



**TÓPICOS DE
GENÉTICA
MOLECULAR
PARA UNIVERSITÁRIOS**

**MARIA DE MASCENA DINIZ MAIA
MARTÍN ALEJANDRO MONTES
ANA CRISTINA LAUER GARCIA
CLÁUDIO GALVÃO DE SOUZA JÚNIOR
FERNANDA CRISTINA BEZERRA LEITE
IÊDA FERREIRA DE OLIVEIRA**

1ª EDIÇÃO

TÓPICOS DE GENÉTICA MOLECULAR PARA UNIVERSITÁRIOS

**MARIA DE MASCENA DINIZ MAIA
MARTÍN ALEJANDRO MONTES
ANA CRISTINA LAUER GARCIA
CLÁUDIO GALVÃO DE SOUZA JÚNIOR
FERNANDA CRISTINA BEZERRA LEITE
IÊDA FERREIRA DE OLIVEIRA**

1ª EDIÇÃO

TODOS OS DIREITOS RESERVADOS. É PROIBIDA A REPRODUÇÃO TOTAL OU PARCIAL DO TRABALHO SEM AUTORIZAÇÃO DA UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO - DEPARTAMENTO DE BIOLOGIA - GENÉTICA E DOS AUTORES



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

Profa. Maria José de Sena

Reitora da UFRPE

Profa. Maria do Socorro de Lima Oliveira

Vice-Reitora

Edson Cordeiro do Nascimento

Diretor do Sistema de Bibliotecas da UFRPE



EDITORA UNIVERSITÁRIA - EDUFRPE

Antônio Marcelo Freitas Athayde Cavalcanti

Diretor da Editora da UFRPE

José Abmael de Araújo

Coordenador Administrativo da Editora da UFRPE

Josuel Pereira de Souza

Chefe de Produção Gráfica da Editora da UFRPE

Projeto gráfico final e arte de capa

Janilson Lemos de Araújo Silva

Dados Internacionais de Catalogação na Publicação (CIP)

Sistema Integrado de Bibliotecas da UFRPE

Biblioteca Central, Recife-PE, Brasil

T674 Tópicos de genética molecular para universitários / Maria de Mascena Diniz Maia ... (et al.) (org.). - 1. ed. - Recife: EDUFRPE, 2024.

109 p. : il.

Inclui referências.

ISBN (físico) 978-85-7946-384-6

ISBN (digital) 978-85-7946-383-9

1.Genética 2. DNA 3. RNA 4. Controle de expressão genética
5. Replicação 6. Transcrição 7. Marcadores moleculares l.
Maia, Maria de Mascena Diniz, org.

CDD 575.1

PREFÁCIO

Este livro nasceu do desejo dos autores de contribuir com a formação dos estudantes da área de Biologia sobre um tema muito atraente no campo da Genética, que é a Genética Molecular. Embora a Genética seja atualmente assunto de grande evidência na mídia, é ainda considerada um campo de muita dificuldade no ensino de Ciências Biológicas como área interdisciplinar. O despertar do interesse para escrever tópicos de Genética Molecular veio da necessidade de oferecer aos estudantes de Biologia temas elaborados de forma clara e direta para melhor compreensão textual dos conceitos abordados em Genética Molecular. O objetivo principal é oferecer aos leitores uma fonte de conceitos básicos de Genética Molecular, descritos em uma linguagem simples, lógica e sintética, com a intenção de contribuir para exercitar e ampliar os conhecimentos em cada tema abordado.

INFORMAÇÕES DOS AUTORES

Coordenadora do Projeto:

Maria Mascena Diniz Maia

Professora Titular da Universidade Federal Rural de Pernambuco
Doutora em Ciências Biológicas pela Universidade Federal de Pernambuco
Mestre em Bioquímica e Fisiologia pela Universidade Federal de Pernambuco
Bacharel em Farmácia pela Universidade Federal de Pernambuco

Demais autores:

Ana Cristina Lauer Garcia

Professora Associada da Universidade Federal de Pernambuco
Doutora em Genética pela Universidade Federal do Rio Grande do Sul
Mestre em Genética pela Universidade Federal do Rio Grande do Sul
Licenciada em Ciências Biológicas pela Universidade Federal do Rio Grande do Sul

Martín Alejandro Montes

Professor Associado da Universidade Federal Rural de Pernambuco
Doutor em Genética pela Universidade Federal do Rio Grande do Sul
Mestre em Genética pela Universidade Federal do Rio Grande do Sul
Licenciado Genética pela Universidad Nacional de Misiones

Cláudio Galvão de Souza Júnior

Professor Titular da Universidade Federal do Agreste de Pernambuco
Doutor em Ciências Biológicas pela Universidade Federal de Pernambuco
Mestre em Genética pela Universidade Federal de Pernambuco
Bacharel em Biologia pela Universidade Federal de Pernambuco
Licenciado em Ciências Biológicas pela Universidade Federal de Pernambuco

Fernanda Cristina Bezerra Leite

Professora Adjunta da Universidade Federal Rural de Pernambuco
Doutora em Ciências Biológicas pela Universidade Federal de Pernambuco
Mestre em Genética pela Universidade Federal de Pernambuco
Bacharel em Ciências Biológicas pela Universidade Federal de Pernambuco

Iêda Ferreira de Oliveira

Bióloga da Universidade Federal Rural de Pernambuco

Doutora em Biologia Animal pela Universidade Federal de Pernambuco

Mestre em Saúde Pública pelo Instituto Oswaldo Cruz

Bacharel em Ciências Biológicas pela Universidade de Pernambuco

Licenciada em Ciências Biológicas pela Universidade Federal de Pernambuco

AGRADECIMENTOS

À Reitoria da UFRPE pelo apoio para realização desse projeto. À Editora Universitária da UFRPE pela encadernação.

Deixo aqui minha homenagem aos estudantes da UFRPE, (ex-estudantes e atuais) que no contato aluno/professor preenchem os espaços desta Instituição na busca pela Educação como processo transformador da sociedade.

Porque DELE, e por ELE, e para ELE são todas as coisas; glória, pois, a ELE (DEUS) eternamente.

Romanos 11:36.

DEDICATÓRIAS

Aos meus filhos, Paula, Bruno e Alexandre, minhas netas, Maria Gabriela e Maria Luísa e meu neto Francisco.

Maria de Mascena Diniz Maia

À nossa amada filha Mariana.

Ana Cristina Lauer Garcia e Martín Alejandro Montes

À minha esposa Daiane, aos meus filhos, Pedro, Matheus, Victor e Davi, à minha filha Camilla e ao meu neto, Arthur. Também aos meus pais e irmãos amados.

Cláudio Galvão de Souza Júnior

À minha imensa família de genes e de coração. Sobretudo, aos amantes da Genética.

Iêda Ferreira de Oliveira

Aos meus pais, Severino e Betânia, meu esposo George e minhas filhas, Maria Fernanda e Maria Clara.

Fernanda Cristina Bezerra Leite

Especial Gratidão

Deixo aqui meus agradecimentos especiais ao Reitor da UFRPE, Professor Dr. Marcelo Brito Carneiro Leão e ao Vice-Reitor Professor Dr. Gabriel Rivas pelo apoio, incentivo e colaboração para produção desse compêndio de Genética Molecular. Obrigada aos colegas da Área de Genética/UFRPE, da Universidade Federal do Agreste de Pernambuco e do CAV/UFPE por suas contribuições valiosas.

Maria de Mascena Diniz Maia

Deixo aqui meus agradecimentos especiais à Doutora Maria de Mascena Diniz Maia, uma irmã. Também aos amigos, Doutor Marcos Antônio Morais Júnior e Doutor Paulo Roberto Eleutério de Souza, pelas oportunidades e acolhimento. À Universidade Federal Rural de Pernambuco, por tudo. Mas, sobretudo, a DEUS.

Cláudio Galvão de Souza Júnior

SUMÁRIO

CAPÍTULO 1: MATERIAL GENÉTICO HEREDITÁRIO	11
1.1 Composição dos ácidos nucleicos	11
1.2 Estrutura dos ácidos nucleicos	13
1.3 Função dos ácidos nucleicos	15
CAPÍTULO 2: FLUXO DA INFORMAÇÃO GENÉTICA	20
2.1 Replicação do DNA	21
2.1.1 Iniciação da replicação do material genético	22
2.1.2 Alongamento de novas fitas	26
2.1.3 Terminação da Replicação	26
2.1.4 Reparo de Replicação	27
2.1.5 Replicação eucariota	27
2.2 Transcrição e Tradução	30
2.2.1 Transcrição do código Genético do DNA	30
2.2.1.1 Estrutura básica dos RNA	33
2.2.1.2 Estrutura do RNA transportador	33
2.2.1.3 Estrutura do ribossomo e do RNA ribossomal	34
2.2.1.4 Etapas da Transcrição	35
2.2.2 Tradução da Mensagem do DNA	38
2.3 Antibióticos de ação no fluxo da informação genética	43
CAPÍTULO 3: MECANISMOS DE CONTROLE DE EXPRESSÃO GÊNICA EM ORGANISMOS EUCARIOTOS	47
3.1 Introdução	47
3.2 Níveis de controle da expressão gênica em eucariotos	52
3.2.1 Regulação Genética	52
3.2.1.1 Regulação Transcricional	54
3.2.1.2 Regulação Pós-Transcricional	60
3.2.1.3 Regulação Pós-Traducional	64

3.2.2	Regulação epigenética	64
3.2.2.1	Composição química da cromatina	65
3.2.2.2	Modificação de histonas.....	66
3.2.2.4	Metilação do DNA.....	72
3.2.2.5	RNAs não codificantes (nc-RNAs).....	73
3.2.2.5.1	Micro-RNAs (miRNAs)	74
3.2.2.5.2	RNAs longos não codificantes (lnc-RNAs)	75
CAPÍTULO 4:	REAÇÃO EM CADEIA DA POLIMERASE	81
4.1	Introdução	81
4.2	Fundamentos da PCR	81
4.3	Transcrição reversa	84
4.4	PCR em Tempo Real	84
4.5	Quantificação Absoluta e Quantificação Relativa por qPCR.....	87
CAPÍTULO 5:	CLONAGEM GÊNICA	90
5.1	Construção de molécula de DNA recombinante.....	90
5.2	Transformação da célula hospedeira	93
5.3	Seleção de células transformadas	94
CAPÍTULO 6:	MARCADORES MOLECULARES	96
6.1	Introdução.....	96
6.2	Definição de marcadores moleculares	97
6.3	Marcadores dominantes e codominantes.....	98
6.4	Principais marcadores moleculares baseados em PCR	100
6.4.1	PCR-RFLP	101
6.4.2	Microsatélites ou SSR.....	103
6.4.3	RAPD	104
6.4.4	ISSR	106
6.5	Considerações finais.....	108

CAPÍTULO 1: MATERIAL GENÉTICO HEREDITÁRIO

Ana Cristina Lauer Garcia
Martín Alejandro Montes

Em alguma reunião familiar você já se perguntou: “por que sou tão parecida/o com os meus parentes (pais, irmãos, avós, tios)?”. Passeando ao ar livre, ao encontrar várias pessoas com seus cachorros, você se questionou: “por que as raças desses animais são tão diferentes?” Você escutou falar que os frangos têm aumentado seu tamanho em um curto intervalo de tempo, não por hormônios, mas por melhoramento genético? Para começar a entender esses e outros assuntos do nosso cotidiano, devemos conhecer o material genético hereditário. Neste capítulo trataremos sobre os ácidos nucleicos responsáveis por gerar o fenótipo dos diferentes seres vivos e pela transmissão das características ao longo das gerações. A informação está organizada em tópicos sobre a composição, estrutura e função do material genético.

1.1 Composição dos ácidos nucleicos

No final de 1860 o médico suíço Friedrich Miescher (1844-1895) chegou à universidade de Tübingen (sul da Alemanha) e se incorporou ao laboratório do químico Felix Hoppe-Seyler. Por sugestão de Hoppe-Seyler, Miescher começou a estudar a química das células do pus. Em um de seus experimentos com estas células, Miescher obteve um precipitado que diferia quimicamente de todas as substâncias proteicas conhecidas. Essa nova substância se concentrava no núcleo celular, organela que nessa época era considerada de pouca importância para o funcionamento da célula. Miescher denominou esta substância de nucleína, pelo fato dela estar concentrada no núcleo das células. A análise química mostrou que as quantidades relativas de hidrogênio (H), carbono (C), oxigênio (O) e nitrogênio (N) presentes na nova substância diferiam das encontradas nas proteínas. Além disso, a nucleína continha fósforo (P), ausente nas proteínas, estes conhecimentos foram publicados em 1871. Em 1889, Richard Altmann obteve preparações altamente purificadas de nucleína e, pelo fato da substância ter caráter ácido, sugeriu que fosse chamada de ácido nucleico.

Entre 1885 e 1901, Albrecht Kossel fez várias descobertas em relação à composição dos ácidos nucleicos, especialmente em relação às bases nitrogenadas, descobrindo a Adenina e Guanina, e posteriormente, a Timina, Citosina e Uracila

(Figura 1). O grupo de pesquisa liderado por Kossel também descobriu que os ácidos nucleicos apresentavam pentose, um açúcar de cinco carbonos. Por suas contribuições, em 1910, Kossel recebeu o Nobel de Fisiologia ou Medicina.

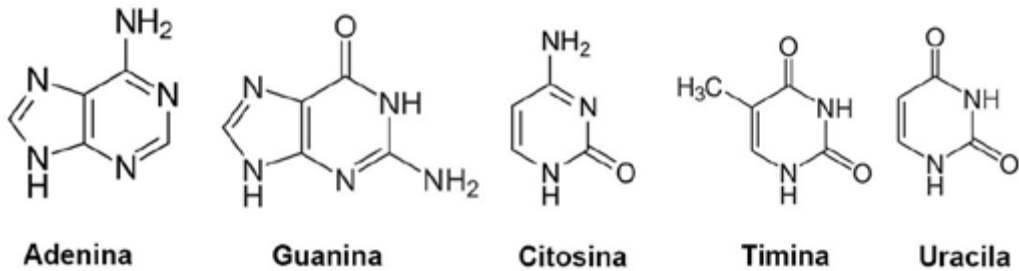


Figura 1. Bases nitrogenadas presentes nos ácidos nucleicos

Em 1909, Phoebis Levine e Walter Jacobs propuseram que os ácidos nucleicos estão constituídos por três unidades: uma base nitrogenada, uma pentose e um fosfato, estrutura que recebeu o nome de nucleotídeo (Figura 2).

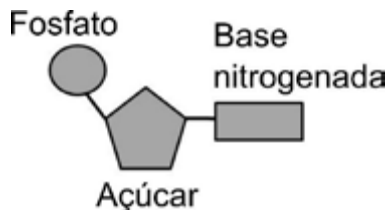


Figura 2. Nucleotídeo formado por um fosfato, um açúcar e uma base nitrogenada.

Phoebus Levene verificou, em 1929, que há dois tipos de ácidos nucleicos que diferem em relação ao tipo de pentose (açúcar) em sua composição. Um deles é ácido ribonucleico (RNA), formado pelo açúcar ribose, que possui o grupo hidroxila (OH) no carbono 2. Já o ácido desoxirribonucleico (DNA) é composto pelo açúcar desoxirribose, que não possui OH no carbono 2 (Figura 3). Outra diferença entre os dois tipos de ácidos nucleicos é em relação às suas bases nitrogenadas. As bases adenina, guanina e citosina são encontradas no DNA e no RNA, já a timina ocorre apenas no DNA e a uracila é exclusiva do RNA.

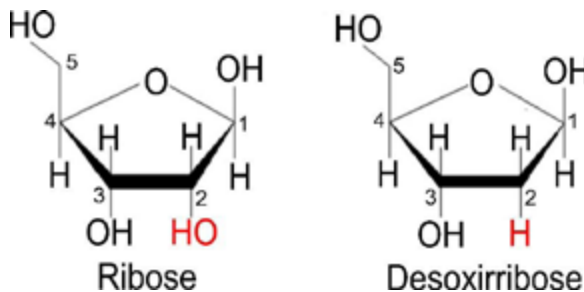


Figura 3. Açúcares presentes nos ácidos nucleicos: a ribose no RNA e a desoxirribose no DNA.

Em 1950, Erwin Chargaff, Stephen Zamenhof e Charlotte Green, analisando o DNA de diferentes espécies, descobriram que a quantidade de Adenina era sempre igual à de Timina ($A=T$), e a quantidade de Citosina era sempre igual à de Guanina ($C=G$) (Tabela 1).

Nome comum	Nome científico	Composição das bases nitrogenadas do DNA em porcentagem			
		Adenina	Timina	Guanina	Citosina
Humano	<i>Homo sapiens</i>	31	31,5	19,1	28,4
Mosca	<i>Drosophila melanogaster</i>	27,3	27,6	22,5	22,5
Milho	<i>Zea mays</i>	25,6	25,3	24,5	24,6
Fungo	<i>Neurospora crassa</i>	23	23,3	27,1	26,6
Bactéria	<i>Escherichia coli</i>	24,6	24,3	25,5	25,6

Tabela 1. Composição das bases nitrogenadas do DNA em diferentes espécies.

1.2 Estrutura dos ácidos nucleicos

No final da década de 1940 iniciou-se uma corrida para conhecer a estrutura do DNA, Rosalind Elsie Franklin, utilizando a técnica de difração de raios X, obteve uma imagem que se tornou crucial para a elucidação da estrutura tridimensional desta molécula (Figura 4).

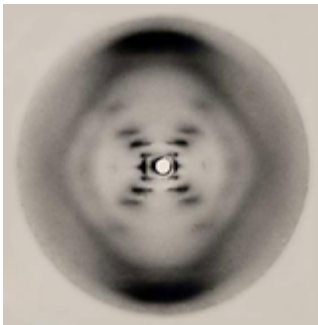


Figura 4. Imagem do DNA feita por difração de raios X por Raymond Gosling, em maio de 1952, sendo ele à época um estudante de doutorado sob a supervisão de Rosalind Franklin no King's College London.

Em 1953, utilizando os dados sobre a composição química do DNA e as imagens de Rosalind Franklin, James Watson e Francis Crick deduziram a estrutura desta molécula. As informações químicas foram importantes para elucidar as interações entre os nucleotídeos e a difração de raios X esclareceu o formato helicoidal da molécula e seu tamanho constante. Watson e Crick desvendaram a estrutura tridimensional do DNA, a qual está formada por duas fitas de nucleotídeos, que se arranjam em uma estrutura de dupla hélice que lembra uma escada em caracol. Os “degraus” da escada correspondem às bases nitrogenadas, unidas por pontes de hidrogênio. Duas pontes de hidrogênio unem a adenina com a timina e três

pontes de hidrogênio unem a citosina com a guanina. O “corrimão” da escada é formado por açúcar (desoxirribose) e fosfato, dispostos alternadamente e unidos por ligações fosfodiéster. Estas ligações conectam o carbono 5’ de uma desoxirribose ao carbono 3’ do açúcar adjacente. As ligações fosfodiéster estão em direções opostas nas duas fitas complementares, uma das fitas está na direção 5’-3’ e a outra na direção 3’-5’, sendo consideradas, portanto, antiparalelas. Isto quer dizer que na extremidade de uma das fitas está livre o agrupamento fosfato do carbono-5 da pentose e na outra ponta desta mesma fita está livre a hidroxila do carbono-3 da última pentose. Na fita complementar este sentido é invertido (Figura 5). O modelo da estrutura do DNA estabelecido por Watson e Crick foi rapidamente aceito pela comunidade científica, porque permitiu compreender como esta molécula se replica, se transcreve e pode sofrer mutações.

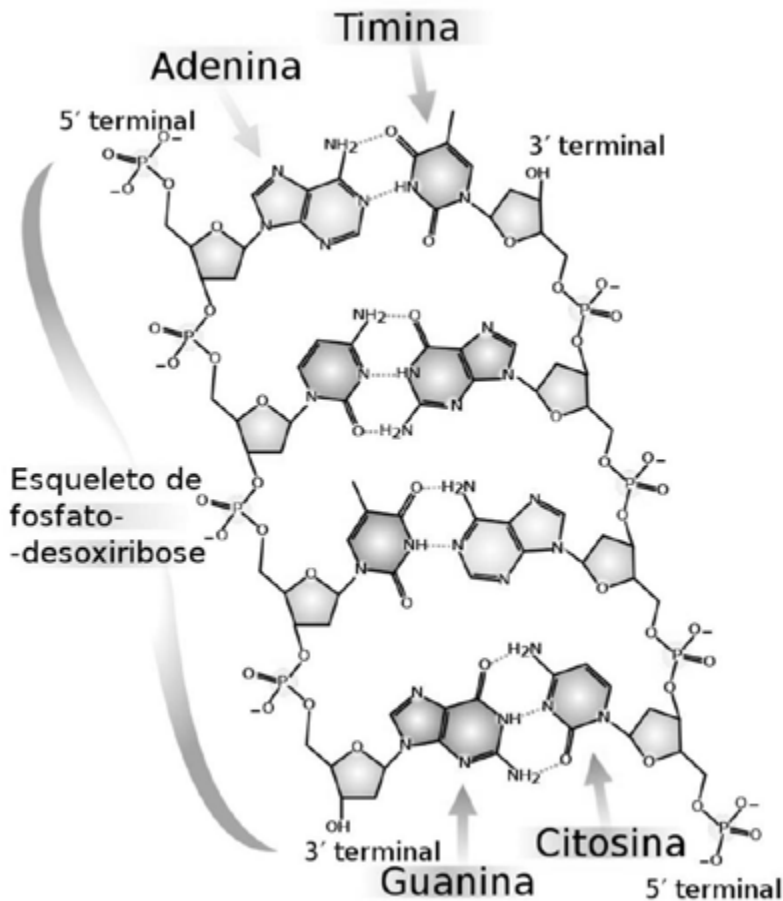


Figura 5. Estrutura da molécula de DNA formada por duas fitas antiparalelas 5’-3’ e 3’-5’. A parte externa de cada fita é formada pelos açúcares e fosfatos dos nucleotídeos unidos covalentemente por ligações fosfodiéster. As bases nitrogenadas ficam para o interior da molécula. As bases de uma fita se unem as bases da outra fita por pontes de hidrogênio, a adenina é ligada à timina por duas pontes de hidrogênio e a guanina se une à citosina por três pontes de hidrogênio.

A estrutura do RNA está formada por uma fita de nucleotídeos ligados covalentemente por ligações fosfodiéster.

1.3 Função dos ácidos nucleicos

O primeiro indicativo de que o DNA é o material hereditário foi obtido por Frederick Griffith em 1928 em experimentos com bactérias pneumococos (*Streptococcus pneumoniae*). Griffith trabalhou com duas cepas destas bactérias: uma contendo uma cápsula de polissacarídeos e causadora da pneumonia, e outra sem essa cápsula e que não produzia a doença. A cápsula apresenta função protetora, evitando a destruição da bactéria pelo sistema imune do hospedeiro. Os experimentos de Griffith se propunham a demonstrar que a cápsula bacteriana era herdada, uma vez que pneumococos capsulados geravam pneumococos com esta mesma condição e os não capsulados originavam bactérias sem a cápsula.

Primeiramente Griffith injetou cepas de pneumococos capsulados em camundongos e observou que os animais morriam. No sangue dos camundongos mortos foi verificada a presença de pneumococos capsulados. Ao injetar cepa de pneumococos sem cápsula nos camundongos, estes sobreviveram e no sangue dos animais não foram encontrados pneumococos. Griffith deu continuidade a seus experimentos matando as bactérias capsuladas com calor e depois as injetando nos camundongos. Neste caso ele observou que os animais não morriam, já que as bactérias estavam mortas e, assim, o camundongo não desenvolvia a doença. Finalmente, Griffith testou uma mistura de pneumococos vivos sem cápsula e bactérias capsuladas mortas pelo calor (lembrando que a aplicação em camundongo desses dois tipos de bactérias separadamente não produzia a doença e não matava os camundongos). Griffith ficou surpreso ao observar que estes últimos camundongos morreram e no sangue destes animais havia bactérias vivas com cápsulas. Inicialmente ele pensou que alguma bactéria viva capsulada poderia estar misturada entre as que foram mortas pelo calor, porém nenhuma infecção foi observada quando eram injetadas somente bactérias mortas pelo calor e, assim, a hipótese foi descartada. A conclusão de Griffith foi que as bactérias sem cápsula tinham se transformado em bactérias capsuladas, a partir de alguma substância das bactérias capsuladas mortas pelo calor, essa substância era o poder transformante (Figura 6).

Alguns anos depois, James Lionel Alloway, demonstrou que a transformação de pneumococos sem cápsula em capsulados poderia ser feita *in vitro* a partir da adição de um extrato de bactérias capsuladas mortas pelo calor. Em 1933 Alloway demonstrou também que adicionando álcool ao extrato de pneumococos capsulados mortos pelo calor era produzido um precipitado esbranquiçado, o qual tinha o poder de transformar pneumococos sem cápsula em capsulados. O que seria essa substância esbranquiçada com tamanho poder de transforma-

ção? Conforme observado por Griffith, a cápsula (ou a falta dela) era uma condição transmitida à progênie bacteriana. Deste modo, a substância que causa a transformação de bactérias sem cápsula em capsuladas deveria ser o material hereditário.

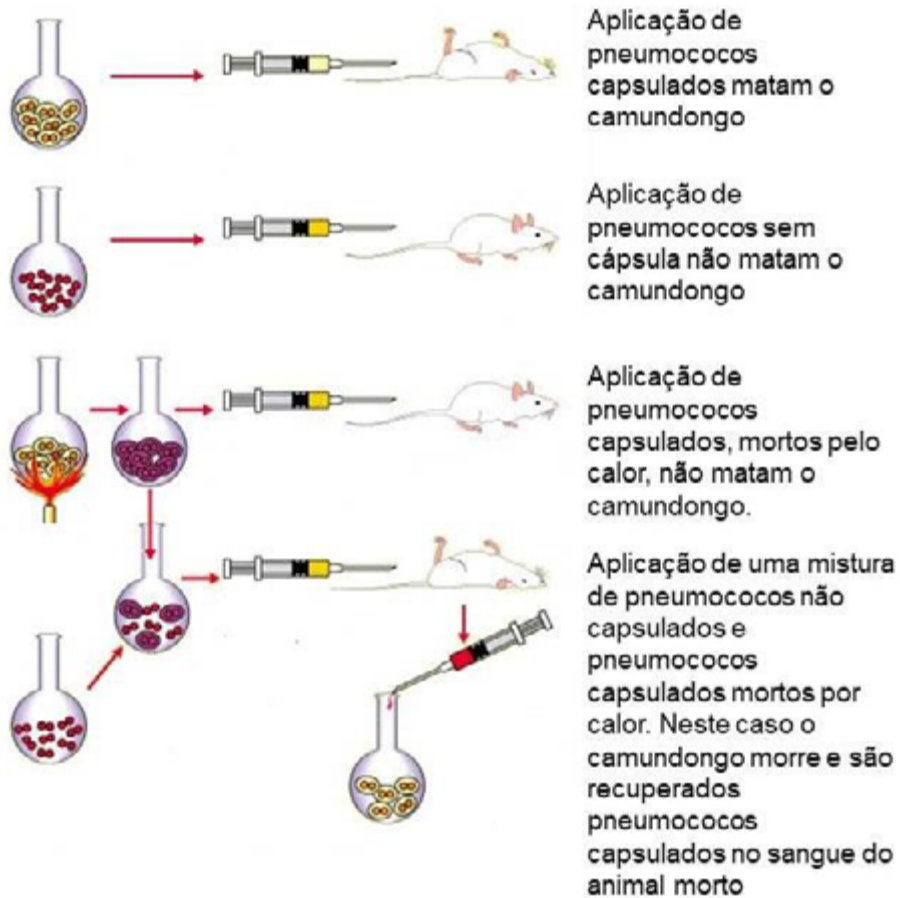


Figura 6. Experimento de Frederick Griffith realizado em 1928.

Em 1944 Oswald Avery, Colin MacLeod e Maclyn McCarty desvendaram o mistério sobre a substância esbranquiçada que causava a transformação bacteriana. Eles trataram essa substância com enzimas que degradam diferentes tipos de moléculas. Ao tratar a substância com enzimas que destruíam polissacarídeos ela mantinha seu poder de transformar pneumococos sem cápsula em capsulados, o mesmo se repetiu ao tratar a substância esbranquiçada com enzimas que degradam lipídeos, proteínas e RNA. Quando, finalmente, os pesquisadores trataram o precipitado esbranquiçado com enzimas que degradam o DNA, eles observaram que a transformação não acontecia, o mistério foi esclarecido: a substância transformante era o DNA (Figura 7).

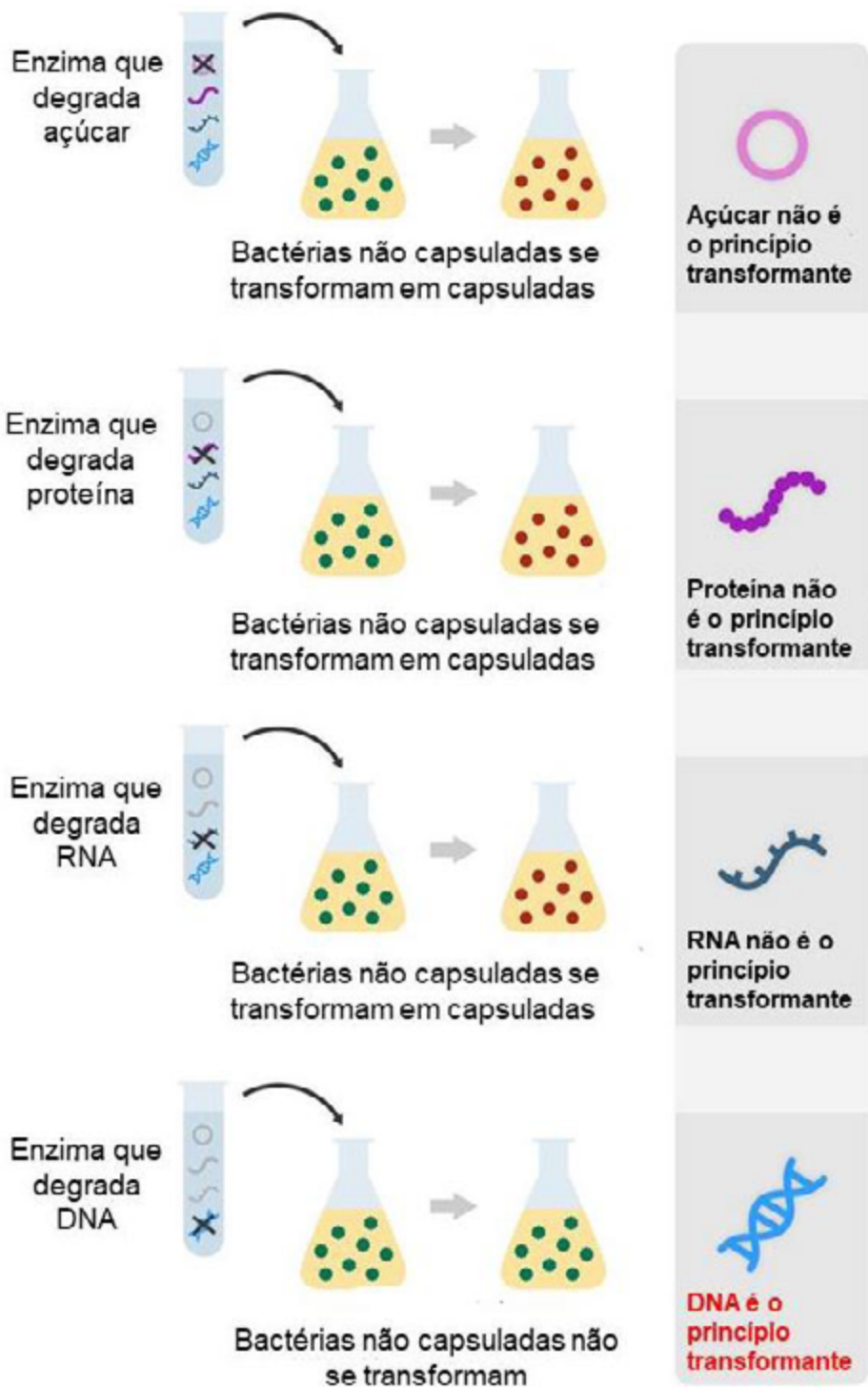


Figura 7. Experimento de Oswald Avery, Colin MacLead e Maclyn MacCarty realizado em 1944.

Mesmo com esses resultados, que mostraram que o DNA é o material hereditário, a comunidade científica ainda tinha dúvidas, já que essa molécula parecia muito simples quando comparada com as proteínas. Assim, em 1952, Alfred Hershey e Martha Chase elaboraram um experimento para elucidar definitivamente qual seria o material hereditário. Para isso, eles realizaram experimentos com vírus que tinha em sua composição DNA e proteínas, assim poderiam saber qual deles é o material hereditário. Hershey e Chase usaram vírus que atacam e matam bactérias, já que sabiam neste processo os vírus injetavam o material hereditário nas bactérias para gerar novas cópias virais. Para diferenciar o DNA das proteínas os pesquisadores utilizaram marcações, sinalizando o DNA com fósforo radiativo ^{32}P (lembre que o DNA tem fósforo, mas não tem esse elemento nas proteínas) e as proteínas foram marcadas com enxofre radiativo ^{35}S (enxofre não existe no DNA). Mas como marcar os vírus? Como não podiam fazer isso, cultivaram bactérias *Escherichia coli* em dois meios separados, uma com fósforo radiativo e outra com enxofre radiativo. Posteriormente, colocaram vírus para atacar as duas linhagens bacterianas, agora radiativas, e assim obter vírus com DNA ou proteínas marcadas radiativamente. Os vírus marcados radiativamente foram colocados separadamente para atacar bactérias sem marcação. A ideia era verificar o que os vírus injetavam nas bactérias para produzir a próxima geração viral. Após um tempo suficiente para que os vírus injetassem seu material hereditário nas bactérias, estes foram desprendidos das bactérias por centrifugação. A seguir outra centrifugação foi realizada para separar as partes leves do sobrenadante (contendo vírus vazios) das partes pesadas do precipitado (contendo bactérias). Isto permitiu observar em que porção do tubo de centrifugação estava a marcação radiativa nos experimentos que utilizaram o fósforo radioativo e nos que utilizaram o enxofre.

Hershey e Chase observaram que as proteínas marcadas radiativamente ficaram na parte leve do sobrenadante, indicando que não haviam sido incorporadas nas bactérias. No experimento com DNA radioativo, a marcação foi detectada no precipitado do tubo de centrifugação, ou seja, o DNA dos vírus tinha sido transferido para as bactérias e era, portanto, o material hereditário (Figura 8).

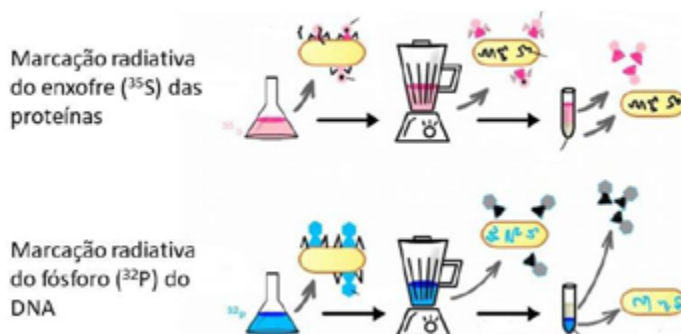


Figura 8. Experimento de Alfred Hershey e Martha Chase realizado em 1952.

Para acabar definitivamente com as dúvidas sobre a identidade do material hereditário, Hershey e Chase permitiram que as duas linhagens de bactérias do precipitado do experimento produzissem uma nova geração de vírus. Apenas as bactérias infectadas por vírus com seu DNA marcado radiativamente geraram vírus com marcações radiativas. Assim, os pesquisadores comprovaram que vírus transferem o DNA para as bactérias, sendo esse o material hereditário.

REFERÊNCIAS

GRIFFITHS, A.J.F., WESSLER, S.R., CARROLL, S.B., & DOEBLEY, J. (2013). Introdução à Genética. 10ª Edição. Rio de Janeiro. Editora Guanabara Koogan
PIERCE, B.A. (2009). Genética: um enfoque conceitual. 3ª Edição. Rio de Janeiro. Editora Guanabara Koogan.

