



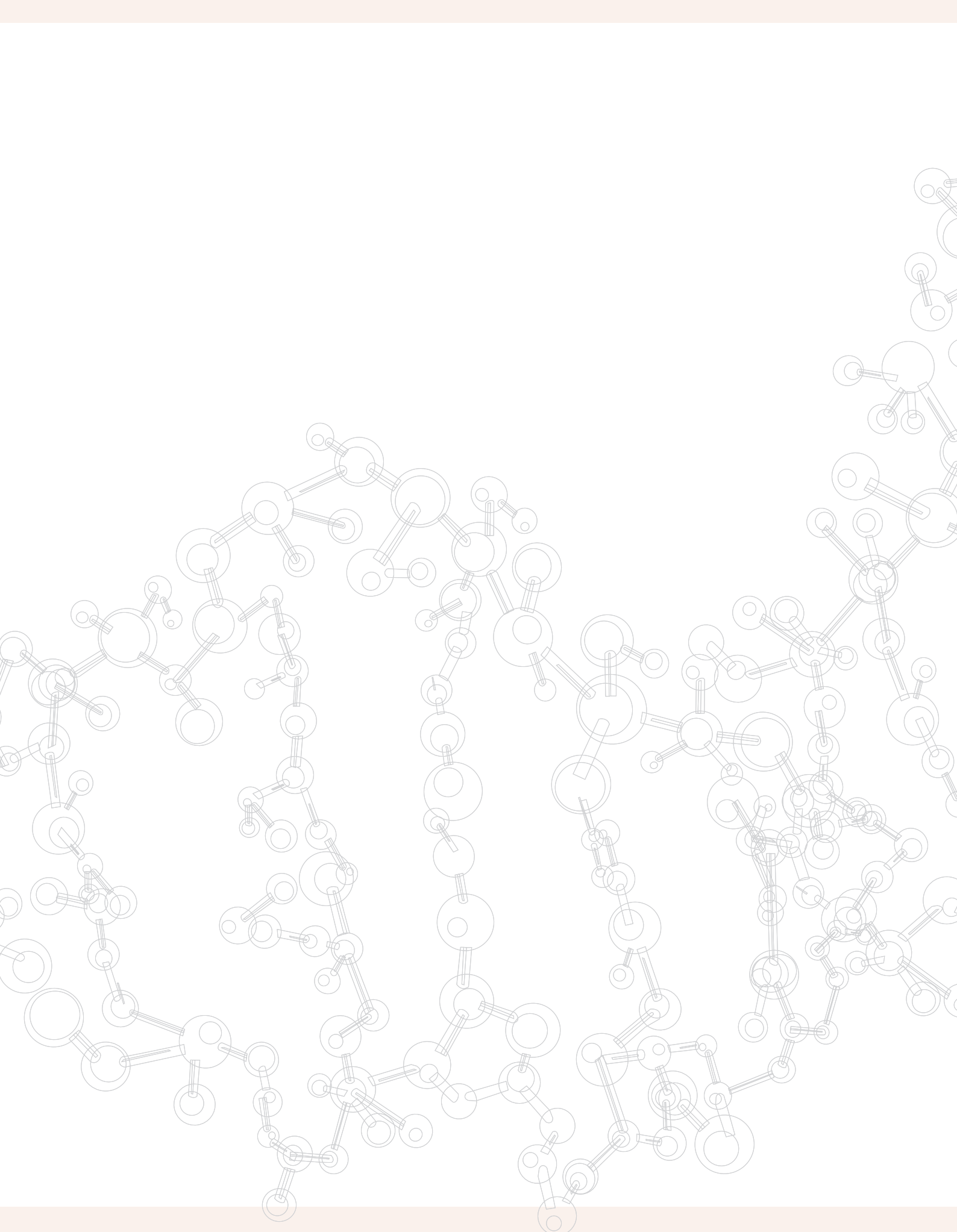
*Universidade Federal de Minas Gerais  
Educação a Distância  
2011*

*Biologia Molecular  
e Noções de Biotecnologia*

*Audrey Heloisa Ivanenko Salgado  
Carlos Renato Machado  
Mônica Bucciarelli Rodriguez*

The background of the page is a repeating pattern of a molecular structure. It consists of various-sized circles connected by thin lines, representing atoms and bonds in a complex, interconnected network. The pattern is light gray and covers the entire page, with a solid light orange horizontal band at the top and bottom.

# Biologia Molecular e Noções de Biotecnologia



The background of the cover is a repeating pattern of light gray molecular structures. These structures consist of circles of various sizes connected by lines, representing atoms and chemical bonds. The pattern is dense and covers the entire page, with a slightly higher concentration of structures in the upper half.

**Audrey Heloisa Ivanenko Salgado**

**Carlos Renato Machado**

**Mônica Bucciarelli Rodriguez**

# Biologia Molecular e Noções de Biotecnologia

Belo Horizonte  
CAED-UFMG  
2011



UNIVERSIDADE FEDERAL DE MINAS GERAIS

**Prof. Clélio Campolina Diniz**

Reitor

**Prof<sup>a</sup>. Rocksane de Carvalho Norton**

Vice-Reitora

**Prof<sup>a</sup>. Antônia Vitória Soares Aranha**

Pró-Reitora de Graduação

**Prof. André Luiz dos Santos Cabral**

Pró-Reitor Adjunto de Graduação

CENTRO DE APOIO À EDUCAÇÃO A DISTÂNCIA

**Prof. Fernando Selmar Rocha Fidalgo**

Diretor de Educação a Distância

**Prof. Wagner José Corradi Barbosa**

Coordenador da UAB/UFMG

**Prof. Hormindo Pereira de Souza Junior**

Coordenador Adjunto da UAB/UFMG

EDITOR CAED-UFMG

**Prof. Fernando Selmar Rocha Fidalgo**

CONSELHO EDITORIAL

**Prof<sup>a</sup>. Ângela Imaculada Loureiro de Freitas Dalben**

**Prof. Dan Avritzer**

**Prof<sup>a</sup>. Eliane Novato Silva**

**Prof. Hormindo Pereira de Souza**

**Prof<sup>a</sup>. Paulina Maria Maia Barbosa**

**Prof<sup>a</sup>. Simone de Fátima Barbosa Tófani**

**Prof<sup>a</sup>. Vilma Lúcia Macagnan Carvalho**

**Prof. Vito Modesto de Bellis**

**Prof. Wagner José Corradi Barbosa**

COLEÇÃO EAD – BIOLOGIA

Coordenadora: Gleydes Gambogi Parreira

LIVRO: Biologia molecular e noções de biotecnologia

Autores: Audrey Heloisa Ivanenko Salgado, Carlos Renato Machado, Mônica Bucciarelli Rodriguez

Revisão: Jussara Maria Frizzera

Projeto Gráfico: Laboratório de Arte e Tecnologia para Educação/EBA/UFMG

Este livro recebeu apoio financeiro do Pró-licenciatura (SEED-MEC) e da UAB/CAPES.

#### Dados Internacionais de Catalogação na Publicação (CIP)

(Biblioteca da Escola de Belas Artes da UFMG, MG, Brasil)

---

S164b Salgado, Audrey Heloisa Ivanenko  
Biologia molecular e noções de biotecnologia / Audrey Heloisa Ivanenko Salgado, Carlos Renato Machado, Mônica Bucciarelli Rodriguez. – Belo Horizonte: CAED-UFMG, 2011.  
111 p. : il. (algumas color.) ; 27 cm.

ISBN: 978-85-64724-02-0

1. Genética molecular. 2. Análise cromatográfica. 3. Eletroforese.  
4. Ensino a distancia. I. Machado, Carlos Renato. II. Rodriguez, Mônica Bucciarelli. III. Título. IV. Universidade Federal de Minas Gerais.

CDD: 574.88

CDU: 577.2

---

Ficha catalográfica elaborada por Luciana de Oliveira M. Cunha, CRB-6/2725  
Este livro recebeu apoio financeiro do Pró-licenciatura (SEED-MEC) e da UAB/CAPES.

# SUMÁRIO

Nota do Editor	07
Apresentação	08
Apresentação	09
<b>AULA 1: MÉTODOS EM BIOLOGIA MOLECULAR 1</b>	<b>12</b>
Autor: Carlos Renato Machado	
1.1 Os vetores de clonagem	13
1.1.1. Plasmídeos	14
1.1.2. Outros vetores	18
1.1.2.1. Fago lambda	18
1.1.2.2. Cosmídeo	19
1.1.2.3. BAC (cromossomo artificial de bactéria)	19
1.1.2.4. YAC (cromossomo artificial de levedura)	19
1.2. Bibliotecas	21
1.2.1. Bibliotecas genômicas	21
1.2.1.1. A história dos íntrons	21
1.2.2. A biblioteca de cDNAs	22
1.3. Clonagem de um gene	23
1.4. Sequenciamento	24
Autoavaliação	27
Boxe 1 – A técnica da PCR	28
<b>AULA 2: MÉTODOS EM BIOLOGIA MOLECULAR 2</b>	<b>30</b>
Autor: Carlos Renato Machado	
2.1. O projeto genoma	31
2.1.1. O Brasil e o projeto genoma	33
2.2. Clonagem	34
2.3. Terapia gênica	36
2.3.1. Células tronco e terapia gênica	40
Autoavaliação	41
<b>AULA 3: LESÕES E REPARO DE DNA</b>	<b>42</b>
Autor: Carlos Renato Machado	
3.1. Lesões no DNA	43
3.1.1. Lesões mutagênicas e letais	46
3.2. As vias de reparo de DNA	47
3.2.1. Reversão direta da lesão	48
3.2.2. Reparo por excisão	48
3.2.2.1. Reparo por excisão de base	49
3.2.2.2. Reparo por excisão de nucleotídeos	50
3.2.2.3. Reparo de erro de pareamento	51
3.2.3. Reparo de quebra dupla	52
3.3. Reparo de DNA e doenças	53
3.4. A quimioterapia e radioterapia	54
Autoavaliação	55
<b>AULA 4: POLIMORFISMOS MOLECULARES: VIVA A DIFERENÇA!</b>	<b>56</b>
autora: Mônica Bucciarelli Rodriguez	
4.1. Definições e classificação de polimorfismos genéticos	57
4.2. Detecção de polimorfismos moleculares	58

4.2.1. Técnicas básicas para o estudo de DNA	59
4.2.2. Técnicas para estudo de polimorfismos de DNA	63
4.3. Utilização da caracterização de polimorfismos genéticos	64
4.3.1. Uso em mapeamento genético e/ou diagnóstico de doenças genéticas	65
4.3.2. Uso de polimorfismos genéticos na identificação de indivíduos ou definição de paternidade	66
4.3.3. Outros usos de detecção de polimorfismos genéticos	67
Autoavaliação	68
<b>AULA 5: PROTEOMA: AS PROTEÍNAS DO MOMENTO</b>	<b>70</b>
Autora: Audrey Heloisa Ivanenko Salgado	
5.1. Conceito de proteoma e importância de seu estudo	71
5.2. Modificações pós-traducionais	72
Autoavaliação	75
<b>AULA 6: COMO ESTUDAR O PROTEOMA: CROMATOGRAFIA</b>	<b>76</b>
Autora: Audrey Heloisa Ivanenko Salgado	
6.1. O que é a cromatografia?	77
6.2. Tipos de cromatografia	78
6.3. Procedimentos experimentais	79
6.3.1. Cromatografia em papel	79
6.3.2. Cromatografia em camada	80
6.3.3. Cromatografia em coluna	80
6.3.4. Cromatografia de filtração em gel, exclusão molecular ou ainda peneira molecular	81
6.3.5. Cromatografia de troca iônica	81
6.3.6. Cromatografia por afinidade ou interação bio-específica	82
6.3.7. Cromatografia gás-líquido	83
6.4. HPLC	83
6.5. Outros conceitos e parâmetros cromatográficos	84
Autoavaliação	85
<b>AULA 7: COMO ESTUDAR O PROTEOMA: A ELETROFORESE</b>	<b>86</b>
Autora: Audrey Heloisa Ivanenko Salgado	
7.1. Introdução	87
7.2. O princípio da técnica de eletroforese	88
7.3. Variações na técnica de eletroforese: a eletroforese bidimensional	91
7.3.1. Focalização isoelétrica	91
7.3.2. Massa molecular	91
7.4. Limitações da eletroforese	94
Autoavaliação	95
<b>AULA 8: ESPECTROFOTOMETRIA E ESPECTROMETRIA DE MASSA E O ESTUDO DO PROTEOMA</b>	<b>96</b>
Autora: Audrey Heloisa Ivanenko Salgado	
8.1. O espectro eletromagnético	97
8.2. Espectros de absorção	98
8.3. Espectrofotometria e quantificação: A lei de Lambert-Beer	100
8.4. Métodos colorimétricos	103
8.5. O equipamento de espectroscopia no ultravioleta-visível	104
8.6. Espectrometria de massa	105
8.7. O estudo do proteoma	107
Autoavaliação	109

## NOTA DO EDITOR

A Universidade Federal de Minas Gerais atua em diversos projetos de Educação a Distância, que incluem atividades de ensino, pesquisa e extensão. Dentre elas, destacam-se as ações vinculadas ao Centro de Apoio à Educação a Distância (CAED), que iniciou suas atividades em 2003, credenciando a UFMG junto ao Ministério da Educação para a oferta de cursos a distância.

O CAED-UFMG (Centro de Apoio à Educação a Distância da Universidade Federal de Minas Gerais), Unidade Administrativa da Pró-Reitoria de Graduação, tem por objetivo administrar, coordenar e assessorar o desenvolvimento de cursos de graduação, de pós-graduação e de extensão na modalidade a distância, desenvolver estudos e pesquisas sobre educação a distância, promover a articulação da UFMG com os polos de apoio presencial, como também produzir e editar livros acadêmicos e/ou didáticos, impressos e digitais, bem como a produção de outros materiais pedagógicos sobre EAD.

Em 2007, diante do objetivo de formação inicial de professores em serviço, foi criado o Programa Pró-Licenciatura com a criação dos cursos de graduação a distância e, em 2008, com a necessidade de expansão da educação superior pública, foi criado pelo Ministério da Educação o Sistema Universidade Aberta do Brasil – UAB. A UFMG integrou-se a esses programas, visando apoiar a formação de professores em Minas Gerais, além de desenvolver um ensino superior de qualidade em municípios brasileiros desprovidos de instituições de ensino superior.

Atualmente, a UFMG oferece, através do Pró-licenciatura e da UAB, cinco cursos de graduação, quatro cursos de pós-graduação *lato sensu*, sete cursos de aperfeiçoamento e um de atualização.

Como um passo importante e decisivo, o CAED-UFMG decidiu, neste ano de 2011, criar a Editora CAED-UFMG como forma de potencializar a produção do material didático a ser disponibilizado para os cursos em funcionamento.

O primeiro desafio foi a publicação dos livros da coleção Educação a Distância, série Biologia. Agradecemos aos autores e à equipe de produção pela competência e dedicação que garantiram, com certeza, o nível de excelência desta obra apresentada à comunidade acadêmica.

Fernando Selmar Rocha Fidalgo  
Editor

## **CURSO DE CIÊNCIAS BIOLÓGICAS NA MODALIDADE “EDUCAÇÃO A DISTÂNCIA”?**

Formação de um Biólogo, e com Licenciatura? IMPOSSÍVEL! Sim, foi esta a minha reação inicial, Bióloga e também licenciada em Ciências. E, para aumentar minha surpresa, fui designada para coordenar a produção dos fascículos, que deveriam ser escritos em uma linguagem em que os alunos não prescindissem da presença física do professor em sala de aula, ou seja, DESAFIADOR! Sim, foi esta a incumbência que a UFMG recebeu do MEC e que nós, professores do Instituto de Ciências Biológicas, abraçamos, com nossas limitações sim, mas acima de tudo com a certeza de que envidaríamos todos os nossos esforços para que os fascículos fossem produzidos.

Na minha primeira reunião com os redatores, me fizeram a seguinte pergunta: “Como se escreve para alunos que estão à distância, sem os recursos acadêmicos?” Minha resposta “... Não sei!” Mas imediatamente me ocorreu que, na verdade, os escritores fazem isto o tempo todo e conseguem, com magnitude, atingir centenas de milhares de leitores.

E no caso do conteúdo da Biologia não seria diferente, pois o nosso “tema” está ao nosso redor, nas plantas, nos animais visíveis a “olho nu” e, se invisíveis, nos levando a outra dimensão do conhecimento. E, acima de tudo, “nós” somos um dos temas da Biologia, concordam? Ou será que a CÉLULA está presente somente no título do livro? Enquanto lemos, falamos, andamos e entendemos o que nos está sendo ensinado, são elas, as células, é que estão funcionando! E se sofremos a invasão de microorganismos, são eles que as invadem e que provocam sua reação e acionam seu metabolismo celular.

E a amplitude que propusemos atingir viaja também entre os conhecimentos macroscópicos desde o estudo do corpo humano como um todo até organismos que habitam o entorno do mesmo, como: parasitas, visíveis ou não a olho nu; insetos, benéficos ou causadores de doenças; plantas, desde aquelas que nos encantam o olhar e acariciam nossa alma até aquelas de que necessitamos para nos alimentar ou para produzir medicamentos que combatem doenças adquiridas por nosso corpo; enfim, todo contexto da nossa vida integrada com a natureza.

E, neste pensamento, descobrimos que escrever um fascículo para alunos que não estejam próximos de nós é muito mais interessante do que pensávamos, pois o nosso leitor é o nosso parceiro, o nosso co-autor!

Bem vindos ao Ensino a Distância!

Dra. Gleydes Gambogi Parreira  
Coordenadora da Produção de Fascículos de Biologia

## APRESENTAÇÃO

Caros alunos, bem-vindos novamente!

Após nossas experiências em Bases Moleculares I e II onde a interdisciplinaridade foi um dos enfoques das nossas aulas, nos encontramos de novo, dessa vez para discutir técnicas de Biologia Molecular, um tema em aberto. Procuramos neste fascículo apresentar os fundamentos da Biologia Molecular e perceber sua importância no cotidiano atual e futuro. A genômica estrutural e a proteômica serão abordados neste fascículo, pois além do genoma é preciso saber quais são e como interagem as proteínas que os genes codificam. Além disso, veremos como as células reagem a lesões em seu material genético e como a eventual ocorrência de mutações pode gerar polimorfismos, que por sua vez são bastante úteis em estudos genéticos, em diagnóstico e em Biologia Forense.

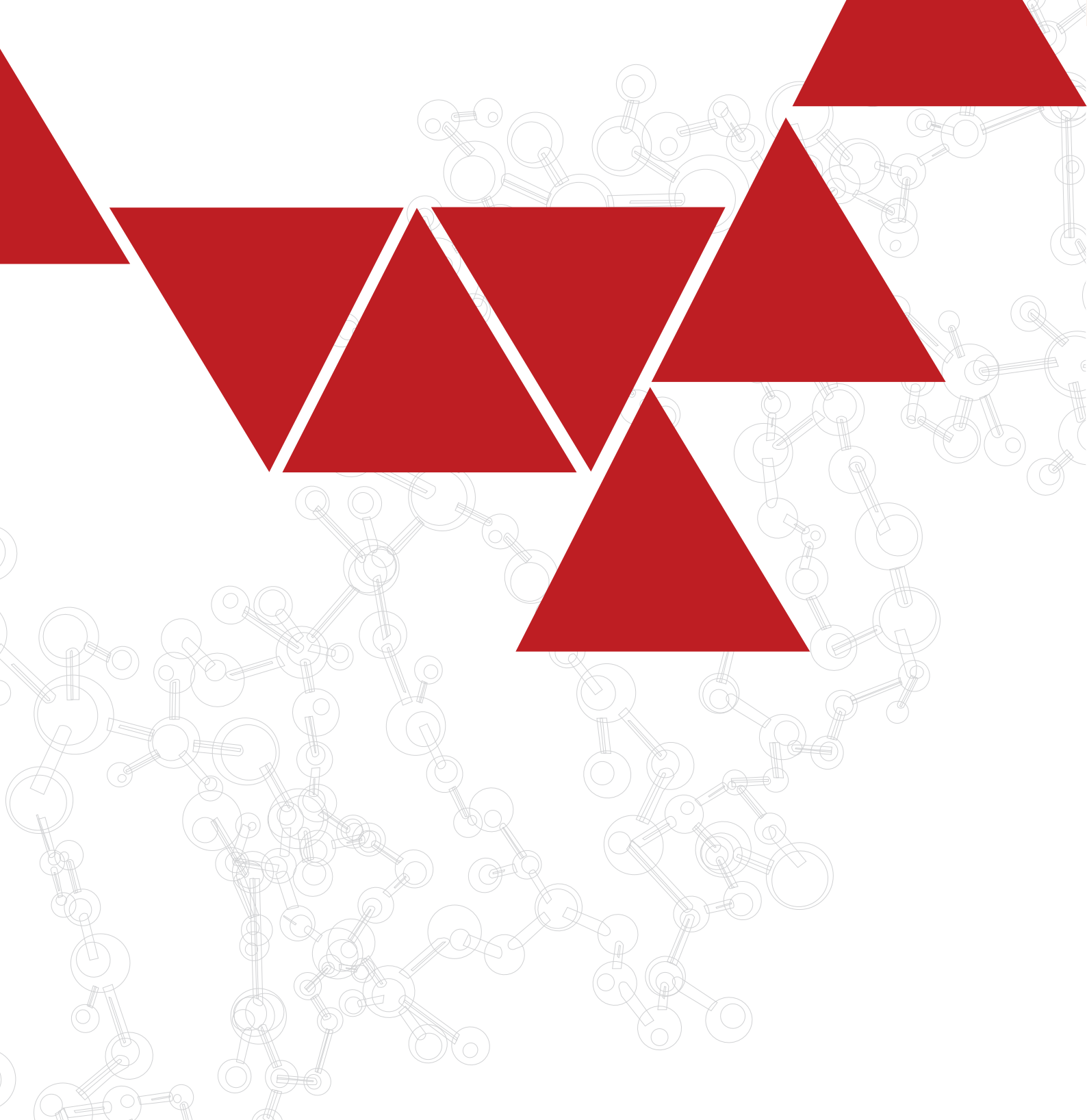
Audrey H. Ivanenko Salgado

Carlos Renato Machado

Mônica Bucciarelli Rodriguez







**1**

*Métodos em  
Biologia Molecular 1*

# AULA 1: MÉTODOS EM BIOLOGIA MOLECULAR 1

Ao final desta aula você deverá saber:

- O que são vetores de clonagem.
- O que é uma biblioteca.
- Como se faz o sequenciamento do DNA.

