Esta classe de moléculas de RNA compõe as subunidades ribossômicas juntamente com proteínas. Como os ribossomos são abundantes na célula, o RNAr corresponde a cerca de 80% do RNA presentes na célula. Apesar de apresentar uma única fita, apresenta estrutura tridimensional pelo dobramento proporcionado pelo pareamento de bases dentro da própria fita.



Estrutura tridimensional do RNA, com pareamento de bases entre uracila e adenina e entre guanina e citosina, na mesma fita.

O RNAr é sintetizado no núcleo, como todos os RNA, mas logo é enviado ao nucléolo para participar da síntese dos ribossomos. Uma vez montados os ribossomos, estes são enviados para o citoplasma. Seu papel parece "ser a ligação do RNAm ao ribossomo no processo de síntese". Também age como ribozima.

Moléculas de RNA denominadas **ribozimas** atuam como catalisadores em várias reações químicas, inclusive em seres humanos. Sabe-se hoje que a molécula que catalisa a formação de uma ligação peptídica nos ribossomos, durante o processo de síntese proteica, é uma ribozima, e não uma enzima.

Existem vários tipos de RNAr, dos quais se destaca aquele cuja massa molecular está em torno de 600000 D, e é o principal componente da subunidade menor do ribossomo e aquele cuja massa molecular é de cerca de 1200000 D e é o principal componente da subunidade maior do ribossomo.

RNA MENSAGEIRO (RNAm)

Esta é a classe das moléculas de RNA que codificam a sequência de aminoácidos de uma proteína em sua sequência de nucleotídeos, servindo como base para a síntese proteica.

Ele também é sintetizado no núcleo, mas pode ser frequentemente encontrado no hialoplasma, para efetivar o processo de síntese proteica. Seu peso molecular depende do tamanho de seu filamento e, conseqüentemente, do número de tamanho da cadeia da proteína a ser sintetizada. Gira em torno de 5×10^4 a 5×10^{16} D. O RNAm representa de 5 a 10% do total do RNA celular.

A estrutura do RNAm é filamentar simples, sem dobramento espacial. O RNAm é formado por várias sequências de três bases nitrogenadas, às quais chamamos códon. Cada códon é responsável pela codificação de determinado aminoácido. Cada códon codifica um único aminoácido, apesar de um mesmo aminoácido poder ser codificado por mais de um códon. Assim, normalmente, existe mais de um códon específico para cada aminoácido.

O RNAm encontra-se ocasionalmente ligado no hialoplasma a ribossomos. Ribossomos isolados são inativos no que diz respeito ao processo de síntese proteica. A estrutura responsável pela síntese proteica é na verdade o **polissomo** ou **polirribossomo**, formado por muitos (cerca de 60 a 80) ribossomos unidos entre si por RNAm. Isto permite que várias proteínas sejam sintetizadas simultaneamente a partir de um único RNAm. O número de ribossomos no polissomo depende do comprimento do RNAm. Uma outra forma ativa do ribossomo acontece quando este se encontra ligado à parede do retículo endoplasmático rugoso.

Tome nota:

RNA TRANSFERIDOR OU TRANSPORTADOR (RNAt)

Este tipo de RNA identifica os aminoácidos no citoplasma e os transporta até os polissomos para participarem da síntese proteica.

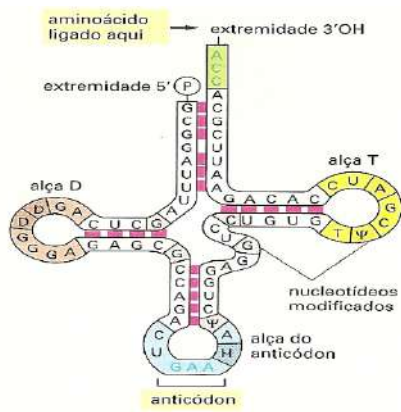
Os RNAt constituem um grupo de pequenos RNA (entre 75 e 85 nucleotídeos) que possuem a importante função de atuar como adaptadores moleculares durante a síntese proteica. Devido aos 20 aminoácidos apresentarem um formato não complementar em qualquer aspecto aos trios de nucleotídeos do RNAm, eles não são capazes de reconhecer os códon por si mesmos. O RNAt possui um trio de nucleotídeos denominado anticódon, que pode estabelecer pontes de hidrogênio com o códon do RNAm, desde que eles sejam correspondentes; ele também pode apresentar o aminoácido correspondente àquele códon em particular ligado a uma de suas extremidades. Assim, o RNAt tem, para determinado códon, um anticódon e o aminoácido correspondente àquele do códon ligado a si. A estrutura composta do aminoácido ligado ao RNAt é chamada **aminoacil-RNAt**, e permite que um aminoácido em particular seja trazido ao ribossomo em resposta ao códon apropriado.