SNUSTAD, D.P., & SIMMONS, M.J. (2013). Fundamentos de Genética. 6ª Edição. Rio de Janeiro. Editora Guanabara Koogan.

SOUZA, P.R.E.D., SILVA, H.D.A.D., LEITE, F.C.B., MAIA, M.D.M.D., GARCIA, A.C.L., & MONTES, M.A. (2015). Genética geral para universitários. Pernambuco. Editora UFRPE.

CAPÍTULO 2: FLUXO DA INFORMAÇÃO GENÉTICA

Cláudio Galvão de Souza Júnior

Dentre os níveis de organização da vida, aquele que diz respeito à informação genética responsável pela formação de proteínas tem recebido destaque pela comunidade científica desde a primeira metade do século 20.

Dos inúmeros esforços para se desvendar os processos presentes no funcionamento celular em diversos centros de pesquisa pelo mundo, muito já se descobriu com o advento da genética molecular. Muito se tem estudado sobre como, em nível celular, se dá a orientação e a realização da produção de enzimas, hormônios, antígenos, anticorpos, membranas, e demais estruturas que compõem desde a mais simples célula até a complexa relação celular de organismos pluricelulares.

Abordaremos neste capítulo, sucintamente, os principais avanços no conhecimento relacionado às estruturas e funções do material genético.

Iniciaremos com o processo de duplicação da informação genética de uma célula, a replicação, que é uma etapa pelo qual uma célula garante que suas células filhas recebam material genético igual ao seu após a divisão celular. Depois, abordaremos os processos de transcrição e tradução, que envolvem a formação de RNAs e proteínas a partir do material genético contido no DNA. Por fim, traremos um exemplo aplicado do avanço tecnológico advindo do conhecimento sobre o fluxo da informação genética, onde você poderá exercitar o conhecimento construído nesta leitura. Ao longo do texto, disponibilizamos endereço de sites onde animações e imagens de domínio público podem ser visitadas para facilitar o entendimento e auxiliar na fixação do conteúdo apresentado.

O Dogma Central da Biologia

As informações para os traços e funções de qualquer organismo vivo conhecido seguem uma linguagem comum, o que permite uma interação intercelular individual e mesmo interespecífica, independente de classificações como ordem, classe ou reinos. Trata-se do **código genético**, que, fundamentado nos nucleotídeos e seus componentes e sua complementaridade de bases, estabelece a comunicação intracelular e entre diferentes células. Essa linguagem comum a toda vida na terra está contida nos ácidos nucléicos promovendo as funções genotípicas, fenotípicas e adaptativas, através da replicação, transcrição, tradução (chamado de dogma central da biologia; Figura 1) e mutação do código genético, respectivamente.

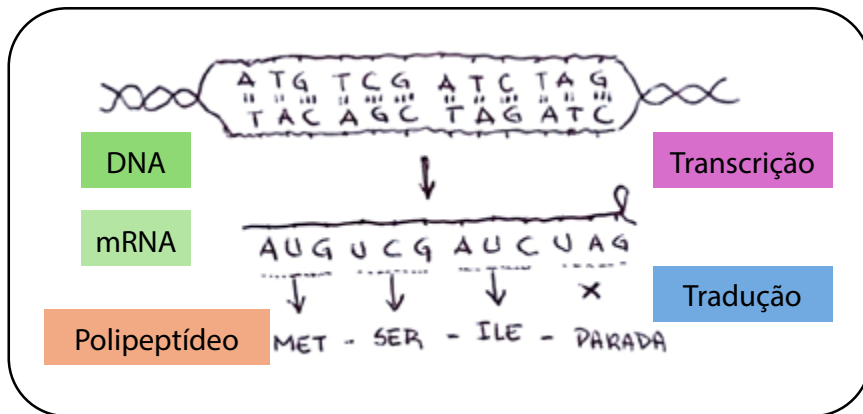


Figura 1. Dogma Central da Biologia. Adaptado de: https://pt.wikipedia.org/wiki/Ficheiro:Dogma_central_-_traduzido.jpg

Vamos tratar, de forma objetiva, do fluxo da informação contida no material genético e sobre o quando, onde e como se dá esse evento fundamental para a manutenção da vida como conhecemos ao longo de milhares de anos.

2.1 Replicação do DNA

O processo de replicação se caracteriza pela duplicação do DNA de uma célula. Sua relevância reside em garantir o mesmo material genético de uma célula-mãe para as duas células filhas resultantes da divisão celular. Sua importância pode ser estudada na função genotípica da biologia celular, que permite a reprodutibilidade do patrimônio genético e tendo relação direta com a herdabilidade de traços e funcionalidades celulares.

São alvos de estudos sobre como se dá uma replicação eficiente a fisiologia da reprodução, a embriologia, crescimento e envelhecimento, teratologia, o câncer e até o processo da morte celular.

Em termos gerais, quanto à replicação, estuda-se a formação de duas duplas fitas de DNA a partir de outra dupla fita pré-existente. Ocorre na fase S (síntese) da divisão celular, quando a célula precisa duplicar seu patrimônio para garantir que, após a divisão celular, as células filhas tenham o mesmo número e tipos de genes de sua célula mãe. É um evento antecessor tanto da mitose, proliferação celular, quanto na meiose, na formação de gametas.

A replicação consiste em três grandes estágios (iniciação, alongamento e término) que acontecem com a ação de um aparato enzimático complexo e eficiente chamado **replissoma**. A enzimologia do replissoma se constitui de uma sequência progressiva e complementar de proteínas responsáveis por:

1. localizar as origens da replicação,
2. realizar a abertura das duplas fitas do DNA-alvo,
3. promover a estabilização das fitas simples resultantes,
4. inserir iniciadores do processo de replicação,
5. polimerizar fita complementar à fita molde,
6. editoria molecular das novas fitas, e
7. finalização, com a desmontagem do replissoma.

Apesar de se parecerem em vários pontos, a replicação entre células eucarióticas e procarióticas têm suas particularidades, o que poderemos tratar de forma objetiva neste capítulo onde apresentamos um estudo comparativo. Para fim de melhor entendimento, apresentaremos como se dá nos procariotos inicialmente, e posteriormente, algumas particularidades dos eucariotos.

2.1.1 Iniciação da replicação do material genético

A iniciação da replicação é a etapa em que o replissoma reconhece um ou mais pontos de partida para duplicação do material genético, localizadas nas chamadas **origens de replicação** dos cromossomos (*oriC*)¹ ricas em sequências AT. A partir dos estudos e registros do modelo de replicação em bactérias *E. coli*, pelos pesquisadores Matthew Meselson e Franklin Sthal, sabemos hoje que os cromossomos circulares bacterianos têm em sua imensa maioria apenas uma origem de replicação.

Uma unidade de replicação ou **replicon** se caracteriza pela região abrangida por uma origem de replicação, sequência de DNA e região finalizadora. Em bactérias, todo seu cromossomo circular constitui o replicon. Toda replicação inicia em uma **origem de replicação** e continua até que o **replicon** inteiro seja duplicado. Diferentemente dos cromossomos bacterianos, cada cromossomo eucariótico, é coberto por várias **origens** de replicação para garantir a eficiência e velocidade de duplicação do material genético contido em seus longos cromossomos.

O DNA cromossômico é helicoidal e se encontra amplamente espiralado e enovelado em torno de histonas² para garantir seu super empacotamento, o que permite que todo DNA seja acomodado dentro de uma minúscula célula. Assim, o acoplamento do aparato de replicação (**replissoma**)³ só se torna possível pela locali-

1 Replicação do DNA circular de bactérias. Acesse: (https://pt.wikipedia.org/w/index.php?title=Especial:Citar&page=Cromossoma_circular&id=66216581&wpFormIdentifier=titleform).

2 Histonas são proteínas envolvidas no empacotamento e regulação de nucleossomos (acesse <https://pt.wikipedia.org/w/index.php?title=Histona&oldid=62098819>)

3 Conjunto de enzimas envolvidas na duplicação do DNA. Acesse: (<https://pt.wikipedia.org/wiki/Replissoma>) .

QUADRO 1 - Componentes necessários para replicação bacteriana (REPLISSOMA)

Componente	Função
Proteína de iniciação	Liga-se à origem e separa as fitas de DNA
Helicase	Desenrola o DNA e abre a bolha
Proteínas aderentes ao DNA de fita única – SSB	Evitam reaproximação das fitas complementares
DNA girase	Age liberando o torque do avanço das forquilha
DNA primase	Sintetiza uma sequência curta complementar à origem da fita molde, que oferece um grupo 3'-OH para inserção de novo nucleotídeo
DNA polimerase II	Inserir novo nucleotídeo (complementar) na extremidade 3'-OH disponível da fita crescente
DNA polimerase I	Remove o iniciador colocado pela primase e substitui por DNA
DNA ligase	Une os fragmentos de polímeros separados presentes na formação da fita complementar descontinuada

Adaptado de: Pierce, 2016

zação de um sítio de iniciação, uma sequência mínima necessária para a exposição da origem de replicação cromossômica. Em bactérias, um sítio de iniciação é composto por uma sequência anterior de 254 pares de bases (254 pb) que permite o reconhecimento pela proteína de iniciação (como a proteína DnaA na *E. coli*), que se liga à sua origem de replicação (chamada *oriC*) e faz um segmento curto de DNA se desenrolar e se tornar acessível ao conjunto de proteínas replicativas, o replissoma.

O Desenovelar e Desespiralizar do DNA

A **estrutura tridimensional do DNA**⁴ (proposta por John Watson e Crick em meados de 1950) proporciona a estabilidade que caracteriza o ácido desoxirribonucleico como a molécula da informatividade biológica de todos os seres vivos, mas por outro lado traz a necessidade de um mecanismo complexo para permitir que a decodificação e replicação dessa informação.

A etapa seguinte à localização da origem de replicação é o desenrolamento

⁴ O DNA é um enovelado de dupla fitas antiparalelas, formadas por polímeros de nucleotídeos ligados por ligações fosfodiéster, e conectadas uma a outra por pontes de hidrogênio que associam bases nitrogenadas adenina (A), timina (T), guanina (G) e citosina (C) a partir da complementaridade A=T e G=C. (acesse: https://pt.wikipedia.org/w/index.php?title=%C3%81cido_desoxirribonucleico&oldid=66902668).

(o alongamento da replicação) para que a síntese de novas fitas filhas possa acontecer a partir de segmentos molde do DNA original. Para tal, um grupo de proteínas e enzimas passam a atuar pela exposição de um segmento curto pós origem de replicação. Sabe-se atualmente que o processo de exposição de sequências a serem replicadas ocorre pela ação determinante de enzimas que rompem as pontes de hidrogênio que ligam as bases nitrogenadas complementares das duplas fitas do DNA, tornando-se lineares e isoladas.

Assim, para que se inicie a replicação é necessário, primeiramente, que a fita dupla de DNA se desenovele, e se apresente como fitas simples e desconectadas uma da outra pela quebra das pontes de hidrogênio. Contudo, como esse desenovelamento da dupla hélice não ocorre espontaneamente, para isso, proteínas especiais chamadas **helicases**⁵ devem se ligar ao segmento de DNA pós origem, promover esse desenovelar e separar das fitas complementares.

Estabilizando as Fitas Simples

Mas, uma vez afastadas, as fitas simples de DNA tendem naturalmente a se reconectar refazendo as pontes. Para que isso não ocorra, proteínas chamadas **SSB**⁶ (do inglês, single-strand binding) atacam as fitas simples e estabilizam a separação ocupando as regiões das pontes, impedindo a reestruturação em fitas duplas, mantendo as fitas separadas.

Uma terceira enzima do replissoma tem papel fundamental na preservação da estrutura das regiões ladeadoras da bolha de replicação, as **DNA topoisomerasas**⁷, como a DNA girase nas bactérias. Quando desespiralizada a dupla fita de DNA sofre em suas regiões próximas, posteriores às forquilhas, um super-enovelamento (enovelamento positivo ou aumentado), causando estresse e acúmulo de energia ao comprimir ainda mais a distância das fitas na dupla hélice nessa região, a semelhança do que ocorre quando se abre uma bolha em uma liga de borracha enrolada em si mesma.

Uma vez que a helicase desespiraliza o DNA, e ela expõe as bases das fitas complementares antiparalelas, uma proteína chamada **primase**⁸ (uma RNA polimerase polimeriza um pequeno fragmento de RNA que tem uma hidroxila (3'-OH) livre

5 Helicases são enzimas responsáveis pelo desenovelamento e abertura da dupla fita de DNA. (acesse: <https://pt.wikipedia.org/w/index.php?title=Helicase&oldid=64157144>).

6 Acesse: (<https://pt.wikipedia.org/wiki/SSB>).

7 Topoisomerasas são enzimas que desempenham papel na preservação da estrutura do DNA durante a abertura da dupla hélice. (acesse: <https://pt.wikipedia.org/w/index.php?title=Topoisomerase&oldid=54433872>).

8 Primase são enzimas que polimerizam os iniciadores ou *primers* de RNA (acesse <https://pt.wikipedia.org/w/index.php?title=Primase&oldid=63967870>).

em sua terminação, o **primer**⁹), adicionando uma série de nucleotídeos complementares às regiões expostas das duas fitas abertas. Esse fragmento de RNA atua como um sinal para que uma terceira proteína chamada **DNA polimerase**¹⁰ (Figura 2) possa se ligar e iniciar uma sequência de inserções de nucleotídeos complementares a cada fita molde.

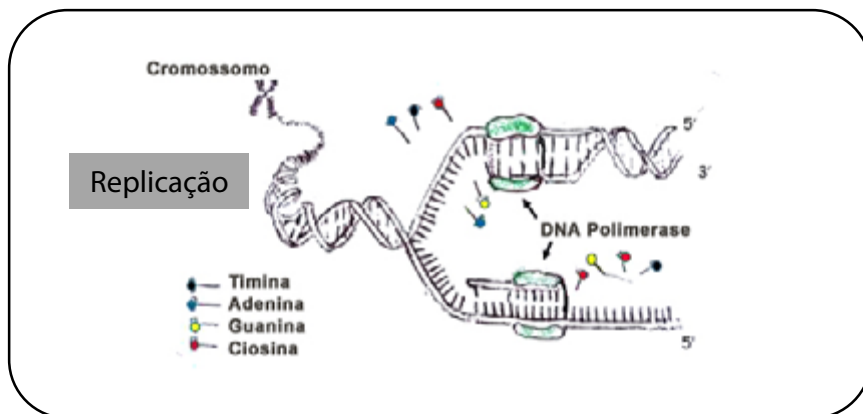


Figura 2. Desenho esquemático de uma DNA polimerase em uma forquilha de replicação. Fonte: <https://tse2.mm.bing.net/th?id=OIP.wsCidbM7K33WQSYaT526HwHaEI&pid=Api&P=0&h=180>

A replicação do DNA é **semiconservativa**, pois se desenvolve a partir da separação da dupla fita inicial e a utilização de cada fita separada como moldes isolados para polimerização de uma nova fita complementar em cada

As fitas simples separadas pela helicase são denominadas **fitas moldes** ou fitas-mãe, e as fitas complementares polimerizadas sobre cada fita-mãe passam a ser chamadas **fitas-filhas**. Assim, cada fita-mãe, fruto da separação da fita dupla inicial, é tomada como fita molde para suas respectivas fitas-filhas (complementares).

Uma vez que o complexo enzimático (o replissoma) localiza a origem de replicação, esta se inicia para as duas direções do comprimento da dupla fita de DNA, promovendo a formação de uma bolha que, à medida que um replissoma segue para cada direção, vai ampliando seu diâmetro e promovendo uma separação gradativa das fitas duplas iniciais (fitas-mãe). A esta bolha formada dá-se o nome de **bolha de replicação** e às regiões que se formam pela separação progressiva, chamam-se **forquilhas** (Figura 2). Daí podemos dizer que cada bolha contém duas forquilhas em suas extremidades que se separam à medida que o processo se desenvolve. Trata-se, portanto, de um processo bidirecional.

⁹ *Primer* de RNA é um fragmento necessário para o início de uma polimerização de nucleotídeos de DNA (acesse <https://pt.wikipedia.org/w/index.php?title=Iniciador&oldid=66286945>).

¹⁰ DNA Polimerases são enzimas envolvidas na formação, edição ou remoção de nucleotídeos. (acesse https://pt.wikipedia.org/w/index.php?title=ADN_polimerase&oldid=64605401).

2.1.2 Alongamento de novas fitas

A enzima DNA polimerase catalisa o acoplamento de nucleotídeos livres a uma fita molde, por complementaridade de bases (A=T ou C=G) promovendo uma ligação fosfodiéster entre os nucleotídeos da nova fita inseridos e liberando pirofosfatos durante o processo. Sabe-se, atualmente, que a DNA polimerase promove as inserções a partir da leitura de cada fita-mãe, leitura essa que se dá na direção 3'-5'. Sendo assim, a polimerização de nucleotídeos da nova fita (fita-filha) só ocorre no sentido oposto (5'-3') por serem as fitas-mãe e fitas-filha antiparalelas e terem seus nucleotídeos dispostos em sentidos opostos.

Esta propriedade de polimerizar uma nova fita de DNA no sentido 5'-3', complementar à uma fita-mãe pela DNA polimerase, trouxe à tona uma estratégia extraordinária do replissoma, isso porque, a forquilha de replicação funciona como um zíper sendo aberto, onde uma fita-mãe se mostra no sentido 3'-5' e a outra fita-mãe, no sentido 5'-3', lembrando que cada forquilha de uma bolha está sendo processada por seu próprio replissoma.

Como cada aparato de replicação disposto nas diferentes forquilhas é constituído por uma DNA polimerase, estas disposições das fitas-mãe exigem que uma fita-filha seja polimerizada continuamente e a outra fita-filha seja polimerizada de forma descontínua. Para este fim, a fita descontínua é polimerizada sobre uma alça móvel da sua respectiva fita-mãe que permite a DNA polimerase inserir nucleotídeos no sentido de abertura da forquilha. Assim, o replissoma forma a alça, insere um *primer*, polimeriza um fragmento de DNA no sentido 5'-3', logo após desfazendo a alça. Essa estratégia produz ao longo da forquilha a formação de uma fita continuada a partir de um único *primer* e outra nova fita, esta descontínua, com vários fragmentos oriundos de vários *primers*¹¹.

Entram em ação outras enzimas do replissoma que seguem a DNA polimerase, desfazendo os *primers* e substituindo-os por sequências de DNA, as exonucleases¹² e as DNA ligases¹³, respectivamente. Isso restabelece o tamanho da fita contínua e estabelece a continuidade à fita-filha descontínua. À medida que as forquilhas se afastam, conseqüentemente a bolha aumenta até que todo DNA (cromossomo) seja replicado.

2.1.3 Terminação da Replicação

A replicação nos procariontes termina quando o replissoma encontra algumas sequências terminadoras, as chamadas sequências Ter, que promovem a disso-

11 Forquilha de replicação e replissoma. Acesse: (<https://pt.wikipedia.org/wiki/Replissoma>)

12 Exonucleases são DNA polimerases com capacidade de desmontar segmentos de uma fita de DNA. (acesse <https://pt.wikipedia.org/w/index.php?title=Exonucleases&oldid=59904942>).

13 DNA ligase é a de enzima responsável pela reconexão de ligação fosfodiéster entre dois nucleotídeos do DNA. (acesse https://pt.wikipedia.org/w/index.php?title=ADN_ligase&oldid=66681329).

ciação do aparato. Ao final da replicação vemos então o surgimento de duas novas dupla-fitas de DNA, formadas de uma dupla-fita antecessora.

Pode-se dizer então que a replicação do material genético, o DNA, se dá de forma semiconservativa, bidirecional e a partir de polimerização simultânea contínua e descontínua em cada forquilha¹⁴.

Comparativamente, a replicação em células eucarióticas segue um processo semelhante por se tratar de material genético o DNA, constituído de nucleotídeos e de estrutura dupla-fita, complementar, antiparalela, helicoidal e super enovelada. Contudo, estudos têm permitido conhecer particularidades como a necessidade de múltiplas origens de replicação (para atender ao extenso comprimento de seus cromossomos), e uma composição de replissoma mais complexo.

2.1.4 Reparo de Replicação

A taxa de erro em uma replicação bacteriana é aproximadamente um erro por bilhão de nucleotídeos inseridos. Estudos demonstram que a DNA polimerase seleciona um nucleotídeo errado a cada 100.000. Nucleotídeos errados ficam mal pareados e são prontamente alvo de reparo por ação exonucleásica 3'-5' da DNA polimerase. Uma vez que um evento de metilação de nucleotídeos ocorre após uma fita estar totalmente concluída, a substituição dos nucleotídeos errados também ocorre por estes ainda não estarem metilados (portando grupo metila -CH₃) e serem mais facilmente identificados pelas polimerases de reparo.

2.1.5 Replicação eucariota

Ao se estudar a replicação do material genético dos eucariotos, sabemos atualmente que um complexo proteico (Quadro 2) toma lugar para que o reconhecimento de origens de replicação e recrutamento de moléculas funcionais envolvidas aconteça, promovendo também o desespiralizar, o polimerizar de uma nova fita complementar às fitas moldes e até mesmo a identificação do fim da replicação.

QUADRO 2 - DNA polimerases de eucariotos

Componente	Função / atividade
alfa (α)	Iniciação de síntese de DNA nuclear; reparo; síntese de <i>primers</i>
delta (δ)	Síntese de fita tardia e reparo de DNA
epsilon (ϵ)	Síntese de fita líder
gama (γ)	Replicação e reparo de DNA mitocondrial

14 Assista vídeo resumitivo da replicação (<https://youtu.be/pin3569Qfgl?si=et4uSFDSCQp9eLuX>)

zeta (ξ)	Reparo de grandes segmentos
eta (η)	Reparo de grandes segmentos
teta (θ)	Reparo de DNA
lota (ι)	Reparo de grandes segmentos
kapa (κ)	Reparo de grandes segmentos
lambda (λ)	Reparo de DNA
mi (μ)	Reparo de DNA
sigma (σ)	Replicação e reparo de DNA nuclear
fi (ϕ)	Reparo de grandes segmentos
Rev1	Reparo de DNA

Fonte: Adaptado de Pierce, 2016

Apesar de bem semelhantes, a replicação em eucariotos difere em alguns pontos com a replicação procariota, seja em função do tamanho bem maior do genoma, seja pela estrutura linear dos cromossomos, seja pela presença das histonas que envolvem os nucleossomos¹⁵, estrutura que necessita ser prontamente refeita após as replicações.

Em comum com a duplicação em procariotas, consideram-se, para efeito de compreensão de todo o processo, três estágios: uma iniciação, um alongamento e uma terminação. No primeiro estágio, chamado iniciação, ocorre a localização dos sítios específicos do DNA cromossômico onde um complexo de enzimas de origem (ORC) capacita para o aportamento das demais proteínas responsáveis pela desnaturação da dupla hélice, inserção de moléculas para polimerização eficiente. O segundo estágio, o alongamento, trata-se da polimerização propriamente dita, protagonizada pela complementaridade das bases, pontes de hidrogênio e ligações fosfodiésteres. Por fim, na terminação, onde reconhecidos os pontos de terminação, se dissocia o complexo de replicação.

A iniciação da duplicação do material genético nos eucariontes tem origem na localização dos sítios de iniciação, também conhecidos como ORI, regiões ricas em nucleotídeos com bases nitrogenadas Adenina e Timina sucessivas, também chamadas regiões AT. Estes sítios estão situados em grande número ao longo da molécula de DNA de cada cromossomo a ser duplicado, para se garantir uma eficiência de replicação do DNA, visto o grande comprimento dos seus cromossomos (salvo não aconteceria em tempo hábil). Nas leveduras, principal modelo genético eucarionte, existem vários sítios de iniciação conhecidos com ARS (sequência de replicação autônoma), localizados ao longo de seu material genético.

Entretanto, a finalização da polimerização em eucariotos difere por ocorrer quando a cadeia do DNA termina na extremidade dos cromossomos, nos telôme-

¹⁵ Nucleossomos são estruturas do cromossomo eucarioto que compreendem DNA envolto por histonas.

ros. Considerando que a extremidade inicial de um cromossomo se caracteriza pela presença de um *primer* de RNA (iniciador), para algumas células, faz-se necessário a síntese de um fragmento de DNA substitutivo pela ação da enzima telomerase, uma enzima que promove uma transcrição reversa a partir da sequência telomérica¹⁶. A ação da telomerase é de ampliar a sequência do cromossomo molde, em sentido 3' - 5', permitindo que uma primase adiciona um *primer* à essa extensão, o que possibilita uma polimerização de uma sequência de DNA complementar ao fragmento do telômero da nova fita, impedindo uma diminuição de seu tamanho.

Resumo do processo de Replicação do DNA

Partindo de conceitos, princípios e fundamentos básicos que nos fazem conhecer características e mecanismos da replicação do material genético, podemos resumir que:

O DNA é o material genético mais estável e funcional das células vivas conhecidas.

A estrutura do DNA é de dupla hélice com dois filamentos de nucleotídeos, antiparalelos, espiralizados entre si e unidos por pontes de hidrogênios entre bases complementares de adenina e timina, e de citosina e guanina.

A replicação de uma dupla fita de DNA original para duas duplas fitas é um processo fundamental para a divisão celular eficiente e que garante a perpetuação das características entre células procariotas ou eucariotas.

A replicação ocorre de modo semiconservativo, onde uma dupla hélice de um DNA é replicada para duas duplas hélices formando dois novos DNAs a partir de um inicial; sendo que cada uma das novas duplas hélices são compostas de uma das fitas originárias e uma nova (formada da original por complementariedade de bases).

Uma quantidade de enzimas participa do processo garantindo o desenrolar do DNA helicoidal, a separação das duplas fitas complementares, a localização de regiões específicas para iniciação da replicação, a polimerização de nucleotídeos da fita complementar de cada fita molde separada, a terminação da replicação e a editoração das novas fitas formadas, conferindo e corrigindo possíveis erros de polimerização.

A polimerização das novas fitas formadas a partir das originais se dá nos dois sentidos, a partir da abertura de duas forquilhas obtidas pela formação de uma bolha de replicação.

A enzima responsável pela polimerização lê a fita molde conservada e insere novos nucleotídeos na extremidade 3' de um nucleotídeo pré-existente formando a fita complementar. Uma vez que a DNA polimerase apenas age no sentido 5'-3', para permitir a disposição de uma -OH livre em uma extremidade 3', uma fita paralela do DNA original é replicada para segmentos curtos de uma complementar nova (fita

¹⁶ Processo de ação da telomerase. acesse (<https://pt.khanacademy.org/science/biology/dna-as-the-genetic-material/dna-replication/a/telomeres-telomerase>).

descontínua), que depois serão unidos por meio de uma enzima chamada ligase.

2.2 Transcrição e Tradução

Os processos de **transcrição e tradução** são fundamentais para a vida, funcionamento e desenvolvimento celular, uma vez que representam o meio de produção de moléculas vitais, as **proteínas**. Por transcrição entende-se o processo de decodificação do DNA para formação dos diferentes tipos de RNA conhecidos. Por tradução, compreende-se o processo de polimerização de aminoácidos constituintes de uma proteína, a partir da leitura da mensagem contida nos RNA mensageiros, o que tem ocorrência em ribossomos. Em procariotos ocorrem em sequência e em mesmo ambiente, enquanto em eucariotos, que têm suas células compartimentalizadas, a transcrição é um evento nuclear e a tradução ocorre em ribossomos encontrados no citoplasma, livres ou nos retículos endoplasmáticos rugosos. Trata-se de processos complexos e dinâmicos, que envolvem enzimas e outras proteínas que promovem descompartimentação, estabilização, decodificação, edição, degradação, modulação e outras atividades em moléculas envolvidas na expressão da informação genética (Figura 3).

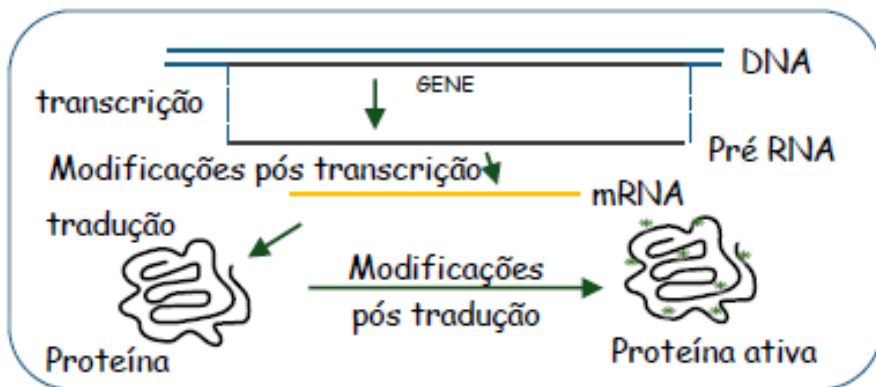


Figura 3. Sucessão de eventos relacionados com o fluxo da informação do material genético celular. Fonte: (https://pt.wikipedia.org/w/index.php?title=Modifica%C3%A7%C3%A3o_p%C3%B3s-traducional_de_prote%C3%ADnas&oldid=65693099) .

2.2.1 Transcrição do código Genético do DNA

Importância da Transcrição

Pode-se dizer que a transcrição é o processo de formação de uma molécula intermediária entre o DNA e a proteína codificada por esse DNA, o **ácido ribonucleico**, ou simplesmente, **RNA**. Uma importância de se existir tal molécula é que

todo RNA mensageiro de uma informação genética é usado pela célula para orientar a formação de toda proteína celular e não é passado para gerações futuras. Isso implica que qualquer dano em um RNA não será herdado e como mais de uma cópia pode ser produzida a partir de uma sequência molde de DNA, moléculas danificadas podem ser substituídas por outras produzidas corretamente em mesmo evento. O processo de transcrição é fundamental para a formação de proteínas em todas as células dos organismos, visto que toda informação para um funcionamento celular está contida em forma de código no DNA e, portanto, precisa ser decodificado antes de ser lido pelos ribossomos.

Pode-se comparar a uma linha de produção a partir de um molde, onde se o molde está intacto, mesmo que um produto saia com alguma avaria ou defeito, este poderá ser alvo de um processo de controle de qualidade, sem prejuízo da produção continuada. Constitui-se, portanto, como uma etapa precursora da polimerização dos aminoácidos na formação de todas as proteínas, que chamaremos de tradução.

Tipos de RNA

Há duas classes de RNA celulares: os **RNA mensageiros** (mRNA), moléculas intermediárias entre o código genético dos DNA e o produto gênico final por carregar a informação codificante de proteínas; e os **RNA funcionais**, atores envolvidos na estabilização de mRNAs e polimerização de aminoácidos das proteínas.

Existem RNA que são sintetizados apenas em organismos procariotos, os CRISPR RNA ou simplesmente **crRNA**.

Existem RNA encontrados apenas em células eucariotas, ditos não codificantes: RNA nuclear pequeno (**snRNA**); RNA nucleolar (**snoRNA**); microRNA (**miRNA**); RNA de interferência (**siRNA**); RNA pré-mensageiro (**pré-mRNA**); e os RNAs de interações com proteínas da subfamília Piwi (**piRNA**).

Alguns RNA são encontrados em células procariotas e eucariotas: RNA mensageiro (**mRNA**); RNA ribossômico (**rRNA**); e o RNA transportador (**tRNA**).

Um RNA mensageiro (mRNA) representa a sequência de nucleotídeos que carrega a informação genética transcrita a partir do DNA. Trata-se de uma sequência linear de trincas de bases (Figura 4) que codificam quais aminoácidos devem ser adicionados na formação de um polímero proteico. Localizam-se tanto no núcleo quanto no citoplasma de células eucarióticas, pois são sintetizadas a partir do DNA nuclear e são alvo da formação de proteínas após saírem do núcleo. Em célula procariota, que não possuem compartimentalização do DNA, a transcrição do RNA e a sua tradução ocorrem concomitantemente no mesmo compartimento celular.

O RNA transportador (tRNA) é a molécula citoplasmática que é responsável por conduzir ao ribossomo o aminoácido equivalente ao prescrito no RNA mensageiro. Em média são constituídos por 80 nucleotídeos e apresentam um formato semelhante a um trevo invertido por apresentar três alças ou grampos.

Cada aminoácido existente no meio citoplasmático é ancorado na extremidade 3' de um tRNA que tem sequência equivalente à presente no códon de um mRNA. Na alça anticódon do tRNA encontra-se uma trinca complementar ao códon do mensageiro a ser lido na polimerização. Promove-se a polimerização de aminoácidos a partir da leitura das trincas da sequência do mRNA.

RNA ribossômico (rRNA) é uma molécula básica constituinte de um ribossomo, isto porque cada ribossomo, é formado por duas subunidades, uma maior e outra menor, e cada uma delas contém um rRNA e várias proteínas. São moléculas predominantemente localizadas no citoplasma.

RNA nuclear pequeno (snRNAs) são tipos de RNA de localização nuclear, relacionados com processamento e edição do pré-RNA em mRNA.

RNA nucleolar pequeno (snoRNA) tem função no processamento e montagem de rRNAs a nível de núcleo celular.

O microRNA (miRNA) é um tipo de RNA de importância regulatória por inibir a tradução de um mRNA de mesma origem pre-RNA comum. São de localização nuclear e citoplasmática.

RNA de interferência pequeno (siRNA) tem função relacionada com o processamento do RNAi (RNA de interferência) envolvido na degradação de outros tipos de RNA agindo, portanto, tanto no citoplasma como no núcleo.

O RNA de interação com Piwi (piRNA) é uma classe de RNA relacionada com a supressão da expressão de elementos transponíveis em células reprodutoras, tendo sido encontrada inicialmente em núcleo e citoplasma de células de testículos de mamíferos.

Recentes estudos identificaram um pequeno tipo de RNA em procariotos que são responsáveis por auxiliar na degradação de DNA exógeno, chamados CRISPR RNAs (crRNA).

Muitos genes eucarióticos possuem regiões codificantes chamadas éxons e regiões não codificantes chamadas íntrons (sequências intervenientes ou espaçadoras de éxons). Ambos, íntrons e éxons são transcritos em pré-RNA, um tipo de RNA em estado bruto, que após lapidado por um processamento ou edição, chamado *splicing*, têm os íntrons removidos e os segmentos éxons unidos formando o RNA maduro (Figura 5).



Figura 4. Representação da estrutura da fita de mRNA. Fonte: Genes to proteins: Central Dogma | BIO103: Human Biology por Autor Desconhecido (adaptado).

2.2.1.1 Estrutura básica dos RNA

A **estrutura de um ácido ribonucleico (RNA) padrão** é de uma fita de nucleotídeos portadores de quatro tipos possíveis de bases nitrogenadas: adenina (A), uracila (U), guanina (G) ou citosina (C), ligadas entre si por ligações fosfodiésteres formadas entre uma hidroxila 3' livre de um nucleotídeo e um grupo fosfato 5' de um nucleotídeo seguinte. Sua mensagem está codificada em códons, que são trincas de nucleotídeos que codificam aminoácidos.

Caracteriza também um RNA a presença de uma hidroxila livre no carbono 2' do açúcar ribose, que o torna uma molécula mais reativa e degradável quando comparada ao DNA.

Mesmo habitualmente se apresentando em fita única (estrutura primária), um RNA pode exibir uma forma de grampo (grampo-alça), em uma região curta, complementar e antiparalela de sua fita. São chamadas estruturas secundárias de RNA estão relacionadas com a função do tipo do RNA, como veremos a seguir.

2.2.1.2 Estrutura do RNA transportador

Todo tRNA tem uma estrutura secundária em forma de trevo em virtude dos pareamentos de nucleotídeos complementares de sua fita, com quatro regiões reativas principais: o braço aceptor, o braço T Ψ C, que tem seu nome pela sequência ribotimidina-pseudo-uridina-citosina; o "braço extra", que possui um número variável de nucleotídeos; o braço DHU, que contém várias dihidrouracilas; e o braço do anticódon. Cada braço é composto por uma haste fita dupla e uma alça formada por uma sequência não pareável (Figura 5).

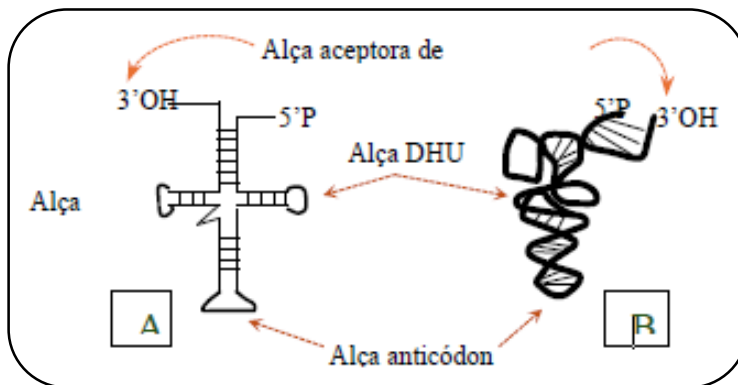


Figura 5. Estrutura bidimensional plana (A) e dobrada (B) de um RNA transportador (tRNA) e suas áreas reativas. Componentes da estrutura de tRNA (C). Fonte: Suzuki et al. 2016 (adaptado).

Um braço funcional para ativação de um aminoácido específico, chamado de braço acceptor, é um sítio de fixação do aminoácido. Inclui as extremidades 3' e 5' do tRNA sendo a extremidade 3' finalizada por uma trinca nucleotídica citosina, citosina, adenina (CCA-) conservada para todos os tRNA conhecidos.

O segundo braço principal é o DHU (dihidrouridina, uma base modificada de uracila), que promove o acoplamento do transportador com a enzima aminoacil-tRNA sintetase, responsável pela formação do complexo tRNA-aminoácido.

O braço que se mostra em polo oposto ao braço acceptor é a **alça anticódon**, onde existe uma trinca complementar ao códon do mRNA para onde deverá ser levado o aminoácido, o que garantirá a ordem correta de aminoácidos para a informação gênica.

Um quarto braço do tRNA apresenta uma sequência em sua alça composta por Timina (T), não comumente presente em RNAs, a pseudouridina (ψ) uma base modificada de uracila, e citosina (C), é o braço T ψ C, importante no ancoramento do tRNA ao ribossomo (Figura 6A).

2.2.1.3 Estrutura do ribossomo e do RNA ribossomal

O ribossomo é a organela onde as informações genéticas de um DNA, transcritas para um mRNA são traduzidas para formação de polipeptídios (as proteínas). Constituem as organelas mais numerosas em uma célula (aproximadamente 20.000 em células bacterianas), e representam mais de 80 % do RNA total celular.

Estruturalmente, se apresentam como uma complexa estrutura de RNA e numerosas proteínas. Um ribossomo funcional em seu estado livre é composto por duas subunidades, uma subunidade menor e uma subunidade maior, que devido a sua velocidade de precipitação em uma centrifugação são denominadas em unidades Svedberg (S). Em células bacterianas o tamanho de um ribossomo é de 70S, constituído por uma subunidade maior de 50S (com mais de 3.000 nucleotídeos e cerca de 30 proteínas) e uma menor de 30S (com 1500 nucleotídeos e 20 proteínas) (Figura 7). Já em células eucariotas, o tamanho de um ribossomo é medido em 80S, distribuídos em 60S na subunidade maior (com 5.000 nucleotídeos e 49 proteínas) e a subunidade menor com tamanho de 40S (1.900 nucleotídeos e 33 proteínas).

Os rRNA variam entre espécies, mas são conservadas em indivíduos de uma mesma espécie. Existem três tipos de rRNA em bactéria, o 23S rRNA, o 16S rRNA e o 5S rRNA, e seus genes são encontrados dispersos em cromossomos procariotos.

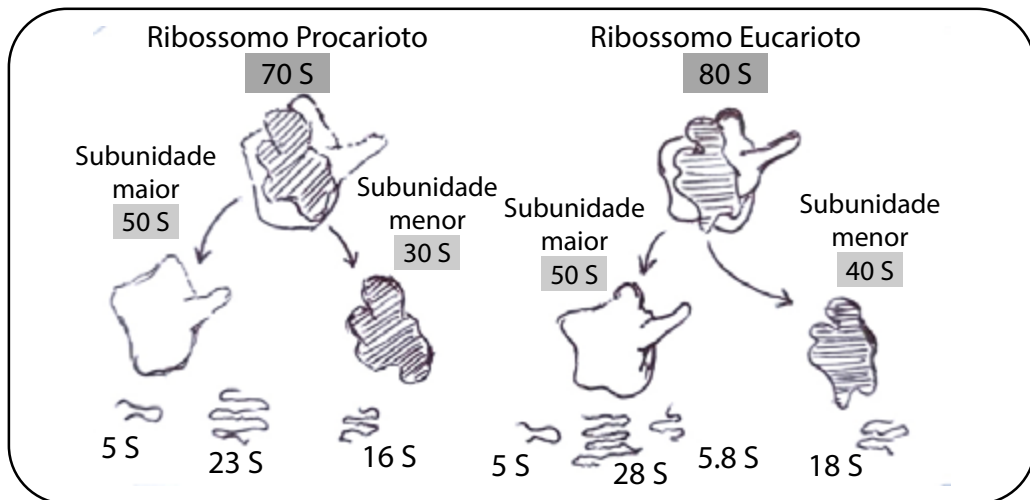


Figura 6. Imagem com constituintes das subunidades ribossomais. Fonte: Pierce, 2016. (adaptado)

Em células de eucariotos são conhecidos 4 tipos de rRNA (28S rRNA, 18S rRNA, 5,8S rRNA e 5S rRNA) codificados por genes que se encontram agrupados e em sequência.

Após transcritos, os rRNA são processados e associados com proteínas na formação dos ribossomos maduros. Em eucariotos o processamento ocorre com ação de pequenos RNA nucleolares (snoRNA). O processamento dos rRNA precursores em procaríotos e eucariotos incluem metilação de bases nitrogenadas, clivagem em sequências menores e terminações aparadas.

2.2.1.4 Etapas da Transcrição

Semelhantemente ao processo de replicação de DNA, a transcrição de DNA em RNA também é estudada em três estágios: **iniciação, alongamento e terminação.**

QUADRO 3 - Etapas e processos de uma transcrição

ETAPA	PROCESSO
Iniciação da transcrição	<ul style="list-style-type: none"> • Localização do sítio de iniciação; • Aporte do complexo transcricional.
Alongamento de um RNA	<ul style="list-style-type: none"> • Polimerização de uma fita de RNA a partir de uma sequência da fita de DNA
Terminação da transcrição	<ul style="list-style-type: none"> • Localização do sítio de terminação; • Desacoplamento do complexo transcricional.

A transcrição é um processo assimétrico uma vez que apenas uma das fitas do DNA é usada por vez. A fita do DNA usada para se obter um transcrito de RNA é chamada fita molde de DNA e a sua fita complementar antiparalela é chamada de fita codificante, uma vez que é semelhante à fita de RNA obtida pela transcrição.

Iniciação da transcrição

A **iniciação da transcrição do DNA em RNA** se dá quando o aparato transcricional, composto por um pool de enzimas chamado de fatores de iniciação de transcrição, percorre a molécula de DNA de um cromossomo ou de uma mitocôndria até localizar o sítio inicial do fragmento do gene a ser transcrito. O sítio de iniciação é localizado pelos fatores de iniciação a partir da sinalização da região promotora, que é uma região do DNA alvo rica em grupos metila, que é conservada em uma espécie, e que precede o fragmento que contém a informação. Assim, pode-se dizer que uma região promotora de um gene é uma sequência do DNA que aponta onde deve se começar uma transcrição de uma sequência do DNA em uma fita de RNA.

Quando os fatores de iniciação do complexo transcricional aportam na região promotora, precursora do gene, inicia-se uma sequência de ativações das enzimas presentes no complexo, promovendo a linearização do fragmento de DNA alvo, a separação das fitas e estabilização da bolha de transcrição com auxílio da RNA polimerase.

Alongamento da fita de RNA

A etapa seguinte, conhecida como polimerização ou **alongamento**, caracteriza-se como o processo em que, a partir da leitura das bases nitrogenadas de uma das fitas de DNA do fragmento alvo, procede-se a **formação de uma fita de RNA complementar ao fragmento de DNA molde** (Figura 4). A inclusão dos nucleotídeos de RNA que formam a fita recém-formada obedece a complementaridade das bases purínicas e pirimídicas, sempre pela combinação de citosina (C) com guanina (G) e de adenina (A) com uracila (U).

Em procariotos a transcrição do DNA em RNA depende da **RNA polimerase**, uma enzima multimérica que contém um fator de ligação ao promotor do DNA (conhecido como fator sigma), um cerne polimerizador de uma fita de RNA complementar e subunidades complementares envolvidas com o desenovelamento e linearização da fita de DNA.

Nos eucariotos, são conhecidas três RNA polimerases envolvidas no processo de transcrição da informação genética contida no DNA, as RNA polimerases I, II e III, cada uma relacionada a um tipo específico de molécula de RNA. A RNA polimerase I transcreve o **RNA ribossomal** (rRNA) relacionado com a formação da organela conhecida como ribossomo, importante na montagem

das proteínas pela correlação de aminoácidos com os códons do RNA mensageiro. A RNA polimerase II está relacionada basicamente com a transcrição do pré-**RNA mensageiro** inicial (pré- mRNA), que é a molécula que contém a informação das proteínas que devem ser produzidas pela célula. A RNA pol II também transcreve outras moléculas de RNA de pequenas sequências, como RNA nuclear (snRNAs), RNA nucleolar (snoRNAs) e microRNA (miRNA). A terceira classe de RNAs transcritos em eucariotos é processada pelas RNA polimerases III, que transcrevem os **RNA transportadores** (responsáveis por carrear os aminoácidos até aos ribossomos), e também outros RNAs de pequeno porte como alguns miRNA, pequenos rRNAs e alguns snRNAs. Em vegetais, são conhecidas duas outras RNA polimerases, a RNA polimerase IV, relacionadas a RNA de interferência (siRNA) e a RNA polimerase V (que participam da formação da heterocromatina em plantas).

Terminação da transcrição do RNA

A etapa final da transcrição é a terminação, caracterizada pelo término da leitura da sequência DNA molde e desacoplamento do complexo transcricional. Isso ocorre quando a RNA polimerase termina de adicionar nucleotídeos à extremidade hidroxila livre da fita de RNA crescente ao chegar a uma sequência de finalização no DNA molde.

Basicamente uma sequência de eventos acontece nessa etapa. A RNA polimerase para de sintetizar a fita, a molécula de RNA é liberada da holoenzima da polimerase, a RNA recém-sintetizada é desacoplada do DNA molde e, por fim, o complexo transcricional se desconecta do DNA.

A transcrição termina quando a RNA polimerase transcreve um terminador e para de sequenciar a fita de RNA. Em bactérias, dois tipos de terminador são conhecidos, o dependente do auxílio da proteína Rho e o tipo independente da proteína Rho para terminar. Nos eucariotos, cada classe de RNA polimerase usa um mecanismo específico. Para o término da transcrição pela RNA polimerase I, ocorre uma ligação desta com uma sequência de DNA de terminação. A RNA polimerase III para após polimerizar uma sequência de nucleotídeos uracila na fita produzida. A ação da RNA polimerase II é interrompida quando uma enzima exonucleásica se liga a uma extremidade 5' exposta da fita recém-sintetizada e alcança a polimerase após degradar fragmento excedente.

Uma etapa pós transcricional importante que ocorre em organismos eucariotes é o **splicing** (processo de edição da fita de RNA recém-sintetizada) que acontece porque como as regiões codificantes (éxons) do fragmento do DNA molde encontram-se espaçadas por longas regiões não codificantes (íntrons), estas precisam ser retiradas antes das moléculas dos RNAs mensageiros seguirem para o citoplasma para serem lidas pelos ribossomos (Figura 7).

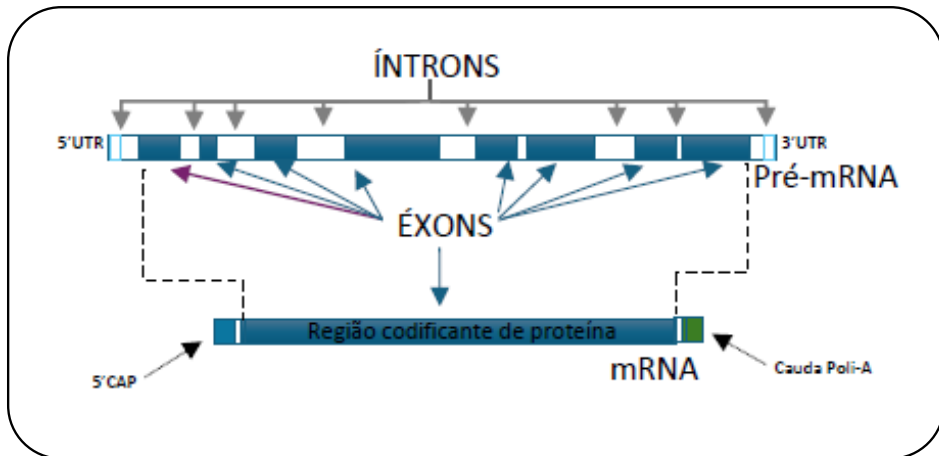


Figura 7. Edição (*splicing*) do pré-mRNA pela extração de íntrons. Adaptado: Precursor mRNA - Wikipedia por Autor Desconhecido (adaptado).

O processo que envolve a escolha das sequências de DNA alvo para transcrição, quais regiões devem ser bloqueadas ou liberadas, é fundamental para a vida celular e conseqüentemente dos organismos vivos. Esse processo é conhecido como regulação gênica e será abordado com detalhes em capítulo seguinte.

2.2.2 Tradução da Mensagem do DNA

A tradução é a etapa que representa a formação de polímero de aminoácidos a partir da leitura dos mRNA feita nos ribossomos. Como os ribossomos estão presentes ativamente livres ou ligados aos retículos endoplasmáticos (RER), conclui-se que diferentemente da replicação ou transcrição, o processo de tradução da mensagem genética em proteínas ocorre no citoplasma celular.

O código genético¹⁷ do DNA é constituído por trincas ou códon, no qual cada três nucleotídeos codificam cada aminoácido de uma proteína.

Da mesma forma dos outros processos moleculares (replicação e transcrição), a **tradução** pode ser classificada em três fases básicas: iniciação, alongamento e terminação. Nestas três fases, cinco etapas são necessárias para formação de proteínas eficientes.

1. Ativação do aminoácido;
2. Iniciação da tradução;
3. Alongamento da proteína (elongação);
4. Terminação da tradução do RNA mensageiro;
5. Processamento pós-traducional (dobramento).

¹⁷ Quadro dos códon do código genético onde constam as trincas e os aminoácidos específicos. Acesse (https://pt.wikipedia.org/wiki/C%C3%B3digo_gen%C3%A9tico).

Ativação do aminoácido - um estágio pré-iniciação

Um importante estágio pré-iniciação é a ligação dos aminoácidos aos tRNAs, que produz o que chamamos de **tRNA carregado ou tRNA aminoacetilado**. Lembremos que todos tRNA têm uma sequência CCA na sua extremidade 3', justamente onde se conecta o grupo carboxila (COO-) do aminoácido equivalente e que cada tRNA tem especificidade para um tipo de aminoácido.

A formação de tRNA aminoacetilado é catalisada pela **aminoacil-tRNA sintetase** (Figura 8). Existem 20 aminoacil-tRNA sintetases descritas, enzimas que garantem a especificidade dos aminoácidos aos seus tRNA exclusivos, uma vez que existe uma para cada aminoácido existente.

Cada sintetase responsável pela união do aminoácido ao seu tRNA específico é uma molécula proteica que tem dois sítios que recebem esses elementos e promovem sua união. A sintetase reconhece o seu aminoácido específico pela carga, tamanho e grupo R, o que o torna único (Figura 8B e 8C) e reconhece o seu tRNA exclusivo pela sequência de nucleotídeos (Figura 8D).

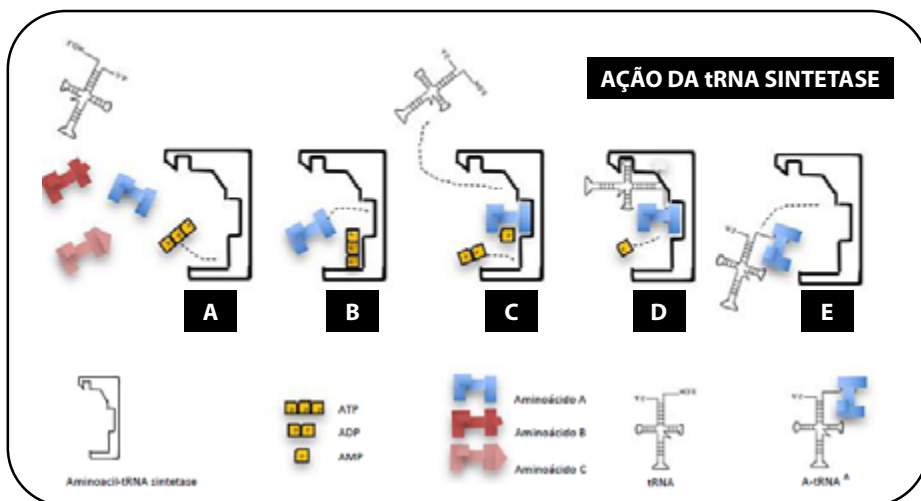


Figura 8. Processo de ligação dos componentes de formação de um tRNA carregado. A enzima é carregada com ATP (A- B) e, na sequência, com seu o aminoácido específico pela liberação de íons fosfato (C). O tRNA específico do aminoácido é ancorado na enzima para formar a Aminoacil-tRNA (D), que então pode ser liberado no citoplasma (E). Fonte: Suzuki et al.a, 2016 (adaptado).

NA aminoacetilado é formado em duas etapas que envolvem um trifosfato de adenosina ou **ATP**, gerador de energia. Primeiramente, o aminoácido reage com a ATP, produzindo a aminoacil-AMP e liberando dois íons fosfato (PPi) (Figura 8C). Em sequência, o aminoácido é acoplado ao seu tRNA específico pela ação da sintetase pela liberação da adenosina mono fosfato (AMP) formando um **aminoacil-tRNA** (Figura 8D). Para fins de nomenclatura, um tRNA carregado com seu aminoácido

cido equivalente é descrito pela sigla do aminoácido ligada por hífen à sigla tRNA acompanhada pela sigla do aminoácido no modo sobrescrito (p.ex. Met-tRNA^{Met}).

Iniciação da tradução da mensagem

A etapa de iniciação da tradução ocorre quando o ribossomo, responsável pela leitura da mensagem pelos tRNA, é acoplado ao mRNA.

São componentes requeridos na fase de Iniciação da Tradução de RNA mensageiro:

- mRNA - portador da mensagem;
- Subunidades Ribossomais - sítio do processo;
- tRNA-iniciador - transportador do aminoácido iniciante;
- proteínas fatores de iniciação - complexo de funcionalidade; e
- trifosfato de guanosina (GTP) - gerador de energia

Montagem do maquinário de tradução

Tanto em procariontes como em eucariontes, o acoplamento se dá inicialmente pela subunidade menor ribossomal ao mRNA, seguido pelo acoplamento da subunidade maior (Figura 9A). Portanto, como os ribossomos estão normalmente disponíveis nas células com suas subunidades acopladas, faz-se necessário a dissociação prévia das subunidades de um ribossomo, que permite a interação primeiro da subunidade menor ribossomal com o mRNA, condição encontrada tanto nos seres procariontes como nos eucariontes.

O RNA mensageiro representa uma cópia da informação gênica (contida no DNA), com a missão de orientar a síntese proteica que acontece no citoplasma.

Em seguida um tRNA ancorado no ribossomo pelo seu braço T ψ C, é acoplado ao códon de mRNA por meio de pareamento de bases e pontes de hidrogênio entre códon iniciador do mensageiro e o anticódon do transportador equivalente. Nos procariontes, esse processo de iniciação é mediado por vários fatores de iniciação como os IF-1, IF-2 e o IF-3.

Na sequência, uma subunidade maior ribossômica junta-se ao aparato, constituindo o complexo traducional (Figura 09A). Inicia-se a tradução quando se forma o complexo traducional (Figura 9B).

Nos procariontes a formação do complexo de iniciação se inicia com o acoplamento da subunidade menor (30S) na extremidade 5' do mRNA, a partir da complementaridade de bases nucleotídicas do rRNA ribossômico com bases do mRNA, em uma região chamada sequência Shine Dalgarno (sequência de aproximadamente 40 nucleotídeos que comporta o códon da metionina, AUG). Uma outra característica procarionte é a realização sequencial da transcrição e tradução, uma vez que não existe compartimentalização nuclear.

Alongamento da proteína

O alongamento de uma proteína é a formação da cadeia polipeptídica de aminoácidos no citoplasma das células.

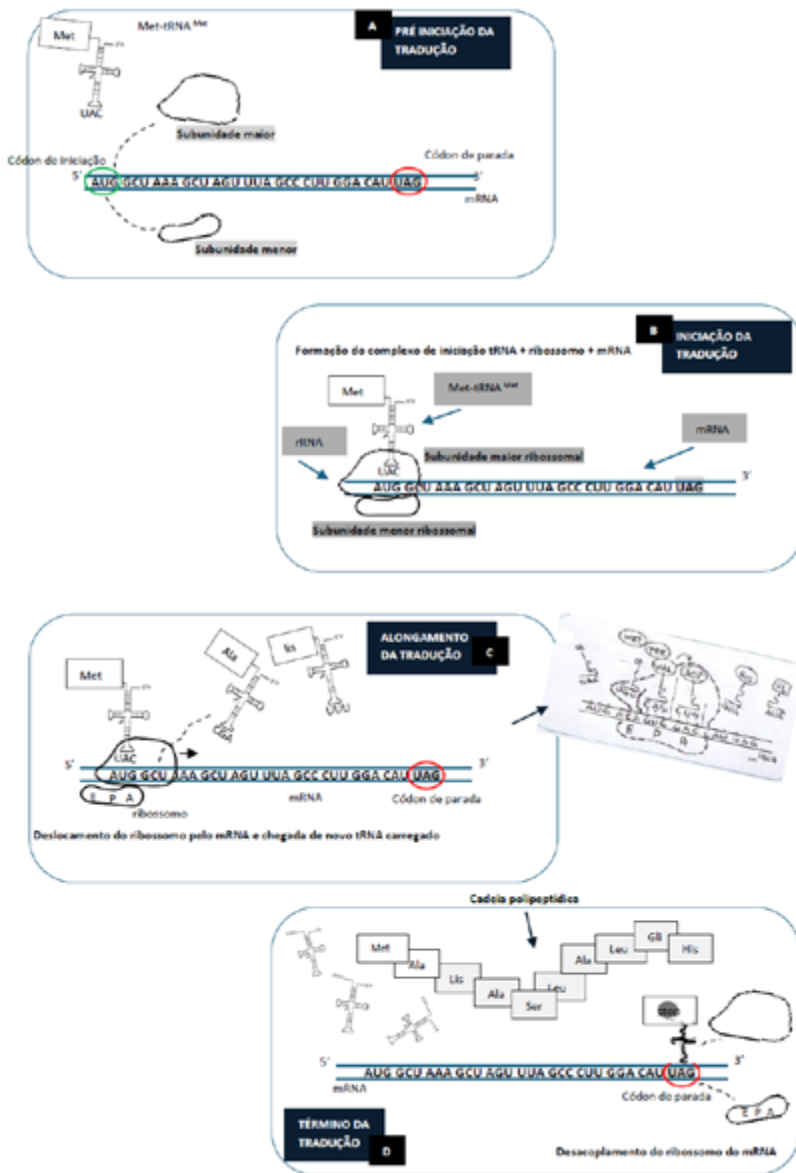


Figura 9. Esquema da tradução bacteriana: A- Componentes do complexo de tradução: Ribossomos, tRNA+aminoácido de iniciação e mRNA. B- Início da formação com complexo C- Esquema da dinâmica da tradução com destaque para a ligação peptídica (desenho); D- Término de uma tradução com desacoplamento do complexo. Fonte Pierce, 2016 (adaptado).

QUADRO 4 - Componentes requeridos na fase de alongamento

Componentes	Função
mRNA	Carreia a informação genética
Subunidades ribossomais	Componentes para interação entre os mRNA e tRNA
tRNA-aminoacil	Transportador que carrega um aminoácido para a reação
Proteínas fatores de alongação	Promovem a polimerização dos aminoácidos
Fatores de dissociação	Responsáveis pela dissociação dos ribossomos ao final da tradução
GTP (guanosina trifosfato)	Gerador de energia para reação de polimerização

A etapa de alongamento (Figura 09 C) ocorre quando o aparato traducional promove a inserção progressiva dos aminoácidos equivalentes aos códons contidos no mRNA maduro ao longo da movimentação ribossomal.

A síntese tem origem pela extremidade amino da proteína e segue crescendo pela adição de novos aminoácidos na extremidade carboxila. O ribossomo se desloca ao longo do filamento de mRNA semelhantemente a um vagão sobre trilhos.

À medida que cada códon do mRNA entra no ribossomo este passa a ocupar um sítio chamado aminoacil (sítio A), onde também um tRNA carregado é inserido. Como o ribossomo continua a se deslocar sobre o mRNA, ocorre um salto para a trinca seguinte, passando o códon inicial (geralmente uma metionina, na maioria dos seres vivos) a ocupar um segundo sítio, chamado sítio peptidil (ou sítio P). Com a consequente ocupação do sítio A por outro tRNA portador de um aminoácido relativo ao códon mRNA, temos os sítios peptidil e aminoacil ambos ocupados (Figura 9C – desenho).

Ocorre a formação de uma ligação peptídica entre os aminoácidos presentes nos sítios ribossômicos P e A, passando o aminoácido presente no tRNA do sítio P para sobre o aminoácido que se encontra no tRNA do sítio A. Uma vez que a dinâmica de deslocamento do ribossomo sobre a fita mensageira é contínua, ao mesmo tempo que ocorre a ligação entre os dois primeiros aminoácidos, os tRNA passam a ocupar novos sítios dentro do ribossomo, ficando o tRNA que ocupava o sítio A, agora no sítio P e carregado com os dois aminoácidos acoplados. No sítio A que agora está sobre o códon seguinte do filamento do RNA mensageiro, será adicionado um novo tRNA carregado com seu aminoácido equivalente. Ao se repetir o processo de ligação peptídica, os dois primeiros aminoácidos que estavam acoplados ao primeiro tRNA inserido, “pulam” para sobre o aminoácido do tRNA

que ocupa o sítio A, que por sua vez é deslocado para o sítio P em virtude da movimentação sentido 5'-3' do tRNA do ribossomo. Essa etapa é também conhecida como **translocação do ribossomo**.

Terminação

A etapa final da tradução se refere ao momento em que o ribossomo encontra o códon de parada do mRNA promovendo o fim da leitura e o desacoplamento do complexo traducional. Com a passagem do(s) aminoácido(s) acoplado(s) ao tRNA do sítio P para sobre o tRNA do sítio A, o tRNA do sítio P passa a ficar descarregado de aminoácidos, sendo desacoplado do complexo mRNA/ribossomo ao ocupar o um terceiro sítio do ribossomo, o sítio de saída ou Sítio E ("Exit"). Esse processo se repete por ação dos fatores de alongamento e uso da energia das GTPs. Visto que não existem tRNAs complementares ao códon de terminação do mRNA, nenhum tRNA é inserido, mas sim proteínas chamadas de fatores de liberação que se ligam ao ribossomo promovendo mudança conformacional e causando a dissociação dos RNA envolvidos.

2.3 Antibióticos de ação no fluxo da informação genética

Talvez a maior contribuição da microbiologia tenha sido a descoberta dos antibióticos, substâncias produzidas por bactérias e fungos que inibem o crescimento de outros microrganismos em sua volta. Vários dos antibióticos isolados e comercializados são usados no tratamento de doenças em humanos e outros animais, em vegetais e até como marcadores genéticos em pesquisas científicas.

Os quimioterápicos antimicrobianos e antibióticos, quanto ao seu mecanismo de ação (Figura 10), podem ser agrupados em cinco classes: os inibidores da síntese da parede celular, os de interferência da membrana celular, os fármacos antimetabólicos, os de interferência na replicação genética e os de interferência na síntese proteica (de ação inibitória na síntese de RNA ou na síntese da cadeia peptídica).

A farmacogenética é parte da farmacologia clínica que pesquisa como diferentes genes podem apresentar respostas frente a uma mesma droga e permite aperfeiçoar a personalização terapêutica. A genética tem contribuído para desenvolvimento de fármacos utilizados tanto na medicina humana quanto na medicina veterinária, com destaque para os antibióticos. Dessa forma, a farmacogenética tem se constituído uma tecnologia que se propõe a pesquisar e produzir novos fármacos de maneira a reduzir os efeitos colaterais e aumentar a eficácia da resposta aos medicamentos.

Os antibióticos de ação molecular podem ser classificados como a) os inibidores das topoisomerasas, b) os inibidores da transcrição e c) os inibidores da tradução.

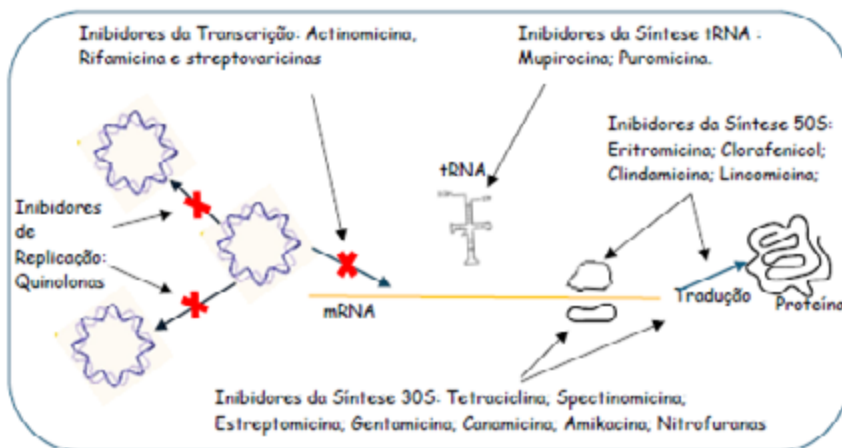


Figura 10. Desenho esquemático sobre mecanismo de ação de antimicrobianos moleculares. Adaptado de: <https://edisciplinas.usp.br/mod/resource/view.php?id=4978429>

Os **fármacos de interferência na replicação de DNA** são as quinolonas e os nitroimidazólicos. As quinolonas atuam através da inibição de uma ou de ambas as topoisomerases tipo II procarióticas, na DNA girase (topoisomerase II) e na topoisomerase IV, enzimas diretamente envolvidas no enovelamento e desenovelamento do material genético, comprometendo a leitura e replicação do DNA. As quinolonas inibem primariamente a DNA girase nos microrganismos Gram-negativos e também inibem a topoisomerase IV nos microrganismos Gram-positivos, como *Staphylococcus aureus*. Uma vez que a replicação do material genético é o que possibilita a reprodução de uma célula por ser a primeira etapa da divisão celular, sua interrupção é letal para os microrganismos sob o efeito desses fármacos.

Dentre os conhecidos como **inibidores da transcrição**, os derivados semisintéticos do antibiótico de ocorrência natural (rifamicina B), exercem sua atividade bactericida através da formação de um complexo estável com a RNA polimerase DNA-dependente bacteriana, inibindo, assim, a síntese de RNA quando bloqueia o alongamento desta molécula após o início da transcrição. Estes fármacos têm alta seletividade pela RNA polimerase bacteriana, resultando em baixa incidência de efeitos adversos por ser bem tolerada em mamíferos.

Os mecanismos moleculares que envolvem a expressão das características e propriedades de um microrganismo dependem do seu material genético e, igualmente, da sua competência transcricional e traducional. Antibióticos como o cloranfenicol, os macrolídeos e as tetraciclinas atuam na interferência da síntese proteica, no evento de tradução do RNA em proteínas. Ao inibir a formação do complexo enzimático responsável pela tradução do código genético presente nos RNA mensageiros, esses quimioterápicos comprometem a sobrevivência dos agentes infecciosos por não permitir se formarem suas proteínas.

Muitos fármacos antimicrobianos são de **ação sobre a tradução** (Quadro 4) pela especificidade sobre agentes infecciosos, isso porque muitos aspectos da tradução das bactérias são significativamente diferentes da tradução em eucariotos. A tetraciclina pertence a uma classe de antibióticos que se ligam ao sítio A dos ribossomos bacterianos, bloqueando a entrada dos tRNA carregados, mas não exercem efeito sobre os ribossomos eucariotos. A neomicina também exerce essa ação bloqueadora em sítio A bacteriano, enquanto o cloranfenicol age sobre a subunidade maior bloqueando a ligação peptídica. A eritromicina tem efeito inibidor na translocação (movimento ribossomal). A estreptomomicina se liga a subunidade menor ribossomal prejudicando o início da tradução.

A puromicina é um antibiótico usado para destruir células cancerosas devido a sua ação inibidora tanto em células bacterianas como em células eucariotas, por se assemelhar a extremidade aceptora 3' do tRNA carregado, uma estrutura conservada tanto em procariotos como em eucariotos.

QUADRO 5 - Locais e mecanismos de ação dos Antibacterianos inibidores da tradução

FÁRMACO OU CLASSE	LOCAL DE AÇÃO	MECANISMO DE AÇÃO
Fármacos dirigidos contra subunidade ribossômica 30S		
Aminoglicosídeos	tRNA 16S	Induzem uma leitura incorreta; interrompem a síntese proteica em concentrações mais altas
Espectinomicina	tRNA 16S	Inibe a translocação
Tetraciclina	tRNA 16S	Bloqueiam a ligação do aminoacil tRNA ao sítio A
Fármacos dirigidos contra a subunidade ribossômica 50S		
Macrolídeos	tRNA 23S	Inibe a translocação
Cloranfenicol	tRNA 23S	Inibe a peptidil transferase ao interferir no posicionamento do tRNA
Lincosamidas	tRNA 23S	Inibe a peptidil transferase ao bloquear a cadeia polipeptídica em crescimento e ao inibir o sítio A e o sítio P
Estreptograminas	tRNA 23S	Inibe a peptidil transferase; provável superposição com o mecanismo de ação dos macrolídeos
Oxazolindionas	tRNA 23S?	Ainda não conhecido
Oxazolindionas	tRNA 23S?	Ainda não conhecido

Fonte: Machado e Souza Júnior, 2019.

REFERÊNCIAS BIBLIOGRÁFICAS

GARDNER, E.J.; SNUSTAD, D.P. Genética. 7ª edição. Rio de Janeiro: Editora Guanabara Koogan. 1987. 497p.

BROWN, T.A. Genética: Um Enfoque Molecular, 3ª edição, Ed. Guanabara Koogan, 1999, 336p.

BURNS, G.W.: BOTTINO, P.J. Genética. 6ª edição, Rio de Janeiro: Ed. Guanabara Koogan. 1991. 381p.

KREUZER, Helen e MASSEY, Adrienne Engenharia genética e biotecnologia. Tradução: Ana Beatriz Gorini da Veiga, [et al.]. – 2.ed. – Porto Alegre: Artmed, 2002.

MACHADO, E.A. e SOUZA JÚNIOR, C.G. Antibióticos de ação gênica mais utilizados pelos produtores rurais no município de Senador Rui Palmeira – Alagoas. Rev. Agr. Acad., v.2, n.3, Mai/Jun (2019).

MAIA, Maria M. D. et al. Genética geral para universitários. Maria de Mascena Diniz Maia, coordenadora; Paulo Roberto Eleutério de Souza ...[et al.]. - 1. ed. - Recife: EDUFRPE, 2015.147 p.

PIERCE, Benjamim A. Genética: um enfoque conceitual. Tradução: Beatriz Araujo do Rosário.- 5.ed. – Rio de Janeiro: Guanabara Koogan, 2016.

RAMALHO, Magno A.P. et al. Genética na Agropecuária. 5.ed. revisada. Editora UFLA. Lavras, MG, 2012. 359p.

SUZUKI, David T.; GRIFFITHS, Anthony J.F.; MILLER, Jeffrey H.; LEWONTIN, Richard C. Introdução à Genética. Tradução: João Paulo campos e Paulo Armando Motta. – 4.ed. - Rio de Janeiro: Guanabara Koogan, 2016.

CAPÍTULO 3: MECANISMOS DE CONTROLE DE EXPRESSÃO GÊNICA EM ORGANISMOS EUCARIOTOS

Maria de Mascena Diniz Maia

3.1 Introdução

Antes de começar a escrever sobre o tema “**controle da expressão gênica em eucariotos**”, começo descrevendo alguns conceitos básicos da biologia molecular que acredito serem fundamentais para entendimento das denominações que envolvem as propriedades dos genes e melhor compreensão dos itens descritos.

Orientação de gene e transcrição

A orientação do gene, por convenção, é determinada pela cadeia de codificação, assim, o início de um gene é chamado de extremidade 5' e o fim de um gene é chamado de 3'. Na Figura 1 pode ser observado que na mesma região existem dois genes que são transcritos em sentidos opostos: o gene WRAP53 para a direita e o gene P53 para a esquerda.

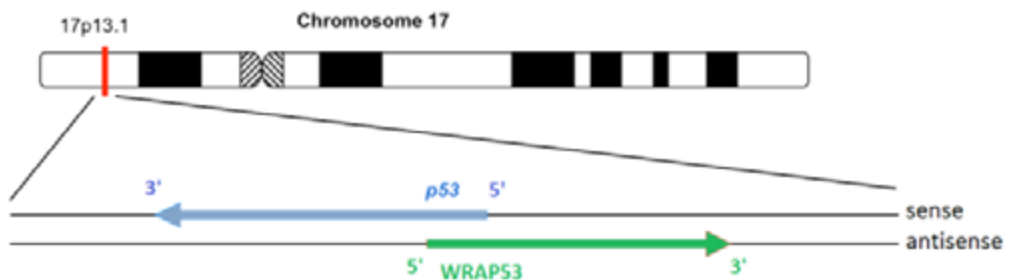


Figura 1. Fonte: Técnicas em Biologia Molecular I. Núcleo de Aprimoramento Científico NAC, 2019.

O que são *exons*?

Exons são partes de genes (DNA) que são convertidas em RNA mensageiro maduro pelo processo de transcrição. Este RNAm maduro é traduzido em proteínas através do RNA transportador (RNAt).

O que são *introns*?

Introns são partes de genes (DNA) que não codificam diretamente para proteínas e são comumente encontrados em organismos eucariotos multicelulares, como os seres humanos, porém, são menos comuns em eucariotos unicelulares, como leveduras e mais raros em bactérias.

Como os *introns* são removidos?

Os *introns* estão presentes no transcrito inicial do RNAm, chamado pré-RNAm, mas devem ser removidos para que o RNAm possa ser traduzido em proteínas. Assim, o pré-RNAm sofre um processo, conhecido como *splicing*, para produzir o RNAm maduro. É fundamental que os *introns* sejam removidos com precisão, caso contrário, pode resultar em uma proteína defeituosa.

Splicing de RNAm

O processamento de RNAm ocorre em locais de emenda especiais. Estes tendem a começar com o dinucleotídeo GU na extremidade 5' e AG na extremidade final 3'. O processo é realizado por pequenas ribonucleoproteínas nucleares (snRNPs) que se ligam às extremidades 5' e 3' do *intron* e fazem com que ele forme um *loop*. Desta forma, o *intron* é removido da sequência e os dois *exons* restantes se ligam entre si (Figura 2).

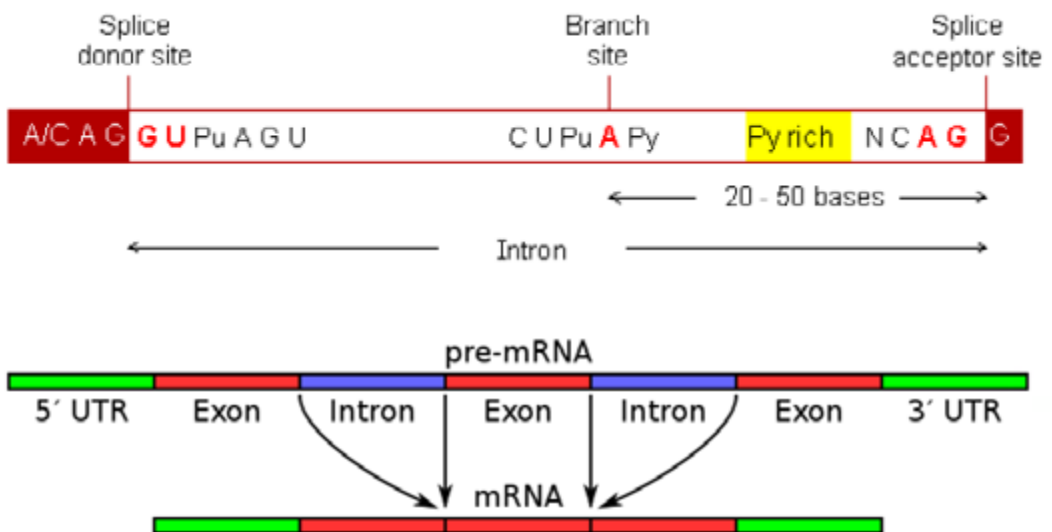


Figura 2. Fonte: https://www.wikidoc.org/images/1/17/Pre-mRNA_to_mRNA.png

Splicing alternativo

O *splicing* alternativo é o processo que se refere à forma como diferentes combinações de *exons* podem ser unidos. Essa ideia foi do físico e bioquímico americano Walter Gilbert em 1932. A proposta dele foi mostrar que as diferentes permutações de *exons* podem produzir diferentes isoformas de proteínas que, por sua vez, teriam diferentes atividades químicas e biológicas. Na Figura 3, podem ser vistas três proteínas diferentes sendo formadas pelo mesmo gene.

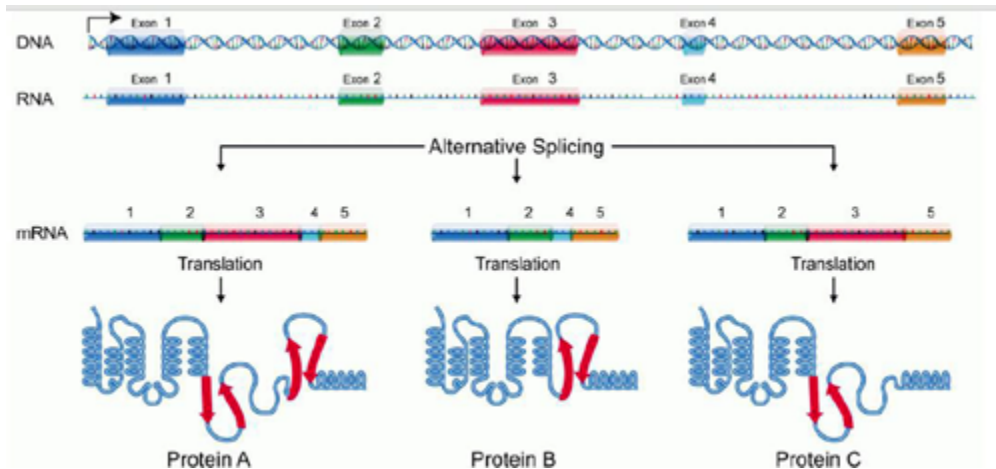


Figura 3. Diferentes combinações de exons no mesmo gene. Fonte: National Human Genome Research Institute.

Promotores

Um promotor é definido como uma região reguladora do gene (DNA) que fica localizado geralmente na região 5' de um gene e fornece um ponto de controle para a transcrição. No promotor existem sequências específicas de DNA que são reconhecidas por fatores proteicos de transcrição. Estes fatores se ligam às sequências do promotor, recrutando a RNA polimerase para realizar a transcrição.

Promotores procarióticos

Nos procariotos, o promotor consiste em duas sequências curtas, a -10 e a -35, ou seja, localizadas a 10 e 35 bases de distância do início da transcrição de um gene.

A sequência em -10 é chamada de caixa Pribnow, e geralmente consiste nos seis nucleotídeos TATAAT. A caixa Pribnow é essencial para começar a transcrição em procariotos.

A sequência, em -35 pares de bases, que fica antes do início da transcrição, geralmente consiste dos seis nucleotídeos **TTGACA** e sua presença permite uma taxa de transcrição muito alta.

Promotores eucarióticos

Os promotores eucarióticos são diversos, complexos e difíceis de caracterizar. Geralmente se encontram antes de um gene e podem ter elementos regulatórios de várias kilobases antes do local de início da transcrição. Lembrando que 1 quilobase é igual a 1.000 bases de DNA ou RNA.

Muitos promotores eucarióticos contêm uma caixa TATA (sequência TATAAA) que se liga a uma proteína de ligação a TATA auxiliando na formação do complexo transcripcional da RNA polimerase II.

Enhancers (intensificadores)

Alguns genes eucarióticos possuem regiões que podem aumentar a taxa de transcrição. Essas regiões são chamadas de intensificadores ou *enhancers*, e não se encontram necessariamente próximos de um gene, mas podem estar localizados dentro da região de codificação, *a montante* (antes) ou *a jusante* (depois) de um gene, ou estar a milhares de nucleotídeos de distância.

Um exemplo de *enhancer* (Figura 4), constituído por pequenas sequências de DNA chamadas elementos de controle distal, interage com proteínas mediadoras e fatores de transcrição. Quando uma proteína de aderência de DNA se liga ao *enhancer*, leva mudança da forma do DNA, o que permite interações entre os ativadores e fatores de transcrição.

Observe que o *enhancer* permite o dobramento da fita de DNA agrupando e acenando a maquinaria básica de transcrição. Entretanto, moléculas como os *insulators* impedem que *enhancers* associados a um gene regulem inadequadamente os genes vizinhos.

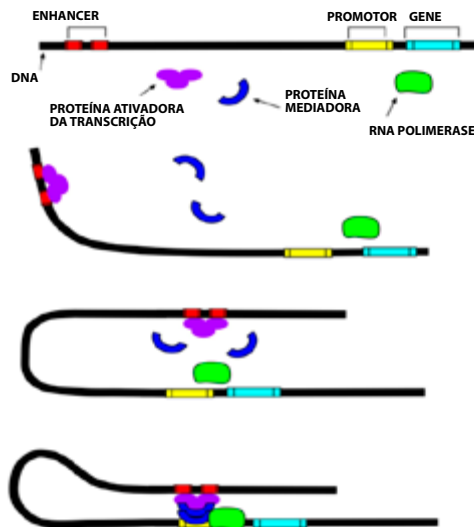


Figura 4. Autor: Jon Cheff.

Silencers (silenciadores)

Os silenciadores são sequências de DNA encontradas a montante (no início), a jusante (no fim), ou dentro do gene de interesse. São locais específicos presentes no DNA (gene) onde grupos de fatores de transcrição se ligam de forma que, quando um silenciador está ativo, significa que a transcrição de um gene está desativada em qualquer célula do organismo.

Em genética molecular, entende-se como um repressor uma proteína que, ao se ligar a silenciadores associados a determinado gene inibe sua expressão. Um repressor de ligação ao DNA bloqueia a ligação da RNA polimerase ao promotor, evitando assim a transcrição dos genes. Um repressor também pode se ligar a um RNAm evitando a tradução para a proteína. Esse bloqueio de expressão é chamado de repressão. Na Figura 5 pode ser vista a disposição desses elementos em um gene. Observe que em ambas as regiões 5' ou 3' pode ser vista a presença de alguns elementos como os *enhancers* e silenciadores.

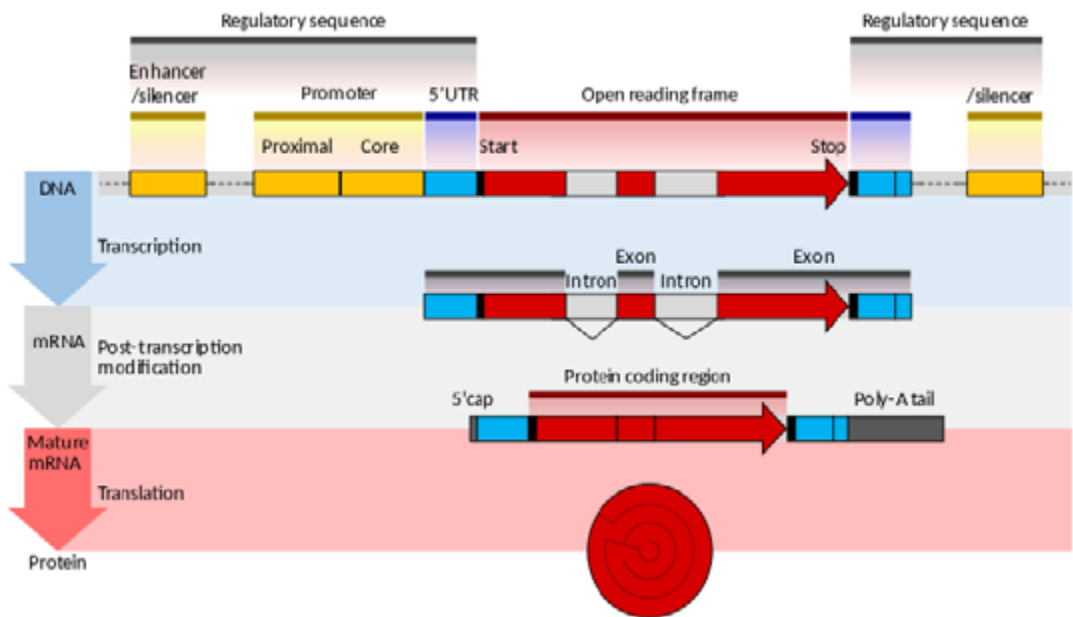


Figura 5. Mecanismo de ação dos silenciadores e repressores de um gene. A estrutura de um operon procarionótico para codificação genética de proteínas. Sequência regulatória controla quando ocorre expressão para a região codificadora de múltiplas proteínas (vermelho). As regiões promotoras, operador e intensificadoras (amarelo) regulam a transcrição do gene em um RNAm. As regiões não traduzidas do RNAm (azul) regulam a tradução nos produtos proteicos finais. Fonte: <https://th.bing.com/th/id/R.3d6b6100ad86d22971777dd8daf09109?rik=BOlgjks%2fZD2UdQ&riu=http%3a%2f%-2fpediaa.com%2fwp>

Insulators (isoladores)

Os *insulators* impedem que *enhancers* associados a um gene regulem inadequadamente os genes vizinhos. Ainda existe uma grande classe de proteínas denominadas de coativadores e co-repressoras que atuam potencializando o efeito de fatores regulatórios.

3.2 Níveis de controle da expressão gênica em eucariotos

Os mecanismos de regulação da expressão gênica em eucariotos compreendem diferentes passos, desde a regulação da estrutura da cromatina até o controle pós- traducional. Estes mecanismos foram divididos em dois grandes grupos: **regulação genética** e **regulação epigenética**. Cada tipo de célula possui um padrão particular de genes expressos. Esta diferenciação em células especializadas ocorre em grande parte como resultado do desligamento da expressão da maioria dos genes na célula, de forma que é muito importante entender os mecanismos de ligar e desligar genes no processo de desenvolvimento do organismo.

3.2.1 Regulação Genética

As células procarióticas e eucarióticas têm mecanismos semelhantes em relação ao controle da expressão gênica, porém as eucarióticas são mais complexas. É só observar que as células procarióticas de uma dada espécie são todas iguais, mas a maioria dos eucariotos são organismos multicelulares complexos, com muitos tipos de células e, portanto, o controle da expressão gênica desses organismos é muito mais profundo e efetuado por vários processos.

Em cada célula humana existem cerca de 20-25 mil genes diferentes, sendo que a maior parte destes já tem funções conhecidas e cerca de 25% ainda não. A maioria desses genes codifica enzimas e proteínas estruturais envolvidas no processamento de outras proteínas, em grandes conjuntos de fatores de transcrição e em diversas vias metabólicas da célula. Entretanto, esses genes não são expressos em todas as células do organismo multicelular em todos os momentos e nem ao mesmo tempo. Assim, é importante entender como é controlada a expressão gênica, quais genes são ativos em diferentes células e tecidos e quais mecanismos estão envolvidos no controle da expressão gênica. Após a fertilização, as células do embrião em desenvolvimento tornam-se cada vez mais especializadas, principalmente ativando alguns genes e desativando muitos outros. Algumas células do pâncreas, por exemplo, são especializadas para sintetizar e secretar enzimas digestivas, enquanto outras células pancreáticas (células β nas ilhotas de Langerhans) são especializadas para sintetizar e secretar insulina.

Como nas bactérias, a transcrição nas células eucarióticas é controlada por proteínas que se ligam a sequências reguladoras específicas e modulam a atividade da RNA polimerase. A tarefa complexa de regular a expressão gênica nos diversos tipos celulares é realizada principalmente pelas ações combinadas de várias proteínas reguladoras transcricionais diferentes. A expressão gênica em eucariotos também pode ser regulada por alterações no empacotamento do DNA, que modula o acesso das enzimas de transcrição da célula por exemplo, a RNA polimerase ao DNA. Sabemos que quando um gene é transcrito gera uma molécula de RNA e quando conhecemos todas as moléculas de RNAs (transcritos) expressas em uma determinada situação da célula (órgãos ou tecidos), significa conhecer o **transcriptoma**, termo que reflete os genes ativos do genoma em um determinado momento da célula. Assim, a expressão gênica pode ser regulada nos seguintes níveis: no início da **transcrição** (descompactação da cromatina), em que o DNA é copiado para uma molécula de RNAm primário, na **pós-transcrição** onde o RNAm primário é processado para se tornar um RNAm maduro, na **tradução** quando o RNAm maduro é transportado para o citoplasma (organismos eucariotos) onde será traduzido e no nível **pós-traducional** que também é um ponto de controle importante da expressão gênica onde a função da proteína vai ser controlada por meio de modificações químicas como fosforilação, metilação, acetilação entre outras ou clivagens proteolíticas para finalmente ser endereçada ao compartimento celular, onde desempenhará sua atividade biológica (Figura 6). É importante saber que a **estabilidade do RNAm** e da **proteína** também vai definir a quantidade da expressão gênica de uma célula.

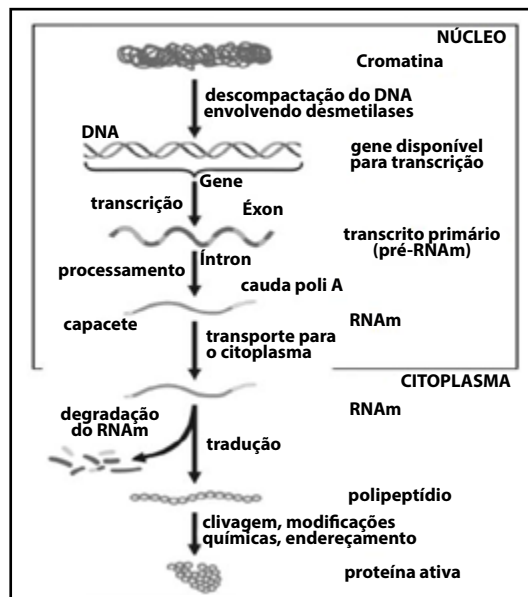


Figura 6. Esquema ilustrando das várias etapas envolvidas na expressão gênica em eucariotos. Fonte: <https://2.bp.blogspot.com/RbrjlqngHxM/WIgPQUfygtI/AAAAAAAAANw/Xg3V4Uddyj8dVdG7L-Z7p5WTto4WLYRcQCLcBGAs/s1600/fig3.png>

3.2.1.1 Regulação Transcricional

O ponto mais importante do controle da expressão gênica é considerado o início do processo da transcrição que envolve dois pontos principais: A acessibilidade da maquinaria de transcrição ao DNA e o início da transcrição propriamente dito.

O mecanismo de regulação transcricional é mediado por vários tipos de proteínas que podem ativar ou reprimir este processo, chamadas de fatores de transcrição, que se ligam a região promotora do gene e recrutam a RNA polimerase II. Também existem muitas sequências regulatórias conhecidas como elementos proximais do promotor, que contém elementos reguladores primários e elementos distais, que podem conter regiões reguladoras adicionais, tais como *enhancers* (intensificadores), *insulators* (isoladores) e *silencers* (silenciadores).

Transcrição é definida como a síntese de RNA a partir de um molde de DNA. Em bactérias, a transcrição de todos os genes é catalisada por uma única Polimerase de RNA, enquanto no núcleo eucariótico, três diferentes polimerases de RNA - Pol I, Pol II e Pol III - realizam a síntese de diferentes classes de RNA. A RNA polimerase I está localizada no nucléolo, uma subestrutura nuclear especializada na qual o RNA ribossômico (RNAr) é transcrito, processado e montado em ribossomos. As moléculas de RNAr são consideradas RNAs estruturais porque têm um papel celular, mas não são traduzidas em proteínas. Os RNAsr são componentes do ribossomo e são essenciais para o processo de tradução. A RNA polimerase I sintetiza o pré-RNA ribossomal 18 S, 5,8S e 28S.

A RNA polimerase II está localizada no núcleo da célula e sintetiza todos os préRNAs nucleares codificadores de proteínas e alguns RNAs especializados. Os préRNAs eucarióticos passam por extenso processamento após a transcrição antes da tradução. A RNA polimerase II é responsável pela transcrição da maioria dos genes eucarióticos, incluindo todos os genes que codificam proteínas e genes para vários tipos de RNAs reguladores, incluindo microRNAs (miRNAs) e RNAs longos não-codificantes (lncRNAs).

A RNA polimerase III também está localizada no núcleo e transcreve uma variedade de RNAs estruturais que incluem o pré-RNAr 5S, pré-RNAs de transferência (pré-RNAs_t) e pequenos pré-RNAs nucleares. Os RNAs_t têm um papel crítico na tradução: servem como moléculas adaptadoras entre o modelo de RNAm e a crescente cadeia polipeptídica. Pequenos RNAs nucleares têm uma variedade de funções, incluindo préRNAs_m de "*splicing*" e fatores de transcrição reguladores. Nem todos os miRNAs são transcritos pela RNA polimerase II, a RNA polimerase III transcreve alguns deles.

A transcrição eucariótica é realizada no núcleo da célula e prossegue em três estágios: **iniciação, alongamento e terminação.**

Iniciação

Diferente da RNA polimerase procariótica que pode se ligar a um molde de DNA por si só, os eucariotos requerem várias outras proteínas, chamadas fatores de transcrição para primeiro se ligarem à região promotora e depois ajudarem a recrutar a polimerase apropriada. Os fatores de transcrição e a RNA polimerase se ligam ao promotor, formando um complexo de pré-iniciação da transcrição. O elemento promotor do núcleo mais estudado nos eucariotos é uma sequência curta de DNA conhecida como caixa TATA, encontrada de 25 a 30 pares de bases a montante do local inicial da transcrição. Apenas cerca de 10 a 15% dos genes de mamíferos contêm caixas TATA, enquanto o restante contém outros elementos promotores principais, mas os mecanismos pelos quais a transcrição é iniciada em promotores com caixas TATA são bem caracterizados. A caixa TATA, como elemento promotor central, é o local de ligação dos Fatores de Transcrição para formar o complexo de pré-iniciação, como mostrado na Figura 7. No entanto, a transcrição não é apenas impulsionada pelo complexo de pré-iniciação, outras proteínas conhecidas como ativadores e repressores, juntamente com quaisquer coativadores ou corepressores associados, são responsáveis por modular a taxa de transcrição. As proteínas ativadoras aumentam a taxa de transcrição enquanto as proteínas repressoras diminuem a taxa de transcrição.

RNA polimerase II

A RNA polimerase II consiste em 12 subunidades, Rpb1 a Rpb12. As Rpb1, Rpb2, Rpb3, Rpb6 e a subunidade Rpb11 formam o núcleo da Pol II. Embora essas subunidades sejam únicas para a Pol II, homólogos delas são encontrados entre todas as polimerases de RNA de múltiplas unidades, incluindo RNA polimerases bacterianas. As subunidades Rpb5, Rpb8, Rpb10 e Rpb12 são compartilhadas entre as Pol I, II e III, conferindo catalisação e outras funções comuns as três polimerases. Rpb4, Rpb7 e Rpb9 são específicos para Pol II, embora sejam encontrados homólogos dessas subunidades em Pol I e Pol III. Rpb4 e Rpb7 são específicas para o estágio da iniciação de transcrição e parecem vincular o núcleo de Pol II com os fatores de transcrição. A RNA polimerase II para catalisar a síntese de RNA necessita dos Fatores de Transcrição Gerais (Figura 7): TBP (ou TFIID e/ou TFIIA, TFIIB, etc (Transcription Fator for Polimerase II) que identificam o promotor e forma uma placa-suporte para a RNAPol II, desempacota o DNA e outras estruturas da cromatina e ajudam a posicionar a RNAPol II corretamente no promotor. Ajudam na atividade do TFIID. TBP (proteína ligadora de TATA) reconhece especificamente a TATA box, TFIIA tem a função de estabilizar a ligação do TFIIB e TBP ao promotor, TFIIB liga-se ao TBP e também recruta o complexo Pol II-TFIIF, TFIIE recruta o TFIID e tem atividade de Helicase e ATPase. TFIIF se liga fortemente a Pol II, que também se liga ao TFIIB impedindo a ligação da Pol II às sequências de DNA

não específicas. TFIID desenrola o DNA no promotor (atividade de helicase), fosforila a Pol II (o interior do CTD); recruta proteínas de reparo de excisão do nucleotídeo. A RNAPol II contém uma cauda proteica chamada CTD (Carboxi-Terminal Domain) localizada no local onde surge o RNA recém-transcrito. A fosforilação dessa cauda está relacionada ao término da fase de início e a transição para a fase de alongamento do transcrito. Esta cauda é muito importante para outras fases de síntese e amadurecimento do RNAm (Figura 7).

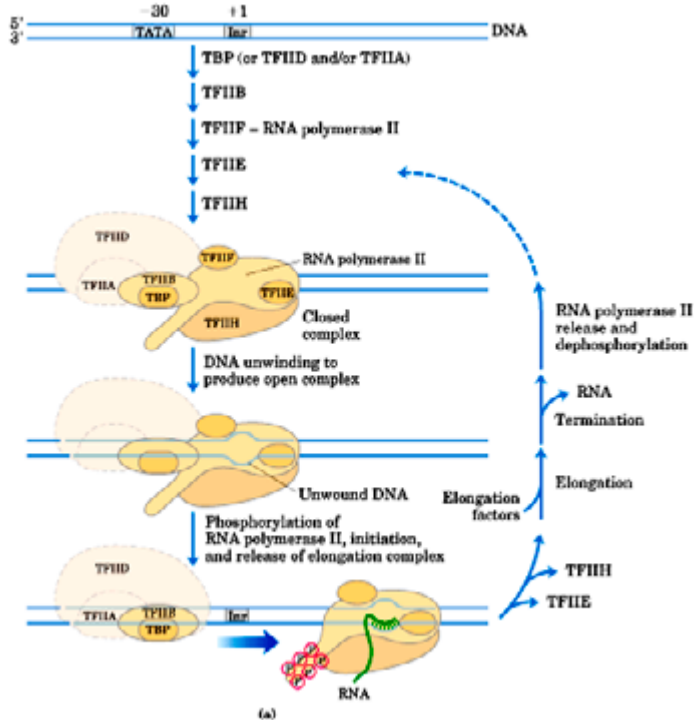


Figura 7. Iniciação da Transcrição Eucariótica: Um promotor generalizado de um gene transcrito pela RNA polimerase II. Os fatores de transcrição reconhecem o promotor e a RNA polimerase II se liga e forma o complexo de iniciação da transcrição. Fonte: <https://slideplayer.com.br/slide/366859/>

Alongamento

O alongamento sintetiza o pré-RNAm na direção de 5' para 3' e a terminação ocorre em resposta a sequências e sinais de terminação.

O TFIID tem duas funções:

1. É uma helicase, o que significa que pode usar o ATP para desenrolar a hélice do DNA, permitindo o início da transcrição.
2. Além disso, fosforila a RNA polimerase II, que faz com que esta enzima mude sua conformação e se dissocie de outras proteínas no complexo de iniciação.

A fosforilação chave ocorre em uma longa cauda C-terminal chamada domínio Cterminal (CTD) da molécula de RNA polimerase II (Figura 8). Curiosamente, apenas a RNA polimerase II que possui um CTD não fosforilado pode iniciar a transcrição, mas apenas uma RNA polimerase II com um CTD fosforilado pode alongar o RNA. A RNA polimerase II agora começa a se mover ao longo do modelo de DNA, sintetizando o RNA, ou seja, o processo entra na fase de alongamento. A síntese de RNA ocorre na direção 5' → 3' com a RNA polimerase catalisando um ataque nucleofílico pelo 3-OH da cadeia crescente de RNA no átomo de alfa-fósforo em um ribonucleosídeo 5trifosfato de entrada. A molécula de RNA produzida a partir de um gene codificador da proteína pela RNA polimerase II é denominada transcrito primária.

Complexo de iniciação/início da transcrição

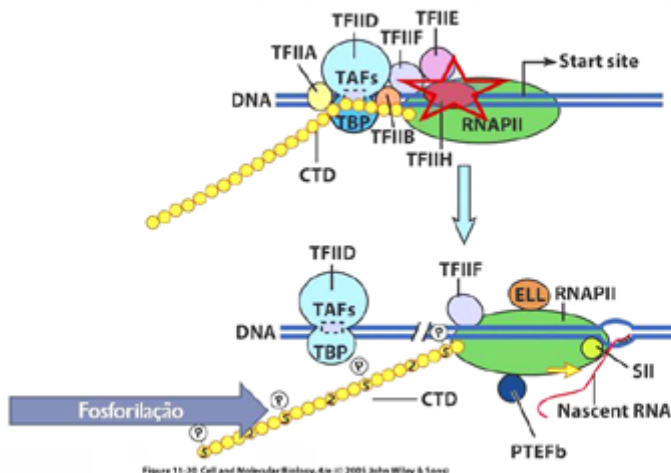


Figure 11-30 Cell and Molecular Biology, 4th © 2001 John Wiley & Sons

▶ 5

MJC-T05 24/out/2013

Figura 8. Mecanismo de alongamento do transcrito do RNAm. Fonte: Cell and Molecular Biology 2005 4ª ed. John Wiley & Sons.

Processos da transcrição a nível dos nucleossomos

Após a formação do complexo de pré-iniciação, a polimerase é liberada dos outros fatores de transcrição e o alongamento prossegue com a RNA polimerase sintetizando o RNA na direção de 5' para 3'. Como a RNA Polimerase II transcreve a maior parte dos genes eucarióticos, vamos nos concentrar em como essa polimerase específica realiza o alongamento e terminação.

Embora o processo enzimático de alongamento seja essencialmente o mesmo em eucariotos e procaríotos, o *template* de DNA eucariótico é mais complexo. Quando as células eucarióticas não estão se dividindo, seus genes existem como uma massa difusa empacotada e compactada de DNA e proteínas, chamada cromatina. Esse complexo DNA-histona, coletivamente chamados nucleossomos, é

regularmente espaçado e inclui 146 nucleotídeos de DNA envolto duas vezes em torno das oito histonas em um nucleossomo como um fio em torno de um carretel.

Para que a síntese de polinucleotídeos ocorra, a maquinaria de transcrição precisa mover as histonas para fora do caminho toda vez que encontrar um nucleossomo. Isso é realizado por um dímero de proteína especial chamado FACT, que significa “facilita a transcrição da cromatina”. O FACT desmonta parcialmente o nucleossomo imediatamente à frente (a montante) de um transcrito da RNA polimerase II, removendo duas das oito histonas (um único dímero das histonas H2A e H2B). Isso presumivelmente afrouxa suficientemente o DNA envolvido em torno desse nucleossomo para que o RNA Polimerase possa transcrever através dele. Experimentos *in vitro* comparando a transcrição em DNA nu e em DNA incorporado na cromatina revelaram que a cromatina bloqueia consideravelmente a transcrição. Esse estudo forneceu o ensaio para identificação de fatores que facilitam a transcrição na presença da cromatina. Assim, um fator chamado FACT (Facilitates Chromatin Transcription-Facilita a transcrição da Cromatina) foi identificado em extratos de células humanas. Este fator torna mais fácil e eficiente a transcrição em moldes de cromatina. FACT é um heterodímero de duas proteínas bem conservadas Spt16 e SSRP1. O FACT funciona assim: os nucleossomos são octômeros constituídos das histonas H2A, H2B, H3 e H4 e DNA. Essas histonas são arranjadas em dois módulos: os dímeros 2HA-2HB e o tetrâmero H3-H4. Então, Spt16 liga a 2HA-2HB e SSRP1 liga a H3-H4. O que foi observado, é que FACT pode desarranjar as histonas removendo um dímero 2HA-2HB e depois remontá-las restabelecendo esse dímero e dessa forma, FACT atua durante o alongamento. À frente de uma RNA polimerase que está transcrevendo, FACT remove um dímero 2HA-2HB permitindo que a polimerase passe esse nucleossomo (*in vitro* demonstrou-se que a remoção de 2HA-2HB do molde permite a transcrição). FACT também possui atividade de chaperona de histona que permite que ela restabeleça o dímero 2HA-2HB ao hexâmero das histonas imediatamente após a passagem da polimerase. Assim, FACT permite que a polimerase alongue, e ao mesmo tempo, mantenha a integridade da cromatina. A RNA Polimerase II continuará a alongar o RNA recém-sintetizado até a finalização da transcrição.

Término

As três RNA polimerases de eucarióticos empregam diferentes mecanismos de terminação. O alongamento da cadeia de RNA continua até o término (Figura 9) e ao contrário da RNA polimerase nos procariontes, a RNA polimerase II não termina a transcrição em um local específico, mas a transcrição pode parar em distâncias variáveis a jusante do gene.

Os genes de RNA transcritos pelo RNA Polimerase II não possuem nenhum sinal ou sequência específica que direcione a RNA Polimerase II para terminar em locais específicos. A RNA polimerase II pode continuar a transcrever o RNA em al-

guns pontos a milhares de pontos após o final real do gene. A transcrição é clivada em um local interno antes que a RNA Polimerase II termine a transcrição. Isso libera a porção a montante da transcrição, que servirá como o RNA inicial antes do processamento (o pré-RNA_m no caso de genes que codificam proteínas). Este local de clivagem é considerado o “fim” do gene. O restante do transcrito é digerido por uma 5'-exonuclease (chamada Xrn2 em humanos) enquanto ainda está sendo transcrito pela RNA Polimerase II.

Quando a 5'-exonuclease alcança a RNA polimerase II, digerindo todo o RNA pendente, ela ajuda a desengatar a polimerase da sua fita modelo de DNA, terminando finalmente essa rodada de transcrição.

Embora ainda pouco se conheça sobre a terminação da transcrição em eucariotos do que em procariotos, são conhecidas algumas características básicas e suas semelhanças e diferenças com a terminação bacteriana. A transcrição de genes pré-RNA_r pela RNA polimerase I é terminada por um mecanismo que requer um fator de terminação específico da polimerase. Essa proteína de ligação ao DNA se liga a jusante da unidade de transcrição, ao contrário do fator E. coli Rho, que é um fator de terminação de ligação ao RNA. Foi observado que a RNA polimerase III purificada termina a síntese de RNA após polimerizar uma série de resíduos U(Uracila), ao contrário da terminação independente de Rho em bactérias, no entanto, a terminação pela RNA polimerase III não requer uma estrutura secundária a montante da haste- alça no transcrito do RNA.

Ao contrário do que ocorre no término da transcrição em procariotos que se dar precisamente na extremidade 3' do RNA_m, nos eucariotos, geralmente esse processo ocorre além da extremidade 3' do RNA_m. Na extremidade 3' da maioria dos RNA_m de eucariotes, é adicionada cerca de 250 nucleotídeos adenina (A), formando uma cauda denominada cauda poli (A): ...pNpNpA(pA)_npA. Este tipo de modificação é chamado poliadenilação e é produzido pela enzima Poli (A) polimerase. Isto sugere que a terminação da transcrição eucariótica envolve mecanismos diferentes da transcrição procariótica. Além disso, é sugerido que o mecanismo de terminação desempenhe um papel na regulação dos genes eucarióticos.

Uma característica única da maioria, mas não de todos os RNA_m eucarióticos, é que o local de adição de poli (A) na extremidade 3' é precedido por uma sequência de consenso conhecida como sinal de poli (A), AAUAAA. A terminação real da transcrição pelo RNA Pol II ocorre além do local poli (A), geralmente heterogêneo ao longo de uma ou mais kilobases de sequência. Um modelo atraente para terminação incorpora esse recurso exclusivo: o mecanismo de processamento de 3' desempenha um papel ativo na seleção dos locais de terminação. As evidências sugerem que o término da transcrição e o processamento de 3' não são de fato independentes, mas acoplados. Para um processamento final eficiente de 3', são necessários dois elementos de sequência separados: a sequência AATAAA e um segundo elemento rico em GT imediatamente a jusante do local de clivagem no qual

poli (A) é adicionado. Embora o processamento de 3' possa iniciar o mecanismo de terminação, acredita-se que a terminação real ocorra nos locais de pausa do RNA Pol II. O local da pausa pode ser uma estrutura em hairpin e / ou um elemento de sequência que se liga a um fator proteico específico. Um desses sinais de terminação foi identificado como a sequência da caixa CCAAT que funciona como um terminador independentemente da posição e dependente da orientação. O mais provável, é que o sinal de terminação seja mediado pela proteína de ligação específica denominada CP1, semelhante ao papel supostamente desempenhado por um fator de ligação à AATAAA.

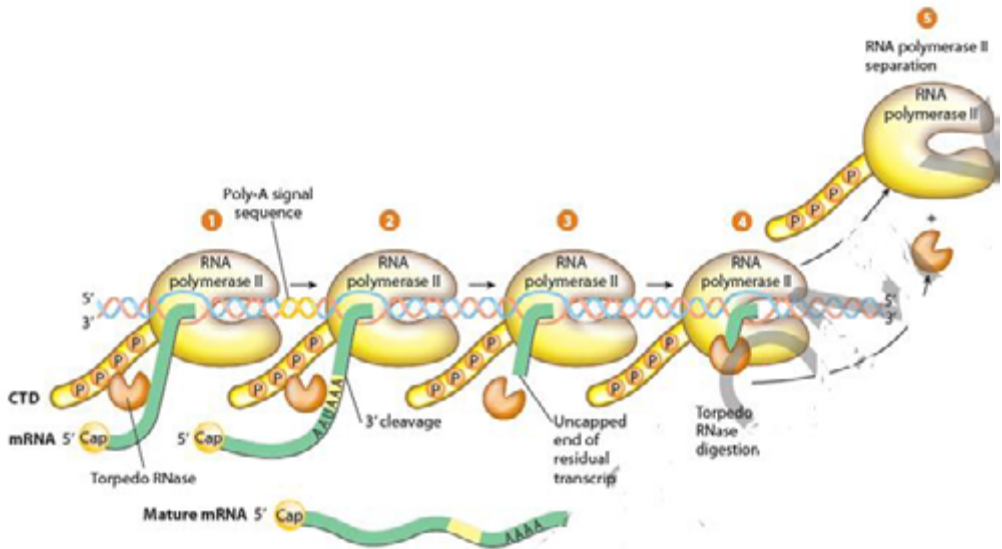


Figura 9. Mecanismo de alongamento e término do RNA em organismo eucarioto. Fonte: <https://microbenotes.com/eukaryotic-transcription/>

3.2.1.2 Regulação Pós-Transcricional

Splicing alternativo

O pré-RNA eucariótico passa por importante processamento antes de estar pronto para ser traduzido de forma que as etapas adicionais envolvidas na maturação do RNA eucariótico criam uma molécula com uma meia-vida mais longa do que um RNA procariótico. Os RNAs eucarióticos tem tempo de duração de várias horas, enquanto o RNA de E.coli não dura mais que alguns segundos. Os pré-RNAs são revestidos primeiro em proteínas estabilizadoras de RNA que protegem o pré-RNA da degradação enquanto é processado e exportado para o citoplasma. Esse mecanismo envolve três importantes etapas: a adição de fatores de estabilização e sinalização nas extremidades 5' (**5'CAP**) (no início do RNA) e de uma cauda 3' Poly-A Tail ao final do RNA e a remoção de sequências não codificantes (íntrons) e ligação dos

exons, sequências codificadoras da cadeia polipeptídica. Em alguns casos, a transcrição do RNAm pode ser “editada” após ser transcrita.

5’CAP

Enquanto o pré-RNAm ainda está sendo sintetizado, uma cap 7-metilguanosi-na é adicionada à extremidade 5’ do transcrito em crescimento por uma ligação fosfato. Essa cap (grupo funcional) protege o RNAm nascente da degradação. Além disso, fatores envolvidos na síntese de proteínas reconhecem a cap para ajudar a iniciar a tradução pelos ribossomos (Figura 10).

3’Poly-A Tail

Uma vez concluído o alongamento, o pré-RNAm é clivado por uma endonuclease entre uma sequência de consenso AAUAAA e uma sequência rica em GU, deixando a sequência AAUAAA no pré-RNAm. Uma enzima denominada poli-A polimerase, adiciona uma sequência de aproximadamente 250 resíduos de A, chamada cauda poli-A. Essa modificação protege ainda mais o pré-RNAm da degradação e sinaliza a exportação dos fatores celulares que a transcrição necessita para o processo de tradução no citoplasma (Figura 10).

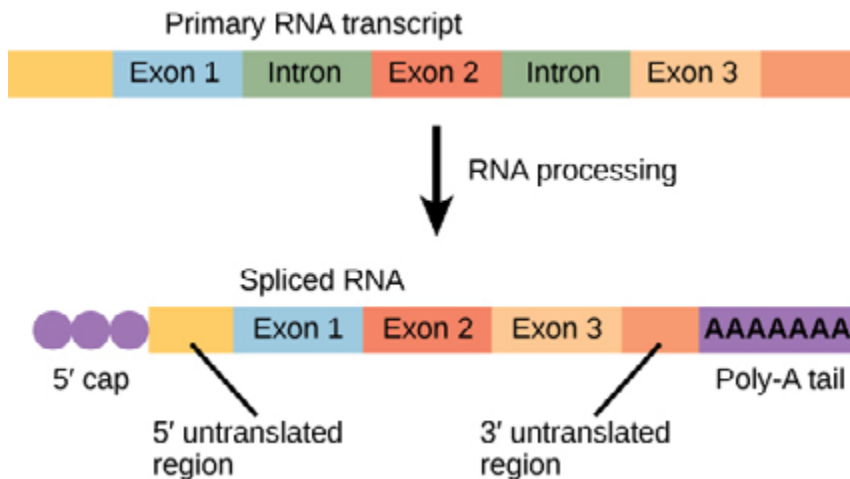


Figura 10. O RNAm eucariótico contém íntrons que devem ser removidos. Uma cap 7-metilguanosi-na 5’ e uma cauda poli-A 3’ também são adicionadas. Fonte: https://philtschatz.com/biology-concepts-book/resources/Figure_09_03_05.jpg

Na etapa da expressão do gene, a molécula de RNAm pode passar por diferentes modificações após o início da transcrição. Na regulação pós-transcricional um processo conhecido como *splicing*, é crítico para essa etapa que consiste na remoção dos íntrons e junção dos éxons da sequência do pré-RNAm. O proces-

so de clivagem dos íntrons é realizado pelo *spliceossomo* e ocorre em sequências conservadas encontradas nas extremidades 5' e 3' dos íntrons designadas por locais de *splicing*.

Os éxons também podem ser unidos de maneira diferente, originando diferentes cadeias polipeptídicas a partir de um mesmo *RNA_m* gerando diferentes isoformas de proteínas no organismo a partir de um mesmo gene com diversas funções e / ou localizações na célula. Esse processo é denominado ***splicing* alternativo**. Dois pontos também importantes dessas regulações, são o capeamento do *RNA_m*, que consiste na adição de uma cauda poly-A na extremidade 3' (poliadenilação) e da adição do CAP na extremidade 5' e o seu transporte para o citoplasma, onde o *RNA_m* é traduzido, originando uma proteína.

Os genes eucarióticos são compostos de exons, sequências que codificam as proteínas, e de íntrons, sequências que não codificam proteínas, mas que podem estar envolvidos na regulação dos genes. Os íntrons são removidos do pré-*RNA_m* durante o processamento tendo em vista que não codificam proteínas funcionais. A descoberta de íntrons na década de 1970, surpreendeu os pesquisadores que esperavam que os pré-*RNA_m* especificassem sequências de proteínas sem processamento adicional, como observado em procariontes. Os genes dos eucariotos superiores geralmente possuem um ou mais íntrons e essas regiões podem ter sequências reguladoras, entanto, o significado biológico de ter muitos íntrons ou ter íntrons muito longos em um gene não está esclarecido. A sugestão é que os íntrons estejam envolvidos na diminuição da expressão gênica, isso porque o processo para transcrever pré-*RNA_m* com muitos íntrons, leva muito mais tempo. Alternativamente, os íntrons podem ser remanescentes não funcionais de sequência que restaram da fusão de genes antigos ao longo da evolução. Na maioria das vezes, as sequências de íntrons podem ser mutadas sem afetar o produto proteico. Todos os íntrons do pré-*RNA_m* devem ser totalmente e precisamente removidos antes da síntese do polipeptídeo. Havendo erro no processo, mesmo com um único nucleotídeo, a estrutura de leitura dos exons reunidos poderá mudar e gerar uma proteína disfuncional. O *splicing*, portanto, corresponde ao processo de remover íntrons e reconectar exons ao longo do *RNA_m* (Figura 11). Os íntrons são removidos e degradados enquanto o pré-*RNA_m* ainda está no núcleo. O *splicing* do pré-*RNA_m* ocorre por um mecanismo específico da sequência que garante que os íntrons sejam removidos e os éxons se unam novamente à exatidão e precisão de um único nucleotídeo. O *splicing* de pré-*RNA_m* é conduzido por complexos de proteínas e moléculas de RNA chamadas spliceossomos. Erros no mecanismo do *splicing* podem levar ao surgimento de cânceres e outras doenças humanas.

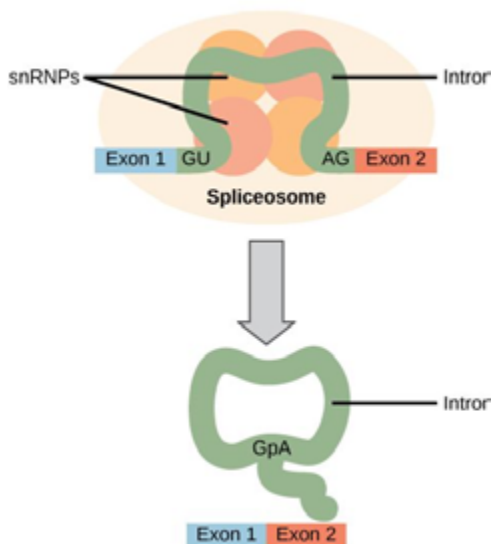


Figura 11. O *splicing* do pré-RNA envolve a remoção precisa dos íntrons do transcrito primário do RNA. O processo de *splicing* é catalisado por complexos de proteínas chamados spliceossomas, que são compostos de proteínas e moléculas de RNA chamadas snRNAs. Os spliceossomas reconhecem sequências nas extremidades 5' e 3' do íntron. Fonte: <https://i.pinimg.com/736x/fd/3a/6f/fd3a6f70f9df4c08fa3bd423b28aefe6.jpg>

Processamento do RNA^t e RNA^r

Os RNA^t e RNA^r são moléculas estruturais envolvidas na síntese de proteínas, porém não são traduzidos. Os pré-RNA^r são transcritos, processados e montados em ribossomos no **nucléolo**. Os pré-RNA^t são transcritos e processados no núcleo e depois migram para o citoplasma, onde funcionam como transportadores de aminoácidos livres para a síntese de proteínas conforme codificado pelo RNA^m.

A maioria dos RNA^t e RNA^r dos eucariotos e procariontes é primeiro transcrita como uma molécula precursora longa que abrange vários RNA^r ou RNA^t. As enzimas específicas separam os precursores em subunidades correspondentes a cada RNA estrutural. Algumas das bases dos pré-RNA^r são metiladas, isto é, um grupo metil funcional –CH₃, é adicionada para sua estabilidade. As moléculas pré-RNA^t também sofrem metilação. Tal como acontece com os pré-RNA^m, a excisão de subunidades ocorre em pré-RNAs eucarióticos destinados a se tornar RNA^t ou RNA^r.

Os RNA^r maduros representam aproximadamente 50% de cada ribossomo. Algumas das moléculas de RNA de um ribossomo são puramente estruturais, enquanto outras têm atividades catalíticas ou de ligação. Os RNA^t maduros assumem uma estrutura tridimensional através da ligação intramolecular de hidrogênio para posicionar o local de ligação de aminoácidos em uma extremidade e o anticódon na outra extremidade (Figura 12). O anticódon é uma sequência de três

nucleotídeos presente no RNAt que interage com um códon do RNAm por meio de emparelhamento de bases do anticódon do RNAt com o códon do RNAm com ligação complementar e **antiparalela**, onde o códon do mRNA é lido de 5' para 3' por um anticódon pareado (RNAt) em orientação invertida (3'-5').

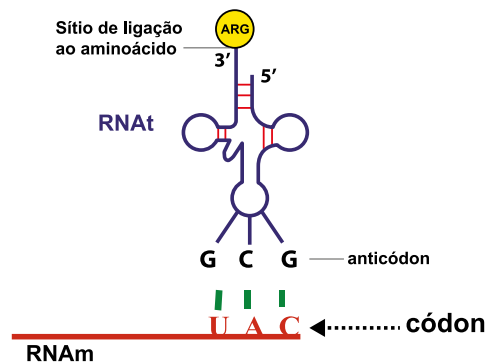


Figura 12. Uma molécula de RNAt adiciona o aminoácido arginina a uma cadeia polipeptídica crescente. O anticodon GCG liga-se ao codon UAC no RNAm. Adaptado de: https://pt.slideshare.net/janaina_leitinho/sintese-de-proteina-2.

3.2.1.3 Regulação Pós-Traducional

A modificação pós-tradução é um processo que ocorre com as proteínas após sua síntese, dobramento e montagem. Foi evidenciado que várias modificações pós-traducionais ocorrem dentro das células, sendo que em geral, essas modificações proteicas podem ser reversíveis, exceto, quando a modificação envolve a degradação da proteína. Diferentes tipos de modificação pós-tradução em geral, envolvem a formação de uma ligação covalente ligando um grupo químico específico a cadeias laterais específicas de aminoácidos na proteína - estes grupos podem variar desde grupos fosfato (fosforilação), um grupo acetato (acetilação), a ligação de lipídeo/hidrofóbico grupos (modificação lipídica) ou carboidratos (glicosilação). Estas modificações pós-tradução em geral são reversíveis, ou seja, uma enzima adiciona o grupo modificador e outra pode removê-lo. Um exemplo, é a fosforilação das proteínas que ocorre pela ação de enzimas conhecidas como proteínas quinases, porém, num processo reversível, as proteínas fosfatases podem remover os grupos fosfatos. As modificações pós-tradução agem de forma semelhante aos efetores alostéricos: modificam a estrutura e, por sua vez, a atividade do polipeptídeo ao qual estão ligados. E também podem modificar as interações de uma proteína com outras proteínas, a localização da proteína dentro da célula ou a sua estabilidade.

3.2.2 Regulação epigenética

Os mecanismos de regulação epigenética compreendem: metilação do *DNA*, modificação de histonas, remodelamento de cromatina e *RNAs* não-codificantes.

3.2.2.1 Composição química da cromatina

A estrutura da cromatina (cromossomo) de organismos eucarióticos é complexa e dinâmica. A cromatina é formada por DNA mais proteínas histonas representadas por cinco tipos principais: H1, H2A, H2B, H3 e H4, todas com alta percentagem dos aminoácidos arginina e lisina que conferem às histonas carga final positiva. A cromatina pode ser de dois tipos: **euromatina**, que passa pelo processo de condensação e descondensação durante o ciclo celular, e **heterocromatina**, que permanece em estado de condensação durante o ciclo celular, mesmo na interfase. Esta é encontrada nos centrômeros e telômeros de todos os cromossomos, em lugares específicos de alguns cromossomos e em todo o cromossomo X inativo nas fêmeas de mamíferos. As proteínas não-histonas constituem cerca de metade da massa de proteína do cromossomo, algumas dessas proteínas não-histonas, estão associadas à cromatina e podem ser divididas naquelas com papéis estruturais e nas que tomam parte nos processos de expressão gênica tais como, replicação e transcrição.

A cromatina tem uma estrutura muito complexa apresentando vários níveis de organização. O nível de estrutura mais simples é a dupla hélice do DNA e no nível mais complexo a molécula do DNA está associada a proteínas e é altamente dobrada para formar o cromossomo. Quando a cromatina é isolada do núcleo de uma célula e vista em microscopia eletrônica, pode ser comparada às contas de um colar (Figura 13). Experimentalmente foi mostrado que se uma enzima nuclease for adicionada a essa estrutura, esta corta o "colar" entre as contas, deixando contas individuais ligadas a 200pb de DNA. Se mais nuclease é adicionada, esta corta o DNA entre as contas e deixa um cerne de proteínas ligado a fragmentos de DNA. Esse experimento mostra que a cromatina não é uma associação aleatória de proteínas e DNA, mas tem uma estrutura essencialmente repetida. Este cerne repetido de proteína e DNA produzidos por digestão com a nuclease é o nível mais simples da estrutura da cromatina que é o **nucleossomo** (Figura 13). O nucleossomo portanto, é uma partícula cerne central que consiste em oito proteínas histonas (duas de cada H2A, H2B, H3 e H4) envolvendo 146 pares de bases de DNA de fita dupla em aproximadamente 1,7 voltas em contato direto com o octâmero de histonas com giro para esquerda e fortemente superelucidado. O DNA dentro de cada nucleossomo é geralmente inacessível aos fatores de ligação ao DNA. A histona H1 não é parte da partícula do cerne, ficando por fora do octâmero e tem papel importante na estrutura do nucleossomo.

A composição dos nucleossomos não é imutável, na verdade, as próprias histonas podem ser substituídas por variantes de histonas ou modificadas por enzimas específicas, tornando assim o DNA circundante mais ou menos acessível à maquinaria transcricional. Os cientistas já entenderam que o nucleossomo, que antes se pensava ser estático, na verdade desempenha um papel fundamental dinâmico no direcionamento de alguns elementos da maquinaria básica de transcrição.

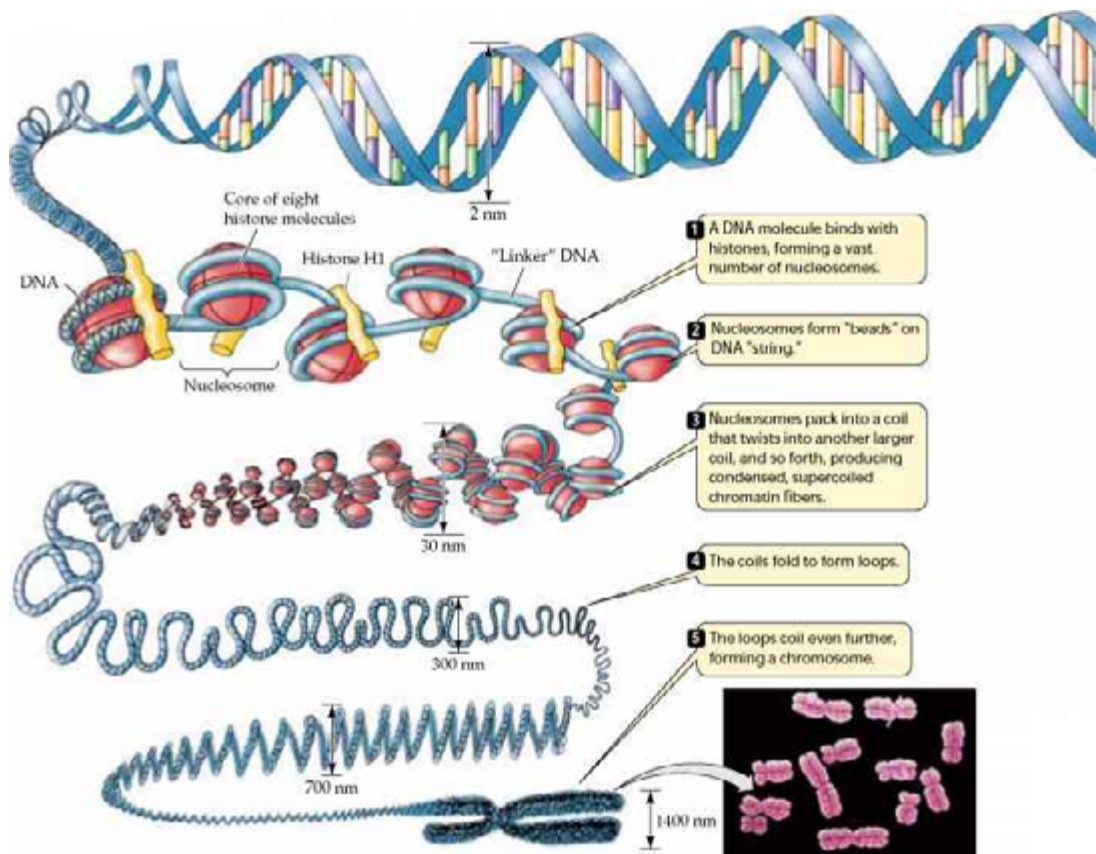


Figura 13. Fonte: <http://www.78stepshealth.us/plasma-membrane/eukaryotic-chromosomes.html>

3.2.2.2 Modificação de histonas

Como o DNA eucariótico está firmemente enrolado em torno dos nucleossomos e as cargas positivas das histonas se ligam firmemente às cargas negativas do DNA, os nucleossomos atuam essencialmente como uma barreira física aos fatores de transcrição que precisam se ligar a certas regiões do DNA. O nucleossomo é a unidade básica da cromatina, formado por um filamento de DNA envolvido em um octâmero de proteínas histonas, numa composição de dois dímeros de cada histona (H2A, H2B, H3, H4) e a histona H1 que tem a função de unir e empacotar os nucleossomos adjacentes. As caudas das proteínas histonas se estendem da superfície do nucleossomo abrigando a maioria dos sítios para modificações pós-traducionais que incluem acetilação, fosforilação,

metilação e ubiquitinação. A acetilação (Figura 14) consta da adição de um grupamento acetil (-COCH₃) aos resíduos de lisina, levando a neutralização da carga positiva da lisina com consequente enfraquecimento da interação DNA-histona, promovendo a abertura da cromatina. As enzimas chamadas de histona acetiltransferases (HATs) catalisam a transferência do grupamento acetil aos resíduos de lisina. Essa reação é reversível, e a remoção do grupamento acetil é feita pelas histonas desacetilases (HDACs). Acetilações específicas podem remover a carga positiva do grupo amino da lisina que é acetilado, de modo que o nucleossomo “perde o controle” do DNA, resultando em um afrouxamento da “bobina” facilitando a transcrição do gene. As histonas do núcleo (H2A, H2B, H3 e H4) têm dois domínios: um domínio de dobra de histonas, que está envolvido nas interações com outras histonas e na quebra do DNA em torno da partícula do núcleo do nucleossomo, e um domínio da cauda amino terminal, que se estende para fora do nucleossomo. Os domínios da cauda amino terminal são ricos em lisina e podem ser modificados por acetilação em resíduos específicos de lisina. A acetilação reduz a carga positiva líquida das histonas e pode enfraquecer sua ligação ao DNA, além de alterar suas interações com outras proteínas. É importante ressaltar que experiências recentes forneceram evidências diretas de que a acetilação da histona facilita a ligação de fatores de transcrição ao DNA nucleossômico, indicando que a acetilação da histona aumenta a acessibilidade da cromatina às proteínas de ligação ao DNA. Além disso, as ligações diretas entre a acetilação da histona e a regulação da transcrição vieram de experimentos que mostram que ativadores e repressores da transcrição estão associados às histonas acetiltransferases e desacetilases, respectivamente. Essa associação foi revelada pela clonagem de um gene que codifica uma histona acetiltransferase do *Tetrahymena*. Foi também observado que a sequência desta histona acetiltransferase revelou que estava intimamente relacionada a um coativador transcricional de levedura conhecido anteriormente chamado Gcn5p, que estimula a transcrição em associação com vários ativadores transcricionais específicos de sequências diferentes. Experiências adicionais revelaram que o próprio Gcn5p possui atividade da histona acetiltransferase, sugerindo que a ativação transcricional resulta diretamente da acetilação da histona. Por outro lado, as histonas desacetilases (que removem os grupos acetil das caudas de histonas) estão associadas a repressores transcricionais nas células de levedura e mamífero. A acetilação da histona é assim regulada por ativadores e repressores da transcrição, indicando que ela desempenha um papel fundamental na expressão do gene eucariótico.

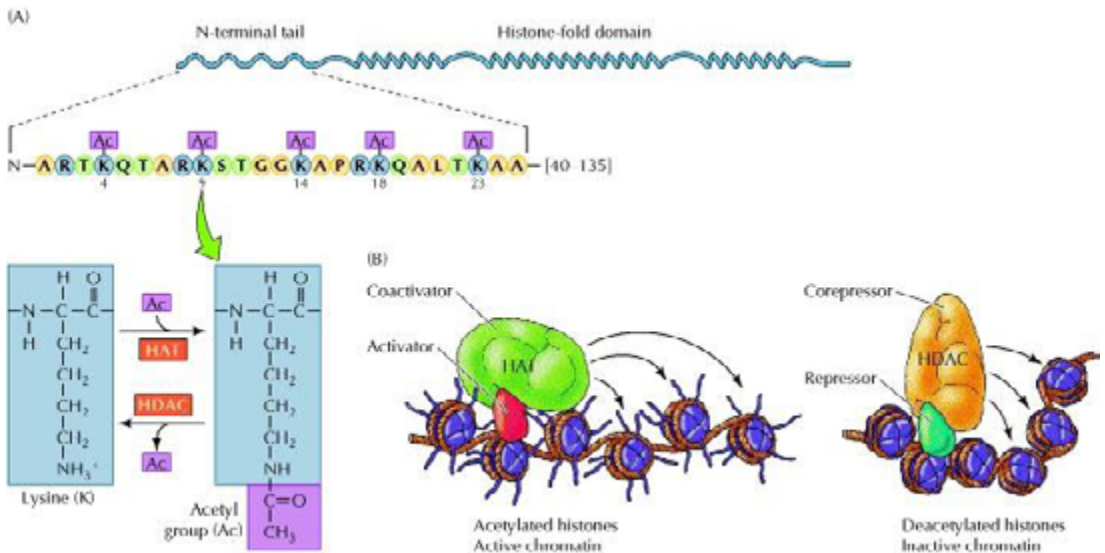


Figura 14. Acetilação de histonas. (A) As histonas do núcleo têm domínios de dobras de histonas, que interagem com outras histonas e com o DNA no nucleossomo, e caudas do N terminal, que se estendem para fora do nucleossomo. As caudas do N terminal das histonas do núcleo (por exemplo, H3) são modificadas pela adição de grupos acetil (Ac) às cadeias laterais de resíduos específicos de lisina. (B) Ativadores e repressores transcricionais estão associados à coativadores e corepressores, que possuem atividades de histona acetiltransferase (HAT) e histona desacetilase (HDAC), respectivamente. A acetilação da histona é característica da cromatina ativamente transcrita e pode enfraquecer a ligação das histonas ao DNA ou alterar suas interações com outras proteínas. Fonte: <https://www.ncbi.nlm.nih.gov/books/NBK9904/>

A metilação de histonas consiste na adição de um radical metil (CH_3 -) nos resíduos de lisina e arginina, numa reação catalisada pelas histonas metiltransferases (HMTs) de forma reversível que pode ser desfeita pela ação de histona demetilases (HDMs).

A arginina pode ser mono-ou di-metilada, a lisina pode ser mono- (me), di- (2me) e tri- (3me) metilada. A metilação pode promover tanto a ativação quanto a repressão gênica, dependendo do resíduo que for modificado.

A fosforilação das histonas também está associada a ativação gênica e a outros mecanismos celulares como o reparo de DNA e a condensação cromossômica. Este processo consiste na adição de um grupo fosfato (PO_4), em resíduos de serinas, treonina e tirosina, modificação essa, catalisada por quinases e fosfatases que adicionam e removem a modificação, respectivamente.

A ubiquitinação compreende a ligação de uma ou mais moléculas de ubiquitina nos resíduos de lisina das caudas das histonas sendo catalisada pelas enzimas E1, E2 e E3 e pode promover tanto a ativação quanto a repressão da expressão gênica, dependendo do resíduo que for modificado (Figura 15).

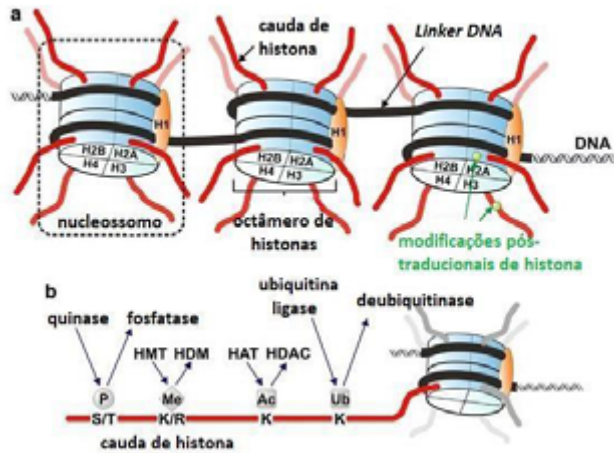


Figura 15. Estrutura da cromatina e enzimas modificadoras de histonas. (a) A cromatina é constituída por unidades repetidas de nucleossomos, formando um octâmero de proteínas histonas envolvidas por um filamento de DNA. (b) As modificações que ocorrem nas caudas das histonas são vistas juntamente com as enzimas que realizam e removem as marcas epigenéticas. Fonte: (Adaptado de FÜLLGRABE; HAJJI; JOSEPH, 2010).

3.2.2.3 Remodelamento da cromatina e acessibilidade da maquinaria de transcrição ao DNA

Em organismos eucariotos, o DNA está fortemente enrolado em um complexo chamado cromatina em unidades básicas chamadas nucleossomos, mas graças ao processo de remodelação da cromatina, este complexo pode ser “aberto” para que genes específicos sejam expressos. Esse enrolamento rígido característico da estrutura da cromatina limita o acesso da RNA polimerase e dos fatores de transcrição ao DNA eucariótico, e assim, para que a transcrição ocorra, a cromatina de uma célula deve “abrir” para que a expressão genética aconteça. Este processo de “abertura” é denominado **remodelação da cromatina** e é muito importante para o funcionamento de todas as células dos organismos eucarióticos. Nos últimos anos, os pesquisadores descobriram muito sobre a remodelação da cromatina, incluindo os papéis que diferentes complexos proteicos, variantes de histonas e modificações bioquímicas desempenham neste processo. No entanto, ainda há muito a ser compreendido antes que a remodelação da cromatina seja totalmente esclarecida. Todos os remodeladores da cromatina contêm uma subunidade com um domínio ATPase conservado. Além da ATPase conservada, cada complexo remodelador também possui proteínas únicas que o especializam por seu papel biológico único. No entanto, como todos os remodeladores movem os nucleossomos e todo esse movimento é dependente de ATP, a mobilização é provavelmente uma propriedade da subunidade conservada da ATPase. As várias moléculas remodeladoras da cromatina fornecem o mecanismo para modificar a cromatina e permitir que os sinais de transcrição cheguem aos seus destinos na cadeia de DNA. A compreen-

são da natureza e dos processos dessas moléculas da construção celular continua sendo uma área ativa de descoberta na pesquisa genética, mas sabe-se que os remodeladores da cromatina são grandes complexos multiproteicos que convertem a energia da hidrólise do ATP em forças mecânicas para mobilizar e reestruturar os nucleossomos e, portanto, são necessários para fornecer acesso ao DNA subjacente para permitir a transcrição, montagem da cromatina, reparo do DNA e outros processos. Ainda não está bem esclarecido como diferentes complexos remodeladores selecionam quais nucleossomos mover e reestruturar.

É importante saber, que os domínios ATPase dos remodeladores são semelhantes em sequência e estrutura às proteínas conhecidas de translocação de DNA em vírus e bactérias. Foi evidenciado que estudos sobre as famílias remodeladoras SWI/SNF e ISWI também revelaram que as ATPases remodeladoras são translocases direcionais de DNA que são capazes de bombear direcionalmente o DNA. Mas como essa propriedade é aplicada aos nucleossomos? Parece que a ATPase se liga a aproximadamente 40 pares de bases dentro do nucleossomo, de onde bombeia o DNA ao redor da superfície do octâmero da histona. Isto permite o movimento do nucleossomo ao longo do DNA, facilitando a exposição do DNA a fatores reguladores.

Os domínios e proteínas adicionais que estão ligados à ATPase são importantes para a seleção do nucleossomo e ajudam a regular a atividade da ATPase. Estas proteínas que estão ligadas à ATPase, ligam-se às histonas e ao DNA nucleossômico, e sua ligação a essas moléculas é afetada pelo estado de modificação das histonas que ajuda a determinar se o nucleossomo é um substrato apropriado para um complexo remodelador. Nesse processo de remodelação da cromatina é possível remover nucleossomos? A resposta pode ser dada baseada em estudos que identificaram dois fatores fundamentais no processo de remoção de nucleossomos: O fator de transcrição GAGA e o fator de remodelagem de nucleossomos (NURF). O fator de transcrição GAGA (proteína constitutiva expressa em *Drosophila* que está ligada a vários promotores com sequências GA) foi identificado como um dos fatores que podem romper nucleossomos. Assim, quando o fator de transcrição GAGA se liga a um nucleossomo contendo a sequência TATA do gene *hsp70* (que codifica a proteína do choque térmico de 70-kDa), o nucleossomo se rompe criando um sítio hipersensível à DNase I no sítio da sequência TATA, sendo evidenciado que esse processo é muito eficiente quando o nucleossomo não tem histona H1. Foi observado também que o fator GAGA não produz esse efeito isoladamente, porém, funciona quando em conjunto com uma proteína que contém quatro peptídeos chamada fator de remodelagem de nucleossomos (NURF). Este complexo foi descrito pela primeira vez em *Drosophila Melanogaster* e compreende as proteínas BPTF, a subunidade ATPásica SNF2L (SMARCA1) e RBBP4, entre outras proteínas como SNF2H (SMARCA5), BAP18 e HMG2L1. O complexo NURF, assim como SWI/SNF, promove o deslizamento do nucleossomo de maneira ATP-dependente promovendo a transcrição gênica (Figura 15). NURF pode hidrolizar ATP e utilizar

a energia para reconfigurar as interações histona-DNA (ou histona-histona). Essas perturbações parecem facilitar a acessibilidade dos fatores de transcrição ao DNA nucleossômico podendo levar a novas modificações na estrutura do nucleossomo.

Na ausência de fatores de transcrição, o NURF pode perturbar o nucleossomo num mecanismo dependente de ATP, permitindo que fatores de transcrição como o GAGA se liguem às regiões promotoras, rompendo os nucleossomos. (Figura 16). O complexo SWI/SNF (*SWItch/Sucrose Non-Fermentable*), também conhecido como BAF, foi primeiro descrito em leveduras e posteriormente evidenciado em *Drosophila* e humanos. Este complexo de remodelamento da cromatina é composto pela ATPase *Swi2/Snf2p*, duas proteínas relacionadas à actina (*Arp7p* e *Arp9*) e outras subunidades envolvidas nas interações com DNA e proteína-proteína. Atua promovendo o deslizamento dos nucleossomos ao longo do DNA, de maneira dependente de ATP, abrindo a cromatina e promovendo a transcrição de genes. Ainda, desempenha papel importante em outros processos, como diferenciação celular e reparo de DNA. Este complexo SWI/SNF é parte da RNA polimerase e está ligado ao seu domínio carboxiterminal (Wilson et al., 1996) e pode ser ativado por fatores de transcrição capazes de romper nucleossomos.