Nas duas próximas aulas iremos discutir aspectos importantes de biologia molecular e biotecnologia. No centro desses processos, temos o processo de clonagem que permitiu desenvolver a tecnologia do DNA recombinante.

A palavra clonagem é originada da palavra grega *Klon* que significa broto de vegetal. O significado atual da palavra se refere a qualquer processo de reprodução/multiplicação que gere um gene, uma célula ou um organismo, fruto de um mesmo DNA. Nessa aula, vamos discutir ferramentas importantes na tecnologia do DNA recombinante como os vetores, as bibliotecas, as técnicas de sequenciamento e da PCR. Durante essas discussões, vamos mostrar aplicações biotecnológicas dessa metodologia.

## 1.1. OS VETORES DE CLONAGEM

Na aula 7 de Bases Moleculares II foi discutida a história do DNA e um dos tópicos foi sobre a primeira clonagem realizada. Nesse experimento, os pesquisadores foram capazes de clonar um gene de resistência a um antibiótico em um plasmídeo. O plasmídeo foi utilizado como vetor; mas o que é isso? Se olharmos no dicionário a definição de vetor nós iremos encontrar:

*Molécula de DNA com porções isoladas do DNA de um organismo vivo e que se introduz em células vivas para que expressem o caráter codificado pelo DNA inserido*

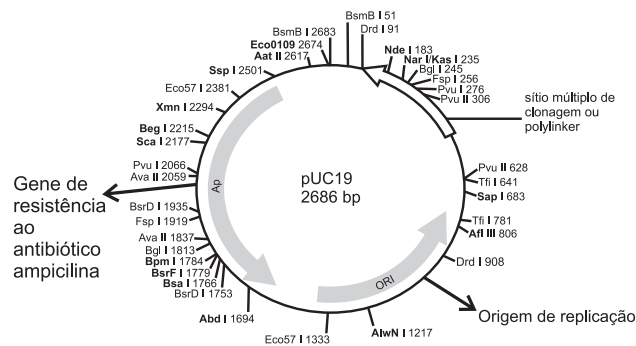
Os vetores que iremos descrever, portanto, devem ser capazes de ter algumas características importantes:

- a. Devem ser uma molécula de DNA;
- b. Devem ser capazes de entrar em uma célula e se multiplicar nessa célula;
- c. Devem ser capazes de transportar material (DNA) exógeno.

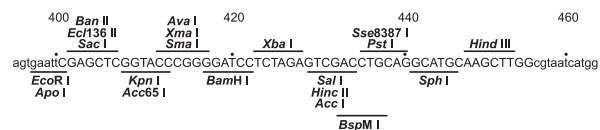
### 1.1.1. PLASMÍDEOS

O plasmídeo utilizado tinha todas essas características, pois como vimos na aula 7 de Bases Moleculares II, era possível fazer a transformação, ou seja, a inserção de um plasmídeo em uma célula bacteriana. O plasmídeo possui uma origem de replicação que permite que o mesmo seja duplicado pela maquinaria bacteriana. E, por fim, é possível inserir um pedaço de DNA exógeno no plasmídeo e este DNA recombinante será replicado dentro da bactéria; dessa maneira, teremos a presença de diversas cópias iguais desse DNA o que é denominado de um clone.

Figura 1: Mapa de um plasmídeo (pUC19) contendo o sítio múltiplo de clonagem SMC, gene de resistência ao antibiótico ampicilina, uma origem de replicação. O SMC está localizado no gene LacZ que permite selecionar a inserção do DNA exógeno pela formação de colônias brancas ou azuis.



#### Região do polylinker ou sítio múltiplo de clonagem



contém várias das características descritas abaixo:

#### a) Sítios múltiplos de clonagem:

Um ponto importante no processo de clonagem é que o DNA exógeno é inserido no plasmídeo, utilizando-se enzima de restrição e enzima ligase (ver aula 7 de Bases Moleculares II). Para facilitar a clonagem, os plasmídeos atuais possuem uma sequência de DNA que apresenta várias regiões que são reconhecidas por enzimas de restrição diferentes. Dessa maneira, fica mais fácil inserir o DNA de interesse, pois existem mais alvos que podem ser utilizados para a inserção.

#### b) Marcadores para inserção de DNA exógeno

O primeiro experimento de clonagem de DNA exógeno realizado utilizou um gene que conferia resistência a um antibiótico. Dessa maneira, foi mais fácil selecionar as bactérias (que foram transformadas com o plasmídeo contendo o DNA exógeno), pois estas se tornaram resistentes a um antibiótico e, portanto,

Desde o primeiro plasmídeo que foi utilizado na clonagem, vários avanços foram obtidos. Hoje temos vários tipos de plasmídeos com características diferentes. A seguir descreveremos as características dos plasmídeos atuais e alguns tipos de plasmídeo. Na figura 1 é apresentado o mapa do plasmídeo pUC19 que

ao colocá-las em um meio contendo o antibiótico em questão só as bactérias transformadas seriam capazes de crescer.

Mas, por exemplo, se fosse clonado o gene que codifica para o hormônio insulina, como a bactéria seria selecionada? Vários plasmídeos que são utilizados atualmente possuem um método interessante de seleção que permite separar os plasmídeos que tiveram DNA inserido daqueles que não tiveram o DNA exógeno inserido. Na região de clonagem do gene existe a expressão de um gene que permite a metabolização de lactose. Ao se utilizar este composto denominado X-gal, nesse processo é liberado um composto de cor azul capaz de obter colônias de bactéria de cor azulada. Mas como ocorre a seleção de plasmídeos que contém DNA exógeno? Como dito acima, o gene que produz a enzima que metaboliza a lactose está presente na região de clonagem e se o DNA exógeno é clonado com sucesso, o mesmo irá interromper o gene de metabolismo da lactose. Dessa maneira, a bactéria será incapaz de metabolizar a lactose e quando o X-gal é adicionado ao meio, o mesmo não será utilizado e não produzirá a cor azul, o que deixará a bactéria com a cor branca. Esse sistema é muito útil para selecionar as bactérias que foram transformadas com o plasmídeo e que contém DNA exógeno.

Esses plasmídeos de seleção só podem ser usados em bactérias modificadas que perderam a capacidade de metabolizar a lactose, as quais são complementadas com o gene que está presente no plasmídeo.

### **c) Plasmídeos de expressão protéica**

Muitas vezes a clonagem de um gene é realizada para que seja possível produzir uma proteína de interesse em grande quantidade. Um exemplo clássico foi a clonagem do gene que codificou o hormônio insulina. A produção de insulina é importante para o tratamento da diabetes.

### **A história da produção da insulina, um exemplo da biotecnologia:**

Já no século XIX os médicos sabiam da importância do hormônio insulina para o consumo de glicose como energia e passou-se a tentar um jeito de ministrar esse hormônio nas pessoas com diabetes tipo 1. As primeiras tentativas foram realizadas com o uso do pâncreas na dieta dos pacientes, o que não surgiu o efeito desejado. Em 1922, o Dr. Frederick Banting foi capaz de extrair insulina do pâncreas de cachorro e utilizou essa insulina em um menino de 14 anos que estava devastado pela diabetes. A primeira injeção de insulina foi capaz de diminuir levemente o nível de glicose no menino, mas não aliviou os sintomas da doença. Na busca de um medicamento mais eficiente, o Dr. Banting aperfeiçoou o mecanismo de purificação de insulina. Com a insulina mais purificada, ele voltou a reinjetar o hormônio no paciente e obteve um ótimo resultado, elevando os níveis de glicose no sangue de aproximadamente 520 mg/dl para aproximadamente 120 mg/dl. A partir desses dados, a purificação de insulina de cavalos e porcos permitiu o tratamento de diversos pacientes portadores de diabetes.

Um novo marco no tratamento da diabetes foi a produção da insulina humana recombinante. Em 1978, uma empresa iniciante de biotecnologia (Genentech) foi capaz de clonar o gene da insulina e produzi-la em bactérias. O hormônio insulina é produzido como pro-insulina que possui os peptídeos A, B e C que são sintetizados por único gene. Após clivagem protéica, ocorre a liberação do peptídeo C e a insulina ativa é formada pela união do peptídeo A e B por pontes dissulfeto. A estratégia dos pesquisadores da Genentech foi clonar o gene da insulina em partes: em um plasmídeo foi clonado a parte do gene da insulina responsável por codificar o peptídeo A, e em outro plasmídeo foi clonado parte do gene da insulina responsável por codificar o peptídeo B. A clonagem desses genes foi realizada em fusão com gene da lactase, ou seja, a bactéria que continha o plasmídeo com o gene da insulina produzia uma proteína quimérica com a parte N-terminal contendo a proteína lactase e a parte C-terminal com o peptídeo da insulina. O plasmídeo utilizado tinha um promotor que permitia a expressão do gene quimérico.

De posse dos dois grupos de bactérias, os pesquisadores as colocaram para crescer em meio que favorecia a expressão do gene quimérico e este era produzido em grande quantidade pelas bactérias. Após a expressão, a proteína quimérica foi purificada em uma coluna de afinidade, utilizando a capacidade da mesma de interagir com a lactose. Após a purificação, com o uso de uma protease, foi feita a clivagem da proteína quimérica, o que permitiu a liberação dos peptídeos A e B. Por fim, em um tampão especial, foram colocados os peptídeos A e B ligadas formando a insulina nativa. Em 1982, a insulina humana recombinante foi autorizada a ser comercializada, sendo o primeiro medicamento produzido pela tecnologia do DNA recombinante.

Como vemos pela história da insulina, um plasmídeo de expressão tem de apresentar um promotor e, muitas vezes, a proteína de interesse é produzida em fusão com uma proteína conhecida.

O promotor é o que permite a transcrição do gene, pois permite que a RNA polimerase reconheça onde deve começar o processo de transcrição. Um promotor constitutivo está sempre ativo, ou seja, permite sempre o acesso da enzima RNA polimerase ao DNA. Por outro lado, um promotor indutivo só estará ativo quando a ele se ligar uma molécula, indicando que é o momento de permitir o acesso da RNA polimerase ao DNA. O uso de promotores indutivos muitas vezes é mais adequado pois permite controlar o momento que a proteína será expressa e, normalmente, a quantidade de expressão obtida é maior que a obtida com um promotor constitutivo.

A expressão com uma proteína de fusão pode facilitar em dois pontos: na expressão de proteínas solúveis e na purificação. Muitas vezes, a expressão de proteínas de outros organismos em bactérias não é muito benéfica para essas células, pois está deslocando a maquinaria de transcrição e tradução para outros fins e produz em excesso uma proteína que não é natural da bactéria. Outras vezes, essa proteína em excesso fica insolúvel dentro da célula e precipita formando corpos de oclusão. A proteína precipita seja por quê está em excesso, seja por não conseguir atingir a conformação ideal quando expressa em bactéria. Muitas vezes a proteína de fusão terá o papel importante de solubilizar a proteína de

interesse e facilitar que ela atinja a sua conformação nativa.

Outro papel importante da proteína de fusão é de facilitar a purificação da proteína de interesse. Como exemplo, podemos citar o uso da proteína ligadora de maltose (MBP – do inglês *maltose binding protein*). Nesses plasmídeos, o gene de interesse é clonado a 3' do gene da MBP e é muito importante que a clonagem do gene de interesse seja realizada em fase com o gene da MBP, ou seja, que os códons que codificam o gene de interesse estejam na mesma fase leitura dos códons que codificam a MBP. Dessa maneira, a proteína de interesse será produzida junto com a proteína MBP. Como o próprio nome diz, a MBP tem a capacidade de se ligar a maltose e, assim, podemos fazer uma coluna de afinidade que apresenta a maltose como ligante; dessa maneira a MBP irá se ligar à coluna e permitirá a purificação da mesma. Como a proteína de interesse está fusionada com a MBP, ela também será purificada nesse processo. Após a purificação, a proteína fusionada poderá ser clivada (como vimos no caso da insulina) utilizando uma protease.

#### **d) Plasmídeos ponte**

Vimos até agora os plasmídeos sendo usados em células bacterianas, mas o mesmo pode ser utilizado em outros tipos celulares, tais como leveduras e células humanas. Todo plasmídeo deve ter uma origem de replicação e ter um marcador que permita que os mesmos sejam selecionados.

Em leveduras, diferente dos plasmídeos bacterianos, usa-se como marcador de seleção genes de auxotrofia de metabolismo. Por exemplo, uma levedura pode ser deficiente em crescer em um meio mínimo sem a presença do nucleotídeo uridina. Essa deficiência ocorre devido a mutações em algum gene da via metabólica deste nucleotídeo. Um plasmídeo que contenha o gene mutado irá permitir que a levedura possa crescer na ausência da uridina, pois a levedura se torna apta a produzi-la a partir de precursores. Diversos genes envolvidos no metabolismo de aminoácidos e nucleotídeos são utilizados em levedura, o que permite que mais de um plasmídeo possa ser inserido em levedura, desde que os mesmos possuam marcas auxotróficas diferentes, ou seja, sejam capazes de complementar a deficiência de mais de uma via metabólica.

Os plasmídeos de leveduras podem utilizar como origem de replicação as sequências ARS que são as origens de replicação normalmente encontradas nos cromossomos das leveduras. Outra origem que pode ser utilizada nos plasmídeos de levedura são denominados de origem 2 micra que são origens encontradas em DNA extracromossômico presentes em levedura.

Os plasmídeos também podem ser utilizados em células humanas e nesse caso eles também precisam de uma origem de replicação e uma marca que permita selecionar as células nas quais os plasmídeos foram inseridos. Como a origem de replicação de células de mamíferos ainda não está estabelecida, os plasmídeos utilizados nessas células são provenientes de vírus, como o SV40 e o EBV. A origem viral permite que o plasmídeo replique normalmente nas células de mamíferos, enquanto que a marca de seleção desses plasmídeos são antibióticos como é utilizado em bactérias.

O uso de plasmídeos em humanos pode ser muito útil para o entendimento da função dos genes e mesmo verificar qual gene pode estar relacionado com doenças genéticas. Por exemplo, como veremos na aula 3, existem pessoas que apresentam defeito na via de reparo de excisão de nucleotídeos e são sensíveis a luz do sol. Uma maneira de se saber qual gene apresenta o defeito que causa a doença é usar um plasmídeo contendo genes candidatos que serão inseridos nas células obtidas de pacientes cultivadas em laboratório. Se a presença desse gene candidato for capaz de reverter a sensibilidade apresentada é uma evidência clara que é uma deficiência nesse gene que causa a doença.

Mas o que são plasmídeos ponte? Para produzir o plasmídeo em grande quantidade a melhor célula para se trabalhar é a bacteriana, pois é possível obtê-las em grande quantidade e, conseqüentemente, os plasmídeos também são obtidos em grande quantidade. Então, mesmo que o plasmídeo seja para ser usado em leveduras ou células humanas, ele é trabalhado inicialmente em bactérias e para que isso ocorra, precisa ter origem na replicação bacteriana e um marcador de seleção para células bacterianas. Portanto, estes plasmídeos apresentam duas origens de replicação, uma bacteriana e outra de levedura ou de células de mamíferos, o que irá permitir que eles se multipliquem em qualquer uma das células. A denominação ponte vem do fato da bactéria ser utilizada como uma ponte para a produção de grande quantidade do plasmídeo que será utilizada em outras células.

### 1.1.2. OUTROS VETORES

Os plasmídeos são muito úteis como descrito acima, mas eles apresentam um grande problema que é o tamanho do DNA exógeno que pode ser inserido nesse vetor. Normalmente, o máximo de 10.000 pares de base é o que pode ser inserido em um plasmídeo e, muitas vezes, os genes que se pretende trabalhar são ainda maiores. Nessa parte da aula iremos discutir alguns outros vetores que são capazes de transportar uma maior quantidade de DNA.

#### 1.1.2.1. FAGO LAMBDA

Outro vetor que foi muito utilizado foi o fago lambda, um vírus que infecta bactérias. O genoma do vírus apresenta 48.500 pares de bases, mas nem todos os genes presentes são essenciais para a multiplicação do vírus. Essa região não essencial pode ser removida e em seu lugar ser inserido o DNA exógeno de interesse.

Após a montagem do DNA recombinante, ou seja, o fago contendo o DNA exógeno, o mesmo é empacotado *in vitro*. O DNA é colocado na presença das proteínas que formam o capsídeo (invólucro protéico que guarda o DNA do vírus) e espera-se que o DNA seja inserido nessa estrutura. O empacotamento *in vitro* não é muito eficiente, pois aproximadamente 10% de fagos são obtidos. Para que o empacotamento funcione é necessário que o fago possua uma sequência de DNA em sua extremidade (denominada COS) que será reconhecida pelas proteínas do capsídeo.

O fago lambda como vetor tem a vantagem de o processo de infecção ser mais eficiente que a transfecção do plasmídeo, uma vez que o fago possui um aparato natural para inserir o DNA dentro da bactéria e poder multiplicá-lo. Apesar dessa facilidade, o tamanho do DNA que pode ser inserido não é muito maior do que se pode inserir em um plasmídeo.

#### **1.1.2.2. COSMÍDEO**

O cosmídeo é um vetor híbrido entre os plasmídeos e os fagos. De forma similar aos plasmídeos ele terá uma origem de replicação e um marcador para seleção como um gene que confere resistência ao antibiótico. Do fago, o cosmídeo mantém basicamente a região de sequência COS que permitirá que o DNA seja empacotado em um fago. Como praticamente todo o DNA do fago foi retirado, a quantidade de DNA exógeno inserido pode ser maior que nos fagos, chegando a aproximadamente 40.000 pares de bases. O limite do tamanho do DNA a ser inserido está relacionado com o quanto que o fago lambda é capaz de empacotar em seu capsídeo. Após o empacotamento *in vitro*, o cosmídeo é utilizado para infectar bactérias. O DNA é inserido como o DNA de um fago e, quando dentro da bactéria, o cosmídeo irá se comportar como um plasmídeo e será replicado a partir da origem de replicação do plasmídeo e a seleção será realizada pela marca do antibiótico presente.

#### **1.1.2.3. BAC (CROMOSSOMO ARTIFICIAL DE BACTÉRIA)**

As bactérias são capazes de realizar conjugação, que é uma forma de sexo entre elas, no qual uma bactéria doadora irá transferir DNA para uma bactéria receptora. A transferência do DNA é realizada pelo fator F, uma espécie de plasmídeo capaz de transferir quase um terço do material genômico da bactéria. Em 1992, foi desenvolvido um novo vetor que usa o fator F como se fosse um plasmídeo. A origem do fator F foi mantida e os BAC têm grande capacidade de transportar DNA exógeno chegando a 300.000 pares de bases. Para manter a estabilidade desse vetor é necessário que o mesmo seja propagado em células de bactérias que não sejam capazes de realizar recombinação. Pelo fato da grande quantidade de DNA transportado e pela facilidade de manuseio, os BACs têm sido a principal escolha para a produção de bibliotecas genômicas para sequenciamento.

#### **1.1.2.4. YAC (CROMOSSOMO ARTIFICIAL DE LEVEDURA)**

A maioria dos vetores que tratamos até agora é para a utilização em bactérias, a não ser alguns plasmídeos que podem ser utilizados em leveduras e células de mamíferos.

Os estudos com levedura permitiram a caracterização de várias sequências de DNA importantes para a replicação e manutenção de um cromossomo nesse organismo. Todo cromossomo deve ter um centrômero, telômeros e origens de replicação.

Os centrômeros têm papel fundamental em várias etapas do processo de

replicação e mitose é através dele que os cromossomos são partilhados igualmente entre as células em divisão.

As origens de replicação, ao contrário das bactérias onde encontramos apenas uma, estão presentes em várias cópias em um cromossomo eucarioto. Na maioria dos organismos não se conhece exatamente qual a sequência de DNA que está presente nas origens de replicação, mas em levedura existe uma sequência consenso, denominada ARS, que é origem de replicação em levedura.

Por fim, a última parte do cromossomo que é importante para a sua manutenção são os telômeros. Diferente das bactérias que apresentam um DNA circular, o DNA das células eucarióticas é linear e isso gera problemas sérios, como veremos a seguir.

Cromossomos lineares podem se juntar se as suas extremidades estiverem livres e isso levaria uma grande instabilidade cromossômica o que é inviável para uma célula. A estrutura dos telômeros que estão nas extremidades dos cromossomos impede que ocorra a junção cromossômica. O telômero é uma sequência de DNA repetitivo que forma um grampo na extremidade do cromossomo que, além de impedir as junções, também torna o DNA resistente a atividade das enzimas éxonucleases que degradam o DNA.

Mas um dos papéis mais importantes do telômero está relacionado com o problema do fim da replicação do DNA. Como vocês já viram na aula sobre replicação do DNA, as enzimas DNA polimerases só são capazes de iniciar a replicação quando um iniciador (pequena sequência de RNA/DNA) fornece uma extremidade OH 3'. Esse processo faz com que seja impossível que o DNA seja totalmente replicado, pois não existe um mecanismo que localize os iniciadores corretamente nas extremidades e, mesmo que isso ocorra, como os iniciadores têm um pedaço de RNA nos mesmos, esta parte tem de ser degradada uma vez que o RNA não pode estar presente na molécula de DNA. Dessa maneira, após cada ciclo de replicação de um DNA linear ocorrerá perdas em sua extremidade com o encurtamento da molécula. Para lidar com esse problema, as células eucarióticas contam com os telômeros que diminuirá a cada replicação. Como o telômero é constituído por uma sequência repetitiva que não apresenta função gênica, a diminuição do telômero não irá levar prejuízo à célula, pois nenhuma informação será perdida. Entretanto, quanto menor ficar o telômero, mais instável ficará a célula, pois mais perto ela ficará de perder funções gênicas, que são importantes para a sua manutenção (veja na próxima aula como os telômeros são feitos e o papel que os mesmos têm no envelhecimento).

Como são conhecidas as sequências de DNA que compõem os centrômeros, origem de replicação e telômeros em leveduras, os pesquisadores foram capazes de construir um cromossomo artificial nesse organismo. Esse cromossomo poderá conter uma grande quantidade de DNA exógeno podendo chegar até 2 milhões de pares de bases.

## 1.2. BIBLIOTECAS

Assim como as bibliotecas tradicionais são formadas por um conjunto de livros, as bibliotecas gênicas são um conjunto de clones que tentam representar o conjunto de um genoma. As bibliotecas podem ser construídas utilizando qualquer vetor e a escolha do mesmo dependerá do que se pretende. Discutiremos a seguir os dois principais tipos de biblioteca que existem: genômica e de cDNA.

### 1.2.1. BIBLIOTECAS GENÔMICAS

Uma biblioteca genômica busca conter todo o genoma de um organismo. Para este fim, é necessário que se extraia o DNA do organismo em estudo. O DNA deverá ser clivado, seja por enzimas de restrição, seja por clivagem mecânica. O DNA digerido é utilizado para ligação com o vetor de interesse. Após a ligação com o vetor, o mesmo é utilizado para transformar a célula hospedeira, seja ela bactéria ou levedura. Várias células transformadas são obtidas e cada uma deve conter um vetor com um fragmento diferente do genoma. As bibliotecas genômicas são muito úteis para o sequenciamento de um genoma e, para este fim, o vetor BAC é um dos mais utilizados atualmente, pois é capaz de transportar um grande fragmento de DNA exógeno em cada clone.

Apesar de ser muito útil, muitas vezes a biblioteca genômica não permite que se estude a expressão de genes. O problema ocorre devido à presença dos íntrons que estão presentes no genoma, mas estão ausentes no RNA mensageiro.

#### 1.2.1.1. A HISTÓRIA DOS ÍNTRONS

Desde que o dogma central da biologia foi estabelecido (DNA – RNA – Proteína) e desde os trabalhos com a transcrição em bactéria, se sabia que o DNA era transcrito em um RNA, que recebeu o nome de mensageiro e era traduzido em uma proteína. A sequência do RNA mensageiro de um gene era complementar ao gene. Dessa maneira, era possível pegar um fragmento genômico de uma bactéria, clonar em um vetor e quando este fosse inserido em uma bactéria, seríamos capazes de obter ao final a produção da proteína de interesse.

Esse processo não se mostrou verdadeiro para vários genes eucariotos e a razão disso foi explicada em 1977. Pesquisadores franceses obtiveram o RNA mensageiro de um gene humano e o hibridaram com o genoma. O que eles viram no microscópio eletrônico foi um dado muito interessante: o RNA era capaz de se ligar em um fragmento do genoma, perdia a capacidade de se ligar a um fragmento adjacente do genoma e voltava a se ligar mais adiante. Esse dado mostrava que o gene de eucarioto era interrompido e que o RNA mensageiro era formado por fragmentos não contínuos do genoma celular. Eles nomearam as regiões do genoma que estavam presentes no RNA mensageiro de éxon e as regiões que estavam ausentes foram denominadas de íntrons. Mais tarde, verificou-se que o DNA era transcrito inteiro com íntrons e éxons, mas que eles só saíam do núcleo quando processados, o que significa a perda dos íntrons e a adição de uma cauda poliA (uma sequência rica em adeninas na região 3' do RNA). O RNA mensageiro maduro, então, é traduzido no citoplasma.

Isso gerou um problema para o estudo dos genes em bactérias, uma vez que esses organismos não possuem íntrons no seu genoma, e não apresentam as enzimas que permitirão o processamento do RNA para sua tradução. Portanto, quando um fragmento do genoma de eucarioto é colocado em uma bactéria, é possível que se faça um RNA mensageiro dele mas estarão presentes tanto os éxons quanto os íntrons; e quando esse RNA for traduzido não levará a produção de uma proteína viável, uma vez que nos íntrons não encontramos os códons corretos para a tradução da proteína nativa. Inclusive, é muito comum encontrar nos íntrons códons de terminação que interrompem a tradução do RNA mensageiro. Como será possível, então, expressar o gene de interesse em bactérias para estudar o produto desse gene?

### 1.2.2. A BIBLIOTECA DE CDNAS

Na década de setenta, já era do conhecimento dos pesquisadores a existência da enzima transcriptase reversa. Essa enzima está presente em alguns vírus que possuem o genoma RNA e uma das etapas do ciclo de vida desses vírus é a formação de um genoma DNA que servirá de molde para a transcrição dos RNAs que traduziram nas proteínas virais. O genoma DNA também formará o genoma RNA que estará presente na progene viral.

A transcriptase reversa é capaz de reconhecer um molde de RNA e a partir dele é capaz de fazer uma sequência complementar de DNA, de onde vem a denominação cDNA (DNA complementar). Como a transcriptase reversa é capaz de transformar o RNA viral em DNA, será que ela também pode formar DNA a partir de RNA mensageiro em DNA? A abordagem adotada pelos pesquisadores foi bem eficiente. Primeiro, pensou-se em um iniciador que fosse capaz de hibridar com o RNA mensageiro e permitisse que o DNA fosse formado a partir do RNA. Como todo RNA possui uma cauda poliA em sua extremidade, o iniciador utilizado foi um formado por timinas que permitiria hibridar com todos os RNA mensageiros. A partir do iniciador, a transcriptase reversa é capaz de sintetizar o DNA complementar e formar uma estrutura híbrida de RNA e DNA. Os passos seguintes são a degradação do RNA e a síntese de outra fita de DNA e assim o RNA mensageiro é transformado em um cDNA que pode ser clonado em um vetor.

No cDNA encontramos todos os éxons unidos com a informação correta para a formação de um RNA mensageiro que será traduzido em uma proteína ativa. Mesmo em bactérias, o cDNA é capaz de formar proteínas nativas pois ele não apresenta íntrons e irá se portar como DNA presente em bactérias.

Um biblioteca de cDNA é muito importante para estudos funcionais ou mesmo quando se quer seqüenciar as partes codificadoras de um genoma.

Uma abordagem que se pode utilizar com a biblioteca de cDNA é a complementação funcional heteróloga. Um exemplo clássico dessa abordagem é a clonagem de genes envolvidos no reparo de DNA. Como veremos na aula 3, as vias de reparo de DNA estão presentes em todos os organismos e muitas vezes apresentam atividade muito similar. Como o trabalho com bactéria é mais fácil que com células humanas, muitas vias de reparo de DNA foram descritas

e estudadas inicialmente neste organismo. De posse de mutantes bacterianos que eram deficientes em uma via de reparo, os pesquisadores tinham em mãos uma boa ferramenta de trabalho. Se existisse homologia funcional entre o gene de bactéria e o humano, seria só usar uma biblioteca de genes humanos na bactéria e ver qual complementava a função perdida com a mutação. Um exemplo clássico dessa estratégia foi a clonagem do gene da 3-metil-Adenina glicosilase. Essa enzima está envolvida com o reparo por excisão de base e é responsável por reconhecer a lesão 3-metil-adenina do DNA e retirá-la. Células deficientes nessa enzima são sensíveis ao tratamento com o agente alquilante metil-metano sulfonato (MMS). A estratégia utilizada para a clonagem do gene humano foi a transformação da bactéria mutante com uma biblioteca humana de cDNA. Após a transformação, as bactérias contendo os genes humanos foram submetidas ao tratamento com MMS e só aquele que tinha o gene humano capaz de complementar a função de 3-metil-adenina glicosilase é que sobreviveu e, assim, o gene humano foi clonado. Esse experimento só foi possível pelo uso de uma biblioteca de cDNA que permite que genes humanos sejam expressos em bactérias e codifiquem proteínas funcionais.

Outro papel importante das bibliotecas de cDNA são que elas permitem seqüenciar os genes que são transcritos em um organismo ou uma célula específica. Por exemplo, se fizéssemos uma biblioteca genômica de diferentes células humanas e seqüenciássemos essas bibliotecas, chegaríamos ao mesmo resultado, uma vez que o genoma é igual em todas as células. Mas se fosse feita uma biblioteca de cDNA dessas mesmas células, o resultado seria diferente já que os genes expressos são diferentes nos diferentes tipos celulares. Portanto, uma biblioteca de cDNA irá mostrar os genes que estão expressos em determinado momento e determinada célula.

### **1.3. CLONAGEM DE UM GENE**

A clonagem de um gene pode acontecer utilizando várias estratégias e abordagens. Pode se buscar um gene em uma biblioteca, como mostrado na complementação funcional em bactérias, ou utilizar outras estratégias como discutiremos a seguir.

Como clonar um gene que codifica um antígeno de parasito? De posse de uma biblioteca de cDNA desse parasito podemos tentar clonar o gene, desde que exista uma boa maneira de pescá-lo. Os diversos plasmídeos ou fagos dessa biblioteca são utilizados para inserção nas células hospedeiras, dessa maneira, teremos diversos clones de bactérias e em cada uma delas um gene diferente estará presente. Portanto, cada bactéria estará expressando um gene diferente do parasito e restará desenvolver uma estratégia para descobrir qual das bactérias contém o gene de interesse. O primeiro passo é colocar uma membrana especial sobre a placa, e, assim, as bactérias são transferidas para a membrana. A membrana é, então, tratada para que a bactéria seja rompida e com isso as proteínas que estavam na célula são expostas. Como verificar qual dessas células possui a proteína que é um antígeno do parasito? Uma boa estratégia é pegar o soro de um paciente infectado com o parasito e ver qual clone de bactéria é reconhecido pelo soro. A bactéria reconhecida na membrana

poderá ser coletada na placa e poderá ser crescida o que permitirá obter grande quantidade do DNA de interesse.

Outra estratégia que pode ser adotada é utilizar uma sonda de DNA como isca que irá hibridar com o clone alvo. Por exemplo, pode-se procurar clonar um gene em um organismo X, mas como obter o DNA desse gene de interesse? Supondo que esse gene já é descrito em diferentes organismos, podemos tentar amplificá-lo no genoma de interesse utilizando iniciadores degenerados e utilizar a técnica da PCR (reação de polimerização em cadeia - ver a história da PCR no boxe 1) para amplificar o DNA que será utilizado como isca. Para construir iniciadores degenerados é importante encontrar uma região conservada na sequência da proteína conhecida nos diferentes organismos; se existe essa região conservada é de supor que a mesma esteja conservada no organismo X. Como sabemos que cada aminoácido possui de um a seis códons (sequência de 3 nucleotídeos que irá determinar o aminoácido), o iniciador degenerado será feito de acordo com a sequência de aminoácidos conservado. Os iniciadores degenerados serão utilizados para amplificar do genoma do organismo X, uma sequência que será utilizada como isca para hibridar com a biblioteca e selecionar o gene de interesse.

O sequenciamento de diferentes genomas tem facilitado o processo de clonagem já que, ao conhecer a sequência de DNA do organismo, não é mais necessário utilizar iniciadores degenerados e já se pode fazer iniciadores específicos que irão permitir amplificar a sequência desejada. Essa sequência será utilizada como isca para buscar o gene em uma biblioteca, ou mesmo, poder-se amplificar o gene completo caso o mesmo não possua íntrons e tenha um tamanho possível de ser amplificado.

Esses foram apenas alguns exemplos de estratégias de clonagem de um gene, mas diversas abordagens diferentes podem ser adotadas de acordo com o gene, sua função e célula que está sendo expresso.

## 1.4. SEQUENCIAMENTO

Uma técnica de biologia molecular muito importante e que teve grande impacto no desenvolvimento biotecnológico foi o sequenciamento de DNA.

Em 1974, o Cientista Frederick Sanger desenvolveu uma metodologia que permitia seqüenciar o DNA que se baseia nos mecanismos que a própria célula usa para duplicar o DNA. Como vocês já viram em Bases Moleculares II, para que a DNA polimerase replique o DNA é necessária a presença dos seguintes componentes: um molde de DNA, um iniciador, desoxiribonucleotídeos, a DNA polimerase e um tampão apropriado que contenha magnésio.

O processo de replicação utilizará o molde para realizar o pareamento de Watson-Crick, ou seja, adenina com timina e guanina com citosina. Por exemplo, se no molde estiver presente uma desoxiguanina, uma citosina desoxiribonucleotídeo trifosfato irá parear, o que permitirá que a DNA polimerase acrescente essa base na fita filha. Essa ligação se dará da seguinte forma: o fosfato da base desoxicitidina será ligado à hidroxila do carbono 3 da ribose presente no nucleotídeo que está na extremidade do iniciador. De onde vem a energia que

permite a replicação do DNA? Como escrito acima, o desoxiribonucleotídeo citosina que está sendo inserido está na forma trifosfato, mas no DNA ele estará na forma monofosfato, o que leva a liberação de um pirofosfato (dois fosfatos ligados) que será imediatamente clivado pela pirofosfatase inorgânica e liberará a energia necessária para que o processo ocorra.

A citosina inserida fará parte do iniciador e fornecerá a hidroxila do carbono 3 da ribose para que o processo possa prosseguir. De posse de todas essas informações, Sanger bolou uma estratégia que permitiu sequenciar o DNA, uma vez que o processo de duplicação do DNA era bloqueado pela inserção de um didesoxiribonucleotídeo (nucleotídeo que não apresenta uma hidroxila no carbono 3 da ribose) que não fornecia uma extremidade 3'OH e permitisse a continuação do processo de replicação do DNA.

O método funciona da seguinte maneira:

Primeiro é preciso extrair e purificar o DNA que se pretende seqüenciar. O DNA purificado é, então, aquecido a 100°C, o que fará que as ligações de hidrogênio (que mantêm a dupla-hélice unida) sejam desfeitas e teremos então um DNA simples fita que servirá de molde. No processo de resfriamento do DNA é acrescentada uma pequena sequência de DNA que servirá como iniciador nesse processo. Com a diminuição da temperatura se permitirá que as pontes de hidrogênio sejam refeitas formando a dupla hélice do DNA. Como existe uma grande quantidade da sequência, que servirá de iniciador, a chance maior é que a dupla que se fará nesse processo seja formada pela fita molde linearizada com a sequência do iniciador. Agora, de posse do molde de DNA e do iniciador, deve-se colocá-los em quatro tubos distintos. Em cada tubo será acrescentado um tampão contendo magnésio, a DNA polimerase e os quatro desoxiribonucleotídeos trifosfatos (adenina, citosina, guanina e timina), sendo que um deles é marcado com material radioativo. Além desses componentes, será em um tubo colocado didesoxiribonucleotídeo adenina; no segundo tubo será acrescentado didesoxiribonucleotídeo citosina; no terceiro tubo será colocado didesoxiribonucleotídeo guanina; e, por fim, no quarto tubo é acrescentado didesoxiribonucleotídeo timina. Os quatro tubos são colocados a 37°C, temperatura ótima de atividade da DNA polimerase que iniciará o processo de replicação do DNA.

Vamos ver o que acontece no tubo 1 que contém didesoxiribonucleotídeos adenina. Nesse tubo, existem várias moléculas de DNA molde ligados com o iniciador que permitirá a atividade da DNA polimerase onde irá começar a sua duplicação, inserindo os desoxiribonucleotídeos que fazem pareamento correto no molde de DNA. Mas nesse tubo temos a presença da adenina didesoxiribonucleotídeo e se ela for inserida no processo de sequenciamento, a duplicação será interrompida e essa fita de DNA não irá mais crescer. Esse processo de inserção da adenina didesoxiribonucleotídeo é aleatório e poderá ocorrer em diferentes momentos da replicação do DNA. Dessa maneira, teremos a presença de sequências de DNA de diferentes tamanhos e que sempre conterà em seu final uma desoxiribonucleotídeo adenina. Como as DNA polimerases duplicam o molde de DNA utilizando o pareamento de Watson-Crick, a adenina só será inserida em frente a uma timina. Outro ponto importante nesse processo

é que o uso de nucleotídeos radioativos permitirá que se verifique o tamanho do segmento de DNA que foi duplicado, como veremos a seguir.

O processo que ocorreu no tubo 1, com a adenina didesoxiribonucleotídeo, também irá ocorrer no tubo 2, só que agora, como temos didesoxiribonucleotídeo citosina, os fragmentos de DNA obtidos nesse tubo terminarão com uma citosina em sua extremidade. O mesmo acontecerá no tubo 3 com a guanina didesoxiribonucleotídeo e no tubo 4 com a timina didesoxiribonucleotídeo. O que precisamos agora é identificar os fragmentos de DNA duplicados e ver o tamanho dos mesmos em cada um dos quatro tubos. Essa visualização será realizada em um gel de eletroforese desnaturante que permitirá a visualização apenas dos fragmentos de DNA que contém material radioativo incorporado, ou seja, da fita recém duplicada.

No gel será inoculado o produto dos quatro tubos em canaletas diferentes. O gel de sequenciamento desnaturante tem duas capacidades especiais: por ser desnaturante ele faz com que a dupla hélice seja desfeita e o DNA corre no gel como DNA simples fita; além disso, o gel é capaz de separar os fragmentos de DNA a cada nucleotídeo mais presente na fita. Após a corrida do DNA no gel, este é exposto em um filme de raio-X e o material radioativo irá marcar o filme e será possível visualizar os fragmentos de DNA e ler a sequência desta molécula. Como isso é feito? Vamos supor que o DNA que pretendemos estudar tem a seguinte sequência:

**AATCCGTAGCTATCAGTACCAT**

No tubo 1, que contém didesoxiribonucleotídeo adenina, a polimerização irá parar sempre que na fita molde conter uma timina (T) e assim, teremos fragmento de DNA com 3, 6, 11, 13, 17 e 22 nucleotídeos. No tubo 2, que contém citosina didesoxiribonucleotídeos, os fragmentos de DNA apresentam um tamanho de 7, 9, 16. Calcule o tamanho dos fragmentos formados no tubo 3 (com guanina didesoxiribonucleotídeos) e no tubo 4 (com timina didesoxiribonucleotídeos). Essa escadinha de fragmentos com um nucleotídeo a mais que é verificado no gel nas 4 canaletas é que permite seqüenciar o DNA. A figura 2 esquematiza o processo de sequenciamento apresentado acima.

O desenvolvimento das técnicas descritas nessa aula foi baseado no desenvolvimento da biologia molecular e um grande passo para o sequenciamento de genomas completos, assim como a clonagem e terapia gênica, como veremos na próxima aula.

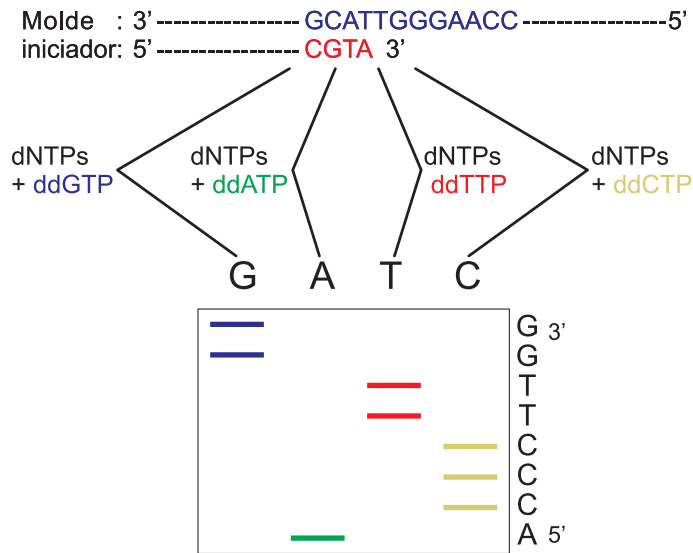


Figura 2: Esquema do método de sequenciamento de Sanger.

### Complemento da aula:

- No moodle vocês encontrarão diferentes animações para maior compreensão da aula.

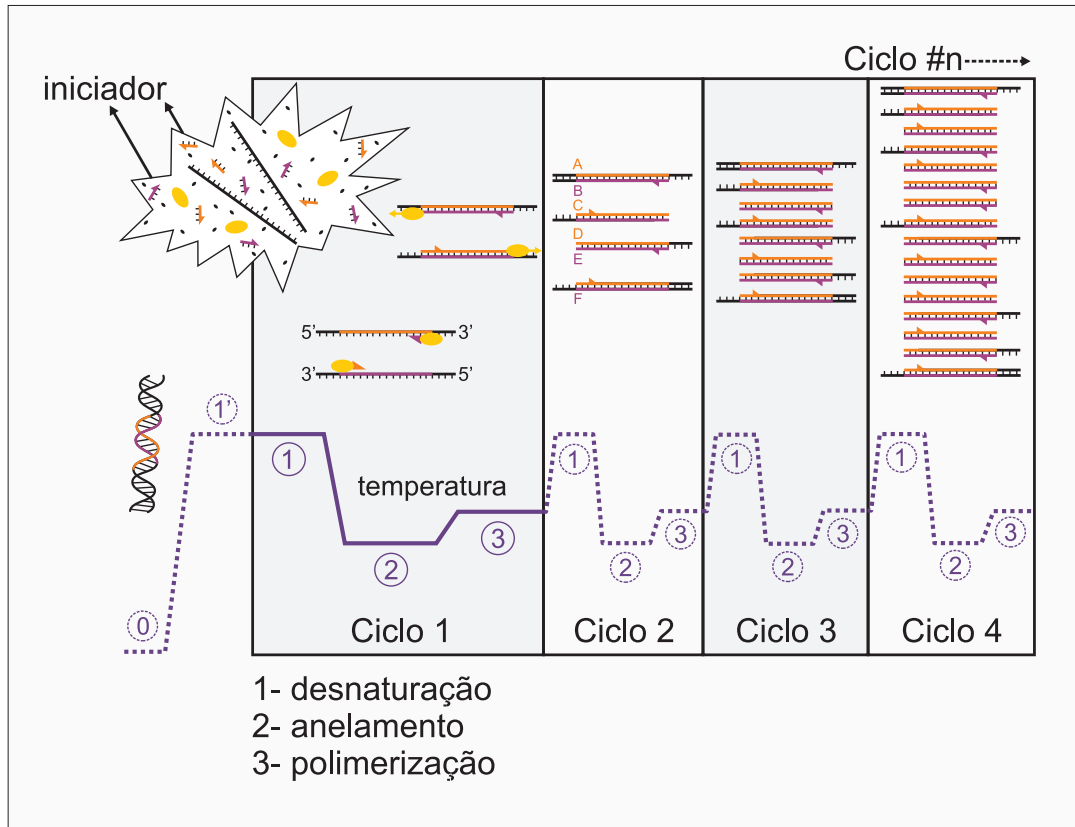
### Autoavaliação

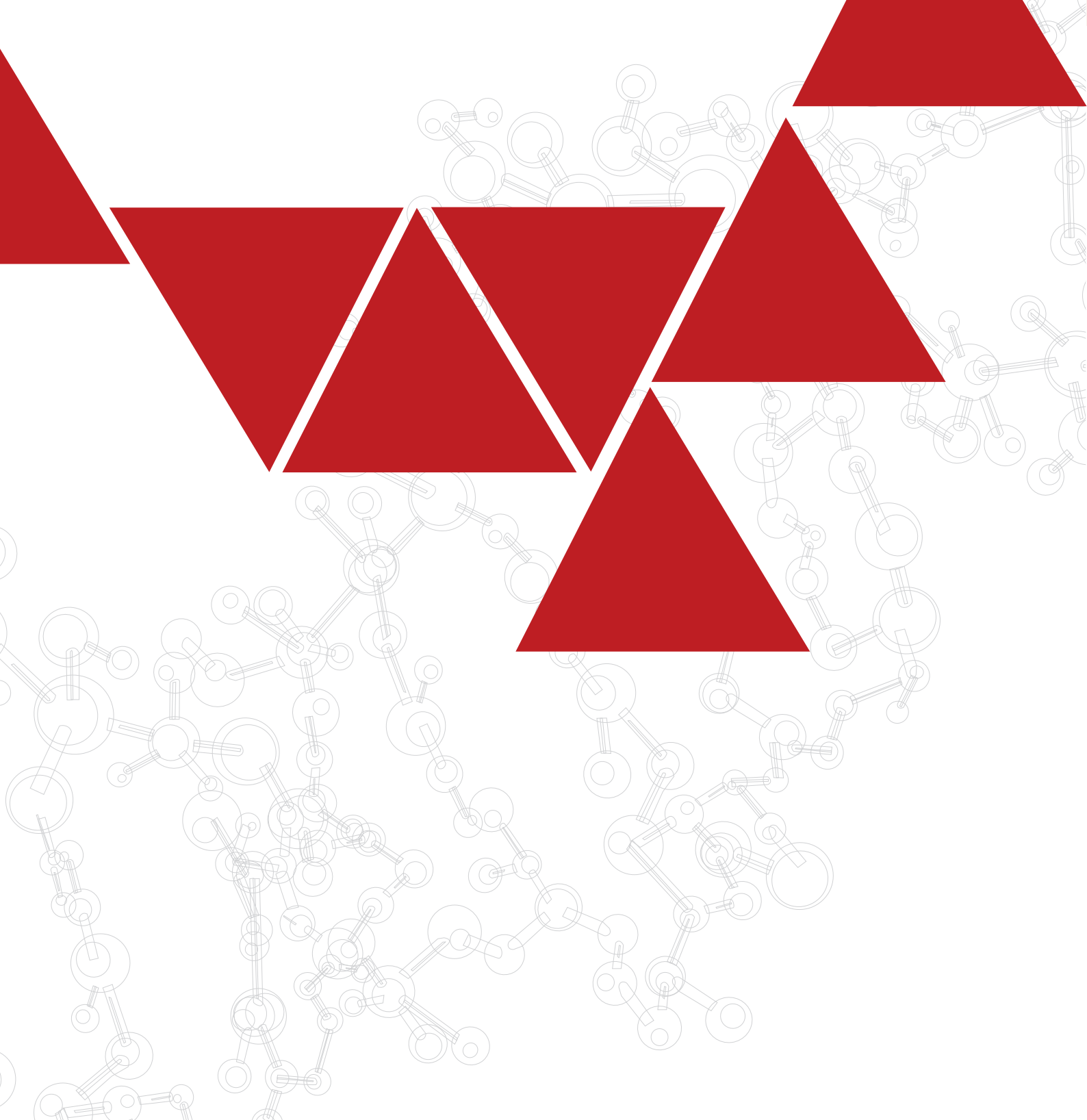
- Quais as características que um vetor deve possuir?
- Quais são os tipos de plasmídeos?
- Quanto de DNA exógeno cabe em cada vetor?
- Qual a diferença de um fago para um cosmídeo?
- Como se faz os BACs e os YACs?
- O que é uma biblioteca?
- O que é um íntron?
- Como se faz um cDNA?
- Descreva como se clona um gene.
- Quais são as funções de um telômero?
- Quais os passos da técnica de sequenciamento?

## BOXE 1 – A TÉCNICA DA PCR

A técnica da PCR (do inglês polymerase chain reaction – reação em cadeia de polimerase) foi desenvolvida pelo pesquisador Kary B. Mullis em 1983. Esta técnica que revolucionou a biologia molecular não nasceu de um trabalho intenso no laboratório, uma vez que foi concebida em uma viagem de moto que Mullis realizava entre São Francisco e Mendocino. Segundo o pesquisador, as suas melhores idéias lhe ocorrem quando está dirigindo. Mas como funciona a PCR? Como o próprio nome diz, ocorre uma reação em cadeia de polimerização que irá permitir amplificar a quantidade de um alvo de DNA. O genoma humano apresenta 3,200,000,000 pares de base; se pretendemos estudar um gene de 10000 pares de base isso irá representar apenas 0,00003125% do genoma. Esse valor mostra como seria difícil obter essa sequência alvo mesmo que se obtenha uma grande quantidade de DNA. A técnica da PCR procura enriquecer a quantidade da sequência alvo amplificando-a. Inicialmente a técnica usava a DNA polimerase Klenow de *E. coli* e funcionava da seguinte maneira:

Em um tubo é acrescentado o DNA de interesse, iniciadores e tampão. O tubo é colocado a 95°C para que o DNA seja desnaturado e forme DNA simples fita. Durante o resfriamento o DNA irá parear com os iniciadores como mostrado na figura abaixo. Quando a temperatura chega a 37°C, é acrescentado a DNA polimerase Klenow que irá polimerizar o DNA a partir dos iniciadores. Os passos descritos acima são repetidos, ou seja, desnaturação, pareamento e acréscimo da DNA polimerase para nova etapa de duplicação. Quando esses ciclos são realizados várias vezes o que se obtém é a amplificação exponencial do DNA contido entre os dois iniciadores. A técnica da PCR teve um grande avanço quando se substituiu a DNA polimerase Klenow pela DNA polimerase da bactéria *Thermus aquaticus*. Essa bactéria é capaz de crescer em altas temperaturas e, por isso, as enzimas dessa bactéria são capazes de atuar em altas temperaturas. Assim, a cada ciclo de desnaturação, renaturação e polimerização não é necessário acrescentar novamente a enzima.





**2**

*Métodos em  
Biologia Molecular 2*

## AULA 2: MÉTODOS EM BIOLOGIA MOLECULAR 2

Ao final desta aula você deverá entender:

- o que é o projeto genoma
- o que é clonagem de um organismo
- o que são as células tronco e a terapia gênica

Na aula anterior verificamos algumas das metodologias que permitiram o desenvolvimento da tecnologia do DNA recombinante. Nessa aula, iremos abordar algumas técnicas importantes que terão grande impacto em nossas vidas. Algumas dessas técnicas ainda estão em estudo e não são amplamente empregadas mas tem grande potencial, tais como a terapia gênica e células tronco.

### 2.1. O PROJETO GENOMA

Como vimos na aula anterior, a metodologia de sequenciamento de DNA já era conhecida desde a década de setenta e permitiu o sequenciamento de vários genes.

Na década de oitenta foi proposto nos Estados Unidos o primeiro projeto que pretendia sequenciar todo o genoma humano. Financiado pelos órgãos de fomento do governo americano, ele pretendia na verdade mapear os genes humanos e verificar possíveis mutações que causam doenças.

Logo depois, se associaram aos Estados Unidos, o Japão, a Austrália e países da Europa. Todos esses países formaram o HUGO (do inglês *Human Genome Organization*) que pretendia até o ano de 2005 terminar a tarefa hercúlea de sequenciar todo o genoma. O presidente do HUGO, H. Van Ommen, afirmou em 1998 que a missão do órgão era facilitar e coordenar a iniciativa global de mapear, sequenciar e analisar funcionalmente o genoma humano e promover a aplicação destes conhecimentos ao melhoramento da saúde humana.

Na década de oitenta, o método de Sanger era realizado manualmente e possibilitava um máximo de 500 bases lidas por experimento. Esses valores praticamente inviabilizavam o sequenciamento de um genoma de 3 bilhões de pares de bases e, por isso, pretendia-se apenas mapear os genes e ver a localização dos mesmos.

A iniciativa do projeto genoma foi fundamental para o avanço técnico que permitiu o desenvolvimento da metodologia do sequenciamento. Um dos marcos importantes nesse avanço foi o desenvolvimento, em 1986, do sequenciador automático de DNA que utilizava marcadores fluorescentes em vez dos marcadores radioativos utilizados anteriormente. Essa máquina permitiu automatizar todo o processo, o que acelerou o sequenciamento de

DNA e permitiu que todo o genoma humano fosse decifrado antes do prazo determinado.

Associado ao projeto humano, iniciou-se o sequenciamento de vários organismos. O primeiro que teve o seu genoma sequenciado foi a bactéria *Haemophilus influenzae* que causa a meningite. Em 1995, o seu genoma de aproximadamente 1 milhão e oitocentos mil pares de base foi completamente decifrado. Também foi verificado que esse genoma codifica para 1813 genes e que, destes, 1737 codificam para proteínas.

Já no ano seguinte, o primeiro genoma de um organismo eucarioto foi finalizado. A *Saccharomyces cerevisiae*, uma levedura utilizada como fermento de pão e cerveja, apresenta um genoma de 12 milhões de pares de bases distribuídos em 16 cromossomos. Esse genoma codifica para aproximadamente seis mil genes.

Em 1997, outro organismo modelo teve todo o seu genoma seqüenciado. A bactéria *Escherichia coli*, que é muito utilizada em todos os experimentos de biologia molecular, apresentando um genoma de aproximadamente 4 milhões de pares de bases e seiscentos mil genes, o qual codifica para cerca de 4 mil genes.

O projeto genoma estava sendo desenvolvido basicamente pelo poder público, mas em 1998 surge uma empresa privada (CELERA) que pretendia sequenciar todo o genoma humano e usaria, para isso, todas as informações obtidas antes pelo projeto HUGO. Um dos dados mais importantes obtidos por este projeto era o mapeamento do genoma, no qual seqüências de DNA eram utilizadas para marcar posições específicas em cada cromossomo e serviam de base para montagem das seqüências obtidas. A estratégia adotada pela CELERA foi de cortar todo o genoma humano e sequenciar todos os fragmentos e, depois montá-los adequadamente para se ter toda a seqüência do genoma humano.

A presença de uma empresa privada na empreitada do sequenciamento do genoma humano levantou a questão da disponibilidade destes dados. Em um congresso dos participantes do projeto genoma nas Bermudas, no ano anterior, os princípios da rápida liberação pública dos dados da seqüência do genoma, sem limitações no uso, foram discutidos e, mais tarde, chamados de “os princípios das Bermudas”.

Em 1998, o primeiro organismo multicelular teve o seu genoma completamente sequenciado. O *Caenorhabditis elegans*, um verme, apresenta várias características similar aos vertebrados, uma vez que possui sistema nervoso, reprodutivo e digestivo. O *C. elegans* também começa com uma única célula fertilizada e segue um processo bem determinado de desenvolvimento até a formação do verme adulto. O genoma de *C. elegans* apresenta aproximadamente 97 milhões de pares de bases em seis cromossomos, que são mais similares aos cromossomos humanos que os de levedura. Esse genoma apresenta aproximadamente 20.000 genes.

O primeiro cromossomo humano a ser completamente sequenciado foi o 22, pois este é relativamente pequeno. A estratégia utilizada foi a de sequenciar clone por clone de uma biblioteca de BAC e obter grande quantidade de informação exata. O braço longo do cromossomo 22 apresenta 33.400.000 pares base e

contém, ao menos, 545 genes. O próximo cromossomo sequenciado foi o 21, cuja trissomia causa a síndrome de Down, que apresentou 33.500.000 pares base e com menos de 300 genes.

Em 2003, dois anos antes do prazo previsto, a sequência do DNA humano foi determinada. Os cerca de 3 bilhões de bases espalhados pelos 23 cromossomos são capazes de codificar 30.000 genes. O número de genes presentes foi uma surpresa para os cientistas que, no início do projeto genoma, esperavam encontrar aproximadamente 60.000 genes.

Os vários genomas apresentados nessa aula seguiram uma tendência de que quão mais complexo o organismo, maior o tamanho do genoma. Entretanto, isso não é sempre verdade, uma vez que genomas de alguns protozoários são maiores que o genoma humano. Essa discrepância é conhecida como paradoxo C, no qual o tamanho do genoma (designado de C) não se correlaciona com a complexidade do organismo.

Mas o que está presente no genoma desses protistas que os tornam maiores? Normalmente, esses genomas contêm muitas sequências de repetidas que não codificam genes. O número de genes e proteínas de um organismo se associa bem com a sua complexidade, sendo quanto maior a complexidade, maior o número de genes e proteínas.

### **2.1.1. O BRASIL E O PROJETO GENOMA**

Os primeiros passos do Brasil no projeto genoma foram realizados na Universidade Federal de Minas Gerais. No início da década de 90, iniciou-se um projeto que buscava identificar os genes expressos do *Schistosoma mansoni*, o causador da esquistossomose. Para chegar aos seus objetivos, os pesquisadores do Laboratório de Genética-Bioquímica, produziram bibliotecas de cDNA das diversas fase de vida do *S. mansoni* e sequenciaram grande parte desses clones. O maior problema das bibliotecas de cDNA é que elas são uma representação dos mRNA presentes em uma célula. Assim, se um gene é muito expresso, haverá uma grande quantidade de mRNA desse gene e, conseqüentemente, haverá um grande número de clones dessa biblioteca que representou esse gene. Por outro lado, genes que são pouco expressos estavam presentes em menor quantidade, o que dificultou o sequenciamento dos mesmos.

Um passo importante no projeto genoma no Brasil foi quando a FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo), em 1997, organizou a rede ONSA (do inglês, Organização para Sequenciamento e Análise de Nucleotídeos), que é um Instituto virtual de genômica e contava inicialmente com 30 laboratórios. O primeiro projeto da rede ONSA foi o sequenciamento da bactéria *Xylella fastidiosa* que causa a praga do amarelinho em laranjas. O projeto foi encerrado em 1999 e, a partir dele, vários outros foram iniciados, como o sequenciamento de genes da cana de açúcar, de células cancerosas e de outras bactérias.

Seguindo o exemplo da FAPESP, o CNPq (o Conselho Nacional de Desenvolvimento de Pesquisa e Tecnologia) também iniciou um projeto com a participação de diversos grupos no Brasil. O primeiro organismo selecionado

para o sequenciamento foi a bactéria *Chromobacterium violaceum* que apresenta grande potencial biotecnológico como a produção de diversos antibióticos.

Todos os projetos genoma tiveram grande impacto no desenvolvimento da pesquisa no País e na capacitação de recursos humanos para trabalhar com biotecnologia.

## 2.2. CLONAGEM

A clonagem de um organismo representa a duplicação de um indivíduo contendo o mesmo DNA, organismos, portanto, que seriam idênticos. Esse processo ocorre normalmente na natureza como é verificado com as bactérias que se dividem após duplicar o DNA.

O mesmo pode ocorrer com protozoários e plantas que, muitas vezes, são propagadas por replicação não sexuada que leva a geração de plantas com o mesmo DNA. Podemos considerar também que os gêmeos univitelinos são clones idênticos, uma vez que são produto de uma divisão de um óvulo fecundado por um espermatozóide. Essa divisão irá produzir dois embriões compostos pelo mesmo DNA, que gerará dois indivíduos idênticos.

As pesquisas sobre clonagem se iniciaram com o zoólogo alemão Hans Dreisch que, em 1894, pegou um embrião com duas células de ouriço do mar e agitou isso em um tubo com água do mar. O embrião foi dividido em duas células, as quais geraram dois ouriços do mar que possuíam o mesmo material genético. O segundo processo de clonagem foi realizado oito anos depois, quando Hans Spemman usou um fio de cabelo de seu filho para dividir um embrião de salamandra, o que produziu dois organismos idênticos.

Os experimentos com clonagem só tiveram outro fato importante em 1952, quando Robert Briggs e Thomas King realizaram uma clonagem de sapo. Esses pesquisadores transferiram o núcleo de células embrionárias para ovócitos anucleados, o que foi capaz de gerar sapos saudáveis.

Esse processo descrito na década de 50 do século passado seria a base de todos os processos de clonagem realizados. Como vimos, a célula doadora foi proveniente de uma célula embrionária. A clonagem usando células de adulto é mais difícil e complicada, uma vez que nas células diferenciadas o padrão de expressão dos genes é específico para o tipo de célula e tecido. Assim, células da pele têm um padrão único de expressão, o que permite que elas se multipliquem e formem outras células da pele iguais às anteriores; mas impede de formar células do fígado ou do coração, muito menos um novo indivíduo.

O grande marco do processo de clonagem foi a famosa Ovelha Dolly que nasceu no ano de 1996 e foi anunciada ao mundo no ano seguinte. O processo para a produção da Dolly consistiu em tirar o núcleo de uma célula somática, no caso da mama (a origem do nome é uma homenagem a atriz Dolly Parton que possuía seios grandes) de uma ovelha, transferindo-o para o óvulo anucleado de outra ovelha. O óvulo fertilizado foi transferido para o útero de uma terceira ovelha que serviu como barriga de aluguel. O processo utilizado para a geração da ovelha Dolly foi extremamente ineficiente, uma vez que quase 300 óvulos

foram produzidos e apenas um gerou um organismo. A maioria dos óvulos não conseguiu se fixar no útero ou morreu precocemente no processo de gestação.

Após o nascimento da ovelha Dolly, a grande pergunta que se fez foi se ela nascia com a idade da ovelha que forneceu o seu DNA ou se ela teria a mesma idade de uma ovelha gerada pelo método tradicional. Essa questão se coloca devido à questão dos telômeros.

Como vimos na aula anterior, os telômeros têm função importante nos cromossomos humanos, pois eles são necessários para impedir a degradação, a junção e o encurtamento do DNA devido ao problema do fim da replicação. Sempre que ocorre a replicação de um DNA linear ocorre perda de DNA nas extremidades e a presença do telômero nessa região previne que a perda ocorra em locais que contenham genes. O telômero é sintetizado por uma enzima denominada telomerase. Essa enzima é uma transcriptase reversa, ou seja, é capaz de sintetizar DNA a partir de um molde de RNA. Essa enzima está presente em todos os organismos que possuem DNA linear. A telomerase possui uma subunidade com função catalítica e outra que é composta por uma molécula de RNA que serve de molde para a atividade de transcriptase reversa. A telomerase quando está ativa é capaz de sintetizar DNA a partir do RNA molde que faz parte da enzima; assim uma fita do telômero é aumentada e a outra é sintetizada pelas DNA polimerases replicativas. Portanto, quando a telomerase está ativa, o telômero não diminui e não há perigo de perda de DNA informativo após a replicação.

Mas então qual poderia ser o problema da Dolly? Nos mamíferos, a telomerase só está ativa na fase embrionária e em alguns tipos celulares específicos. Assim, ao nascer, um indivíduo já tem o tamanho do seu telômero definido. Quando as células do indivíduo começam a se replicar, o encurtamento do telômero começa a acontecer. As células irão duplicar indefinidamente? Não. A cada duplicação o telômero é encurtado. Chega a um determinado momento que, para evitar que a célula comece a perder material genético importante, as células recebem uma sinalização indicando que as mesmas devem parar de se duplicar. Quando isso ocorre a célula entra no estado de senescência no qual ela está ativa metabolicamente, mas é incapaz de se duplicar.

Será que a senescência tem algo a ver com o envelhecimento? Várias evidências mostram que sim. Por exemplo, células de indivíduos jovens quando replicadas no tubo de ensaio demoram mais a entrar em senescência do que de indivíduos idosos. Outra evidência foi que células de pessoas que possuem uma doença que causa envelhecimento precoce entram em senescência mais rapidamente. Foi verificado também que tecidos de pessoas idosas apresentam mais células em senescência do que em indivíduos jovens.

Mas o encurtamento do telômero está mesmo envolvido com o processo de senescência? A resposta a essa pergunta veio quando se realizou um experimento de transformação de células senescentes. Os pesquisadores utilizaram células que já estavam em senescência e transformaram-nas com um plasmídeo que expressava o gene da telomerase. As células senescentes foram capazes de sintetizar telômeros e, assim, voltaram a se multiplicar. Esses dados mostram

que o encurtamento do telômero é um fator importante para a senescência e o envelhecimento.

A célula que originou a ovelha Dolly já havia sofrido várias multiplicações na mama; será que o telômero dela estava menor? Assim que a Dolly nasceu vários estudos com as suas células foram realizados e entre eles verificou-se o tamanho dos telômeros nessas células. Realmente, o tamanho do telômero das células da ovelha Dolly era menor do que o das ovelhas originadas de uma gestação normal. Portanto, a atividade da telomerase durante a embriogênese não foi suficiente para recompor o tamanho normal do telômero. O telômero menor indicava que a ovelha Dolly nasceu “mais velha” e isso se confirmou com a morte dela, seis anos depois com doenças relacionadas ao envelhecimento.

Após a clonagem da ovelha Dolly, vários experimentos de clonagem foram realizados e vários animais foram clonados. O mesmo grupo de pesquisadores clonou uma segunda ovelha, mas esta apresentava uma característica especial que não estava presente na Dolly e mostrando o potencial que a clonagem poderia ter para a biotecnologia. A ovelha Polly teve o seu DNA modificado para que contivesse um gene capaz de produzir uma proteína utilizada no combate à doença fibrose cística. Hoje, já temos vários clones que produzem produtos de interesse humano, como na Argentina, onde já existe um clone de vaca que é capaz de produzir insulina no leite.

Entretanto, o primeiro clone de vaca foi relatado em 1998 por cientistas japoneses. Curiosamente, os clones de vaca apresentam uma diferença significativa para os de ovelha: os telômeros dos clones são maiores que os das vacas que foram geradas por uma gestação natural. Será que uma vaca clonada irá sobreviver mais que uma vaca normal? A resposta a essa pergunta terá que aguardar algum tempo, mas já se sabe que as células da vaca clonada entram em senescência mais tardiamente que as vacas normais.

Vários animais já foram clonados como ratos, porcos, cachorros, gatos, macacos entre outros. No Brasil, pesquisadores da Embrapa já foram capazes de clonar vacas a partir de células embrionárias e células adultas

A grande pergunta que se faz é se é possível fazer a clonagem do ser humano. Vários grupos trabalham com a clonagem para a obtenção de células embrionárias para terapia gênica, mas como vimos no trabalho da Dolly, a eficiência do processo de clonagem é muito reduzido, com quase trezentas tentativas realizadas até se obter um resultado positivo. É inimaginável ser feito isso com humanos, além de todo o debate ético e moral que esse tipo de experimento levanta.

### **2.3. TERAPIA GÊNICA**

No início dos anos setenta, com o desenvolvimento da tecnologia recombinante, cientistas propuseram o que eles chamaram de “cirurgia gênica” para tratamento de doenças genéticas. Pretendia-se corrigir os defeitos genéticos causadores de certas doenças.

Em 1983, cientistas americanos propuseram que um dia a terapia gênica poderia

ser utilizada para tratar a doença Lesch-Nyhan, uma doença neurológica rara causada pela deficiência de uma enzima envolvida com o metabolismo de nucleotídeos. Eles realizaram experimentos nos quais o gene que codificava a enzima era inserido em um grupo de células, as quais acreditava-se, poderiam ser inseridas nos pacientes.

Como vimos, a terapia gênica consiste em “consertar” um gene que causa uma doença ou deficiência. Critérios para que um protocolo de terapia gênica seja estabelecido devem levar em conta os seguintes princípios:

- Disponibilidade do gene a ser corrigido.
- Eficiente introdução de genes nas células alvos.
- Célula alvo acessível ao procedimento de transferência gênica.
- A terapia não deve ser danosa ao paciente.
- A terapia gênica deve mostrar incremento significativo no quadro do paciente.

O primeiro protocolo realizado com um paciente seguiu todos os princípios acima. O procedimento foi realizado em 1990 em uma paciente que sofria de uma doença do sistema imune. Essa menina vivia em uma bolha e não podia ter contato com nenhum antígeno, uma vez que seu corpo não era capaz de desenvolver uma resposta imune adequada. Essa doença é causada por uma deficiência no gene que codifica a enzima adenosina desaminase que está envolvida no metabolismo de nucleotídeos.

Como descrito acima, o gene para a terapia gênica dessa doença era conhecido e poderia ser utilizado no tratamento. A célula alvo era factível, uma vez que os leucócitos podem ser obtidos do sangue do paciente e manipulados em laboratório antes de serem novamente reintroduzidos nele. Esse tipo de terapia que retira a célula e depois reintroduz é denominado de *ex-vivo*.

As células do sangue da menina foram trabalhadas *in vitro* e nelas foram inseridas um gene da adenosina desaminase que codificava a proteína sem defeito. As células modificadas foram selecionadas e utilizadas para reintroduzir na paciente. O resultado foi significativo e permitiu que a criança fosse capaz de passar algum tempo fora da bolha. Esse procedimento foi repetido várias vezes com aumento da eficiência do tratamento.

Como o DNA do gene da adenosina desaminase foi inserido no DNA da célula da paciente? Este é um ponto chave da terapia gênica que é a escolha do vetor adequado para a inserção do DNA. Iremos discutir a seguir os tipos de vetores utilizados que servem de “Cavalo de Tróia” nesse processo.

Na terapia gênica descrita acima, o vetor foi um retrovírus. Esses vírus apresentam genoma de RNA que é transformado em DNA e é capaz de ser inserido no genoma da célula. O que os cientistas fizeram foi modificar o genoma do vírus e inserir nele o gene da adenosina desaminase. Como as células da paciente estavam sendo crescidas no laboratório, os cientistas puderam verificar o local em que ocorreu a inserção do DNA do vírus e se esta inserção estava causando

algum problema. Depois que a checagem foi realizada, os cientistas puderam verificar se as células estavam produzindo a enzima de interesse e, após a confirmação do resultado, as células transformadas puderam ser reintroduzidas na paciente sem maiores problemas.

Os retrovírus são vetores muito eficientes na inserção de DNA exógeno no genoma humano, mas eles apresentam alguns problemas, tais como: a inserção realizada pelo genoma do vírus é randômica, ou seja, ela pode acontecer em qualquer lugar no genoma o que pode acarretar sérios riscos para o paciente. Por exemplo, ao ser inserido no genoma de uma célula, o vírus pode inativar um gene essencial causando sua morte, ou mesmo, inativar um gene envolvido no controle do ciclo celular da célula, que uma vez inativado, fará com que a célula perca o controle do ciclo celular e inicie um processo de duplicação contínuo que pode levar a um tumor. Além desses sérios problemas, o retrovírus só é capaz de infectar células que estão em multiplicação. Devido a suas limitações, os retrovírus só são utilizados em protocolos *ex vivo*, nos quais as células do paciente são extraídas, trabalhadas *in vitro* e depois reintroduzidas no paciente.

Mas como poderíamos fazer um tratamento de uma doença pulmonar, por exemplo? Não é possível retirar células do pulmão, trabalhar com elas *in vitro* e após a modificação reintroduzi-las. Os retrovírus não seriam um bom vetor para esse tratamento, uma vez que eles poderiam inserir o DNA em qualquer parte do genoma o que seria extremamente danoso para o paciente. Uma alternativa é o uso do adenovírus como vetor, com grandes vantagens, tais como: ele é capaz de infectar uma ampla gama de células *in vivo*, atingindo diferentes tipos de células com as quais não é possível fazer tratamento *ex-vivo*; aceita a inserção de grande quantidade de gene exógeno o que permite o uso de genes que tenham um grande número de pares de base; é capaz de expressar proteínas em células não mitóticas; e, por fim, não insere o seu DNA no genoma e, com isso, não leva o risco de inativar genes que sejam importantes para a célula.

O fato de não inserir o seu genoma no DNA humano é uma vantagem, mas tem uma desvantagem que é o de ter de se fazer várias infecções com o vírus para o prosseguimento do tratamento; pois o DNA do vírus não permanece muito tempo na célula e, dessa forma, a expressão da proteína é transiente. O grande problema das várias infecções virais é que podem gerar uma resposta imunológica do paciente que, mesmo branda, leva a uma deficiência no tratamento e causar a morte, em alguns casos.

Foi utilizando vetores adenovirais em protocolos de terapia gênica que se notificou um dos casos mais famosos de erro nos procedimentos adotados, o que levou a morte de um indivíduo. Como consequência, vieram vários questionamentos éticos e uma moratória nos procedimentos de terapia gênica.

Um pesquisador da Universidade da Pensilvânia nos Estados Unidos elaborou um procedimento para o tratamento de um distúrbio metabólico de origem genética denominada de deficiência de ornitina transcarbamilase (OTC). Esse gene está envolvido no ciclo da uréia e pessoas com essa doença não são capazes de eliminar este material. Normalmente, pacientes com essa doença entram em coma com 72 horas de vida e apresentam graves sequelas neurológicas. O

interesse do protocolo desenvolvido era tratar bebês com essa doença e fazer com que o fígado das mesmas fosse capaz de produzir a enzima necessária. O processo seria realizado com o uso de um adenovírus capaz de infectar células hepáticas e produzir a enzima nessas células. O propósito, nessa fase da pesquisa, era verificar se os adenovírus seriam capazes de serem direcionados para as células hepáticas.

O comitê de ética do Instituto da Pensilvânia considerou que solicitar a autorização de pais fragilizados com a gravidade do estado dos seus bebês não seria conveniente, e decidiu que o teste fosse realizado somente em adultos capazes de autorizar tais experimentos. Em virtude disso, um rapaz de 18 anos que apresentava uma forma mais branda da doença, foi convidado a participar do experimento. Ele apresentava um desenvolvimento compatível com a sua idade e sua deficiência era controlável através de uma dieta com uma drástica redução de proteínas além da utilização de medicamentos (32 comprimidos por dia).

A família do jovem não foi alertada sobre os problemas anteriores dessa pesquisa, como a morte de três macacos Rhesus e problemas hepáticos que tinham acometido em três outros pacientes. Sem essas informações, o adolescente se submeteu a uma infusão, na artéria hepática, de adenovírus contendo o gene para a correção da OCT. Infelizmente, quatro dias depois do tratamento, o adolescente faleceu devido a uma reação imunológica causada pelo adenovírus.

Outro vírus que também tem sido utilizado como vetor para o transporte de genes é o adenovírus associado. Diferente do adenovírus, o adenovírus associado é capaz de inserir o seu genoma no DNA humano e o faz em uma única posição do genoma, localizada no cromossomo 19. A inserção nessa região não é deletéria e o uso dos adenovírus associado tem grande potencial para a terapia gênica. O grande problema para o uso desse vírus é que ele é incapaz de transferir grande quantidade de DNA, o que impossibilita a transferência de vários genes humanos maiores do que pode ser inserido no genoma do vírus.

Outros vetores que podem ser utilizados na terapia gênica são os lipossomos. Estes são formados por lipídeos que constituem micelas em cujo interior está o DNA de interesse. Os lipossomos têm a grande vantagem de não possuírem nenhum gene viral, mas apresentam a desvantagem de serem menos eficientes na transferência gênica.

Além dos lipossomos, injeções de DNA nu também podem ser utilizados para a transferência do DNA. O DNA nu tem sido utilizado em protocolos de vacinação. O DNA de interesse normalmente contém um gene que codifica um antígeno para qual se pretende gerar uma resposta imune e espera-se que a expressão desse antígeno seja capaz de deflagrar a resposta imune.

A principal busca da biotecnologia para o desenvolvimento da terapia gênica é a produção de vetores mais adequados ao uso. Um vetor ideal terá que ter as seguintes características:

- Desenvolvimento de sistemas com altas taxa de transferência gênica e altos níveis de biossegurança.
- Capacidade de fazer o direcionamento adequado para a célula alvo.

- Capacidade de controlar a expressão do gene de interesse, ou seja, que o gene só seja expresso na célula alvo e no tempo adequado.
- Controle da expressão do gene, pois muitas vezes uma quantidade maior da proteína também pode levar a fenótipos deletérios para os pacientes.

Até agora vimos a terapia gênica como arma para tentar tratar doenças que ocorrem devido a perda de função de um gene, mas existem protocolos para tratar também outros tipos de doenças. Um exemplo clássico é o tratamento do câncer, sendo que a maioria de protocolos implementados de terapia gênica são direcionados para essa doença.

Várias estratégias podem ser adotadas para o tratamento do câncer utilizando a terapia gênica como arma. Pode-se procurar expressar antígenos de histocompatibilidade nas células cancerosas, o que facilitaria o reconhecimento dessas células pelo sistema imune. Muitos genes são capazes de controlar a replicação celular e impedir que a célula duplique. Muitas vezes ocorrem mutações nesses genes nas células cancerosas e uma abordagem utilizada é a expressão desses genes supressores de tumor na tentativa de bloquear o crescimento do tumor. Outra abordagem pode ser a expressão de genes suicidas nas células tumorais, que produzem uma proteína capaz de ter uma atividade que leva a morte celular.

### **2.3.1. CÉLULAS TRONCO E TERAPIA GÊNICA**

Além do processo de terapia gênica que busca consertar ou inativar um gene na célula, podemos pensar também em um processo celular, ou seja, a introdução de células em pacientes que possam melhorar a doença. Por exemplo, o infarto é a morte de várias células musculares cardíacas levando a um mau funcionamento do coração. Assim, se formos capazes de implantar células musculares cardíacas saudáveis no indivíduo, o coração poderá voltar a ter as funções normais.

Outro exemplo clássico de uma terapia celular é o tratamento de pessoas com paralisia que têm problemas com as suas células da medula. Espera-se que o implante de células do sistema nervoso saudáveis possa permitir que o paciente seja capaz de readquirir os movimentos perdidos.

Muitas pesquisas têm sido realizadas com células tronco, pois essas têm a capacidade de se diferenciar nas diferentes células de um tecido. Normalmente, um tecido é composto por diferentes tipos de células que possuem funções distintas. Um grupo especial de células do tecido tem a capacidade de se diferenciar nos diferentes tipos celulares e estas são as células tronco. A obtenção dessas células são importantes para tentar se obter todos os tipos celulares desse tecido.

O que as células embrionárias têm a ver com esse processo de terapia? As primeiras células do embrião têm a capacidade de desenvolver qualquer célula de qualquer tecido e são designadas de totipotentes. Após algumas divisões, as células do embrião já sofrem algumas modificações e umas irão formar a placenta e outras conservam a capacidade de formar as células de qualquer outro tecido. Essas células são chamadas de pluripotentes. Como vimos na

parte sobre clonagem, as células da pele já estão diferenciadas e não são capazes de formar outras células, uma vez que o seu padrão de expressão de genes já está definida e dificilmente é revertido. Portanto, os usos de células embrionárias podem ser de grande valia, uma vez que elas permitem formar células de diferentes tecidos.

A questão das células embrionárias levanta discussões religiosas uma vez que se discute se o embrião é um ser vivo ou não. Porém, considera-se que o estudo de células embrionárias poderá ter um grande impacto para diferentes doenças.