Será posteriormente discutido que a síntese de proteínas depende da ligação do aminoácido correto à molécula do RNAt. Esta importante atividade é exercida por enzimas específicas denominadas

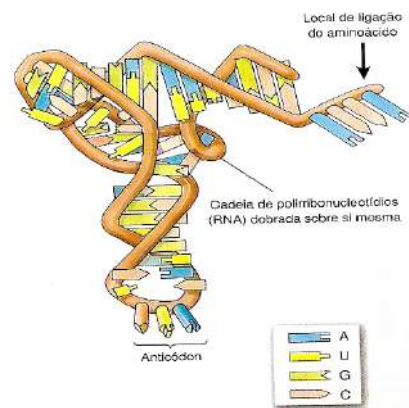
aminoacil-RNAt sintetases ou enzimas de ativação. Se um RNAt for mal colocado, haverá incorporação de um aminoácido incorreto à cadeia de proteína, alterando-lhe a estrutura primária.

Todos os RNAt partilham características comuns, das quais a mais notável é o pregueamento que forma a estrutura espacial desse tipo de RNA: a "**folha de trevo**". Outras características comuns incluem a **sequência CCA** na **extremidade aceptora ou livre**, que estabelece a ligação covalente com o aminoácido correspondente àquele determinado pelo anticódon, e a presença de bases nitrogenadas não presentes em nenhuma outra forma de RNA, fixos em determinadas posições da folha de trevo, como a pseudo-uridina, o ácido inosínico, a metilcitosina, a metilguanina, a ribotimidina e outros. Estas bases não usuais aparecem por modificação das bases normais após a transcrição, por ação de enzimas de transcrição e outras.

Como existem 20 aminoácidos participando de proteínas, é importante que haja um mecanismo de reconhecimento de qual aminoácido está sendo transportado. Este mecanismo equivale à **alça do anticódon**, localizada na extremidade oposta à aceptora e que possui as três bases que reconhecem e estabelecem as pontes de hidrogênio com o códon do RNAm, para garantir a precisão da síntese proteica, ou seja, a colocação do aminoácido correto na posição correta na cadeia polipeptídica.



Estrutura plana do RNAt em "folha de trevo".



Estrutura espacial do RNAt.

MicroRNA

Mais recentemente, novos tipos de RNA têm sido descobertos, como os **microRNA**. Esses microRNA são complementares a determinados trechos de RNAm e se ligam a eles formando fitas duplas de RNA, que não podem ser traduzidas pelos ribossomos, impedindo a expressão do gene que codificou o RNAm. Esse mecanismo é conhecido como silenciamento de RNA e auxilia na regulação da atividade gênica.

QUADRO RESUMO DE DIFERENÇAS ENTRE DNA E RNA

	DNA	RNA
Pentose	Desoxirribose	ribose
Base nitrogenada	adenina, citosina, guanina e timina	adenina, citosina, guanina e uracila
Número de cadeias	2 (bicatenária)	1 (monocatenária)

DNA FITA SIMPLES E RNA FITA DUPLA EM VÍRUS

Em vírus, as regras são às vezes diferentes. Podem ocorrer vírus com DNA fita simples e vírus com RNA fita dupla. Como reconhecê-los? Num vírus de DNA fita simples, não necessariamente se obedece à relação de Chargaff, e pode ser que o teor de A não seja igual ao de T, bem como o teor de G não seja igual ao de C. Já num vírus de RNA fita dupla, a relação de Chargaff é obedecida com U no lugar de T, e assim o teor de A é igual ao de U e o teor de G é igual ao teor de C.

Vírus de DNA ou RNA fita dupla	Vírus de DNA ou RNA fita simples
A = T ou U; G = C	A ≠ T ou U; G ≠ C

TRANSCRIÇÃO REVERSA EM RETROVÍRUS

Na natureza, a informação gênica caminha do DNA para o RNA.

Existe uma exceção a este fluxo unidirecional da informação genética, onde, o RNA pode ser algumas vezes copiado ou transcrito em DNA. Isto acontece com certos vírus que têm genoma de RNA (**retrovírus**, como o vírus HIV da AIDS). Este RNA atua como molde para a síntese de DNA viral promovida pela **enzima transcriptase reversa** do RNA, uma DNA polimerase RNA dependente, que tem a capacidade de, a partir de uma única cadeia de RNA, sintetizar uma molécula completa de DNA de cadeia dupla e complementar ao RNA original.

O tratamento para AIDS hoje se baseia num conjunto de medicamentos denominado de **coquetel anti-HIV** ou **terapia antirretroviral**. O coquetel é composto por três categorias de drogas, sendo a mais importante a categoria dos inibidores da transcriptase reversa. Estes foram as primeiras drogas desenvolvidas, como o AZT e o DDC. Eles combatem o vírus HIV inibindo a transcriptase reversa e impedindo a formação de DNA viral, para que a célula seja parasitada e o vírus se reproduza.

Perceba que os **inibidores de transcriptase reversa** não matam o vírus HIV, mas impedem sua reprodução, evitando a proliferação do mesmo e, assim, aumentando o período assintomático da doença e retardando o surgimento dos sintomas de imunodepressão.

Tome nota: