

# MECANISMOS DE EXPRESSÃO GÊNICA EM EUKARIOTOS

Emerson A. Castilho Martins<sup>\*</sup>, Paulo Roberto Maciel Filho

*Departamento de Fisiologia, Instituto de Biociências, USP*

*Apoio: FAPESP; CAPES; CNPq*

*Recebido 19out09 / Aceito 14jan10 / Publicação inicial 15abr10*

*martinseac@gmail.com*

**Resumo.** Com raras exceções, as células de eucariotos pluricelulares apresentam o mesmo material genético. Porém, com o decorrer das fases de desenvolvimento do organismo ou em diferentes tecidos, as exigências metabólicas são diferenciadas, e diferentes genes são ligados e desligados, expressando um conjunto distinto de proteínas. Existem vários mecanismos responsáveis por controlar a ativação e desativação de genes (controle da expressão gênica), em diferentes momentos da vida celular. Apresentamos nesta revisão alguns passos que são passíveis de controle, bem como uma breve descrição e exemplos ilustrativos de mecanismos de regulação de expressão gênica.

**Palavras-chave.** Controle de expressão, regulação gênica, adaptação fisiológica

## MECHANISMS OF GENE EXPRESSION IN EUKARYOTES

**Abstract.** With rare exceptions, all cells of multicellular eukaryotes have the same genetic material. However, metabolic requirements are different over the stages of their development or in different tissues. These requirements are satisfied by gene expression control in these organisms. In the present review we discuss some steps that are likely to be controlled, and a brief description and examples of mechanisms of gene expression.

**Keywords.** Expression control, gene regulation, physiological adaptation

### Introdução

Basicamente, expressão gênica pode ser descrita como o conjunto de processos que ocorrem para que um organismo, tecido ou célula inicie, aumente, diminua ou cesse a produção de produtos finais de seus genes, proteínas e/ou RNAs. No início dos anos 60, o grupo de Nirenberg começou a decifrar o código genético, permitindo correlacionar a sequência de aminoácidos das moléculas proteicas com as sequências de nucleotídeos do RNA mensageiro (mRNA), que por sua vez é determinada pela sequência dos nucleotídeos no DNA. No entanto, os mecanismos e estímulos que levam à produção das diferentes proteínas e RNAs, em diferentes células, são tão importantes quanto a própria sequência codificadora.

Com raras exceções, o material genético dos eucariotos pluricelulares é idêntico em todas as células. Isso decorre do fato de todas serem originadas da única célula-ovo. Os mecanismos de controle de expressão devem, então, começar por aí: as exigências das células, durante seu desenvolvimento, mudam. Consequentemente, os produtos de seu metabolismo devem responder a essas mudanças. Além disso, as células se diferenciarão em diferentes tecidos, que responderão a diferentes estímulos e apresentarão atividades específicas. Com isso, é possível notar quão importante é a regulação gênica.

A molécula de DNA é um polímero de nucleotídeos, unidades essas constituídas de uma molécula de açúcar (desoxirribose), um grupo fosfato e uma base nitrogenada, que pode ser adenina, guanina, timina ou citosina. A molécula é formada por uma fita dupla em hélice, unidas por pontes de hidrogênio entre as bases

nitrogenadas (Watson e Crick, 1953). Quando analisamos a sequência de bases nitrogenadas de uma fita, encontramos regiões que determinarão a sequência de aminoácidos. Essa região é, portanto, codificante. Como não se conhece produtos (RNA ou proteínas) de uma parte do genoma, se poderia pensar que essas não codificam nada, não têm função. Sabemos, no entanto, que não é o que acontece. Essas sequências são importantes porque nelas encontramos sinais pelos quais é possível regular a expressão, uma vez que apresentam sequências específicas para interação de proteínas que fazem a transcrição (no DNA) ou tradução (no mRNA). Há ainda espaçadores onde a sequência em si não é importante, mas seu tamanho torna acessível as regiões a serem expressas.

### A organização nuclear

O núcleo apresenta um domínio cromossômico com alta concentração de genes, chamado domínio central, e uma região chamada domínio periférico, com concentração menor de genes. Esse tipo de organização, contudo, não é estática, já que há reorganização dos cromossomos nas células de acordo com o nível de expressão dos seus genes (Verschure, 2004). Dessa forma, uma maneira de se controlar a própria expressão seria endereçar o cromossomo em regiões onde seus genes poderiam ser mais ou menos expressos (Singer e Green, 1997; Elias e col., 2002)

### Aumentando o número de genes

Um mecanismo que garante o aumento da expressão de um determinado gene é aumentar o número de suas cópias no genoma. Dessa forma,

maior quantidade de DNA molde, ao sofrer transcrição, gera uma quantidade maior de mRNA e, conseqüentemente, da respectiva proteína. Em alguns raros casos, acontece a amplificação gênica: a dupla fita de DNA se abre no gene e em seguida, a DNA polimerase sintetiza as fitas complementares, resultando então em duas duplas-fitas de DNA no local do gene. O processo pode se repetir, obtendo-se o crescimento exponencial do número de cópias do mesmo: de 2 para 4, 8 e assim por diante. A amplificação acontece somente naquelas células e não será transmitida para as gerações futuras. Alguns genes de produção de células foliculares de *Drosophila* são um exemplo desse tipo de estratégia (Claycomb e col., 2004).

Outro mecanismo que garante o aumento do número de genes é chamado de duplicação gênica. A sequência do gene é incorporada novamente ao genoma ou através de recombinação desigual, na qual o pareamento não é completamente homólogo, fazendo com que parte do cromossomo seja duplicada, ou através da ação de elementos transponíveis ou retrotransponíveis, que podem levar consigo parte da informação do hospedeiro e incorporá-lo em outra parte do genoma. Com isso, a cópia do gene poderá ficar incorporada ao genoma definitivamente. Quando mantida, essa cópia pode mutar, ou manter a estrutura original e ser transmitida aos descendentes. Apresenta a vantagem de conferir maior plasticidade ao organismo. Uma das cópias do gene se mantém sem alterações funcionais enquanto a outra pode sofrer mutações que podem levar ao aparecimento de novas formas da proteína com diferenças funcionais vantajosas ao organismo. Um caso de sucesso é observado em peixes antárticos da espécie *Trematomus bernacchii*, que apresentam três cópias do gene de pepsina A, decorrentes de dois eventos de duplicação. Esses genes passaram por mutações que geraram três formas diferentes de pepsina, proporcionando melhor funcionamento da digestão dos peixes na temperatura baixa em que vivem (Carginale e col., 2004).

### **A atividade das polimerases**

A transcrição dos genes é catalisada por enzimas chamadas RNA polimerases. Essas enzimas reconhecem sequências no DNA às quais se ligam, chamadas de regiões promotoras, ou simplesmente promotores. No entanto, diferentes complexos proteicos podem estar ligados a esses promotores, interferindo indiretamente na expressão de genes.

Os genes que codificam o RNA ribossômico (rRNA), o RNA que constitui o ribossomo, organela responsável pela tradução, são transcritos pela enzima RNA polimerase I, comumente chamada de RNA polI. Essa enzima apresenta uma série de fatores dos quais seu funcionamento depende, e a presença desses

conhecidamente altera a sua atividade. Além disso, alguns desses fatores são capazes de reconhecer sequências à montante do sítio de ligação da enzima - região conhecida como Espaço Intergênico (IGS) - fazendo com que a eficiência da atividade enzimática aumente na presença de sequência específica na região não transcrita anterior ao sítio promotor (de Andrade Stempljuk e Floeter-Winter, 2002).

Sabe-se ainda que essa enzima apresenta associação com as proteínas motoras actina e miosina nucleares, o que interfere no funcionamento da mesma. A miosina nuclear participa do início da transcrição, enquanto a actina participa do alongamento do transcrito. No entanto, aparentemente o processo não depende de atividades motoras dessas proteínas, já que se consegue observar a importância das mesmas tanto *in vitro* quanto *in vivo* (Philimonenko e col., 2004).

Os mRNA são transcritos pela RNA polimerase II, chamada de RNA polII. Essa enzima também apresenta fatores ligados a ela que permitem sua atividade. Além disso, é fato conhecido, a interferência na atividade da RNA polII por proteínas que se ligam à montante do promotor e podem aumentar ou diminuir sua atividade (Kadonaga, 2004).

### **Splicing**

Chamamos de "splicing" o processamento do pré-mRNA em que sua sequência é alterada, seja pela retirada de partes da molécula, seja pela adição de sequências específicas. No primeiro caso, denominamos de "cis-splicing" ou somente "splicing", porque todo o processo envolve uma única molécula. Quando se adiciona uma sequência transcrita a partir de outro gene a um pré-mRNA, denominamos o processo de "trans-splicing" (Alberts e col., 2002).

### **Cis-splicing**

Durante a transcrição, ainda no núcleo, um complexo chamado spliceossomo liga-se ao pré-mRNA. Esse complexo lê a sequência do pré-mRNA e, caso encontre marcações para isso, retira parte da molécula, liga a parte anterior ao resto da mesma e continua o processo. As partes retiradas são chamadas de íntrons, enquanto que as regiões que farão parte do mRNA maduro e que, portanto, serão expressas, são chamadas de exons. As marcações para a retirada de íntrons são bastante variáveis, embora apresentem algumas características conservadas. Essa variabilidade confere ao sistema uma maior plasticidade de ação no processamento do mRNA (Abril e col., 2005).

Por algum tempo, acreditou-se que os íntrons não continham informações, e que faziam parte de resquícios evolutivos que perderam a função com o tempo. No entanto, hoje se reconhece que esse pensamento era equivocado: é sabido que a expressão pode depender da

presença de íntrons, e ainda que esses íntrons têm sítios de ligação para reconhecimento indireto de condições ambientais. É o caso, por exemplo, do gene para a produção de arylalkylamina N-acetiltransferase (AA-NAT), que controla a produção do hormônio melatonina. Uma região no íntron 1 desse gene apresenta sítio de ligação de um elemento responsivo a AMP cíclico (AMPC). Esse elemento é chamado de CREB e aumenta a tradução do mRNA quando ocorre um aumento do AMPC, ou seja, quando a célula é estimulada. Quando estímulos chegam às células da glândula pineal, ocorre um aumento do AMPC, que permitirá a ligação do CREB no íntron 1 e levará a um aumento de mais de 5 vezes na quantidade de melatonina produzida na glândula. O mesmo efeito é muito diminuído com a retirada do íntron 1, como ocorre nas outras células do corpo (Baler e col., 1999). Com isso, a glândula pineal pode regular a produção de melatonina por responder de maneira indireta à presença de luz, o que torna possível regular o ritmo circadiano através de condições luminosas do ambiente.

#### **Acoplamento de fatores de transcrição**

Outra característica interessante da RNA polII é a presença de estruturas que permitem o acoplamento de fatores que agirão pós-transcrição, chamados fatores de processamento do RNA. Esse processamento transforma o mRNA recém transcrito em mRNA maduro, sem íntrons, com cauda de poli-A e CAP, que ajudam a manter a molécula mais estável. O CAP é uma guanidina metilada adicionada na extremidade 5' do mRNA, e a cauda de poli-A é a adição de dezenas de adenosinas na extremidade 3' da molécula (Alberts e col., 2002). A estrutura da RNA polIII aumenta a eficiência de todos esse processos porque concentra os fatores necessários para o processamento próximo ao local da transcrição (Maniatis e Reed, 2002; Zorio e Bentley, 2004).

#### **Splicing alternativo**

O número de genes encontrado no genoma humano foi estimado em cerca de 25.000 genes, no entanto, esses guardam informações para a produção de um número muito maior de proteínas (Venter e col., 2001). Uma das explicações para essa discrepância está na flexibilidade dos sítios de reconhecimento de "splicing", o que permite ao sistema de processamento a criação de diferentes formas de proteínas a partir de uma mesma sequência de DNA. É o evento chamado "splicing" alternativo. Segundo Sharp (2009), mais da metade do DNA de células de vertebrados é expresso segundo diferentes formas de "splicing". É o caso, por exemplo, da imunoglobulina M, que, em peixes, tem a sequência de DNA expressa em linfócitos B no início de desenvolvimento sob a forma de proteína de membrana, enquanto que nos linfócitos B maduros o gene é expresso sob a

forma de proteínas plasmáticas. (Wilson e col., 1995). Outro exemplo conhecido é a expressão do gene de beta-tropomiosina do músculo esquelético, que ocorre de maneira diferenciada em fibroblastos e músculo liso por meio de "splicing" alternativo (Helfman e col., 1988).

O "splicing" alternativo também pode ser a chave para processos patológicos, dando origem a transcritos que não são funcionais ou tem função exacerbada. É o caso, por exemplo, do insulinoma humano, onde o mRNA da insulina apresenta como variante de "splicing" uma forma com a inclusão de parte do primeiro íntron na região 5'UTR. Com isso, tanto o RNA sem o íntron, variante mais comum, quanto o RNA com o íntron darão origem à mesma proteína viável, no entanto, a variante do mensageiro da insulina com a parte do íntron 1 é traduzida mais vezes do que a variedade nativa. Em células pancreáticas normais, a quantidade de variante com íntron não chega a 0,5%. No entanto, em insulinomas, as células do tumor aumentam significativamente a porcentagem de variantes do mensageiro com o íntron 1, chegando a produzir mais de 13% dos mRNA de insulina com o íntron conservado. Com isso, essas moléculas, por serem mais traduzidas, acabam gerando muito mais insulina por célula, gerando o quadro de hiperinsulinemia e, conseqüentemente, hipoglicemia (Minn e col., 2004).

#### **Trans-splicing**

O "trans-splicing" consiste na união de exons de diferentes transcritos em um único mRNA. Alguns organismos apresentam transcrição policistrônica, isto é, apresentam um promotor no cromossomo e em seguida a ele diversos genes. Esses organismos precisam, então, individualizar a informação, ou seja, processar o pré-mRNA em RNA com informação para a síntese de proteína. Assim, o "trans-splicing" permite a formação desses mRNA maduros. Esse tipo de "trans-splicing" é conhecido como SL "trans-splicing", porque adiciona uma sequência líder (SL) no início dos mRNA. Essa sequência vem acompanhada do CAP, o que permite a proteção dessas moléculas altamente instáveis. Esse mecanismo é responsável pelo controle de expressão em organismos com transcrição policistrônica, quando a individualização da informação só é possível com a adição da sequência SL. Embora o mecanismo seja comum em protozoários, o processo é também encontrado em metazoários, como em cnidários, platelmintos, nemátodes, insetos e inclusive em mamíferos (Mayer e Floeter-Winter, 2005).

#### **Edição de RNA**

Outro processo pós-transcricional é a edição do RNA, também chamado de RNA *editing*. O processo consiste na modificação do transcrito, e foi descrito em protozoários

cinetoplastídeos. Esses organismos apresentam uma única mitocôndria, com seu material genético organizado na forma circular, concatenados em minicírculos e maxicírculos. Os maxicírculos contêm a informação primária para a produção das proteínas mitocondriais, enquanto que os minicírculos codificam os chamados gRNA, que se associam aos transcritos dos maxicírculos e, junto com uma maquinaria chamada editossomo, farão o processo de edição.

Basicamente, esse processo consiste na inserção e ocasionalmente na retirada de uridinas (U) dos pré-mRNA ou na troca de uma citidina (C) por uridina. Essas alterações provocam, muitas vezes, a alteração da fase de leitura do RNA, de forma que apenas os mRNA que passarem por esse *editing* serão lidos corretamente. Com isso, as moléculas de mRNA formadas são, na verdade, híbridos de sequência transcritas a partir de DNA de maxicírculos com sequências transcritas dos minicírculos (Simpson e col., 2000).

### Estabilidade do mRNA

As moléculas de mRNA apresentam, como citado, estruturas nas extremidades 5' e 3' não traduzidas que aumentam sua estabilidade. No entanto, há uma ciclagem dessas moléculas na célula, através de degradação controlada por marcas presentes na própria molécula. Dentre os mecanismos para degradação, encontramos complexos enzimáticos capazes de detectar RNA dupla-fita. Esses complexos reconhecem as duplas-fitas de RNA e as cortam em fragmentos com aproximadamente 21 bases. Esses fragmentos são, então, usados como moldes na detecção de moléculas complementares de RNA. Quando uma molécula de mRNA se hibrida a esses fragmentos, ocorre sua degradação. Com isso, é possível inibir a expressão de um gene através da expressão de sua sequência complementar. Esse processo vem sendo explorado atualmente para o *knockdown* de genes em fases específicas do organismo (Fjose e col., 2001; Shoji e col., 2005).

### Considerações finais

A disposição de se obter o sequenciamento do genoma humano estava baseada na ideia de que o conhecimento do conjunto de bases que compõem o DNA humano seria revelador de grande parte de seu funcionamento. Com o anúncio da composição do genoma, no entanto, se percebeu que, na verdade, a complexidade fisiológica envolvia muito mais do que apenas a predição de produtos codificados na sequência de bases. A regulação da expressão desses produtos e o entendimento da relação dos produtos e mecanismos alternativos associados à regulação são os pontos que fazem dos seres vivos em geral, incluindo o homem, organismos com diversos níveis de organização e complexidade.

Cada vez mais se reconhece a importância de diversas regiões do genoma, de sistemas enzimáticos de controle de transcrição e tradução e de sistemas pós-transcricionais de edição. Conceitos como o DNA "lixo", de que sequências não codificadoras seriam apenas resquício evolutivo sem função nenhuma, estão, senão acabando completamente, ao menos diminuindo em grande escala com relação ao que se acreditava ainda há poucos anos, e o estudo da expressão gênica vem contribuindo para entendermos processos fisiológicos e suas variantes patológicas.

### Bibliografia

- Abril J.F., Castelo R. e Guigó R. (2005). Comparison of splice sites in mammals and chicken. *Genome Research*. 15, 111-119.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. e Walter, P. *Molecular Biology of the Cell*. 4ed., Nova Iorque, Garland Sciences, 2004.
- Baler R., Covington S. e Klein D.C. (1999). Rat arylalkylamine N-acetyltransferase gene: upstream and intronic components of a bipartite promoter. *Biology of the Cell / Under the Auspices of the European Cell Biology Organization*. 91, 699-705.
- Carginale V., Trinchella F., Capasso C., Scudiero R. e Parisi E. (2004). Gene amplification and cold adaptation of pepsin in Antarctic fish. A possible strategy for food digestion at low temperature. *Gene*. 336, 195-205.
- Claycomb J.M., Benasutti M., Bosco G., Fenger D.D. e Orr-Weaver T.L. (2004). Gene amplification as a developmental strategy: isolation of two developmental amplicons in *Drosophila*. *Developmental Cell*. 6, 145-155.
- de Andrade Stempluk V. e Floeter-Winter L.M. (2002). Functional domains of the rDNA promoter display a differential recognition in *Leishmania*. *International Journal for Parasitology*. 32, 437-447.
- Elias, M.C., Marques-Porto, R., Freymüller, E. e Schenkman, S. (2001). Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. *Molecular Biochemistry Parasitology*, 112, 79-90.
- Fjose A., Ellingsen S., Wargelius A. e Seo H.C. (2001). RNA interference: mechanisms and applications. *Biotechnology Annual Review*. 7, 31-57.
- Helfman D.M., Ricci W.M. e Finn L.A. (1988). Alternative splicing of tropomyosin pre-mRNAs in vitro and in vivo. *Genes & Development*. 2, 1627-1638.
- Kadonaga J.T. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell*. 116, 247-257.
- Lewin B.. *Genes VII*. Oxford University (Ed.), 2000.
- Maniatis T. e Reed R. (2002). An extensive network of coupling among gene expression machines. *Nature*. 416, 499-506.
- Mayer M.G. e Floeter-Winter L.M. (2005). Pre-mRNA trans-splicing: from kinetoplastids to mammals, an easy language for life diversity. *Memórias do Instituto Oswaldo Cruz*. 100, 501-513.
- Minn A.H., Kayton M., Lorang D., Hoffmann S.C., Harlan D.M., Libutti S.K. e Shalev A. (2004). Insulinomas and expression of an insulin splice variant. *Lancet*. 363, 363-367.
- Neumann M., Sampathu D.M., Kwong L.K., Truax A.C., Micsenyi M.C., Chou T.T., Bruce J., Schuck T., Grossman M., Clark C.M., McCluskey L.F., Miller B.L., Masliah E., Mackenzie I.R., Feldman H., Feiden W., Kretzschmar H.A., Trojanowski J.Q. e Lee V.M. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 314, 130-133.
- Philimonenko V.V., Zhao J., Iben S., Dingová H., Kyselá K., Kahle M., Zentgraf H., Hofmann W.A., de Lanerolle P.,

- Hozák P. e Grummt I. (2004). Nuclear actin and myosin I are required for RNA polymerase I transcription. *Nature Cell Biology*. 6, 1165-1172.
- Sharp P.A. (2009). The centrality of RNA. *Cell*. 136, 577-580.
- Shoji M., Chuma S., Yoshida K., Morita T. e Nakatsuji N. (2005). RNA interference during spermatogenesis in mice. *Developmental Biology*. 282, 524-534.
- Simpson L., Thiemann O.H., Savill N.J., Alfonzo J.D. e Maslov D.A. (2000). Evolution of RNA editing in trypanosome mitochondria. *Proceedings of the National Academy of Sciences of the United States of America*. 97, 6986-6993.
- Singer R.H. e Green M.R. (1997). Compartmentalization of eukaryotic gene expression: causes and effects. *Cell*. 91, 291-294.
- Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G. e cols. (2001). The sequence of the human genome. *Science*. 291, 1304-1351.
- Verschure P.J. (2004). Positioning the genome within the nucleus. *Biology of the Cell / Under the Auspices of the European Cell Biology Organization*. 96, 569-577.
- Watson, J.D. E Crick, F.H.C. (1953) Molecular structure of nucleic acids. *Nature*. 171, 737-738.
- Wilson M.R., Ross D.A., Miller N.W., Clem L.W., Middleton D.L. e Warr G.W. (1995). Alternate pre-mRNA processing pathways in the production of membrane IgM heavy chains in holostean fish. *Developmental and Comparative Immunology*. 19, 165-177.
- Zorio D.A.R. e Bentley D.L. (2004). The link between mRNA processing and transcription: communication works both ways. *Experimental Cell Research*. 296, 91-97.