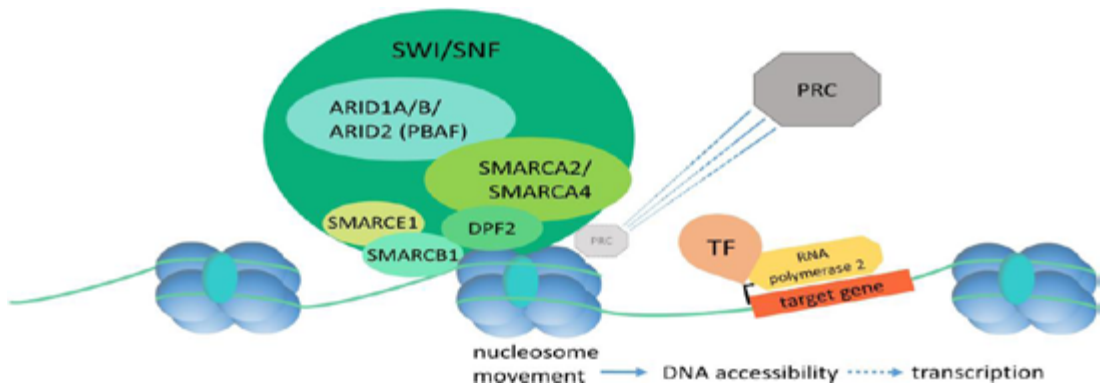


Figura 16. Complexos SWI/SNF e NURF. Os complexos SWI/SNF e NURF ligam-se ao DNA e histonas efetivando o deslocamento do nucleossomo, melhorando a acessibilidade do DNA e facilitando a instalação da maquinaria de transcrição. Ambos atuam de maneira ATP-dependente. Fonte: Adaptado de BÖGERSHAUSEN; WOLLNIK, 2018 e KATO; KOMATSU, 2015.

A atividade de remodelação da cromatina por SWI/SNF ou outras máquinas de remodelação, também, pode ser necessária para recrutar atividade adicional de remodelação da cromatina para locais adicionais a jusante. As modificações em um promotor podem ocorrer em múltiplas etapas que são reguladas independentemente, e modificações adicionais podem acontecer passo a passo a partir do ponto da primeira modificação estendendo-se ao longo da fita de DNA em uma direção a jusante na direção do promotor.

A modificação das histonas pode abrir a cromatina, permitindo assim a ligação seletiva de fatores de transcrição que, por sua vez, recrutam a RNA polimerase

II. Vários níveis e tipos de modificações de histonas estão correlacionados com os níveis de ativação da cromatina. Por exemplo, estudos de imunoprecipitação foram realizados baseados em experimentos usando anticorpos para determinar que a acetilação da histona H3 e a metilação no resíduo de lisina K4 pareciam coincidir entre si e coincidem durante com a ativação transcricional em embriões de galinha. A metilação no resíduo de lisina K9 marcou a cromatina inativa.

3.2.2.4 Metilação do DNA

Outro mecanismo de controle da expressão gênica em eucariotos é conhecido como metilação do DNA ou modificações químicas ligadas à estrutura da cromatina. As bases nitrogenadas citosinas no DNA dos eucariotos, podem ser modificadas pela adição de grupos metil na posição 5-carbono. Os resíduos de citosina alterados são em geral próximos a um nucleotídeo de guanina, resultando em dois resíduos de citosina metilados colocados diagonalmente um ao outro em fitas opostas de DNA (Figura 17). A metilação da cadeia do DNA não deve ser confundida com a metilação das histonas, a metilação do DNA envolve a conversão de bases citosinas do DNA eucariótico em 5-metilcitosina, resultando na repressão da transcrição, particularmente em vertebrados e plantas. O DNA é metilado especificamente nas Cs(citosinas) que precedem as Gs(guaninas) na cadeia do DNA (dinucleotídeos CpG ou ilhas CpG). Esta metilação está associada com a atividade transcricional reduzida de genes que contêm altas frequências de dinucleotídeos CpG nas proximidades de seus promotores. A metilação inibe a transcrição desses genes através da ação de uma proteína, a MeCP2 (Methyl-CpG Binding Protein 2), que se liga especificamente ao DNA metilado e reprime a transcrição. A MeCP2, se liga ao DNA metilado e recruta o complexo Sin-3-HDAC para a cromatina. O HDAC então, remove grupos acetil das histonas, tornando a cromatina compacta, deixando a estrutura do nucleossomo mais fechada sem acesso aos fatores de transcrição ao DNA resultando na transcrição gênica silenciada.

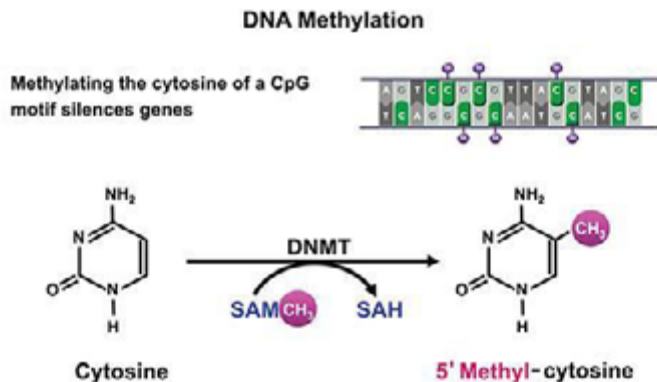


Figura 17. Metilação do DNA e Ilhas CpG. Um grupo metil é adicionado à posição 5-carbono dos resíduos de citosina no DNA. Fonte: <https://www.ncbi.nlm.nih.gov/books/NBK9904/>

3.2.2.5 RNAs não codificantes (nc-RNAs)

Os RNAs das células dos organismos eucariotos estão presentes em dois grupos: RNAs codificantes e RNAs não-codificantes. Os RNAs codificantes são transcritos e traduzidos originando proteínas, enquanto, os não-codificantes são transcritos, mas não originam proteínas. As pesquisas científicas tem evidenciado que aproximadamente 2% do genoma humano codifica proteínas, e por muito tempo se considerou que os 98% restantes do genoma, eram material inativo, ou seja, regiões de vários megabases sem qualquer função, chamado “DNA lixo”. Porém, com os avanços no desenvolvimento de refinadas técnicas para avaliação de transcritos, como *microarrays*, sequenciamento de última geração (NGS), bem como ferramentas de análises de bioinformática, foi observado que o genoma é amplamente transcrito. Assim, foi constatado que a transcrição de genomas eucarióticos produz uma infinidade de transcritos não-codificantes, que são atualmente reconhecidos como fundamentais reguladores envolvidos em diversos processos celulares incluindo, diferenciação e desenvolvimento celular, *imprinting* genômico, regulação da expressão gênica, controle do ciclo celular e adaptação a mudanças ambientais. A seguir estudaremos os grupos de RNAs não codificantes.

Os RNAs não codificantes (ncRNAs) são representados em duas classes: *housekeeping* ncRNAs e ncRNAs reguladores. Os *housekeeping* ncRNAs são os RNAs de transferência (tRNAs) e ribossômicos (rRNAs). Os ncRNAs reguladores se dividem em duas subclasses de acordo com sua função e seu tamanho: os pequenos RNAs não codificantes (snc-RNA), com menos de 200 nucleotídeos e RNAs longos não codificantes (lnc-RNA), com tamanho maior que 200 nucleotídeos. snc-RNA incluem microRNAs, pequenos RNAs nucleolares (snoRNAs), pequenos RNAs de interferência (siRNAs), RNAs nucleares (snRNAs) e *PIWI-iinteracting* RNAs (piRNAs) (Figura 18).

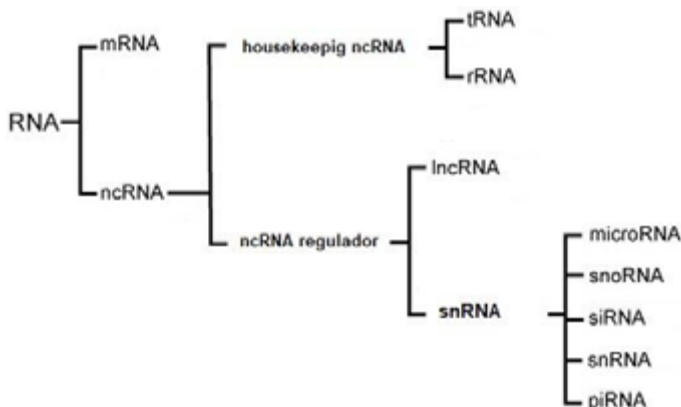


Figura 18. Classes de RNA. Os RNAs são divididos em duas classes principais: RNA mensageiro (mRNA), codificador de proteína e RNA não codificante (ncRNA). Os ncRNAs incluem os ncRNAs (housekeeping) e os ncRNAs (reguladores). Esse último grupo, é dividido em lncRNAs e sncRNAs. Os sncRNAs são subclassificados em snc-RNA representados por microRNAs, pequenos RNAs nucleolares (snoRNAs), pequenos RNAs de interferência (siRNAs), RNAs nucleares (snRNAs) e *PIWI-iinteracting* RNAs (piRNAs). Fonte:(Adaptado de INAMURA, 2017).

A seguir estudaremos os grupos de RNAs não codificantes.

3.2. 2. 5.1 Micro-RNAs (miRNAs)

Os miRNAs são RNAs não codificantes (de 19 a 22 nucleotídeos), de fita simples, transcritos no núcleo da célula eucariótica pela RNA polimerases II ou pela RNA polimerase III e que tem função de regular a expressão de diferentes genes no organismo em nível pós-transcricional, como por exemplo, durante o desenvolvimento. Em humanos, a maior parte dos microRNAs identificados exerce o efeito pelo processo de repressão da tradução. Aproximadamente 50% dos microRNAs identificados no genoma humano estão em sítios genômicos associados ao desenvolvimento de tumores. Foram inicialmente descritos no verme *Caenorhabditis elegans* e, a partir daí, milhares de miRNAs foram identificados, compreendendo o grupo mais bem estudado dos snRNAs (RNAs nucleares).

Os genes codificadores dos miRNAs estão localizados em todo o genoma e uma grande proporção é encontrada organizada dentro clusters compreendendo vários miRNAs. Os miRNAs são codificados dentro ou se sobrepõem a genes que codificam proteínas ou não codificantes, estando sua expressão ligada à transcrição e processamento desses genes.

Na biogênese dos miRNAs observa-se que após a transcrição são formados precursores conhecidos como pri-miRNA de dupla fita com uma alça que tem função de estabilizar o pri-miRNA tendo em vista que essa molécula é uma estrutura muito instável devido principalmente ao seu tamanho e deve ser protegido das enzimas endonucleases até chegar ao citoplasma. (Figura 19). Assim como o mRNA, o pri-miRNA passa por modificações pós-transcricionais que incluem o capeamento da extremidade 5' e a poliadenilação em 3'. Os pri-miRNA exibem uma estrutura em grampo com uma cauda de nucleotídeo, que em seguida é processada por dois importantes eventos: primeiro, o pri-miRNA deve ser metabolizado pela ação do complexo proteico nuclear enzimático, formado pela proteína DGCR8 e pela endonuclease DROSHA, que faz a clivagem da cauda do pri-miRNA, originando o pré-miRNA (Figura 19). Na sequência, o pré-miRNA é transportado do núcleo para o citoplasma através da proteína de membrana, a Exportin5 e a RAN-GTP. O segundo evento, acontece no citoplasma, onde pré-miRNA vai ser convertido em micro-RNA processo mediado pela ação do complexo enzimático chamado DICER que, interagindo com TRBP, cliva o pré-miRNA perto do loop terminal liberando duplex de RNA (para manter a estabilidade). Esse micro-RNA duplex, em seguida se separa em fitas simples para se tornar um micro RNA maduro. Esse micro-RNA maduro, pode controlar (impedir) o processo de tradução do RNA mensageiro de duas maneiras: Primeiro, se ligando ao RNAm pelo processo de pareamento incompleto de bases reprimindo a tradução por impedir que o ribossomo se desloque ao longo do RNAm (Figura 19). Segundo

o micro-RNA degrada o RNA mensageiro se o pareamento de bases entre o micro-RNA e o RNA mensageiro for completo. Nos dois casos, não ocorre a tradução da proteína. Em humanos, a maior parte dos micro-RNAs já identificados exerce o efeito pelo processo de repressão da tradução pelo pareamento incompleto de bases. O processo de interação entre o miRNA-RISC induz à repressão da tradução por dificultar o acesso da maquinaria básica de tradução ou pode favorecer a degradação do mRNA. A degradação do mRNA alvo, se dá pela retirada do cap 5' e da cauda poli-A em 3'. Os miRNAs podem também influenciar a biogênese de outros miRNAs ou outras espécies de RNAs, podem também ser incorporados em vesículas ou micropartículas, atuando como moléculas de sinalização para transferência de informação genética entre células ou tecidos.

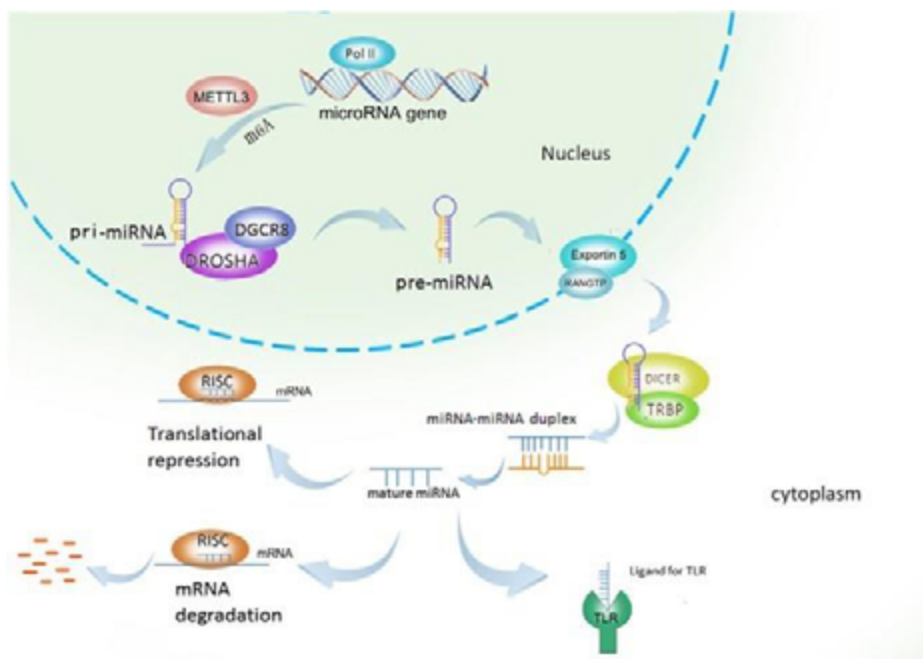


Figura 19. Biogênese e função de microRNAs (miRNAs). Os miRNAs são transcritos como precursores (pri-miRNA) ou derivados de íntrons. O processo de interação entre o miRNA-RISC induz à repressão da tradução por dificultar o acesso da maquinaria básica de tradução ou pode favorecer a degradação do mRNA. Os miRNAs também podem ser secretados para o meio extracelular interagindo com proteínas ou através de vesículas, atuando como moléculas de sinalização. Fonte: www.nature.com/sigtrans

3.2.2.5.2 RNAs longos não codificantes (lnc-RNAs)

RNAs longos não codificantes (lncRNAs) são uma classe de RNAs com tamanhos superiores a 200 nucleotídeos que não possuem potencial codificante, mas atuam como reguladores da expressão gênica tanto na transcrição quanto na pós-transcrição através de diversos mecanismos moleculares. Isso os distin-

que de pequenos RNAs não codificantes, como microRNAs (miRNAs) e pequenos RNAs interferentes (siRNAs). Os lncRNAs foram por muito tempo considerados “lixo” transcricional, mas, nos últimos anos, estudos crescentes revelaram que estas moléculas desempenham papéis cruciais em vários processos fisiológicos e patológicos, como regulação da expressão gênica, desenvolvimento embrionário e carcinogênese.

Com o desenvolvimento do sequenciamento de RNA e da bioinformática, um grande número de lncRNAs foi identificado. Algumas classes deles são geradas a partir de elementos distintos de DNA nos genomas, incluindo promotores, intensificadores, regiões intergênicas e a fita oposta de genes codificadores de proteínas, enquanto outros são produzidos a partir de longos transcritos.

Os lncRNAs são moléculas típicas de RNA transcritas pela RNA polimerase II e polimerase III e abrigam uma capa de metil-citosina 5' e uma cauda 3' poli (A). De acordo com suas diferentes características, os lncRNAs são classificados em muitos tipos diferentes. Por exemplo, os lncRNAs podem ser divididos em cinco classes com base em suas origens genômicas distintas, incluindo sense, antisense, bidirecional, intrônico e intergênico (Figura 20). Além disso, de acordo com a sua localização celular, os lncRNAs também podem ser divididos em lncRNAs nucleares, lncRNAs citoplasmáticos e lncRNAs mitocondriais.

A biogênese dos lncRNAs ocorre no núcleo das células e podem ser transcritos pelas RNA Pol II ou III. Suas regiões promotoras podem ser epigeneticamente reguladas por modificações de histonas/remodelamento de cromatina bem como por ligação e regulação de fatores de transcrição que favorecem ou dificultam a expressão gênica. O processamento pós-transcricional é outra particularidade que os lncRNAs compartilham com genes codificadores de proteínas: uma grande proporção de lncRNAs também é capeada com 7-metilguanilato à extremidade 5' terminada por trifosfato não modificada de uma molécula de RNA e poliadenilada na sua extremidade 3'. Também podem sofrer splicing (normal ou alternativo), processos de edição de RNA e padrões de ativação transcricional. Os lncRNAs podem ser encontrados em diferentes compartimentos da célula, podendo permanecer no núcleo, serem transportados para o citoplasma ou até mesmo exportados para fora das células por vesículas.

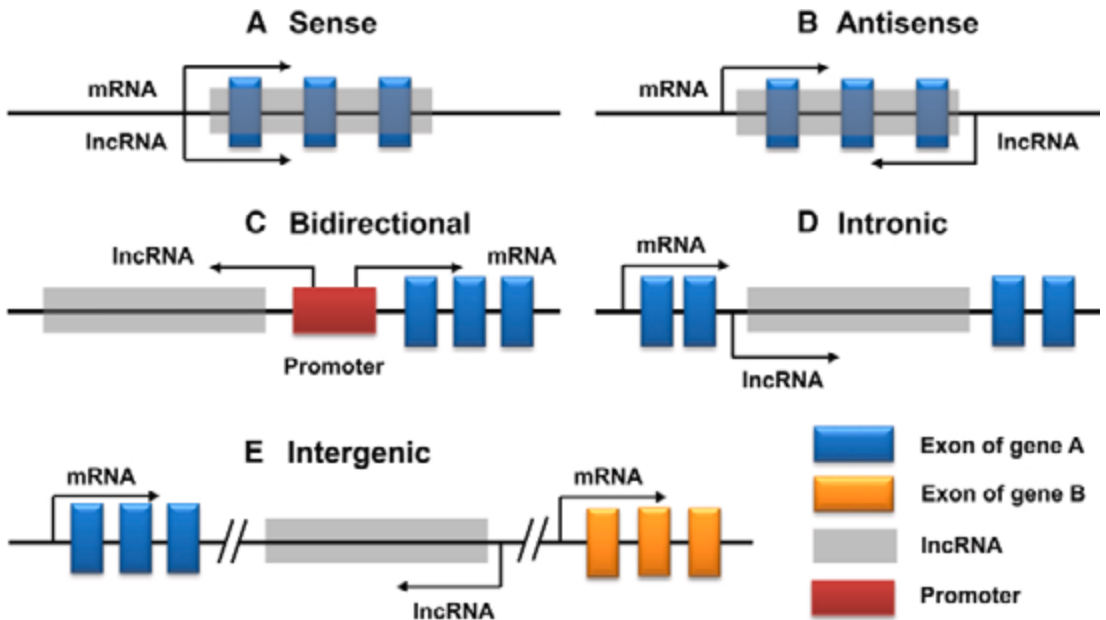


Figura 20. De acordo com sua origem genômica os lncRNAs podem ser classificados em cinco classes: (A) lncRNA sense ou (B) antisense que se localiza dentro ou se sobrepõe aos exons do gene codificador de proteína associado na mesma fita ou na fita oposta. O lncRNA antisense transcreve na direção oposta do gene codificador de proteínas. (C) O lncRNA bidirecional localiza próximo ao promotor (dentro de 1 kb) do gene codificador de proteína associado e transcreve na direção oposta. (D) O lncRNA intrônico surge de íntrons longos e transcreve de dentro de um íntron do gene codificador de proteína associado. (E) O lncRNA intergênico origina-se do segmento intergênico de dois genes codificadores de proteínas. Fonte: Mol Ther Oncolytics. 2021 23: 458–476.

REFERÊNCIAS

Bermúdez M., Aguilar-Medina M., Lizárraga-Verdugo E., Avendaño-Félix M., Silva-Benítez E., López-Camarillo C., Ramos-Payán R. lncRNAs as regulators of autophagy and drug resistance in colorectal cancer. *Front. Oncol.* 2019;9: 1008.

Wu H., Yang L., Chen L.L. The diversity of long noncoding RNAs and their generation. *Trends Genet.* 2017;33: 540–552

Zhang X., Hong R., Chen W., Xu M., Wang L. The role of long noncoding RNA in major human disease. *Bioorg. Chem.* 2019;92: 103214.

Ying Liu, Wei Ding, Wanpeng Yu, Yuan Zhang, Xiang Ao, and Jianxun Wang. Long non-coding RNAs: Biogenesis, functions, and clinical significance in gastric cancer. *Mol Ther Oncolytics.* 2021, 17; 23: 458–476.

Phillips, T. & Shaw, K. (2008) Chromatin Remodeling in Eukaryotes. *Nature Education* 1(1):209.

Nina Bögershausen and Bernd Wollnik. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. REVIEW article. Frontiers in Molecular Neuroscience Volume 11 - 2018.

Maia, Maria de Mascena Diniz Conceitos básicos de epigenética para universitários / Maria de Mascena Diniz Maia, Isaura Isabelle Fonseca Gomes da Silva. 1ª. ed. - Recife: EDUFRPE, 2020. 87 p.: il. Inclui referências. ISBN: 978-65-86547-17-7.

Benjamin A. Pierce. Genética – Um Enfoque Conceitual. 5ª. Ed-2017.

GUO, X. et al. Advances in long noncoding RNAs: identification, structure prediction and function annotation. Briefings in Functional Genomics, v. 15, n. 1, p. 38–46, 2016.

BEERMANN, J. et al. Non-coding RNAs in development and disease: background, mechanisms, and therapeutic approaches. Physiological Reviews, v. 96, p. 1297–1325, 2016.

Bruna Aguiar do Amaral, Cassiano Francisco Weege Nonaka, Roseana de Almeida Freitas, Lélia Batista De Souza, Leão Pereira Pinto. MicroRNAs – Biogenesis, functions and its potential role in oral carcinogenesis. Odontol. Clín.-Cient., Recife, 9 (2) 105-109, abr./jun., 2010 www.cro-pe.org.br

Singh SK, Bhadra MP, Girschick HJ, Bhadra U. MicroRNAs – micro in size but macro in function. FEBS Journal, 2008; 275: 4929-4944.

Schmittgen TD. Regulation of microRNA processing in development, differentiation and cancer. J Cell Mol Med, 2008; 12: 1811-1819.

Zhang W, Dahlberg JE, Tam W. MicroRNAs in tumorigenesis: a *primer*. Am J Pathol, 2007; 171: 728-738.

Yong Peng and Carlo M Croce. REVIEW ARTICLE: The role of MicroRNAs in human cancer. Signal Transduction and Targeted Therapy (2016) 1, 15004; doi:10.1038/sigtrans.2015.4 www.nature.com/sigtrans

OZSOLAK, F. et al. Chromatin structure analyses identify miRNA promoters. Genes & Development, v. 22, p. 3172–3183, 2008.

Inamura Kentaro. Review: Major Tumor Suppressor and Oncogenic Non-Coding RNAs: Clinical Relevance in Lung Cancer. Cells 2017, 6, 12; doi:10.3390/cells6020012.

ANDRES-PABLO, A.; MORILLON, A.; WERY, M. LncRNAs, lost in translation or licence to regulate? Current Gene Therapy, v. 63, p. 29–33, 2017.

<https://www.ncbi.nlm.nih.gov/books/NBK9904/> VISITADO EM 14 DE MARÇO DE 2024.

BÖGERSHAUSEN, N.; WOLLNIK, B. Mutational landscapes and phenotypic spectrum of SWI/SNF-related intellectual disability disorders. *Frontiers in Molecular Neuroscience*, n. 252, v. 11, p. 1–18, 2018.

KATO, A.; KOMATSU, K. RNF20-SNF2H pathway of chromatin relaxation in DNA double-strand break repair. *Genes*, v. 6, p. 592–606, 2015.

LU, C.; ALLIS, C. D. SWI/SNF complex in cancer: ‘remodeling’ mechanisms uncovered. *Nature Genetics*, v. 49, n. 2, p. 178–179, 2017.

T Tsukiyama and C Wu. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, 15;83(6):1011-20. 1995.

Modificações pós-traducionais das histonas (1library.org) visitado em 14 de março de 2024

KOLUDROVIC, D. et al. Chromatin-Remodelling Complex NURF is essential for differentiation of adult melanocyte stem cells. *PLoS Genetics*, v. 11, n. 10, p. 1–29, 2015.

J Füllgrabe , N Hajji, B Joseph. Cracking the death code: apoptosis-related histone modifications. *Cell Death Differ*, 2010;17(8):1238-43.

SAWAN, C.; HERCEG, Z. Histone modifications and cancer. *Advances in Genetics*, v. 70, p. 58–85, 2010.

CAO, J.; YAN, Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response and cancer. *Frontiers in Oncology*, v. 2, n. 26, p. 1–9, 2012.

<https://www.ncbi.nlm.nih.gov/books/NBK9904/> visitado em 12 de março de 2024.

Andrew J Bannister and Tony Kouzarides, Review: Regulation of chromatin by histone modifications. *Cell Res*. 201,21(3):381-95.

<http://www.78stepshealth.us/plasma-membrane/eukaryotic-chromosomes.html> visitado em 12 de março de 2024.

https://pt.slideshare.net/janaina_leitinho/sintese-de-proteina-2. visitado em 12 de março de 2024.

ALBERTS, B. et al. *Molecular biology of the cell*. 6. ed. United States: Garland Science, 2017, p. 1463.

BARALLE, F. E.; GIUDICE, J. Alternative splicing as a regulator of development and tissue identity. *Nature Reviews: Molecular Cell Biology*, v. 18, p. 437–451, 2017.

RAMANOUSKAYA, T. V.; GRINEV, V. V. The determinants of alternative RNA splicing in human cells. *Molecular Genetics and Genomics*, v. 292, p. 1175–1195, 2017.

Cell and Molecular Biology 2005 4^a ed. Jonh Wiley & Sons.

LEE, R. C.; FEINBAUM, R. L.; AMBROST, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, n. 75, p. 843–854, 1993.

CAPÍTULO 4: REAÇÃO EM CADEIA DA POLIMERASE

Fernanda Cristina Bezerra Leite
Lêda Ferreira de Oliveira

4.1 Introdução

A PCR (Polymerase Chain Reaction - Reação em Cadeia da Polimerase) foi descrita em meados da década de 1980 por Mullis e Faloona. É uma técnica poderosa e relativamente simples que permite a síntese enzimática *in vitro* de milhões de cópias de um segmento específico de DNA na presença da enzima DNA polimerase (DNApol).

Anteriormente, a obtenção de fragmentos de DNA dependia de técnicas de clonagem *in vivo*, que são laboriosas, dependem de grandes quantidades de amostras e necessitam de dias para atingir os resultados. Nesse contexto, a PCR revolucionou o campo da Biologia Molecular, visto que trouxe consigo alta especificidade (uma simples cópia do gene pode ser extraída de uma mistura complexa da sequência genômica de forma rotineira) e sensibilidade (é capaz de amplificar a partir de pequena quantidade de DNA), além de apenas horas para a obtenção dos produtos.

Esses fatores contribuíram para sua popularização e amplo uso, não somente na ciência básica, como também na pesquisa aplicada e uso comercial, como em exames genéticos, diagnósticos moleculares de patógenos, genética forense, controle de qualidade industrial e melhoramento vegetal e animal. Dada sua importância, em 1993, o prêmio Nobel de Química laureou os seus desenvolvedores.

4.2 Fundamentos da PCR

A PCR é uma técnica básica em laboratórios de Biologia Molecular. O método tradicional, também conhecido como PCR específica, amplifica fragmentos de DNA, a partir de uma sequência molde extensa, usando um par de *primers* específicos (*forward* e *reverse*) que flanqueiam a região-alvo. Os *primers* ou iniciadores são pequenas moléculas de DNA (oligonucleotídeos) fita simples sintetizadas artificialmente que possuem sequência complementar a uma região específica que delimitam o local de interesse. Eles fornecem o grupo 3' hidroxila (3'-OH) livre para o início da síntese da nova cadeia de nucleotídeos pela DNA polimerase.

Atualmente, com a disponibilidade de sequências genômicas em bancos de dados públicos, é possível isolar genes ou outros tipos de sequências de DNA de interesse pelo emprego da PCR, já que os *primers* podem ser prontamente de-

senhados pelo uso de softwares, muitos disponíveis em versões gratuitas e *online*. Comumente, esses programas seguem alguns critérios essenciais para o desenho dos *primers*, como: a) presença de uma extremidade hidroxila livre no carbono 3'; b) conteúdo GC bem distribuído em todo o *primer* (35% a 60%); c) complementaridade com a fita molde (é tolerável algum mau-pareamento ou *mismatch*, exceto na extremidade 3'); d) temperatura de anelamento de 52 °C a 58 °C e não ultrapassar a diferença de 5 °C entre eles); e) tamanho do *primer* de 10 a 30 nucleotídeos; f) distância mínima e máxima entre os *primers* de 20 pares de bases (pb) a 20 kilobases (kb) (média de 100 pb e 5 Kb); g) evitar o auto-pareamento (*self-annealing*) e formação de dímeros (*dimers*). Ao final, a análise visual detalhada pelo pesquisador é primordial para a escolha do par a ser testado experimentalmente.

Numa reação, além do DNA molde, par de *primers* e DNA polimerase, são adicionados: solução tampão, cloreto de magnésio ($MgCl_2$) e uma mistura equimolar de desoxirribonucleotídeos trifosfatos (dNTPs). O primeiro reagente propicia o pH ótimo para a reação, enquanto o íon magnésio (Mg^{+2}) atua como cofator, ativando a DNA polimerase. Por fim, os dNTPs representam os nucleotídeos necessários à síntese da cadeia, sendo do tipo dATP, dTTP, dCTP e dGTP. Comumente, esses reagentes são transferidos com o auxílio de micropipetadores para microtubos de 200 microlitros (μL) de capacidade. Portanto, uma PCR é realizada em pequenos volumes, variando de 10 a 100 microlitros.

Uma vez preparadas, as reações são depositadas no termociclador, um equipamento que as expõe a uma série de diferentes temperaturas durante intervalos de tempos pré-determinados. Esses ajustes de tempo e temperatura correspondem ao ciclo de amplificação. Basicamente, cada ciclo de PCR ocorre em três etapas: 1) desnaturação ou *melting*, que consiste na separação da dupla-fita de DNA a ser amplificado, ocorrendo com a elevação da temperatura (90°C a 95°C); anelamento, *annealing* ou hibridização, quando os iniciadores ou *primers* ligam-se por complementaridade de bases à sequência de interesse, tornando o molde de DNA parcialmente dupla-fita (45°C a 65°C); 3) extensão ou alongação, caracterizada pela ligação da DNA polimerase à extremidade 3' do *primer*, iniciando a síntese do fragmento de interesse, cuja temperatura ótima é cerca de 72°C (Figura 1).

Sendo uma reação em cadeia, essas etapas são programadas para serem repetidas algumas dezenas de vezes. Teoricamente, ao final de cada ciclo de PCR são formadas o dobro de moléculas iniciais e, portanto, considera-se que a reação ocorre de modo exponencial (2^n , onde n corresponde ao número de ciclos). Logo, no final da reação, cerca de 20 ciclos são suficientes para amplificar na ordem de mais de um milhão de vezes a região de interesse, enquanto 30 ciclos amplificam mais de um bilhão de fragmentos. Na prática, como a eficiência da reação não é de 100% e devido à alteração nas concentrações de substratos e produtos ao longo do processo, esse número é menor. Todavia, a quantidade final é suficiente para ser analisada de modo satisfatório.

Além do uso de termocicladores, uma importante condição na automatização do processo foi a substituição da DNA polimerase originalmente utilizada (fragmento Klenow da bactéria *Escherichia coli*) por outra termoestável extraída da bactéria termofílica *Thermus aquaticus* (*Taq* DNA polimerase – *Taq* DNA pol), a qual resiste às diferentes temperaturas utilizadas nas três etapas que compreendem um ciclo de PCR. Os experimentos costumam ser otimizados de acordo com a complexidade do DNA molde e da combinação de *primers*. Além disso, recomenda-se incluir nos ensaios controles negativos e positivos, a fim de descartar problemas de execução e da qualidade dos reagentes. Neste caso, o controle negativo representa uma reação de PCR isenta de DNA molde. Ao final, os produtos amplificados podem ser diretamente analisados por eletroforese quanto ao tamanho e quantidade. Também podem ser empregados como componentes de outros procedimentos moleculares, como na clonagem vetorial e reações de sequenciamento de DNA.

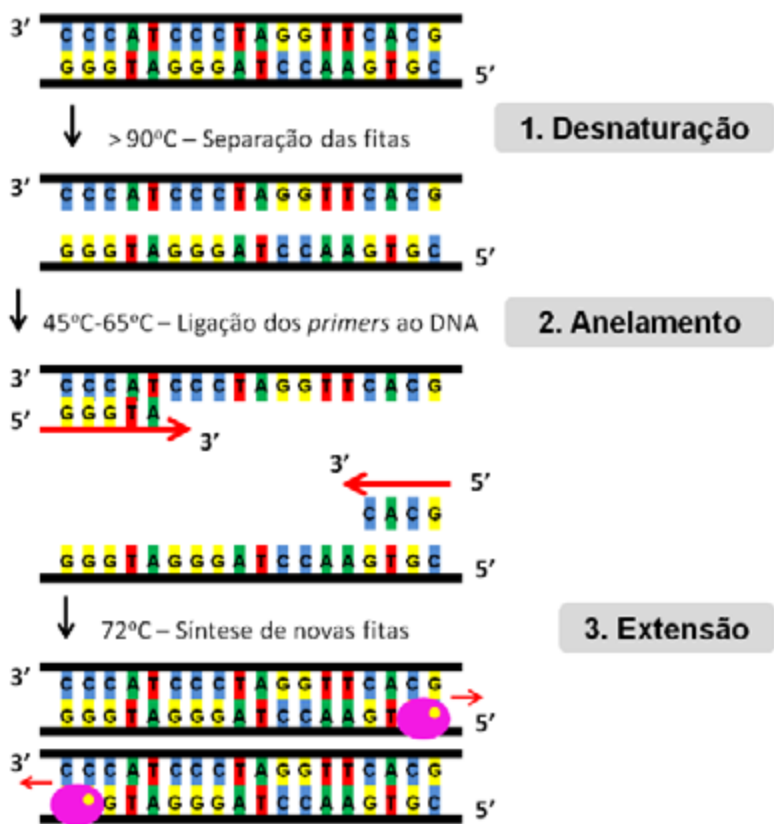


Figura 1. Ciclo de PCR. Na etapa de desnaturação (1), as altas temperaturas promovem o rompimento das pontes de hidrogênio, separando as duas fitas de DNA. Na etapa de anelamento (2), cada *primer* hibridiza em uma das fitas do DNA molde, de modo que a orientação da síntese do DNA ocorra na região localizada entre eles. Na etapa de extensão (3), a DNA polimerase (representada pelo círculo cor de rosa), uma vez ativada pelos íons magnésio (círculo em amarelo), promove a adição dos nucleotídeos no sentido 5'-3'. Ao final, ocorre a duplicação do segmento de DNA.