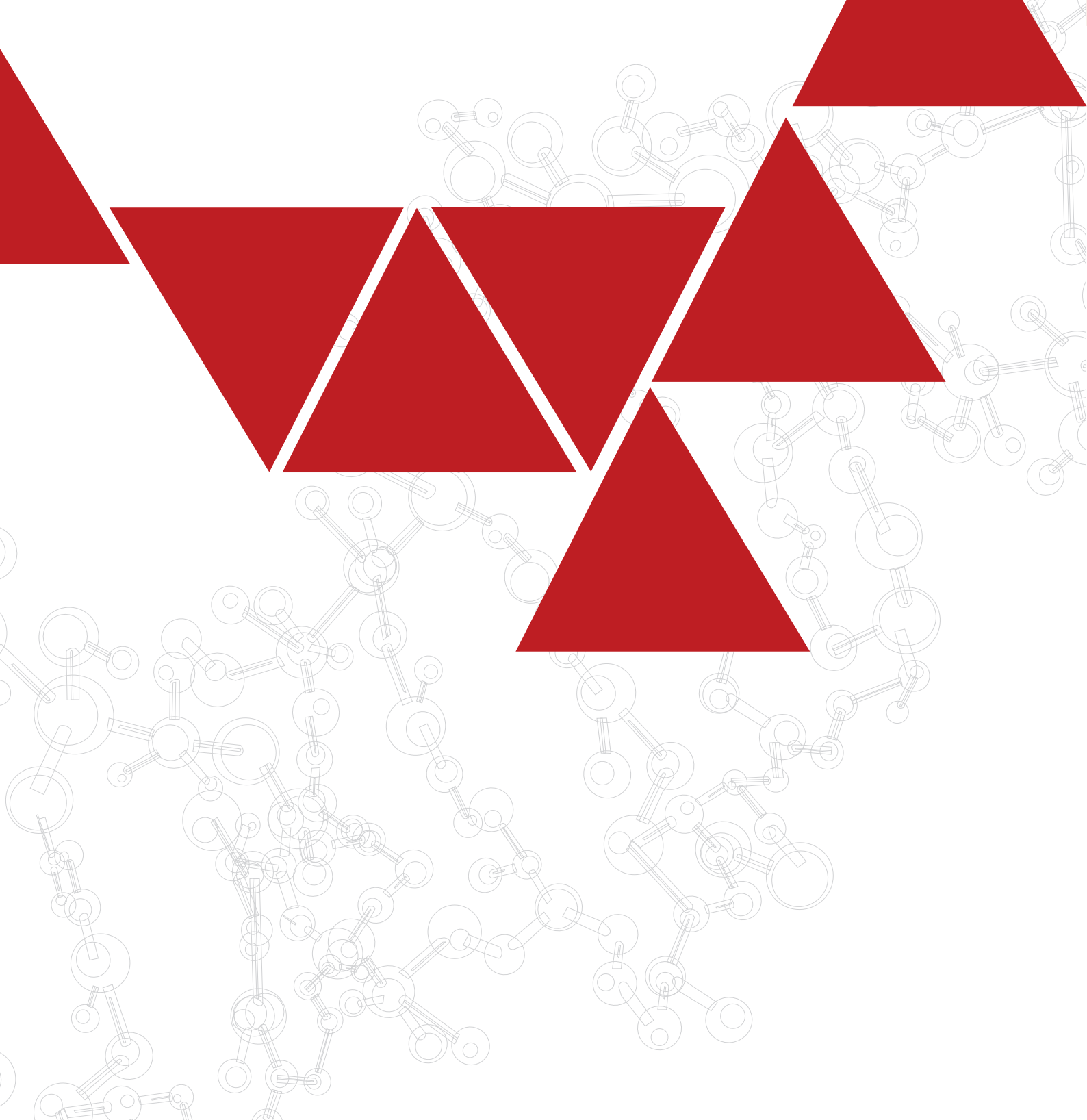
O estudo com células embrionárias ainda busca saber quais as condições que as células embrionárias têm de ser crescidas para que se obtenham os diferentes tipos celulares. Em um meio de cultura, quais são os fatores de crescimento devem ser acrescidos para que estas células sofram modificações que levem a formação da célula desejada? Essa é uma questão chave para o desenvolvimento dessa terapia.

Um grande avanço no estudo de células embrionárias foi realizado no ano de 2007. Pesquisadores foram capazes de modificar células de fibroblasto de tal forma que elas se tornaram pluripotentes. Os estudos com células embrionárias foram a chave para que os cientistas soubessem quais os genes deviam estar expressos em outras células para que estas se tornassem pluripotentes e fossem capazes de se modificarem em outras células. Apesar de muito promissora, essa abordagem ainda guarda alguns problemas. Um dos genes utilizados nesse processo é um oncogene, ou seja, um gene que é capaz de promover a replicação das células e a formação de tumores. Esse fato é um inconveniente para a utilização dessas células em processos terapêuticos.

Portanto, nessa aula vimos vários aspectos da biotecnologia e como elas podem interferir em nossa vida.

### Autoavaliação

- a. Cite exemplo de organismos que já foram sequenciados e qual o tamanho de seu genoma e número de genes.
- b. O que é o paradoxo C?
- c. Quais as etapas de um processo de sequenciamento de um organismo?
- d. O que é clonagem?
- e. Com DNA de que tipos celulares a clonagem é possível?
- f. Por que a Dolly nasceu “velha”?
- g. O processo de clonagem de um organismo é eficiente?
- h. O que deve nortear um protocolo de terapia gênica?
- i. Quais os tipos de vetores são utilizados? Quais as vantagens e desvantagens de cada um?
- j. Como seria um vetor ideal?
- k. Como seria uma terapia celular?
- l. O que é uma célula tronco?
- m. Quais as diferenças entre células troncos e células embrionárias?
- n. Qual a diferença entre um célula totipotente e a pluripotente?



3

*Lesões e  
Reparo de DNA*

## AULA 3: LESÕES E REPARO DE DNA

Ao final desta aula você deverá:

- Saber o que causa lesões no DNA.
- O que é uma lesão mutagênica e uma lesão letal.
- Quais são as vias de reparo de DNA.

Como já vimos em várias aulas, a informação presente no DNA é essencial para o desenvolvimento e manutenção de um organismo. Portanto, qualquer modificação na estrutura do DNA pode ser deletéria para um organismo. Mas o DNA é uma molécula inerte que não sofre alterações? Não, ao contrário, o DNA está exposto a uma série de agentes que são capazes de reagir com essa molécula e alterar a sua estrutura. Para que o organismo possa manter sua atividade normal é essencial que seja possível reparar as lesões às quais o DNA está sujeito. Nessa aula, iremos discutir o que causa e quais são os tipos de lesões que o DNA sofre, assim como quais são os mecanismos que permitirão reparar essas lesões.

### 3.1. LESÕES NO DNA

O DNA está sujeito ao ataque de uma série de agentes que são capazes de gerar lesões nessa molécula. Esses agentes podem ser endógenos, ou seja, produzidos pelo próprio organismo, ou exógenos. Além disso, podemos dividir os agentes genotóxicos entre químicos (estresse oxidativo, por exemplo) e físicos (luz ultra-violeta, por exemplo). Veremos a seguir alguns tipos de lesões que ocorrem no DNA.

O que é considerado uma lesão no DNA? Lesão é tudo que modifica a estrutura dessa molécula, desde pequenas modificações a grandes alterações são consideradas lesões. O DNA é uma dupla hélice composta pelos desoxiribonucleotídeos adenina, citosina, guanina e timina, e uma lesão pode afetar as bases nitrogenadas, ou mesmo, o esqueleto desoxiribose-fosfato. Algumas alterações na estrutura do DNA são fisiológicas e, assim, não são consideradas lesões. Um exemplo é a metilação (adição de um grupo  $\text{CH}_3$ ) que desempenha papel importante na regulação do DNA, mas se a metilação ocorrer em algumas regiões será considerada uma lesão.

A bactéria *Escherichia coli* metila o nitrogênio 6 da adenina em seu DNA em regiões que contém a sequência de nucleotídeos GATC (guanina, adenina, timina e citosina). Essa metilação tem um papel importante na biologia da bactéria, pois ela é uma assinatura de que aquele DNA pertence a bactéria. Caso um vírus infecte a bactéria, o seu DNA não conterá essa assinatura e a bactéria saberá que aquele DNA não lhe pertence e o mesmo deve ser degradado, o que será realizado por enzimas de restrição.

Nas células humanas e de outros organismos, a metilação do carbono 5 da citosina tem um papel muito importante na regulação da expressão gênica. Genes que não devem ser expressos, muitas vezes tem a citosina metilada na região promotora dos genes. Esses dois exemplos mostram que o DNA pode ser modificado por metilação e essa alteração é importante para o bom funcionamento celular. Mas nem toda metilação tem papel fisiológico e algumas podem formar lesões importantes no DNA. A metilação é realizada por enzimas específicas que são capazes de reconhecer o local a ser metilado; entretanto, uma atividade descontrolada das enzimas pode levar a metilação em locais indevidos e formar lesões. Por exemplo, a metilação da guanina na posição 6 (o número significa qual o átomo metilado na estrutura do anel da base nitrogenada) é uma lesão importante que tem conseqüências danosas para a célula, como veremos depois. A metilação da adenina na posição 3 também é uma lesão importante no DNA, mas resulta em conseqüências bem diferentes da lesão O-6-metilguanina.

A metilação no DNA pode ser feita tanto por uma ação inadequada das DNA metilases, como também podem ser feitas por agentes químicos como o metilmetanosulfonato (MMS).

Como vimos na aula 6 de Bases Moleculares II, sobre o metabolismo de aminoácidos e nucleotídeos, sempre que se troca uma ligação C-O por uma C-N é necessário o gasto de ATP. A necessidade de fosforilar uma molécula antes de ocorrer o processo de ligação de um grupo amina, ocorre devido a ligação carbono-nitrogênio ter mais energia que a ligação carbono-oxigênio. Portanto, para que essa troca possa ocorrer é necessário o gasto de energia que virá pela ligação fosfato rica em energia. Para a síntese das bases nitrogenadas esse processo ocorre em vários momentos. Se para a ligação do grupo amina é necessário o gasto de energia, a quebra dessa ligação liberará energia o que tornará essa reação espontânea. A água será responsável por atacar o DNA e liberar o nitrogênio e com isso teremos a troca de um nitrogênio por um oxigênio, o que denominamos de desaminação. Esse processo é capaz de modificar diferentes bases, mas a principal modificação é a transformação de citosina em uracila.

A uracila é uma base nitrogenada que está presente no RNA, mas não está presente no DNA e isso é muito importante para a estabilidade dessa molécula. A desaminação da citosina é um processo comum na célula e se a uracila fosse a base presente no DNA seria muito comum a troca da base citosina pela uracila, o que diminuiria as possibilidades de informação dessa molécula. Assim, a troca de uracila pela timina no DNA foi fundamental para a manutenção da sua integridade. A timina é sintetizada a partir da uracila e a diferença entre estas bases nitrogenadas é um grupamento metil presente no carbono 5. Se a citosina sofrer desaminação irá formar uracila e esta base será vista como lesão no DNA.

Assim como a desaminação é capaz de gerar a uracila, outra lesão importante que ocorre proveniente de um processo de hidrólise (quebra pela água) e que rompe uma ligação de carbono com o nitrogênio, é a depurinação. A ligação da base nitrogenada com o açúcar desoxiribose é feita entre o carbono 1 do açúcar

com um nitrogênio da base. Essa ligação pode ser quebrada e é gerado um sítio apurínico ou apirimidínico, ou seja, sem uma pirimidina ou purina

Outro processo espontâneo que ocorre em uma célula é o estresse oxidativo. As células aeróbias obtêm energia no processo de oxidação de biomoléculas com a consequente redução do oxigênio em água. Nesse processo, são retirados elétrons da glicose, ácidos graxos e aminoácidos, os quais são passados para o oxigênio. As etapas finais desse processo ocorrem na cadeia de transporte de elétrons, sendo que o oxigênio recebe os elétrons no complexo IV dessa cadeia. Entretanto, algumas vezes o oxigênio pode receber os elétrons em outros pontos da cadeia e muitas vezes essa transferência não é adequadamente realizada e acaba formando o que são conhecidas como espécies ativas de oxigênio (EAO). O oxigênio para formar água tem que receber quatro elétrons e quatro H<sup>+</sup> de uma vez; caso ele receba menos de quatro elétrons irão se formar as EAOs. Como elas possuem elétrons desemparelhados, se tornam extremamente reativos; e para se estabilizar é necessário que elas reajam com uma molécula e transfiram o elétron desemparelhado.

Uma das principais moléculas alvo das EAOs é o DNA, onde as bases nitrogenadas são os mais oxidados. Das várias lesões que são causadas pelo estresse oxidativo as principais são a 8-oxoguanina e a timina glicol. Na 8-oxoguanina, como o nome diz, um oxigênio é ligado no carbono 8 da guanina. Já na timina glicol duas hidroxilas são ligadas à base nitrogenada. Outras bases também podem ser oxidadas, tais como: a adenina (formando a 8-oxoadenina), a uracila (formando uracila glicol e 5-hidroxiuracila) e a citosina (formando a 5-hidroxicitosina). Mas a 8-oxoguanina é a principal lesão gerada e existem estimativas que ocorre a formação de aproximadamente 10000 8-oxoguaninas por célula por dia. Esses valores mostram como a molécula do DNA não é estável e como as lesões são frequentes.

Além de estar sujeito a sofrer lesões devido ao metabolismo celular (metilação e estresse oxidativo) ou ao ambiente em que se encontra (desaminação e depurinação devido à hidrólise), o DNA está sujeito a lesões causadas por agentes químicos e físicos que estão presentes no ambiente.

A luz ultravioleta é um exemplo clássico de agente físico que é capaz de gerar lesões no DNA. Desde o século passado já se sabia que a luz UV era um potente germicida, mas só na década de 50 descobriram que essa atividade é devido à formação de lesões no DNA. As bases nitrogenadas, principalmente as pirimidinas (citosina e timina), são capazes de absorver a energia da luz UV e isso faz com que essas moléculas fiquem energeticamente ativas e permite que as mesmas façam ligações covalentes entre elas formando os dímeros de pirimidina. A luz UV pode ser dividida em três grupos diferentes (UV-A, UV-B e UV-C) de acordo com o comprimento de onda. A luz UV-C é a mais energética e seu comprimento de onda está entre 250 nm até 280 nm. O comprimento de onda da luz UV-B vai de 280 nm até 320 nm, já a luz UV-A que tem menor energia tem um comprimento de onda acima de 320 nm. A luz com maior potencial germicida é a luz UV-C que induzirá a formação dos dímeros de pirimidina. A luz UV-B e, principalmente a luz UV-A, irão levar a lesões oxidativas no DNA. O sol emite os três espectros de luz UV, mas apenas a luz UV-A e UV-B chegam até

à terra, pois a luz UV-C é bloqueada na camada de ozônio.

Vários agentes químicos também podem causar lesões no DNA. A aflotoxina, muitas vezes presentes no amendoim, é capaz de se ligar covalentemente ao DNA (principalmente na guanina) formando o que é conhecido como aducto (qualquer composto químico ligado covalentemente a uma base nitrogenada que altere a sua estrutura). A cisplatina, uma droga utilizada no tratamento do câncer, também é capaz de se ligar ao DNA e formar aductos nessa molécula. A carne queimada em churrasco pode formar o que é conhecido como aminas heterocíclicas aromáticas que são substâncias que podem causar lesões no DNA e formar aductos no mesmo. Além desses exemplos, a vida moderna nos coloca em contato com diferentes agentes químicos que são capazes de gerar lesões no DNA.

Outros agentes externos capazes de causar lesões no DNA são a radiação gama e os raios-X. A energia gerada pela irradiação será absorvida pela molécula de DNA o que leva a quebra das ligações do esqueleto desoxiribose-fosfato. Por isso, esses agentes levam a formação de quebras simples (quando só uma fita do DNA é clivado) ou duplas (quando as duas fitas do DNA são clivadas). As quebras duplas do DNA são consideradas uma das lesões mais severas ao que DNA pode estar sujeito.

### **3.1.1. LESÕES MUTAGÊNICAS E LETAIS**

O que pode ocasionar as lesões no DNA? A função do DNA é de armazenar a informação genética. Nesse processo ele terá de ser duplicado para que as duas células contenham a mesma informação e, também, terá de ser transcrito em RNA para que a informação seja decodificada em proteínas. As duas enzimas envolvidas nesses processos são a DNA polimerase e a RNA polimerase e essas duas enzimas possuem um mecanismo básico de atuação que é usar uma fita de DNA como molde, seja para a duplicação, seja para a transcrição. Então, para que o processo de duplicação e transcrição ocorra normalmente é necessário que o DNA esteja íntegro para que o pareamento de Watson-Crick (adenina-timina e citosina-guanina) esteja correto. Qualquer alteração nessa informação pode levar a problemas nesses processos. Quais poderiam ser esses erros? As lesões no DNA podem levar a dois erros principais: mutação e letalidade. Se a lesão permitir a passagem da DNA polimerase, mas nesse processo levar a DNA polimerase a cometer um erro de pareamento, a lesão será considerada mutagênica. Por outro lado, se a lesão não permitir que a DNA polimerase faça o pareamento, isso levará a uma parada da replicação o que acarretará a morte da célula.

Um exemplo clássico de lesão mutagênica é a O-6-metil-guanina. Normalmente, a guanina faz três pontes de hidrogênio com a citosina no pareamento normal de Watson-Crick. O problema que acontece é que o oxigênio 6 da guanina é um dos átomos envolvidos na formação de pontes de hidrogênio e a presença do grupamento metil impede que isso ocorra. Na verdade, a base modificada O-6-metil-guanina parecia mais adequadamente com a timina, formando duas pontes de hidrogênio. Dessa maneira, quando a DNA polimerase encontra a

O-6-metil-guanina no DNA ela conseguirá fazer a duplicação do DNA, mas ela só conseguirá fazer o pareamento com timina. Em um próximo passo de replicação, a timina que foi colocada irá parear normalmente com adenina e, assim, teremos uma mutação na qual trocamos um pareamento C-G por um pareamento T-A. Portanto, esta é uma lesão mutagênica, pois será capaz de levar a troca de pareamento de base no DNA. É muito importante vermos que a lesão em si não é uma mutação e até poderá levar a uma mutação, mas desde que seja capaz de induzir a DNA polimerase a um erro durante o processo de duplicação.

Uma lesão será letal caso ela não permita que o processo de duplicação ocorra, ou seja, que impeça a atividade da DNA polimerase. Um exemplo clássico de uma lesão desse tipo são os dímeros de timina provocados pela ação da luz UV. Os dímeros de timina são ligações covalentes entre duas bases timina adjacentes. O sítio catalítico das DNA polimerases envolvidas no processo de replicação é bem estrigente, ou seja, o espaço do sítio ativo permite apenas a presença adequada de um nucleotídeo da fita molde e o nucleotídeo que será acrescentado na fita filha. Essa estringência é importante para que a DNA polimerase seja o mais preciso possível na hora da duplicação e passe a informação correta para gerações futuras. No sítio catalítico só são bem acomodados e permitem a atividade da enzima os nucleotídeos que fazem o pareamento de Watson-Crick. Qual o problema do dímero de timina? Como as timinas estão ligadas covalentemente, quando uma é colocada no sítio ativo da enzima ela trará consigo a base com que ela está ligada; assim a fita molde não conseguirá se encaixar no sítio ativo da enzima e a DNA polimerase não terá a capacidade de continuar a duplicar o DNA e, com isso, esse processo é bloqueado e pode levar à morte celular. Qualquer lesão que impeça a progressão da DNA polimerase será letal para a célula.

Para o processo de transcrição é muito mais importante as lesões letais do que as mutagênicas. As lesões mutagênicas não impedem a produção do RNA, apesar de poderem levar a alterações na informação dessa molécula; mas como isso estará restrito apenas a essa célula e não será passado para a descendência (só a duplicação de DNA está envolvida nesse processo) é mais tolerado. Por outro lado, lesões que bloqueiam a RNA polimerase poderão ter um impacto maior para a célula e têm de ser rapidamente reparadas, como veremos no processo de reparo preferencial de genes transcritos.

## 3.2. AS VIAS DE REPARO DE DNA

Como vimos acima, as lesões no DNA são de vários tipos e podem ser causadas por diferentes agentes; além disso, elas têm um grande impacto no metabolismo do DNA podendo gerar mutações e mesmo a morte das células. Em vista disso, não é surpresa que todos os organismos vivos possuam diferentes vias de reparo que são capazes de detectar e reparar os diferentes tipos de lesão. As vias de reparo de DNA devem ter surgido muito cedo na evolução da vida, uma vez que estão presentes em todas as células estudadas até o momento e que elas apresentam características comuns que foram preservadas. Veremos

a seguir as diferentes vias de reparo e os tipos de lesões que elas são capazes de detectar e reparar.

### 3.2.1. REVERSÃO DIRETA DA LESÃO

A via de reparo mais simples que existe é denominada de reversão direta da lesão. Nesse processo, uma enzima será capaz de reconhecer a lesão e será capaz de, simplesmente, retirar a lesão deixando o DNA intacto. Nesse processo não é preciso a síntese de nova molécula de DNA. Temos dois exemplos desse processo que são muito importantes: a enzima O-6-Metilguanina metil transferase (MGMT) e a fotoliase.

Como o próprio nome diz, a enzima MGMT é responsável em retirar o grupamento metil da guanina e transferi-lo para um aminoácido cisteína presente na enzima. Na verdade, essa proteína não pode ser considerada uma enzima, uma vez que quando a cisteína está metilada a proteína perde atividade, pois não é capaz de repassar o grupo metil. Portanto, para cada lesão O-6-metilguanina presente uma proteína MGMT será utilizada e inativada. Como descrito, a MGMT simplesmente retira o grupamento metil e refaz o nucleotídeo guanina em sua estrutura normal. Como essa lesão é mutagênica, a ausência da proteína MGMT fará com que a taxa de mutação da célula seja maior. É interessante verificar que vários tipos de câncer são deficientes nessa proteína.

A outra enzima clássica envolvida na reversão direta da lesão é a fotoliase. Essa enzima é capaz de desfazer os dímeros de timina. Na presença de luz branca, a fotoliase é capaz de absorver energia (proveniente da luz branca) e usar essa energia para desfazer as ligações covalentes dos dímeros de timina. Uma curiosidade sobre a fotoliase é que ela está presente em quase todos os organismos, mas não é encontrada nos mamíferos. Por que isso ocorreu? Acredita-se que os mamíferos primitivos eram organismos com hábitos noturnos e, portanto, não estavam expostos a radiação solar, que é a principal fonte de luz UV e que gera os dímeros de timina. É curioso verificar que existe uma enzima em mamíferos que apresenta homologia com a fotoliase, mas essa enzima tem papel no ciclo circadiano, ou seja, está envolvida com a resposta, a presença e a ausência de luz. É possível que essa enzima possa ter evoluído da fotoliase.

### 3.2.2. REPARO POR EXCISÃO

Um tipo importante de reparo de DNA é o reparo por excisão. O princípio básico desse mecanismo de reparo é a retirada do nucleotídeo que contém a lesão, o que gera um espaço no DNA. Para preencher esse vazio, a fita de DNA que não contém lesão será utilizada como molde e a DNA polimerase irá sintetizar um novo nucleotídeo no lugar onde estava o nucleotídeo que continha a lesão. Esse tipo de reparo pode ser dividido em três grupos principais: reparo por excisão de base, reparo por excisão de nucleotídeo e reparo de erro de pareamento.

### 3.2.2.1. REPARO POR EXCISÃO DE BASE

No reparo por excisão de base, como o próprio nome diz, apenas o nucleotídeo que contém a lesão é removido. Nesse processo, as enzimas responsáveis em reconhecer a lesão são as DNA glicosilases. Existem diferentes DNA glicosilases que irão reconhecer lesões específicas; por exemplo a 8-oxoguanina glicosilase (OGG) irá reconhecer a lesão 8-oxoguanina. A enzima uracil DNA glicosilase irá reconhecer uracila presente no DNA, já a 3-metil adenina glicosilase irá reconhecer a adenina metilada na posição 3.

A uracila pode surgir devido à desaminação da citosina ou pode ser incorporada durante o processo de replicação pelas DNA polimerases, uma vez que existe na célula desoxiribonucleotídeo uracila e este pareia perfeitamente com a adenina e pode enganar a DNA polimerase. Por essa razão, possuímos diferentes Uracil DNA glicosilases, uma que reconhece uracila pareada com adenina (reconhece as uracilas inseridas no processo de replicação) e outra que reconhece uracila pareada com guanina (reconhece as uracilas que surgem com a desaminação das citosinas). Como descrito mais acima, a citosina pode ser metilada para regulação da expressão gênica, e caso essa citosina sofra um processo de desaminação ela irá formar timina ao invés de uracila. A timina é uma base normal presente no DNA, mas nesse caso ela estará pareada com guanina e existe uma DNA glicosilase que é capaz de reconhecer a timina pareada com guanina.

A enzima OGG é capaz de reconhecer a 8-oxoguanina pareada com citosina mas é incapaz de reconhecer a 8-oxoguanina pareada com a adenina. Isso é importante para evitar o surgimento de uma mutação. A 8-oxoguanina é uma lesão que surge com o estresse oxidativo e, se ela está pareada com a citosina, deve ser retirada para que uma guanina normal possa ser inserida. Mas se a 8-oxoguanina não for retirada antes da replicação, a DNA polimerase poderá inserir em frente a essa base lesionada o nucleotídeo adenina, pois a 8-oxoguanina é capaz de parear com essa base eficientemente. Se a OGG fosse capaz de reconhecer a 8-oxoguanina em frente à adenina, ela tiraria a lesão do DNA e a DNA polimerase iria colocar a timina em frente à adenina. Dessa maneira, ocorrerá uma mutação com troca do pareamento G-C pelo pareamento T-A. Por esse motivo, a OGG não reconhece 8-oxoguanina pareada com adenina. Existe uma DNA glicosilase (denominada MutY) que é capaz de reconhecer a adenina pareada com a 8-oxoguanina e retirar não a lesão, mas a base que foi incorporada erroneamente (a adenina). Essa atividade dará uma segunda chance para a DNA polimerase inserir uma citosina em frente a 8-oxoguanina, e se isso ocorrer a OGG poderá retirar a 8-oxoguanina e permitir a incorporação de uma guanina normal.

As DNA glicosilases retiram somente a base nitrogenada que contém a lesão, pois quebram a ligação glicosídica entre a desoxiribose e a base nitrogenada. A ação dessas enzimas deixa outra lesão no DNA que é o sítio abásico, ou seja, sem uma base nitrogenada. Essa lesão é letal, pois a DNA polimerase não tem nenhuma informação do que deve parear em frente e, por isso, ocorre o bloqueio da replicação. O próximo passo do reparo de excisão de base é a

quebra da ligação do esqueleto da fita de DNA. A enzima AP endonuclease é a responsável por essa atividade e gera uma quebra simples na fita do DNA.

Apesar de já possuir um extremidade 3'OH que servirá de iniciador para a atividade da DNA polimerase, a polimerização não ocorrerá, pois restam no DNA uma desoxiribose e um fosfato do nucleotídeo que perdeu a base. Em células humanas, a DNA polimerase beta tem dois sítios catalíticos distintos, e um desses sítios tem a capacidade de clivar a ribose-fosfato que gera um gap de um nucleotídeo. Esse espaço será preenchido pela atividade da DNA polimerase beta.

O último passo do reparo de excisão de base é a ligação do nucleotídeo incorporado com fita do DNA, selando a quebra ainda existente. A ligação é realizada pelas DNA ligases e para isso elas gastam energia que oriunda do ATP ou NADH.

### **3.2.2.2. REPARO POR EXCIÇÃO DE NUCLEOTÍDEOS**

Outra via de reparo por excisão é o de nucleotídeos, no qual a lesão é reconhecida e não apenas o nucleotídeo que contém a lesão é removido, mas sim vários nucleotídeos são excisados e entre eles aquele que contém a lesão. Diferente do reparo por excisão de base em que diferentes enzimas reconhecem diferentes lesões, no reparo por excisão de nucleotídeos é reconhecida qualquer lesão que leve a uma distorção na hélice do DNA, tais como os dímeros de timina e os aductos do DNA.

As enzimas XPC e HHR23B são responsáveis por percorrer o DNA em busca de lesões, e ao reconhecer as lesões, irão recrutar diversas proteínas que irão participar do reparo. Entre estas enzimas encontramos as enzimas XPB e XPD que têm atividade de helicase. Essas enzimas têm a capacidade de desfazer as ligações de hidrogênio que são feitas pelos nucleotídeos em uma dupla hélice do DNA o que irá abrir a fita do DNA. As proteínas XPA e RPA são importantes para estabilizar as proteínas do reparo e para proteger o DNA de fita simples gerado pela atividade das helicases. O próximo passo do reparo é feito pelas enzimas XPG e XPF/ERCC1 que são endonucleases. A XPG é capaz de clivar o DNA na extremidade 3' da lesão. Esse corte entre a ligação da desoxiribose com o fosfato é realizado a aproximadamente 10 nucleotídeos da lesão. Já a enzima XPF/ERCC1 é responsável por clivar a ligação desoxiribose com o fosfato a aproximadamente 12 nucleotídeos a 5' da lesão. Como as helicases já tinham desfeito as pontes de hidrogênio entre as bases e as endonucleases clivaram as ligações ribose fosfato, um segmento de nucleotídeos é liberado do DNA, sendo que a lesão sai junto com os nucleotídeos. Dessa forma, é gerado um espaço do DNA em que temos uma fita molde e um 3' OH que será utilizado como iniciador por uma DNA polimerase. Esse processo de polimerização será realizado pelas DNA polimerases replicativas e o processo será terminado novamente por uma DNA ligase que selará a quebra entre a fita recém polimerizada com a que já era existente.

O reparo por excisão de nucleotídeos reconhece alterações na estrutura do DNA e muitas vezes essas distorções impedem a atividade da DNA polimerase. Elas também são capazes de impedir a passagem das RNA polimerases o que inviabilizaria a transcrição gênica. Portanto, uma RNA polimerase parada é um sinal de que aquela região do DNA contém uma lesão que deve ser reparada. A esse processo chamamos de reparo preferencial dos genes transcritos. A parada da RNA polimerase irá sinalizar a lesão e todas as proteínas acima, com exceção da XPC e HHR23B que atuam no reconhecimento da lesão, irão se deslocar para a região em que RNA polimerase está parada e iniciarão o processo de reparo.

### 3.2.2.3. REPARO DE ERRO DE PAREAMENTO

O outro reparo por excisão que existe é reparo por erro de pareamento. Quando a DNA polimerase duplica o DNA ela pode cometer um erro a cada  $10^4$  nucleotídeos incorporados. Esse erro acontece porque as bases nitrogenadas podem adotar configurações raras nas quais o pareamento de Watson-Crick é modificado. Por exemplo, a guanina normalmente pareia com a citosina, mas se ela está em sua forma enol (que é rara) ela é capaz de parear com a timina. Se esse erro não fosse corrigido, em uma próxima replicação a DNA polimerase iria inserir uma adenina em frente a timina, o que levaria à troca de um par G-C para um par A-T.

Para aumentar a fidelidade do processo de replicação existe o reparo de erro de pareamento que atuará logo após a duplicação do DNA. Em bactérias, o reparo de erro de pareamento funciona da seguinte maneira. Um dímero da proteína MutS caminha atrás da DNA polimerase e essa proteína é capaz de reconhecer erros de pareamento. Após reconhecer o erro de pareamento, a proteína MutS recruta um dímero da proteína MutL que irá recrutar a endonuclease mutH. No erro de pareamento as duas bases presentes são bases normais. Como saber qual das duas bases foi colocada de forma errônea? Para descobrir isso é preciso saber qual é a fita filha, ou seja, aquela que foi recentemente sintetizada pela DNA polimerase. Como já vimos no início da aula, o DNA de bactéria apresenta uma adenina metilada que serve como uma marca que aquele é o seu DNA. No processo de replicação a fita mãe (molde) está metilada, mas a fita filha ainda não está metilada; assim a ausência da metilação é a marca para identificar qual é a fita recém sintetizada. A endonuclease H irá reconhecer a região hemimetilada e irá clivar o DNA na fita filha. O DNA clivado será degradado por um endonuclease e, assim, a DNA polimerase terá uma segunda oportunidade para polimerizar corretamente o DNA.

Em eucariotos, o reparo de erro de pareamento segue as mesmas características do descrito para o de bactérias. Entretanto, o DNA de eucarioto não apresenta um padrão de metilação que permita distinguir entre as fitas mãe (molde) e filha. Então, como é feita a discriminação entre as duas fitas? No processo de replicação, a fita filha apresenta várias quebras oriundas dos fragmentos de Okasaki e de outros processos. Estas quebras são a indicação de qual fita deve ser reparada.

Outra função importante do reparo de erro de pareamento em eucariotos é a estabilidade de microssatélites, que são regiões repetitivas no genoma. Quando as DNA polimerases estão replicando, essas regiões podem “derrapar” e aumentar ou diminuir o número de repetições gerando alças de inserção ou deleção. O reparo de erro de pareamento é responsável por corrigir esses erros e manter a estabilidade do genoma.

### 3.2.3. REPARO DE QUEBRA DUPLA

Um tipo de lesão muito complicada para fazer o reparo é a quebra dupla do DNA. Essa lesão não pode ser corrigida pelas vias apresentadas acima, uma vez que elas precisam de uma fita molde não lesada para fazer o reparo. Na quebra dupla as duas fitas estão danificadas e não podem servir como molde.

Existem duas vias de reparo que são responsáveis por reparar a quebra dupla no DNA: a recombinação e a junção de extremidades não coesivas.

Na recombinação, o DNA quebrado será sujeito à atividade de uma éxonuclease que degradará uma das fitas do DNA. Esse processo irá gerar uma fita simples com uma extremidade 3'OH. A proteína RecA (em procariotos) ou RAD51 (em eucariotos) irá se associar com a fita simples do DNA que irá se deslocar em direção ao cromossomo homólogo e procurar a região de homologia dessa região simples com o cromossomo homólogo. Ao encontrar essa região, teremos o pareamento da fita simples com o cromossomo homólogo e, para que isso pudesse ocorrer, o DNA que estava intacto teve que desfazer as pontes de hidrogênio que os unia. A fita que foi deslocada para que pudesse ocorrer o pareamento poderá parear com a outra fita quebrada. Essas ligações da fita quebrada com o cromossomo homólogo servem de substrato para o DNA polimerase duplicar o DNA e reconstituir o que foi quebrado. Por fim, é necessário que ocorra a resolução da junção de Holliday (ver filme no Moodle que mostra como ocorre a recombinação em *E. coli*). O processo de recombinação é livre de erro e restaura a informação presente no DNA.

A junção de extremidades não coesivas é a maneira mais simples de reparar as quebras duplas do DNA. Quando elas ocorrem, a primeira reação é a ligação das proteínas Ku nas extremidades do DNA. Essa ligação é importante para proteger o DNA da ação de éxonucleases que poderiam degradar o DNA. Além disso, a ligação da proteína Ku impede que as extremidades se liguem em qualquer DNA. Após a ligação das proteínas Ku, várias proteínas são recrutadas para a extremidade do DNA e promovem a ligação dessas extremidades. Esse processo nem sempre restaura a informação genética, pois a quebra pode liberar fragmentos de DNA que serão perdidos nesse processo.

Vários tipos celulares, inclusive as células humanas, dão preferência em reparar quebras duplas de DNA pelo método de junção de extremidades não coesivas. Por que será que estas células dão preferência para um tipo de reparo que pode levar à perda da informação genética? Como já discutimos em aulas anteriores, o genoma de alguns organismos possuem muitas repetições e isso pode gerar

um grande risco no processo de recombinação, pois o processo de recombinação implica na busca de sequências homólogas. Se a quebra ocorrer em uma região repetitiva, a homologia poderá ser feita com uma região repetitiva que esteja em outro cromossomo que não o homólogo. Dessa maneira, a recombinação poderia levar a alterações cromossômicas que podem ser muito deletérias. Essa é a razão de se preferir fazer um mecanismo de reparo que pode levar a perder alguma informação genética, mas que seria menos deletéria do que a instabilidade cromossômica gerada pelo processo de recombinação.

### 3.3. REPARO DE DNA E DOENÇAS

Os mecanismos de reparo de DNA são muito importantes para as células e deficiências nessas vias estão associadas com doenças humanas.

A primeira doença associada com problemas em vias de reparo de DNA foi a xeroderma pigmentosum. As pessoas que sofrem dessa doença têm como sintoma principal a sensibilidade à luz UV e são denominadas “crianças da lua”, pois as áreas dos seus corpos expostas ao sol apresentam uma grande incidência de câncer de pele. Em 1967 foi a primeira vez que se mostrou que células de pacientes com *xeroderma Pigmentosum* apresentavam deficiência na via de reparo de excisão de nucleotídeos e são sensíveis à exposição da luz UV. Os pesquisadores verificaram que essas células não realizavam a polimerização de DNA que está associada ao reparo de DNA. A doença xeroderma pigmentosum pode ser causada por deficiências nos diferentes genes envolvidos no reparo por excisão de nucleotídeos. Por estar associada a uma doença humana, essa via de reparo foi por muito tempo a mais estudada.

Só na década de 90 do século passado, outra doença humana foi associada com as vias de reparo de DNA. Cientistas americanos verificaram que algumas pessoas que tinham predisposição ao câncer de cólon apresentavam deficiência no reparo de erro de pareamento. Uma característica desses indivíduos era a instabilidade de microssatélites, que variavam de tamanho nas células devido a ausência do reparo de erro de pareamento. Depois de verificarem esse fenótipo nesses pacientes, os cientistas mostraram que várias células cancerosas também apresentavam instabilidade de microssatélite, o que mostra o envolvimento da via de reparo de erro de pareamento em diversos tipos de tumores.

Hoje já se conhece vários genes que levam a suscetibilidade ao câncer. Um exemplo clássico é o gene BRCA1 que está envolvido no processo de recombinação e se sabe que deficiência desse gene predispõe ao câncer de mama.

Algumas vias de reparo por serem essenciais ao desenvolvimento não estão associadas a nenhuma doença. Esse é o caso da via de reparo por excisão de base, pois esta via é responsável por reparar várias das lesões que surgem durante o metabolismo normal de uma célula (estresse oxidativo, desaminação e metilação) e, portanto, se esta via está inativa ocorrerá várias lesões durante a embriogênese o que levará a morte antes do nascimento.

### 3.4. A QUIMIOTERAPIA E RADIOTERAPIA

Se a deficiência no reparo de DNA pode ser letal ou causar doenças, as lesões também podem ter um lado benéfico para o homem. Uma célula cancerosa está em um ritmo acelerado de duplicação e para tanto precisa replicar o seu DNA constantemente. Como as lesões do DNA podem impedir a replicação, uma boa alternativa para o tratamento dessa doença seria a geração de lesões no DNA, que bloqueassem a replicação e levassem a morte da célula. Na verdade, essa abordagem no tratamento é muito utilizada com os quimioterápicos e radioterápicos.

Vários quimioterápicos utilizados levam a lesões no DNA que bloqueiam a replicação deste. A cisplatina é uma droga muito utilizada no tratamento de diversos tipos de câncer e ela é capaz de se ligar covalentemente ao DNA e bloquear a replicação. Diversas pesquisas estão sendo realizadas em busca de agentes químicos que sejam capazes de causar lesões no DNA e bloquear a replicação dessa molécula.

O processo de radioterapia também leva a lesões no DNA, sendo a principal delas a quebra dupla do DNA que é uma lesão muito letal. O tratamento das células cancerosas tem efeitos colaterais, pois as lesões irão atuar nas células que estão replicando. Como ainda não somos capazes de direcionar o tratamento apenas para as células cancerosas, células do intestino, da pele, do sistema imune e dos pelos também serão afetadas pelo tratamento. Assim, um indivíduo tratando do câncer terá como efeito colateral os seguintes sintomas: diarreia, ressecamento da pele, imunodeficiência e queda de cabelos.