Devido à sua versatilidade, ao longo do tempo, a PCR passou por diversas modificações. A maioria envolve alterações nos tipos e número de *primers* usados nas reações. Algumas dessas variações são descritas no capítulo sobre marcadores moleculares.

4.3 Transcrição reversa

Na década de 1970, um novo tipo de enzima foi identificado como parte do ciclo de vida de retrovírus. Denominadas de transcriptases reversas, são DNA polimerases dirigidas por RNA capazes de sintetizar uma molécula de DNA complementar (DNAc) a partir de um RNA mensageiro (RNAm), *primer*, dNTPs e íons magnésio. Posteriormente, o DNAc obtido pode ser amplificado por PCR convencional com a *Taq* DNA Pol.

A técnica de RT-PCR (Reverse Transcription-Polymerase Chain Reaction - Transcrição Reversa-Reação em Cadeia da Polimerase) permite avaliar a presença de moléculas de RNA numa determinada amostra, permitindo tanto a detecção como também a quantificação da expressão gênica. Além disso, é fundamental para a análise do material genético de vírus de RNA, sobretudo na detecção molecular de patógenos desse grupo. Também possibilita a identificação de variantes de *splicing* durante o processamento do RNAm.

4.4 PCR em Tempo Real

O ensaio de PCR em tempo real é um método quantitativo capaz de monitorar passo a passo o progresso da PCR, ou seja, é capaz de detectar e quantificar as moléculas de ácidos nucleicos amplificadas (DNA ou DNAc) durante cada ciclo de reação. Isto deve-se ao uso de compostos fluorescentes capazes de ligar-se à dupla-fita de moléculas de DNA e ao uso de termocicladores capazes de detectar esta fluorescência. Consequentemente, o acúmulo de fluorescência é monitorado durante o processo, ciclo a ciclo da reação, e não ao final, permitindo o acompanhamento e a obtenção de dados com mais rapidez, pois dispensa a necessidade de realizar a eletroforese após a PCR.

Entre os kits comerciais disponíveis, destacam-se aqueles que empregam SYBR Green I, um corante fluorescente (fluoróforo) altamente específico que se liga ao DNA fita-dupla, detectando os produtos de PCR acumulados durante os ciclos. Como cada vez mais fragmentos são sintetizados ao longo da PCR, o resultado é o aumento da fluorescência, cuja intensidade é proporcional à quantidade de produto de PCR produzido durante a fase exponencial (Figura 2). Desse modo, SYBR Green I pode ser usado para monitorar a amplificação de qualquer sequência de DNA fita-dupla (Figura 3) e, como não requer o uso de sondas, facilita o planejamento dos ensaios e diminui seus custos. Sua principal desvantagem consiste na

geração de sinais falso-positivos, devido à detecção de fragmentos inespecíficos e dímeros de *primers*.

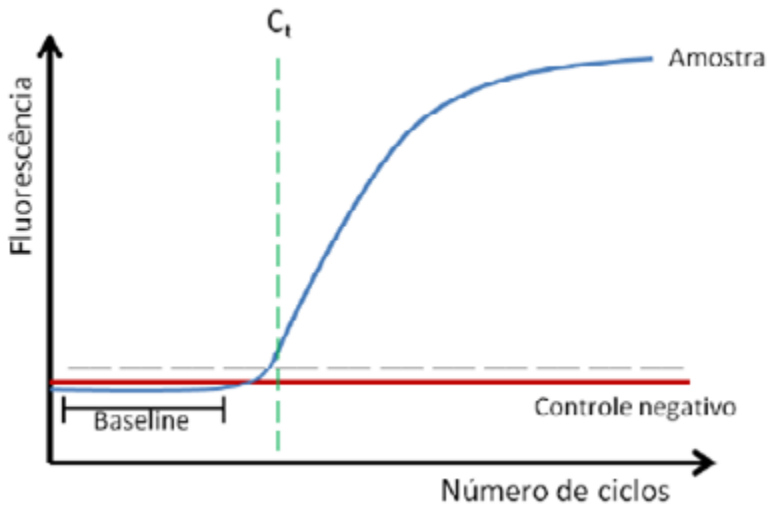


Figura 2. Representação de uma curva de amplificação por PCR em Tempo Real. A partir do gráfico, observa-se a relação entre o aumento da fluorescência detectada e o aumento no número de ciclos na amostra amplificada (linha azul). Nos ciclos iniciais de PCR, há pouca detecção no sinal fluorescente (baseline). O *cycle threshold* ou C_t (linha vertical pontilhada) corresponde ao número de ciclos a partir dos quais a fluorescência gerada pela reação é detectada acima do *background* (linha horizontal pontilhada). No controle negativo, não há amplificação e, portanto, ausência de emissão de fluorescência.

Com a finalidade de aumentar a especificidade de ensaios desenvolvidos com este método, o uso de sondas com alta especificidade ao alvo de amplificação pode ser adotada. Neste caso, além de desenhar os *primers* específicos, o pesquisador deverá também desenhar uma sonda de alta especificidade a uma região interna ao alvo de amplificação, a qual deverá anelar-se ao alvo sequência-específica, assim como os *primers* da reação (Figura 4). Esta sonda, uma curta sequência de nucleotídeos, possui em suas extremidades 5' e 3' dois componentes químicos. Ligada à extremidade 3', há um componente emissor de fluorescência (ou Repórter - R) e, à extremidade 5', há um componente capaz de capturar esta fluorescência (Quencher - Q). Desta forma, a sonda íntegra não é capaz de emitir tal fluorescência, devido à proximidade entre os compostos emissor e capturador. Uma vez que a amplificação tenha sido iniciada pela DNA polimerase, a partir da sequência de *primers*, ao encontrar a sonda pareada ao filamento-molde, a atividade exonuclease da DNA polimerase irá degradar a sonda, excisando nucleotídeo a nucleotídeo e, portanto, removendo o componente capturador da extremidade 5', distanciando-o do componente emissor de fluorescência. Desta forma, a fluorescência emitida pode ser capturada pelo equipamento termociclador. Assim, este método garante maior especificidade aos fragmentos detectados.

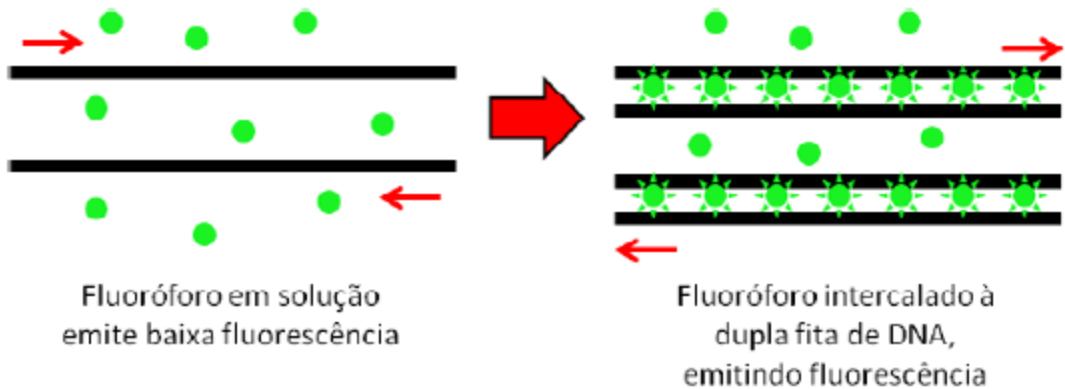


Figura 3. Sistema SYBR Green para detecção de amplificação por qPCR. Quando o DNA (barras pretas) está desnaturado, o fluoróforo (pontos verdes) está liberado e a fluorescência é drasticamente reduzida. Após a polimerização, o fluoróforo liga-se ao DNA fita-dupla, resultando no aumento da fluorescência, que é proporcional à quantidade de produtos amplificados, que são detectados pelo instrumento. Setas em vermelho indicam sentido da amplificação.

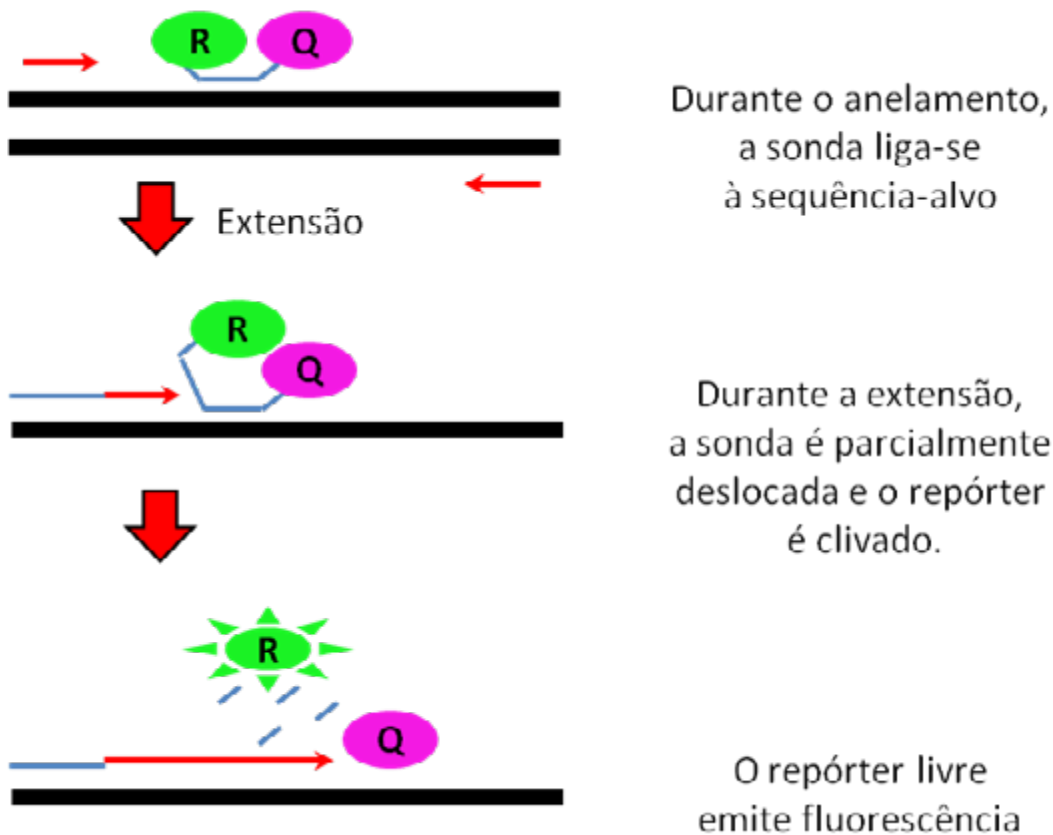


Figura 4. Sistema de sonda TaqMan para detecção de amplificação por qPCR. A atividade exonucleásica da DNA polimerase hidrolisa a sonda, permitindo a emissão de fluorescência. As barras pretas representam o DNA. Setas em vermelho indicam sentido da amplificação. R: Repórter. Q: Quencher.

A transcrição reversa seguida por PCR em tempo real ou RT-PCR quantitativa (RT-qPCR) é uma poderosa ferramenta que pode amplificar e detectar transcritos de baixa abundância, RNA parcialmente degradado e amostras limitadas de tecido. Essa técnica é empregada em vários estudos, como na análise do padrão de expressão diferencial em importantes processos fisiológicos, eficácia de tratamentos e na validação de ensaios com microarranjos de DNA. Entre suas vantagens, comparadas aos métodos convencionais, estão a alta sensibilidade e especificidade, boa reprodutibilidade e amplo intervalo de quantificação. O Quadro 1 resume as principais diferenças entre os sistemas SYBR Green e Sonda Taqman.

QUADRO 1 - Comparação entre os métodos SYBR Green e Sonda Taqman

Características	SYBR	Taqman
Especificidade	Inespecífico	Específico
Custo	Menor	Maior
Permite <i>multiplex</i> ?	Não	Sim
Distingue alelos?	Sim, mas de tamanhos diferentes	Sim

4.5. Quantificação Absoluta e Quantificação Relativa por qPCR

A Quantificação Absoluta (QA) e a Quantificação Relativa (QR) são dois diferentes métodos utilizados para analisar dados de qPCR. A Quantificação Absoluta determina o número absoluto de cópias do transcrito de interesse, interpolando o sinal de PCR de amostras cuja concentração é desconhecida a uma curva padrão absoluta. Por exemplo, QA pode ser usada para correlacionar o número de cópias de um vírus com o estado de uma doença. Nesse caso, o interesse é conhecer o número exato de cópias do ácido nucleico alvo em uma amostra biológica, a fim de monitorar o progresso de uma doença.

Em outras situações, é desnecessário determinar o número absoluto de cópias de um transcrito, mas é importante avaliar a mudança relativa na expressão gênica em comparação a uma condição referência. Por exemplo, saber que um dado tratamento aumentou a expressão de um determinado gene em 2,5 vezes, em relação ao seu controle, é mais relevante do que saber que o tratamento aumentou a expressão desse gene de 1.000 para 2.500 cópias. Desse modo, a Quantificação Relativa ou Quantificação Comparativa analisa mudanças na expressão de um gene alvo em uma amostra biológica em relação a um grupo de referência, como um controle não tratado ou uma amostra no tempo zero em um estudo realizado ao longo do tempo. QR não requer curva de calibração e a amostra de referência pode ser qualquer trans-

crito (gene referência ou controle endógeno), que tenha expressão gênica estável, não variável, dentre as condições teste e referência analisadas, ou seja, um gene cuja expressão seja constitutiva. Alguns exemplos desses genes de referência são: β -actina, gliceraldeído 3- fosfato desidrogenase (GAPDH) e RNA ribossomal.

No entanto, para estudos de expressão relativa por PCR em tempo real, vários controles são necessários para garantir a validade e precisão dos dados obtidos. Isto inclui verificação da qualidade de extração, rendimento, grau de pureza e integridade do RNA, eficiência da transcrição reversa e da reação de PCR, quantidade uniforme de RNA adicionado em cada reação, entre outros. Mesmo com todo este ajuste refinado, a normalização entre as amostras biológicas continua a ser etapa crítica na utilização desta técnica, sendo, o tratamento mais comumente adotado, a normalização da quantidade de RNA total com relação a um gene referência, como citado anteriormente. Alguns modelos matemáticos já foram propostos para cálculo da expressão relativa de um determinado gene-alvo em relação ao gene referência a citar: método $2^{-\Delta\Delta Ct}$, proposto por Livak e colaboradores em 2001, e que apesar de ter sido bastante aceito, este método de análise assume que as eficiências de amplificação dos genes teste e referência sejam ótimas e idênticas e que a expressão do gene referência seja perfeitamente estável para o conjunto de amostras em questão; e o método GeNorm, proposto por Vandesompele e colaboradores em 2002, o qual baseia-se na média geométrica de um número de candidatos a genes de referência para uma determinada condição. Este método de análise resulta num fator de normalização (FN) para cada uma das amostras a serem testadas.

Esse método pode ser aplicado para investigar pequenas mudanças fisiológicas na expressão gênica, como na resposta a um agente tóxico ou resposta a uma infecção, na determinação de padrões de circuitos regulatórios, na caracterização de genes envolvidos em respostas a drogas e na descrição de padrões temporais e espaciais de expressão gênica. Com a finalidade de padronizar a quantidade da amostra de RNA ou DNA adicionada à reação, executa-se a amplificação de um gene de expressão constitutiva ou gene de referência (controle endógeno), que pode ser co-amplificado no mesmo tubo, em um ensaio multiplex, ou ser amplificado em tubos separados.

REFERÊNCIAS

AZEVEDO, M. O. et al. Técnicas básicas em Biologia Molecular. Brasília: Editora da Universidade de Brasília, 2003. 211 p.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods*, v. 25, n. 4, p. 402-408, 2001.

MATIOLI, S. R. Ribeirão Preto: Holos Editora, 2001. 202 p.

PCR amplification. Disponível em: <https://www.promega.com.br/resources/guides/nucleic-acid-analysis/pcr-amplification/>. Acessado em: 15 de abril de 2024.

VANDESOMPELE, J.; DE PRETER, K.; PATTYN, F.; POPPE, B.; VAN ROY, N.; DE PAEPE, A.; SPELEMAN, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*, v. 3, n. 7, p. 34, 2002.

CAPÍTULO 5: CLONAGEM GÊNICA

Fernanda Cristina Bezerra Leite

Na década de 1970, a biologia experimental passou por uma grande revolução a partir do desenvolvimento de diversas ferramentas moleculares que, em conjunto, constituem o que chamamos de Tecnologia do DNA recombinante. Estas ferramentas possibilitaram que o pesquisador pudesse, em laboratório, criar e/ou modificar moléculas de DNA e ainda transferir um segmento de DNA de um organismo doador para outro organismo hospedeiro.

A clonagem gênica pode ser utilizada para diversos fins, por exemplo:

- a) clonagem de um determinado segmento de DNA com o objetivo de posterior sequenciamento;
- b) manipulação de uma determinada sequência de DNA, resultando na modificação da sequência codificante, visando engenharia de proteínas;
- c) expressão de proteínas heterólogas em sistemas celulares hospedeiros.

De forma didática, a clonagem gênica convencional pode ser explicada em três etapas:

5.1 Construção de molécula de DNA recombinante

Para transferir uma sequência codificante de um organismo doador para um organismo hospedeiro, os cientistas utilizam pequenas moléculas de DNA como vetores de clonagem, a exemplo dos plasmídeos, para carrear esta sequência. Os plasmídeos são moléculas DNA dupla-fita circulares, capazes de armazenar segmentos de DNA de interesse e de serem reconhecidos pela célula hospedeira. Originalmente, os plasmídeos foram identificados em bactérias e um dos seus principais papéis é de armazenar sequências de DNA codificantes que conferem resistência a antibióticos. Portanto, para que um plasmídeo tenha a função de vetor de clonagem, esta molécula precisa conter três importantes características (Figura 1):

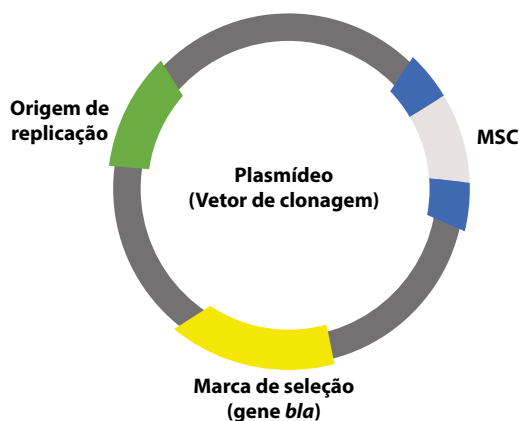


Figura 1. Plasmídeo (Vetor de clonagem). MSC. Múltiplo Sítio de Clonagem. Gene bla. Gene que confere resistência à ampicilina.

a) Origem de replicação: sequências que os tornam reconhecíveis pelo sistema de replicação da célula hospedeira, denominadas origem de replicação, permitindo a replicação desta molécula de DNA acessória e sua propagação por sucessivas gerações;

b) Marca de seleção: um gene que confere à célula hospedeira uma nova característica fenotípica que permitirá distinguir as células transformadas (que receberão o plasmídeo) das células não transformadas (que não receberão o plasmídeo). Como exemplo de genes marcadores podemos citar: gene que confere resistência a um determinado antibiótico, gene que resulta na produção de um pigmento, gene que promove complementação metabólica relacionada a uma carência presente na célula hospedeira, dentre outros. Desta forma, ao utilizar um plasmídeo, cujo gene marcador confere resistência à ampicilina, é possível distinguir as células transformadas de não transformadas por sua habilidade em crescer em meio contendo ampicilina.

c) Múltiplo sítio de clonagem (MSC): representam um conjunto de sequências que são reconhecidas por endonucleases denominadas enzimas de restrição. Estas enzimas são capazes de promover a quebra do duplo-filamento da molécula de DNA circular, tornando-o um duplo-filamento linear. As sequências reconhecidas constituem palíndromos, podendo estar localizadas em ambas as fitas do duplo-filamento. O sítio de quebra da ligação fosfodiéster é altamente conservado e, portanto, resulta em extremidades altamente reprodutíveis. Outra característica importante da atividade das enzimas de restrição é que podem resultar em extremidades coesivas, permitindo que diferentes segmentos de DNA, ao serem expostos à mesma enzima de restrição, tornem-se coesivos, permitindo unir fragmentos de diferentes moléculas de DNA.

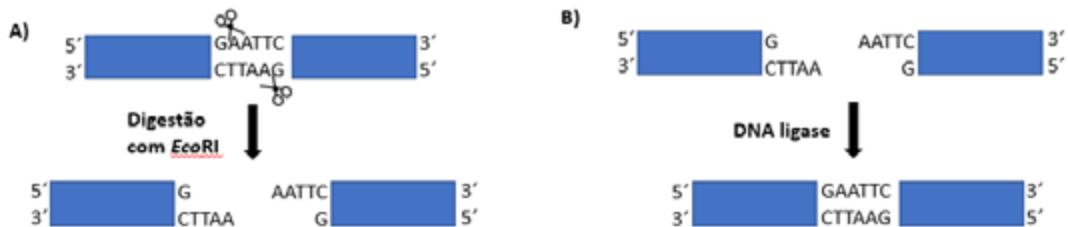


Figura 2. Enzimas utilizadas na construção de DNA recombinante. A) Atividade da enzima de restrição (endonuclease) EcoRI ao reconhecer o sítio de clivagem 5´- GAATTC-3´, resultando em extremidades coesivas. B) Atividade da enzima DNA ligase, que permite restaurar a ligação fosfodiéster entre nucleotídeos de extremidades coesivas resultantes da digestão por endonucleases.

Uma vez que um vetor de clonagem tenha sido escolhido por suas características, é necessário então isolar a sequência de DNA de interesse a ser clonado. No geral, este isolamento é realizado por meio da Reação em Cadeia da Polimerase (PCR), técnica discutida em capítulo anterior, utilizando *primers* específicos que ladeiam a região genômica a ser amplificada. Neste caso, visando à clonagem em vetor, usa-se o artifício de adicionar à extremidade 5´ dos *primers* as sequências de reconhecimento para as enzimas de restrição que serão utilizadas para a clonagem deste segmento de DNA na molécula do plasmídeo. Desta forma, os produtos da reação de PCR resultam na sequência desejada ladeada por sítios de restrição (Figura 3). Assim, por meio dos sítios de clonagem, é possível inserir a sequência do DNA de interesse no vetor de clonagem mediante a atividade da DNA ligase, resultando na molécula de DNA recombinante (Figura 4). A enzima DNA ligase é ATP-dependente para realizar a ligação fosfodiéster entre os nucleotídeos de segmentos de DNA, devendo-se adicionar ATP à reação *in vitro*.

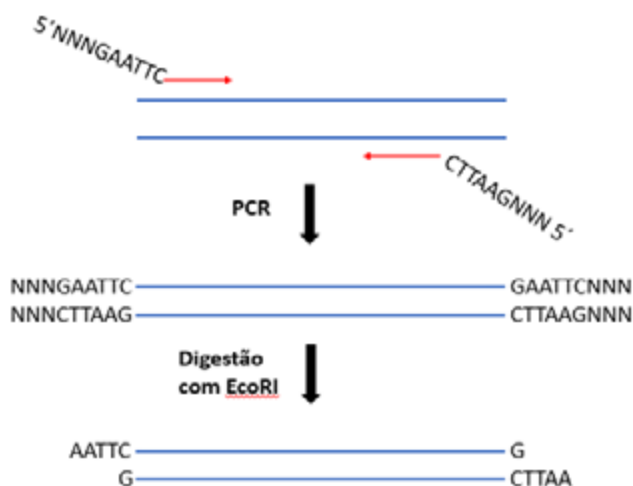


Figura 3. Amplificação por PCR e digestão enzimática de uma sequência nucleotídica de interesse pela endonuclease EcoRI. Primeiramente, a sequência-alvo é amplificada por *primers* contendo, além da sequência complementar ao molde, uma “cauda” com sequência correspondente ao sítio de restrição para EcoRI. Posteriormente, o produto de amplificação é clivado pela enzima de restrição.

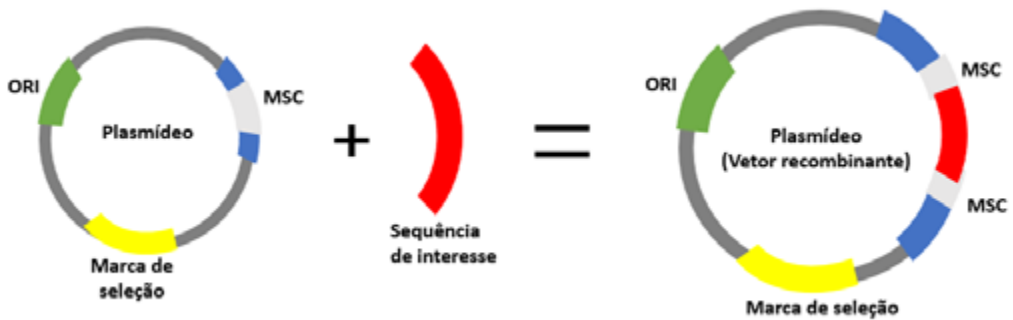


Figura 4. Construção de vetor recombinante. Após a clivagem por uma enzima de restrição no sítio MSC, a sequência de interesse é integrada ao vetor com o auxílio da DNA ligase.

5.2 Transformação da célula hospedeira

A depender da célula hospedeira, diferentes métodos de transformação celular podem ser utilizados com o objetivo de tornar a membrana e/ou parede celular mais permeável, permitindo a entrada da molécula de DNA recombinante no interior da célula.

A bactéria *Escherichia coli* é um microrganismo amplamente utilizado por sua facilidade de manipulação genética e rápido cultivo. Dois métodos são comumente utilizados para promover a transformação de células recombinantes: choque térmico e eletroporação. Uma célula, ao adquirir capacidade de receber DNA exógeno, é chamada de célula competente.

O método por choque térmico é precedido pela exposição das células a uma solução de cloreto de cálcio em presença de gelo, cujos íons reduzem a repulsão entre as cargas da membrana celular e da molécula de DNA recombinante. Um rápido choque térmico é realizado, ao transferir as células expostas ao DNA para uma banho-maria com temperatura de 42°C, por um breve tempo de 30 segundos, provocando a abertura de poros na membrana celular e permitindo a entrada do DNA exógeno na célula. Este método é relativamente simples, de baixo custo e de rápida realização.

Por outro lado, o método por eletroporação consiste em expor as células a um breve choque elétrico de milissegundos, por meio de um aparelho denominado eletroporador, com o objetivo de também promover a abertura de poros na membrana celular. O método de transformação por eletroporação resulta em maior número de células transformantes, ou seja, apresenta uma maior eficiência de transformação. É o método de escolha principalmente quando se deseja transformar células utilizando grandes moléculas de plasmídeos. Apesar disso, a possibilidade de promover quebras cromossômicas e consequentes rearranjos e a necessidade de aquisição de equipamento para seu uso são as suas principais desvantagens.

Vale ressaltar que nem todas as células expostas serão transformadas, ou seja, irão receber o plasmídeo recombinante. Assim, faz-se necessário que métodos de seleção se-

jam utilizados para isolar as células transformadas (desejáveis) das células não-transformadas (não desejáveis).

5.3 Seleção de células transformadas

O isolamento das células transformadas está intrinsecamente relacionado ao gene marcador presente no vetor de clonagem (plasmídeo) utilizado. A partir desta escolha, o meio de cultivo seletivo é também definido.

Por exemplo, se o plasmídeo contém o gene *bla* como gene marcador, o qual confere mecanismo de resistência ao antibiótico ampicilina, as células pós-transformação deverão ser plaqueadas em meio de cultivo contendo todos os nutrientes necessários ao crescimento da *E. coli* acrescido do antibiótico ampicilina. Uma vez que as células sejam colocadas em placa contendo este meio, apenas aquelas células que receberam o plasmídeo serão capazes de crescer e resultar numa colônia visível. Células não transformadas (sem plasmídeo) são susceptíveis ao antibiótico ampicilina e, portanto, não conseguem prosperar neste meio de cultivo (Figura 5A e 5B).

Além disso, é comum em vetores comerciais para uso em bactérias como células hospedeiras, a presença do gene *LacZ*, codificador da enzima beta-galactosidase. Quando expressa, esta enzima é capaz de hidrolisar o substrato comercial cromógeno, análogo à lactose, denominado X-gal (5-bromo-4-cloro-3-indoxil- β -D-galactopiranosídeo), resultando na liberação de um pigmento azul.

Nestes vetores de clonagem, a sequência referente ao MSC encontra-se interna à sequência do gene *LacZ*, de forma que a clonagem de sequências gênicas nos sítios de clonagem presentes no MSC interrompem o quadro de leitura desse gene, impedindo a sua expressão. Assim, este sistema permite diferenciar, por inspeção visual da cor da colônia, clones que possuem o vetor recombinante desejável, cuja sequência do gene *LacZ* foi interrompida pela clonagem do segmento de DNA no múltiplo sítio de clonagem, os quais resultam em colônias brancas (Figura 5C). Por outro lado, as colônias azuis, indicam células transformadas cujos vetores plasmidiais não são recombinantes.

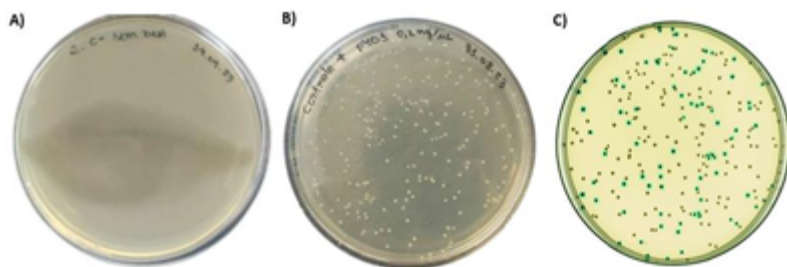


Figura 5. Placas ilustrando o plaqueamento de células de *Escherichia coli* após o ensaio de transformação utilizando vetor plasmidial. Todas as placas contém meio LB acrescido de ampicilina. (A) Controle negativo: células submetidas ao protocolo de transformação na ausência de plasmídeo. (B) Placa sem X-gal: células submetidas ao protocolo de transformação evidenciando clones. (C) Placa com X-gal: células submetidas ao protocolo de transformação evidenciando clones, porém colônias azuis ou brancas, indicando a presença de vetores sem e com sequência de interesse clonada, respectivamente.

Ao utilizar plasmídeos que não contenham o sistema *LacZ* para a inspeção de clones, é possível confirmar a correta construção do vetor recombinante utilizando o método de PCR de colônia (Figura 6). Para isso, uma pequena porção da colônia é adicionada diretamente a uma reação de PCR contendo *primers* que la-deiam a região MSC de inserção da sequência de interesse. O resultado é facilmente verificado pela diferença no tamanho, em pares de bases, do fragmento amplificado. O fragmento maior representa o vetor recombinante (plasmídeo + sequência de interesse), enquanto o fragmento menor representa o vetor não-recombinante (apenas o plasmídeo).

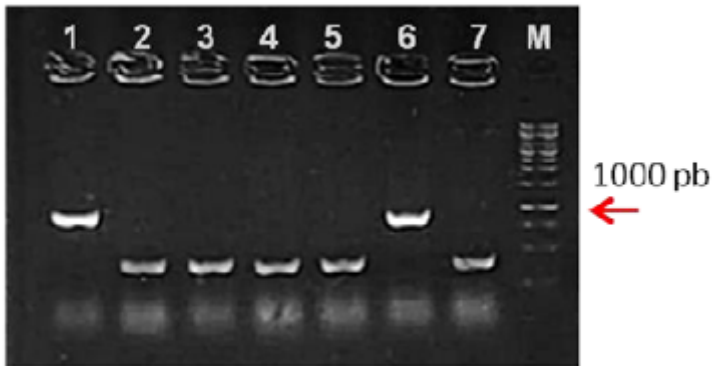


Figura 6. Eletroforese em gel de agarose de um ensaio de PCR de colônia para confirmação da transformação em *Escherichia coli*. Pistas 1 e 6: clones contendo vetor recombinante. Pistas 2, 3, 4, 5 e 7: clones contendo vetores não-recombinantes. M: Marcador molecular GeneRuler 1Kb.

REFERÊNCIAS

BROWN, T.A. Clonagem Gênica e análise de DNA. 4º Ed. EDITORA ARTMED. 2003

SNUSTAD, D.P.. Fundamentos da genética. . GUANABARA KOOGAN. 2008

CAPÍTULO 6: MARCADORES MOLECULARES

lêda Ferreira de Oliveira

A variação genética dos seres vivos pode ser analisada de diferentes modos. Por muito tempo, avanços nas pesquisas em Genética básica e aplicada foram alcançados pelo entendimento da diversidade biológica em nível morfológico. Características comuns visíveis, tais como cor dos olhos, presença ou ausência de estruturas, entre outros, foram usados como marcadores fenotípicos. No entanto, a disponibilidade de marcadores controlados por genes associados a caracteres morfológicos é bastante reduzida. Além disso, eles descrevem apenas uma pequena parte da variação genética subjacente, são sensíveis a fatores ambientais e são de difícil contextualização com processos biológicos.

6.1. Introdução

Os marcadores moleculares são valiosos porque são capazes de distinguir genótipos. Embora metabólitos secundários (antocianinas e compostos fenólicos, por exemplo) tenham sido as primeiras moléculas utilizadas como marcadores moleculares, sua instabilidade e disponibilidade limitada restringem seu amplo uso.

Por um curto período, com o desenvolvimento de novas metodologias bioquímicas, na década de 1960, os marcadores de proteínas tornaram-se bastante populares. Conhecidos como marcadores isoenzimáticos ou aloenzimáticos, a técnica consiste na identificação de diferentes alelos de um determinado gene por meio da análise de um perfil eletroforético de isoenzimas, obtidas de um extrato de tecido, separadas através de uma matriz de gel de amido, de agarose ou de poli(acrilamida). Duas consequências importantes decorrentes do emprego dessa tecnologia foram o aumento no número de marcadores genéticos disponíveis em pelo menos uma ordem de magnitude e sua aplicabilidade a potencialmente qualquer espécie.

Alguns anos mais tarde, com o surgimento dos novos sistemas de marcadores moleculares baseados em DNA, maior nível de polimorfismo pode ser revelado, uma vez que mutações em íntrons ou mesmo nos códonos de um gene podem potencialmente fornecer maior informação do que em nível de proteínas. De modo geral, comparados aos marcadores morfológicos, os marcadores moleculares são abundantes, já que são encontrados em maior número de locus e estes geralmente apresentam muita variabilidade; são neutros em relação a efeitos fenotípicos, com mínimo ou nulo efeito epistático ou pleiotrópico; são

versáteis, pois podem ser analisados a partir de amostras de células e/ou tecidos de diversos estágios do desenvolvimento; são objetivos, uma vez que sua identificação costuma ser de definição mais simples; permitem o estudo dos seres vivos em diferentes níveis, de indivíduos a domínios da vida.

6.2. Definição de marcadores moleculares

Apesar de ter descoberto o “gene”, Mendel desconhecia a natureza celular e molecular dos conceitos que estava elaborando. Atualmente, sabemos que os genes geralmente funcionam por meio das proteínas e que a informação genética está fisicamente localizada nos cromossomos. Por sua vez, cada cromossomo é composto por uma única molécula de DNA e proteínas que atuam em seu empacotamento. Portanto, levando em consideração o Dogma Central da Biologia Molecular e que diferenças na molécula de DNA podem refletir em diferentes fenótipos, podemos acessar e analisar a biodiversidade molecular em diferentes níveis (DNA, RNA ou proteínas) utilizando técnicas moleculares.

Segundo Ferreira e Grattapaglia (1998), marcador molecular é todo e qualquer fenótipo molecular oriundo da expressão de um gene, como é o caso das isoenzimas, ou de segmentos específicos de DNA (correspondentes a regiões expressas ou não do genoma). Quando se verifica seu comportamento de acordo com as leis básicas da herança Mendeliana, um marcador molecular é definido como marcador genético. É importante enfatizar que o simples fato de um marcador ser de DNA, um produto da transcrição ou da tradução de uma sequência de DNA, não implica que se constitua em um marcador “genético”, ou seja, que tenha segregação mendeliana.

As diferenças genéticas existem naturalmente entre os organismos e são produtos primários da mutação. Sua ocorrência está condicionada à detecção de, ao menos, duas formas alélicas diferentes para uma determinada característica. Quando elas são observadas em indivíduos de uma mesma espécie, são conhecidas como polimorfismos. Se estão acumuladas entre as espécies, recebem o nome de divergência genética. Por outro lado, quando a variação não é observada, trata-se de um monomorfismo. Já o conceito quantitativo de polimorfismo genético (P) está relacionado à proporção de loci variáveis em determinado genoma e é um importante indicador em estudos de diversidade genética populacionais.

Dados experimentais gerados por métodos moleculares podem ser usados para estudar polimorfismos genéticos. Por exemplo, após a separação de fragmentos de DNA de uma região específica por eletroforese em gel de agarose, as bandas monomórficas serão aquelas que mostram todas o mesmo tamanho, em pares de bases, para todos os indivíduos analisados. Por outro lado, quando ocorre variação no tamanho das bandas, a região será polimórfica (Fi-

gura 1). Logo, nenhuma análise genética é possível sem a existência de variantes (mutantes).

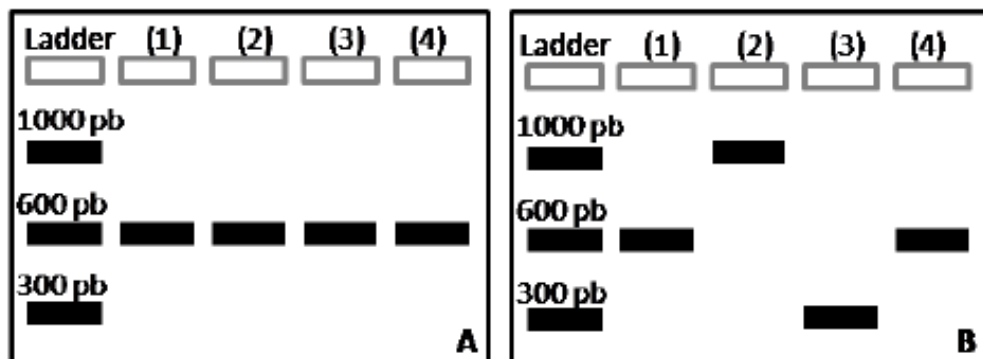


Figura 1. Representação de perfis de marcadores moleculares monomórficos (A) e polimórficos (B) após separação por eletroforese em gel. Em A, observa-se que todas as amostras possuem fragmentos de mesmo tamanho, com aproximadamente 600 pares de bases (pb). Em B, observa-se que as amostras apresentam diferentes tamanhos, variando de 300 pb a 1000 pb. O monomorfismo é resultante da falta de variação na região considerada (loco) para todas as amostras, enquanto o polimorfismo é decorrente de sua presença. Ladder: marcador de peso molecular. (1), (2), (3), (4): amostras.

6.3. Marcadores dominantes e codominantes

Com relação ao poder discriminatório dos genótipos analisados, os marcadores moleculares podem ser do tipo dominantes ou codominantes.

Os resultados de algumas técnicas moleculares são visualizados como a presença (1) ou ausência de uma banda (0) de tamanho específico, que são interpretadas como um loco gênico. Nesse caso, observam-se duas classes fenotípicas, dominante e recessiva, pois apenas uma de duas variantes alélicas existentes é expressa. Independentemente da causa para a ausência de banda, essa condição é interpretada como homozigótica recessiva. Por outro lado, na presença da banda, não é possível determinar se o loco corresponde a um genótipo homozigoto ou heterozigoto. Sendo assim, marcadores moleculares dessa natureza são chamados de dominantes (Figura 2). Em geral, tais marcadores são multilocos, com apenas dois alelos por loco, e são interpretados como dados binários (1 e 0), dada a presença ou ausência do alelo. Alguns exemplos são as técnicas de RAPD (Random Amplified Polymorphic DNA - Polimorfismo de DNA Amplificado Randomicamente), AFLP (Amplified Fragment Length Polymorphism - Polimorfismo de Comprimento de Fragmento Amplificado), DAF (DNA Amplification Fingerprinting - Impressão Digital de DNA Amplificado) e ISSR (Inter Simple Sequence Repeats - Região Entre Repetições de Sequência Simples).

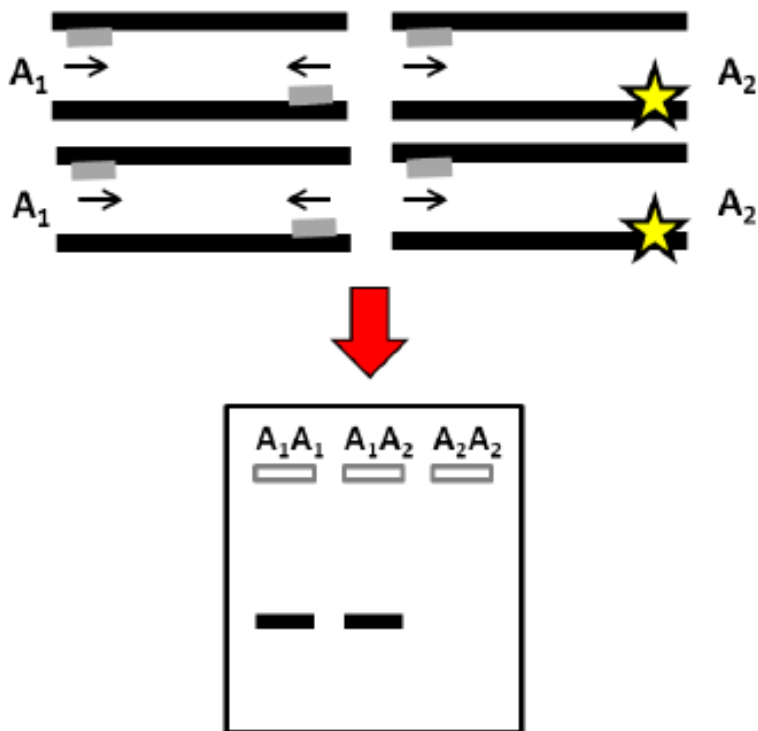


Figura 2. Base genética e detecção de marcadores dominantes. No exemplo, A_1 representa a forma alélica dominante, uma vez que a região-alvo (barras pretas) é amplificada por PCR e pode ser visualizada como uma banda no gel. A_2 representa a forma alélica recessiva, já que seu produto não pode ser visualizado no gel. A causa para isso é decorrente de alguma mutação (símbolo de estrela) no sítio de anelamento do *primer* (barra de cor cinza), ocasionando a não amplificação da região por ele flanqueada. Considerando indivíduos diplóides, da combinação desses dois alelos podem resultar três diferentes genótipos: A_1A_1 (homozigoto dominante) A_1A_2 (heterozigoto) e A_2A_2 (homozigoto recessivo). Observando-se o perfil no gel, como o fenótipo “presença de banda” pode corresponder aos genótipos A_1A_1 ou A_1A_2 , portanto, não sendo possível distingui-los um do outro, os marcadores são do tipo dominante. O fenótipo “ausência de banda” corresponde ao genótipo homozigoto recessivo (A_2A_2). As setas indicam a direção da amplificação.

Quando na análise molecular é possível visualizar a expressão fenotípica independente de cada variante alélica de um indivíduo, o marcador molecular é do tipo codominante. Portanto, estes marcadores permitem a distinção entre genótipos homozigotos e heterozigotos (Figura 3). Comumente, são interpretados como um único loco com vários alelos e, portanto, permitem o cálculo das frequências alélicas de uma população. Isoenzimas, RFLP (Restriction Fragment Length Polymorphism - Polimorfismo de Comprimento de Fragmento de Restrição), microssatélites ou SSR (Simple Sequence Repeats - Repetições de Sequência Simples) são alguns dos principais marcadores codominantes.

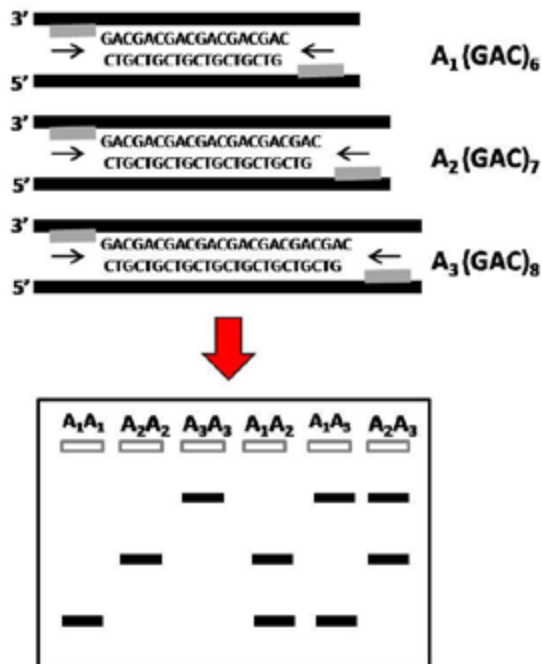


Figura 3. Base genética e detecção de marcadores codominantes. No exemplo, A_1 , A_2 e A_3 representam três diferentes variantes alélicas, cujas regiões-alvo (barras pretas) são reconhecidas pelos *primers* (barras de cor cinza) que as flanqueiam e são amplificadas por PCR. Porém, elas diferem quanto ao tamanho interno da região amplificada. Quando seus produtos são separados em gel, é possível identificar cada uma dessas variantes por tamanho, em pares de bases. Os genótipos homocigotos (A_1A_1 , A_2A_2 e A_3A_3) são determinados pela presença de uma única banda no gel de tamanho correspondente a sua variante, enquanto os heterocigotos são aqueles com duas bandas de tamanhos diferentes, cada uma correspondente ao seu respectivo alelo (A_1A_2 , A_2A_3 e A_1A_3). As setas indicam a direção da amplificação.

6.4. Principais marcadores moleculares baseados em PCR

Os avanços decorrentes da tecnologia do DNA recombinante e das técnicas de PCR (Polymerase Chain Reaction - Reação em Cadeia da Polimerase) e de sequenciamento de nucleotídeos permitiram um enorme desenvolvimento científico a partir da informação genética acessada. Atualmente, existem inúmeros marcadores moleculares de DNA, principalmente baseados em PCR e sequenciamento direto. Cada técnica possui suas próprias vantagens e limitações. Geralmente, elas diferem no poder de detecção das diferenças genéticas, nos tipos de dados gerados e na sua aplicabilidade a níveis taxonômicos em particular. Portanto, o cientista deve ter muito claro qual a pergunta a ser respondida ao optar pela escolha de determinado marcador molecular em detrimento de outro. Além disso, ele deve considerar se a infraestrutura disponível, os recursos financeiros, o tempo, a disponibilidade de recursos humanos com treinamento apropriado e conhecimento genético da espécie estudada são adequados para a execução do trabalho. Quanto à infraestrutura básica necessária, um laboratório básico

de Biologia Molecular precisa contar com insumos e equipamentos para a realização de extração de DNA, amplificação de material genético via PCR, separação por eletroforese horizontal em gel de agarose, fotodocumentação e análise estatística dos dados.

Nos últimos anos, a gama de métodos disponíveis para a detecção da variabilidade genética a nível de sequência de DNA e RNA, com o desdobramento das -ômicas, permitiu a obtenção de um número virtualmente ilimitado de marcadores, cobrindo todo o genoma do organismo. Apesar de poderosas, as técnicas moleculares baseadas em sequenciamento ainda são caras e oferecem melhor resolução nos níveis acima de intragênero, enquanto os dados gerados pelos marcadores clássicos de DNA são capazes de identificar a variação genética dentro da espécie, podendo discriminar indivíduos, com um custo e tempo reduzidos. A seguir, algumas das principais técnicas moleculares baseadas em PCR são brevemente descritas, com o intuito de esclarecer seu funcionamento básico.

6.4.1. PCR-RFLP

Por algum tempo, os marcadores RFLP (Restriction Fragment Length Polymorphism - Polimorfismos de Comprimento de Fragmento de Restrição) foram amplamente utilizados na análise da variabilidade genética. A metodologia consiste na digestão de amostras de DNA com enzimas de restrição, que reconhecem sítios específicos e fazem cortes. Os fragmentos de DNA são posteriormente separados de acordo com o tamanho por eletroforese em gel, transferidos para uma membrana e hibridizados com sondas sequência-específicas marcadas. Após revelação por técnicas de autorradiografia ou fluorescência, os polimorfismos são identificados e analisados. As variações observadas são interpretadas como eventos de mutação que criam ou eliminam sítios de reconhecimento para as enzimas de restrição. Portanto, dependendo da constituição genética do indivíduo analisado, diferentes tamanhos de bandas podem ser observados no gel.

Com o surgimento da PCR, a técnica RFLP tradicional foi modificada, eliminando muitas desvantagens. Desse modo, a metodologia PCR-RFLP (Polymerase Chain Reaction-Restriction Fragment Length Polymorphism - Reação em Cadeia da Polimerase-Polimorfismos de Comprimento de Fragmento de Restrição) passou a permitir a determinação de genótipos de indivíduos a partir da amplificação de uma sequência específica de DNA, seguida pela digestão enzimática por uma endonuclease de restrição e separação dos fragmentos por eletroforese em gel. Quando ocorre o reconhecimento de um único sítio de corte para a enzima de restrição utilizada, interno à sequência amplificada, uma das formas alélicas é identificada. Ela é evidenciada por duas bandas no gel, cuja soma corresponde ao tamanho total da sequência original amplificada em pares de bases. Caso não existam sítios de restrição, o produto será visualizado como uma única banda de tamanho esperado no gel, correspondendo ao outro alelo (Figura 4).

Entre suas principais vantagens, com relação à metodologia RFLP tradicional, estão: não necessitar de grandes quantidades de tecido para análise, permitindo seu

emprego em indivíduos muito pequenos ou em quantidade ínfimas de amostras; não requer o uso de sondas marcadas e, portanto, é mais segura, rápida e de menor custo; é um método alternativo ao sequenciamento direto de DNA; além disso, uma vez a sequência-alvo esteja disponível, os *primers* alelo-específicos podem ser desenhados e prontamente usados combinados às centenas de enzimas de restrição comercialmente disponíveis. Entre as desvantagens estão: cobrir uma região restrita do genoma, a possibilidade de não existência de polimorfismos identificáveis por esse método nos loci estudados e o alto custo de algumas enzimas de restrição.

Os marcadores produzidos por PCR-RFLP são do tipo codominantes, ou seja, permitem distinguir entre genótipos homocigotos e heterocigotos (Figura 4). Desse modo, esta técnica é amplamente utilizada na análise da diversidade genética entre indivíduos ou populações, identificação de associações entre variantes genéticas e características fenotípicas e diagnóstico de doenças genéticas, entre outros.

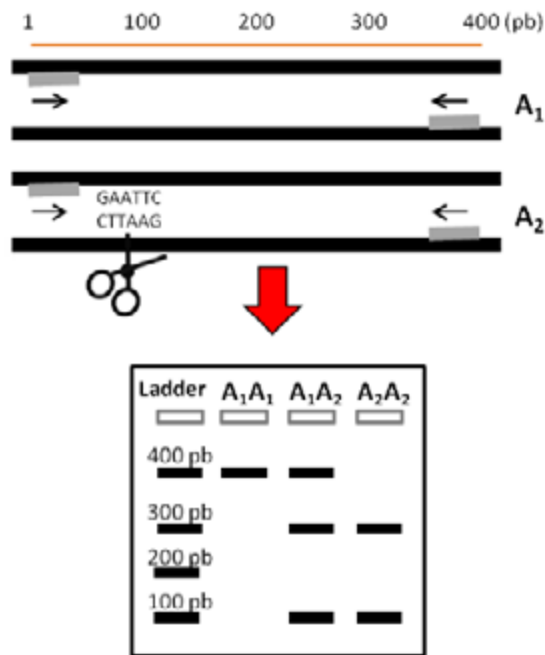


Figura 4. Base genética e detecção de polimorfismos de marcadores PCR-RFLP. No exemplo, primeiramente, amostras de DNA foram amplificadas por PCR pelo uso de um par de *primers* específicos (barras de cor cinza) que flanqueiam uma região de 400 pares de bases (pb) de comprimento (linha superior de cor laranja). Quando os produtos dessa reação são digeridos pela endonuclease EcoRI, se a variante alélica A_1 estiver presente, não ocorrerá clivagem. Por outro lado, na presença da variante A_2 , devido à existência do sítio de restrição para EcoRI (GAATTC), ocorre clivagem aproximadamente na região correspondente aos 100 primeiros pares de nucleotídeos (símbolo de tesoura). Quando seus produtos são separados em gel, é possível identificar cada uma das variantes. O fragmento de 400 pb corresponde ao alelo A_1 , enquanto os fragmentos de 100 pb e de 300 pb juntos correspondem ao alelo A_2 . Consequentemente, indivíduos com genótipo A_1A_1 são aqueles que mostram apenas uma única banda no gel de 400 pb; indivíduos com genótipo A_2A_2 são aqueles que mostram ambas as bandas de 100 pb e 300 pb e os indivíduos com genótipo A_1A_2 são aqueles que mostram as três bandas (100 pb, 300 pb e 400 pb) simultaneamente. Como a técnica permite identificar os indivíduos heterocigotos, este marcador é do tipo codominante. As setas indicam a direção da amplificação. Ladder: marcador de peso molecular.

6.4.2. Microssatélites ou SSR

Além das sequências codificantes, os genomas possuem longos trechos de sequências repetitivas. Uma dessas classes, conhecidas como microssatélites ou SSR (Simple Sequence Repeats - Repetições de Sequência Simples), são caracterizadas por apresentar repetições em *tandem* de di-, tri- ou tetranucleotídeos. Tipicamente, os microssatélites estão dispersos entre os diferentes cromossomos dos organismos, seja em regiões que codificam ou não proteínas. Possuem alta taxa de mutação e, portanto, são importantes para evolução no fornecimento e manutenção da variação genética. Acredita-se que tais eventos ocorram por erros durante a replicação do DNA ou por sua recombinação desigual.

Para a obtenção de marcadores microssatélites ou SSR, um par de *primers* específicos (20 a 30 nucleotídeos de bases), complementares à região de sequência única que flanqueiam a região repetitiva, são usados em uma PCR. A seguir, as diferenças no comprimento de cada segmento repetitivo que houver na amostra podem ser detectadas pela separação dos produtos por eletroforese (Figura 5). Considera-se que a região amplificada pelo par de *primers* corresponde a um loco gênico e que cada um dos fragmentos de diferentes tamanhos visualizados são suas variantes alélicas.

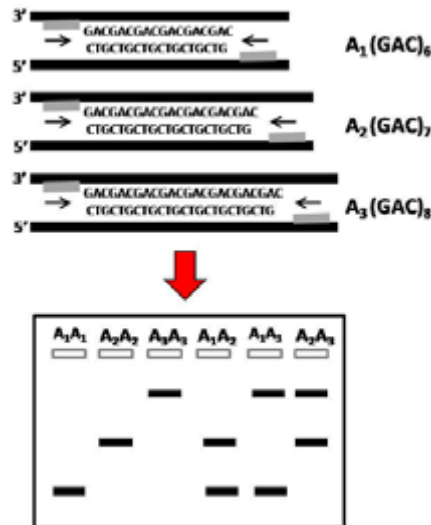


Figura 5. Base genética e detecção de polimorfismos de marcadores microssatélites ou SSR. No exemplo, uma região genômica (barras pretas) que compreende uma região microssatélite com motivo (GAC/CTG) foi amplificada por PCR pelo uso de um par de *primers* específicos (barras de cor cinza) que anelam na região de sequência única que delimita a região-alvo. Observa-se que as regiões podem variar quanto ao número de elementos repetitivos, os quais representam diferentes variantes alélicas, que podem ser identificadas por eletroforese como bandas de tamanhos correspondentes ao número de repetições. Como o alelo A_1 possui seis blocos de repetição, enquanto A_2 possui sete e A_3 possui oito, em um gel, eles são representados como bandas em ordem crescente de tamanho, em pares de bases, respectivamente. Desse modo, indivíduos com genótipo A_1A_1 são aqueles que exibem uma única banda no gel e esta possui menor tamanho. Indivíduos com genótipo A_2A_2 revelam uma única banda de tamanho intermediário, enquanto indivíduos A_3A_3 mostram uma única banda e é a maior. Os genótipos heterocigotos exibem duas bandas, cada qual correspondente a um dos diferentes alelos. Como a técnica permite distinguir genótipos homocigotos e heterocigotos, o marcador é do tipo codominante. Embora raros, podem ocorrer alelos nulos (falha na amplificação). As setas indicam a direção da amplificação.

Alternativamente, algumas melhorias na metodologia convencional já foram propostas. Uma delas foi aperfeiçoar a resolução obtida na separação dos fragmentos pela substituição da eletroforese em gel de poliacrilamida por eletroforese capilar, bem mais eficiente em detectar pequenas variações de tamanhos. Outra foi adotar a estratégia de PCR multiplex, que consiste na utilização, numa mesma amplificação, de pares de *primers* específicos para diferentes *loci* genômicos, desde que os alelos de cada *loco* possuam tamanho suficientemente diferente para migrar em zonas separadas no gel.

Geralmente, *loci* microssatélites são abundantes nos genomas (com alta cobertura) e seus alelos são altamente variáveis (multialelos). Apresentam padrão de herança codominante, são de análise simples e podem ser transferíveis para outras espécies, visto a existência de conservação de sítios microssatélites entre espécies relacionadas.

A principal limitação dessa técnica é decorrente da necessidade de muito trabalho, recursos e pessoal qualificado para seu desenvolvimento. A obtenção desses marcadores requer a construção de bibliotecas genômicas, seleção e posterior sequenciamento dos clones microssatélites, desenho de *primers* e validação das amplificações. Porém, com a crescente disponibilidade de genomas completos em bancos de dados públicos, o desenvolvimento de tais marcadores pode ser abreviado para as espécies descritas ou relacionadas.

Os microssatélites são a classe mais variável entre os marcadores moleculares, permitindo a estimativa das frequências alélicas nas populações. Além disso, são considerados marcadores neutros, ou seja, estão sob baixa pressão seletiva. Por isso, são valiosos nos estudos de diversidade genética. Também são amplamente usados no mapeamento físico e genético, no teste de paternidade, na análise clonal, entre outros.

6.4.3. RAPD

RAPD é uma técnica que detecta polimorfismos por meio da amplificação por PCR de uma sequência de DNA utilizando um único *primer* de sequência arbitrária, ou seja, sua sequência-alvo no genoma é desconhecida. Se houver duas sequências opostas, em cada uma das fitas complementares, que hibridizam com esse *primer*, dispostas de modo palindrômico, a uma distância igual ou menor que 2 Kb, haverá amplificação do segmento interno (Figura 6). Devido ao uso de *primers* curtos (8 a 10 nucleotídeos de comprimento) e de sequência arbitrária ou aleatória, elimina-se a necessidade de conhecer previamente a sequência-alvo, amplificando, simultaneamente, vários locos anônimos do genoma. Portanto, os *primers* usados em reações RAPD amplificam sequências de DNA simples ou DNA repetitivo. Diferentes *primers* usados com o mesmo DNA genômico originam produtos de PCR com diferentes números e tamanhos e vice-versa.

Uma característica fundamental dos marcadores RAPD é seu comportamento como marcador genético dominante. Logo, um indivíduo diplóide homocigoto dominante para aquele loco RAPD e um outro indivíduo heterocigoto são considerados como uma mesma classe fenotípica (presença de banda no gel) (Figura 6). A ausência de banda representa, na realidade, o conjunto de todos os outros alelos daquele loco que não podem ser amplificados devido a uma alteração na sequência de DNA no sítio de anelamento dos *primers* (mutação do tipo inserção, deleção, entre outros). Conseqüentemente, algumas premissas devem ser aceitas para interpretar tal marcador: os alelos recessivos são idênticos em estado; os alelos segregam de acordo com as Leis Mendelianas e os genótipos estão em Equilíbrio de Hardy-Weinberg (EHW).

Alterações nas concentrações de reagentes e/ou condições de PCR influenciam no padrão de bandas observado. Por isso, além do mecanismo de herança dominante, outra significativa limitação da técnica diz respeito à falta de reprodutibilidade dos resultados e à produção de artefatos. No entanto, o RAPD pode ser otimizado para cada espécie e as bandas confiáveis podem ser validadas pela realização de testes de reprodutibilidade. Entre as vantagens do RAPD, destacam-se: é uma técnica relativamente simples, barata (se comparada a outras técnicas moleculares) e eficiente; permite o trabalho com genomas desconhecidos; requer pouca quantidade de DNA (da ordem de nanogramas); produz grande número de locus a ser examinado; é facilmente adaptável ao laboratório.

Os marcadores RAPD são dos mais versáteis métodos disponíveis para analisar a diversidade genética, porque eles podem ser usados para determinar genótipos ou a identidade taxonômica de espécies; estimar a paternidade e relações de parentesco; analisar a estrutura de populações e a especiação de híbridos; gerar sondas específicas.

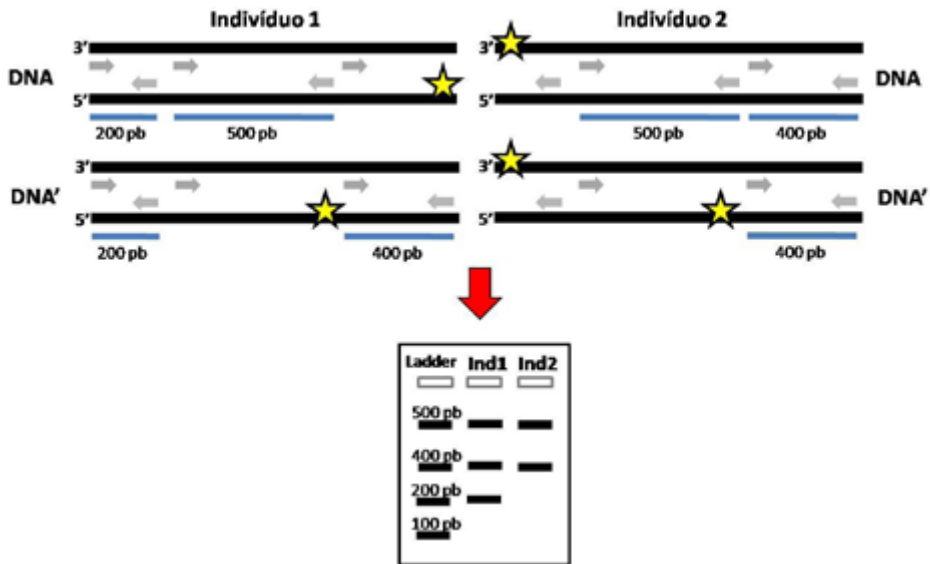


Figura 6. Base genética e detecção de polimorfismos de marcadores RAPD. No exemplo, são representados um par de cromossomos homólogos (DNA e DNA') de dois indivíduos diplóides e seus respectivos sítios de anelamento com um *primer* RAPD (setas de cor cinza). As estrelas representam mutações, impossibilitando o reconhecimento pelo *primer* e consequente amplificação. As linhas azuis representam os produtos de amplificação e seus respectivos tamanhos em pares de bases (pb). No indivíduo 1, observam-se duas mutações em diferentes cromossomos e regiões gênicas, inviabilizando as amplificações nesses alvos. Genotipicamente, o indivíduo 1 é homocigoto para a presença de amplificação no loco 200 pb e heterocigoto (presença e ausência de amplificação) para os loci 400 pb e 500 pb. Usando o mesmo raciocínio, o indivíduo 2 é homocigoto para o loco 200 pb, porém, para a ausência do produto de amplificação; heterocigoto para o loco 500 pb e homocigoto para a presença de ambos os amplicons para o loco 400 pb. Por sua vez, o fenótipo molecular, revelado por eletroforese em gel, não permite diferenciar genótipos homocigotos de heterocigotos, pois o produto da amplificação de uma das variantes alélicas é suficiente para ser evidenciada no gel. A análise é qualitativa, ou seja, só é capaz de discriminar entre a presença ou ausência dos fragmentos. Portanto, os marcadores RAPD são do tipo dominantes, já que tanto os indivíduos homocigotos para a presença de banda (homocigoto dominante) quanto os heterocigotos são capazes de expressar uma banda no gel. Também é possível observar que a técnica é capaz de reproduzir, num único experimento de PCR, a informação de diferentes locos gênicos (multilocos) e que estes estão distribuídos ao longo da molécula de DNA.

6.4.4. ISSR

A metodologia ISSR (Inter Simple Sequence Repeats - Região Entre Repetições de Sequência Simples) é muito semelhante à metodologia RAPD (Figura 6). Ela envolve a amplificação de um segmento de DNA localizado entre duas regiões microssatélites idênticas, orientadas em direções opostas, com distância suficiente para ser amplificada. Para isso, a técnica utiliza regiões microssatélites, com cerca de 16 a 25 pares de bases de comprimento, como *primers*. As repetições podem ser do tipo di-, tri-, tetra- ou pentanucleotídeos. Os *primers* podem ser não-ancorados ou comumente ancorados nas extremidades 3' ou 5' com uma a quatro bases degeneradas nas sequências flanqueadoras, acarretando em diferenças nos resultados (Figura 7).

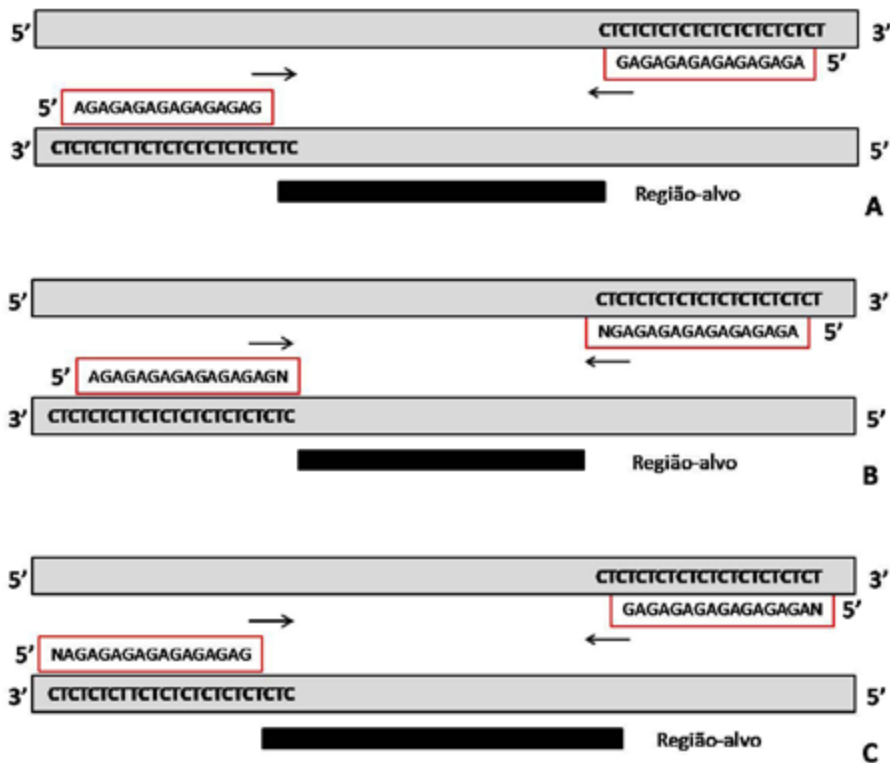


Figura 7. Representação de *primers* ISSR de sequência (AG)₈ do tipo não-ancorado (A) e ancorados na extremidade 3' (B) ou extremidade 5' (C). Um único *primer* costuma ser usado por PCR e ele é de sequência complementar a duas regiões de orientação oposta, localizadas nas diferentes fitas da mesma molécula de DNA, ricas em repetições CT, que flanqueiam a região-alvo de sequência única. Em A, o *primer* não-ancorado pode se ligar em diferentes trechos da repetição CT em cada ciclo da PCR, resultando em produtos de amplificação de diferentes tamanhos, observados em gel como bandas borradas. Em B, *primers* ancorados na extremidade 3' tendem a produzir produtos específicos, visualizados como bandas nítidas no gel. Em C, *primers* ancorados de modo específico na extremidade 5' podem amplificar trechos da repetição, gerando bandas de tamanho maior. Adaptado de Reddy et al. (2002).

Para a realização de uma reação de PCR, um único *primer* é comumente usado. Visto que ao longo de uma molécula de DNA existem diversas regiões repetitivas complementares ao *primer*, o resultado é a amplificação de diversos fragmentos de diferentes tamanhos, que são separados por eletroforese em gel e interpretados cada um como um loco. As mutações que ocorrem no sítio de anelamento do *primer*, impedindo a amplificação do fragmento, explicam a ausência da banda para determinado loco.

Desse modo, os resultados são observados como presença ou ausência de bandas, sendo os marcadores considerados dominantes com segregação segundo o padrão mendeliano de herança. Ocasionalmente, algum polimorfismo de variação de comprimento pode ocorrer devido ao uso de *primers* ancorados na região 5', que podem amplificar a região interna repetida do microssatélite com número maior ou menor de nucleotídeos presentes.

De modo geral, a técnica ISSR é considerada de simples execução, rápida, eficiente, de baixo custo e segura, pois dispensa uso de radiação. Os marcadores ISSR são mais reprodutíveis do que os RAPD, provavelmente, devido ao maior comprimento dos *primers*, que possibilitam maior estringência na etapa de anelamento. Além disso, os *primers* podem ser desenvolvidos pelo próprio pesquisador sem deter conhecimento prévio da sequência, permitindo definir comprimento, motivo e ancoragem. A partir de um conjunto de *primers* disponíveis no laboratório, é possível testar sua aplicabilidade para diferentes espécies. O uso de poliacrilamida como matriz do gel melhora a resolução da eletroforese, podendo aumentar, consequentemente, o número de bandas polimórficas observadas. A combinação de dois diferentes *primers* também pode ser testada para gerar número maior de produtos. Entre suas principais limitações, estão a natureza dominante dos marcadores, impedindo algumas análises genéticas, e a escassez de regiões microssatélites em alguns grupos de seres vivos.

Por serem abundantes, multilocos e exibirem alto polimorfismo, os marcadores ISSR são muito aplicados na obtenção de *fingerprints* de DNA, a fim de caracterizar germoplasmas, identificar clones, estabelecer parentesco entre indivíduos. Além disso, são usados em estudos sobre a análise da diversidade genética de populações e mapeamento genômico.

6.5. Considerações finais

Os marcadores moleculares de DNA podem ser meios mais fáceis e menos laboriosos de obter informações genéticas de diferentes organismos. As técnicas disponíveis são capazes de responder diversas questões (evolucionárias, taxonômicas, parentesco, melhoramento genético e conservação da biodiversidade), que não eram previamente possíveis apenas com o emprego dos métodos morfológicos tradicionais. Todavia, eles são metodologias alternativas e devem ser vistas como recursos complementares a outras abordagens, a fim de resolver problemas biológicos.

REFERÊNCIAS

FERREIRA, M. E.; GRATTAPAGLIA, D. Introducción al Uso de Marcadores Moleculares en el Análisis Genético. 1. ed. Brasília: EMBRAPA-CENARGEN, 220 pp, 1998.

REDDY, M. P.; SARLA, N.; SIDDIQ, E. A. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica. v. 128, p. 9-17, 2002.

Este livro nasceu do desejo dos autores de contribuir com a formação dos estudantes da área de Biologia sobre um tema muito atraente no campo da Genética, que é a Genética Molecular.

Embora a Genética seja atualmente assunto de grande evidência na mídia, é ainda considerada um campo de muita dificuldade no ensino de Ciências Biológicas como área interdisciplinar. O despertar do interesse para escrever tópicos de Genética Molecular veio da necessidade de oferecer aos estudantes de Biologia temas elaborados de forma clara e direta para melhor compreensão textual dos conceitos abordados em Genética Molecular. O objetivo principal é oferecer aos leitores uma fonte de conceitos básicos de Genética Molecular, descritos em uma linguagem simples, lógica e sintética, com a intenção de contribuir para exercitar e ampliar os conhecimentos em cada tema abordado.