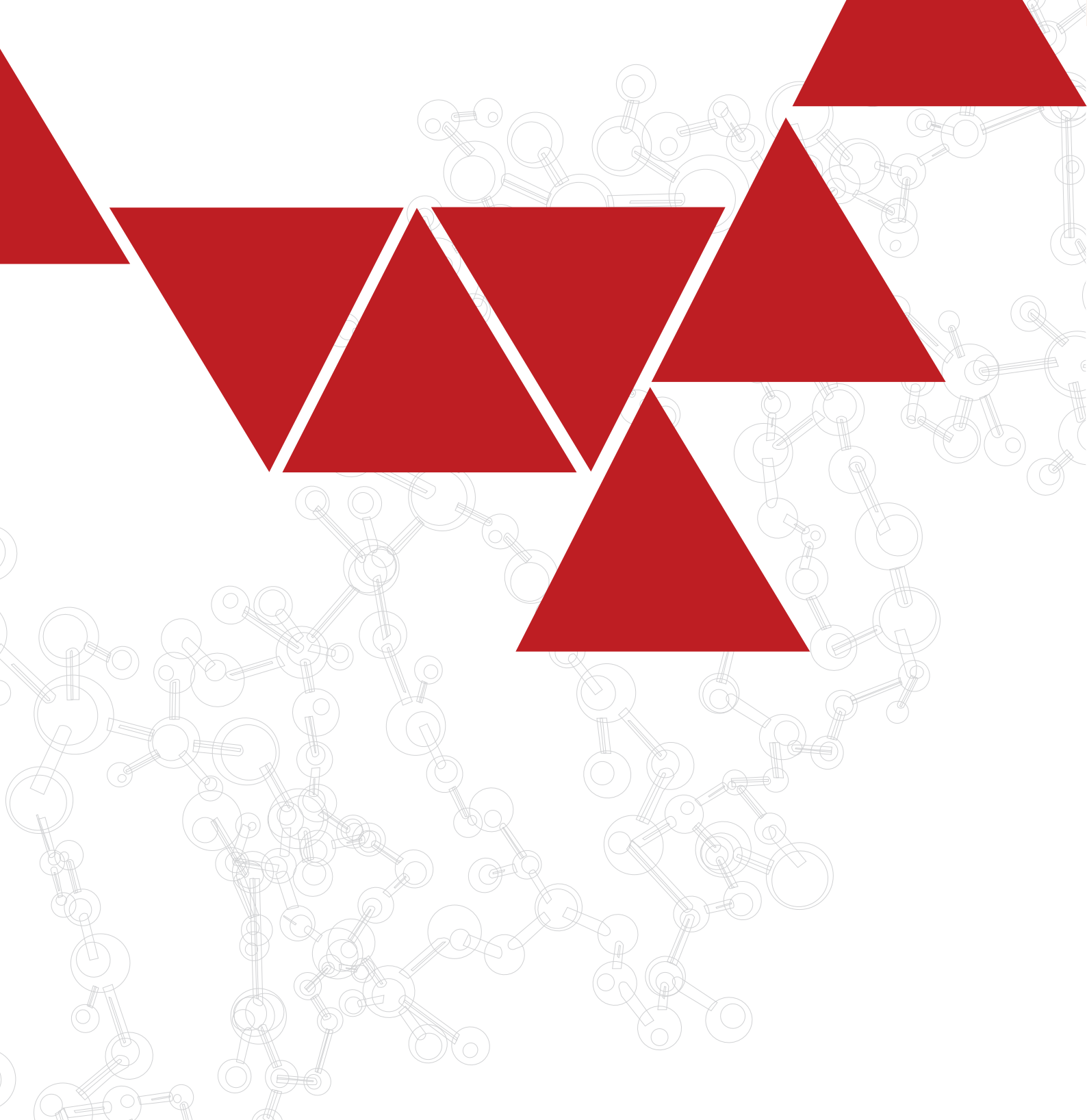
Nessa aula vimos que o DNA é uma molécula sujeita a uma série de lesões que podem ser mutagênicas e letais. Existem várias vias de reparo de DNA que são capazes de reconhecer diferentes tipos de lesão e deficiências dessas vias pode causar doenças ou mesmo a morte prematura do organismo.

#### **Complemento da aula:**

- No moodle vocês encontrarão diferentes animações e figuras para maior compreensão da aula.

## Autoavaliação

- a. O que pode causar lesões no DNA?
- b. O que é uma lesão? Dê exemplos.
- c. O que é uma lesão mutagênica?
- d. O que é uma lesão letal?
- e. No que a lesão pode interferir?
- f. Quais são as vias de reparo de DNA?
- g. Quais as evidências que apontam para a importância do reparo de DNA para a vida?
- h. Como funciona a reversão da lesão?
- i. Como funciona o reparo por excisão de base?
- j. Como funciona o reparo por excisão de nucleotídeos?
- k. Como funciona o reparo de erro de pareamento?
- l. Como funciona a recombinação?
- m. Como funciona junção de extremidades não coesivas?
- n. Quais doenças estão associadas ao reparo de DNA?
- o. Como podem atuar os quimioterápicos e radioterápicos?



**4**

***Polimorfismos moleculares:  
viva a diferença!***

## AULA 4:

# POLIMORFISMOS MOLECULARES: VIVA A DIFERENÇA!

Ao final desta aula o aluno deverá:

- Compreender o conceito de polimorfismos.
- Compreender as técnicas que detectam polimorfismos moleculares.
- Compreender as aplicações da caracterização de polimorfismos.

### 4.1. DEFINIÇÕES E CLASSIFICAÇÃO DE POLIMORFISMOS GENÉTICOS

O termo polimorfismo já sugere diferenças, pois é formado por “poli”, que significa muitas, e “morfos”, que significa formas. Comparem, então, as duas definições abaixo de polimorfismo genético:

- *Polimorfismo genético é a ocorrência de variação de natureza genética em uma população para um ou mais locos cujos alelos têm frequência inferior a 0,99. Oposto de monomorfismo, ausência de variação ou predominância de um dos alelos cuja frequência é = 0,99.” (Glossário de Biotecnologia, edição 2005)*
- *Polimorfismo genético é a ocorrência em uma população (ou entre populações) de várias formas fenotípicas associadas a alelos de um gene ou homólogos de um cromossomo.” (Introdução a Genética, Griffiths e cols. Oitava edição, 2006)*

Pode-se notar que essas duas definições são muito semelhantes, sendo que a única diferença é a necessidade de frequência maior do que 1% para o alelo polimórfico menos frequente. Essa distinção separa alelos mutantes (que ainda ou nunca atingirão a frequência de 1%) de alelos polimórficos, e isso é levado em consideração por diversos autores. Assim, para esses autores, mesmo que os polimorfismos sejam originados de mutações (com possíveis

recombinações posteriores, formando um haplótipo<sup>1</sup>), só serão considerados como tal se não forem eliminados da população por forças evolutivas, como seleção ou deriva genética (recorde os conceitos de forças evolutivas relendo as aulas de Genética de Populações).

Os polimorfismos podem ser classificados em diversos tipos, como morfológicos, imunológicos, bioquímicos (variantes em atividade enzimática, afinidade de

<sup>1</sup> **haplótipo:** inicialmente usado para denominar um grupo de dois ou mais genes extremamente próximos que, portanto, serão herdados conjuntamente, já que a probabilidade de haver crossing over entre eles é bastante remota. Com a possibilidade de análise molecular do DNA esse termo passou a ser usado também para grupos de duas ou mais mutações no mesmo trecho da mesma molécula de DNA, que também será herdado em conjunto.

ligação, carga, etc.), cromossômicos (tamanho, presença de rearranjos, tamanho da heterocromatina centromérica etc.), e de DNA (podem ocorrer em regiões codificantes ou não).

Os **polimorfismos morfológicos** englobam as diferenças de fenótipos morfológicos, sejam eles visíveis a olho nu ou com auxílio de aparelhos (como lupas e microscópios). Os demais tipos de polimorfismos podem ser chamados de **polimorfismos moleculares**, pois necessitam de técnicas analíticas para serem detectados. Os **polimorfismos bioquímicos**, sejam metabólicos ou protéicos, requerem técnicas de detecção de presença de metabólitos, atividade enzimática ou caracterização protéica. Boa parte dessas técnicas foi discutida nas aulas anteriores dessa disciplina.

Os **polimorfismos de DNA** são detectados por técnicas de análise de DNA. Assim como as mutações, esses polimorfismos são gerados por troca de bases, inserções ou deleções, que podem ocorrer em regiões codificantes ou não, e podem ser relacionadas ou não com fenótipos bioquímicos ou morfológicos. Mesmo quando o polimorfismo é resultante em uma região não codificante, ele é chamado de **alelo** em relação às demais possibilidades de sequência de nucleotídeos da região de DNA em que ele se encontra. Ampliamos agora o conceito de alelo proposto inicialmente por Mendel, que era vinculado ao conceito de gene. Um alelo agora pode estar (o que não quer dizer que ele necessariamente esteja) em uma região que nunca é transcrita e que não interfere em nada com a regulação gênica. Classificamos os polimorfismos de DNA, em função de suas características, em:

- **SNPs** (lê-se *snipes*), que são polimorfismos de um nucleotídeo, ou seja, resultantes da troca de bases.
- **Indels**, que são inserções ou deleções, geralmente de um ou poucos nucleotídeos.
- **VNTR**, repetições de sequências variadas, de 20 a 200 nucleotídeos (minissatélites)
- **STR**, repetições de sequências simples, (ou **SSLP**, polimorfismo de comprimento de sequências simples), com número baixo de nucleotídeos em cada unidade de repetição (microsatélites)
- Inserções de elementos transponíveis

## 4.2. DETECÇÃO DE POLIMORFISMOS MOLECULARES

As técnicas para detecção dos polimorfismos protéicos, como a eletroforese de proteínas e “*Western blot*”, serão discutidas em aulas posteriores. Mas vale salientar que a análise de proteínas pode detectar:

- variação de tamanho, carga ou imunogenicidade.
- variação na quantidade da proteína produzida (região reguladora).

Mas não detecta:

- variações por mutações silenciosas.
- aquelas substituições que não alteram a carga elétrica da proteína.
- que ocorrem nas regiões não codificadoras.

Essas variações indetectáveis quando se estuda proteínas podem ser reveladas pela análise do DNA, cujas técnicas principais serão descritas adiante.

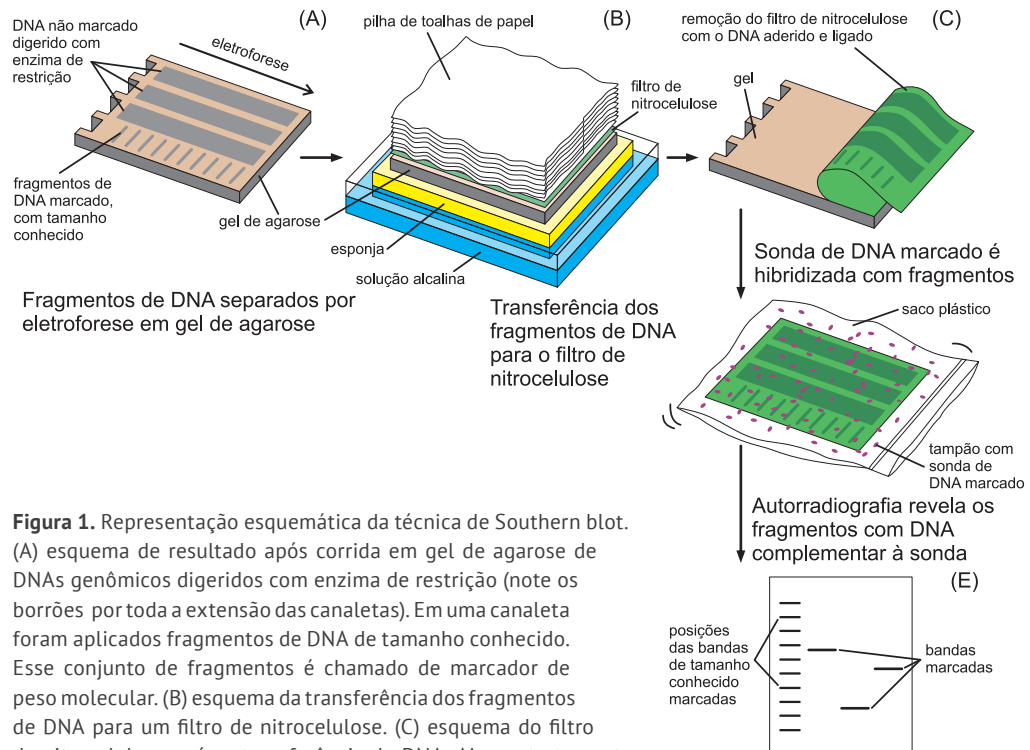
#### 4.2.1. TÉCNICAS BÁSICAS PARA O ESTUDO DE DNA

As técnicas de análise de DNA muitas vezes se baseiam na capacidade de pareamento entre sequências complementares. Além disso, uma ferramenta bastante importante na análise e manipulação de moléculas de DNA são as **enzimas de restrição**, que são endonucleases bastante específicas e já foram discutidas em Bases Moleculares II (aula 7).

Entretanto, é importante lembrar que as moléculas de DNA dos organismos são muito longas. Quando um DNA é digerido com uma determinada enzima de restrição cada molécula gerará certo número de fragmentos, de tamanhos variados, porque os sítios de restrição não se distribuem homoganeamente nas moléculas de DNA. Esses fragmentos podem ser separados por eletroforese, em geral em **gel de agarose** (um gelificante inerte), que forma poros maiores que os encontrados em um gel de poliacrilamida (eletroforese em gel de poliacrilamida será discutida em aulas posteriores). O gel de poliacrilamida só é usado para DNA quando analisamos fragmentos muito pequenos, com menos que 500 pares de bases. Após a corrida da eletroforese a posição das bandas de DNA (que são formadas por fragmentos de mesmo tamanho) pode ser visualizada pela coloração com corantes fluorescentes que se intercalam na dupla-hélice e iluminação com luz ultravioleta.

Quando digerimos com enzima de restrição o DNA de um genoma grande (como o de uma planta ou o de um mamífero), a quantidade de bandas de tamanhos diferentes e próximos é tão grande que o gel de agarose vai mostrar um borrão por toda a canaleta. Assim, se estamos interessados em um trecho específico de DNA não vamos conseguir distingui-lo dos demais. Para superar esse problema, outra técnica básica na análise de DNA é a **transferência de Southern** (*Southern blotting*), que consiste em visualizar a posição de uma ou mais bandas de DNA específicas pela hibridização com uma sequência de DNA marcada (chamada de sonda).

Ficou difícil de entender? Vamos ver se com os passos da técnica ficará mais claro. Após a corrida em gel de agarose, os fragmentos de DNA (todos) são transferidos para uma membrana de nitrocelulose ou nylon, como esquematizado na figura 1. O filtro contendo o DNA aderido é então tratado para que o DNA se



**Figura 1.** Representação esquemática da técnica de Southern blot. (A) esquema de resultado após corrida em gel de agarose de DNAs genômicos digeridos com enzima de restrição (note os borrões por toda a extensão das canaletas). Em uma canaleta foram aplicados fragmentos de DNA de tamanho conhecido. Esse conjunto de fragmentos é chamado de marcador de peso molecular. (B) esquema da transferência dos fragmentos de DNA para um filtro de nitrocelulose. (C) esquema do filtro de nitrocelulose após a transferência do DNA. Alguns tratamentos são feitos para imobilizar o DNA desnaturado no filtro. (D) esquema da etapa de hibridização. A sonda marcada (e desnaturada) é incubada com o filtro de nitrocelulose. Após a hibridização são feitas lavagens para retirar a sonda que não está especificamente ligada. (E) esquema de resultado do autorradiograma. A presença de bandas marcadas radioativamente sensibiliza o filme de raio X. Modificado de <http://employees.csbsju.edu/hjakubowski/classes/ch331/dna/southernblot.jpg>

desnature (desfaça as pontes de hidrogênio) e é incubado com uma solução que contém o segmento de DNA de interesse, também desnaturado, e marcado radioativamente ou com outro tipo de marcação não radioativa. Esse segmento de DNA é chamado de **sonda**, pois é ele que vai descobrir onde o DNA de interesse está. Para tanto, a incubação é feita em condições que permitem a renaturação, ou seja, a formação de pontes de hidrogênio novamente entre as cadeias complementares. Assim, a sonda irá se hibridizar com os fragmentos de DNA que tiveram sequência complementar à sua e, com isso, revelará onde esses fragmentos estão. Após a hibridização a membrana será lavada para retirar todas as moléculas de sonda que não hibridizaram especificamente, e se procederá ao processo de revelação, que varia conforme a marcação usada na sonda. Se a sonda for marcada radioativamente a revelação será a exposição de um filme de Raio X à membrana.

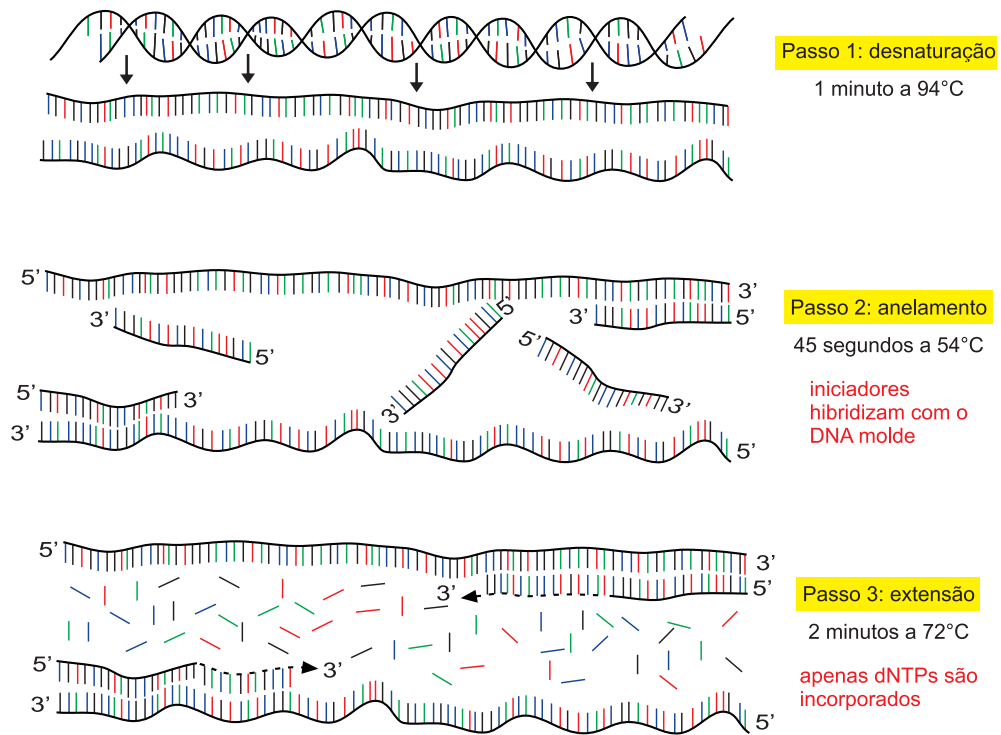
Uma variação de hibridização é usada para avaliar a quantidade de RNAs específicos em um conjunto de células ou tecido. Ela é chamada de **transferência de Northern** (*Northern blotting*) e difere do *Southern*, porque o que é corrido no gel e transferido para a membrana de nitrocelulose é o RNA das células e não o seu DNA digerido com enzima de restrição. A sonda usada continua sendo de DNA, com o qual se trabalha mais facilmente por ser mais estável. A sonda de

DNA marcado irá se hibridizar com as moléculas de RNA que tiverem regiões complementares a ela.

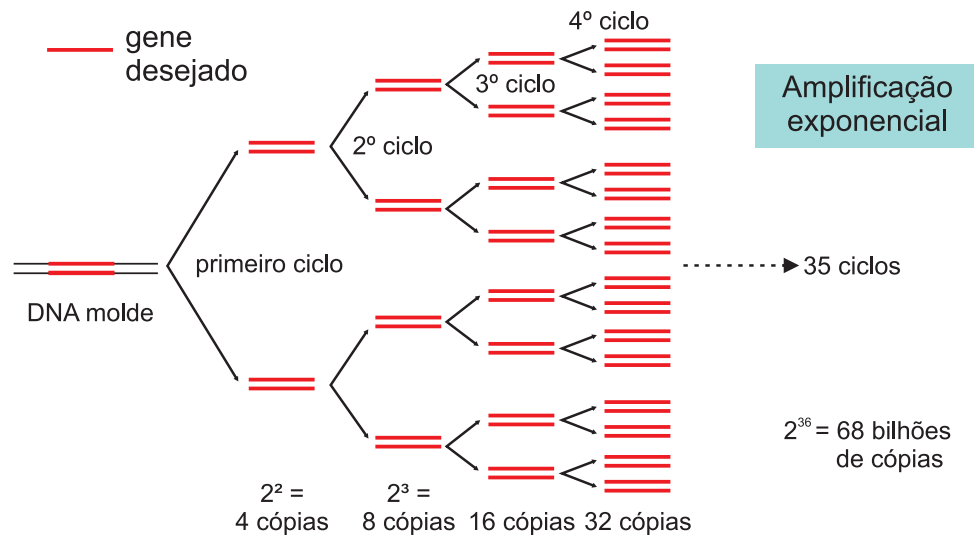
Atualmente, uma das técnicas mais utilizadas é a **PCR**, sigla para **reação em cadeia da polimerase** (*polymerase chain reaction*). Como foi visto rapidamente na aula 1, essa técnica consiste em amplificar uma região específica do DNA, produzindo um número imenso de moléculas desse fragmento. O tamanho do fragmento é determinado pelos iniciadores usados. Esses iniciadores, também chamados pela palavra inglesa *primers*, são oligonucleotídeos sintetizados quimicamente. Os iniciadores são utilizados para que a síntese de DNA seja começada pela DNA polimerase, porque, como vimos anteriormente, nenhuma DNA polimerase descrita até agora pode iniciar a síntese de DNA na ausência de iniciadores (e, recordando aulas anteriores, os iniciadores são produzidos naturalmente pela primase durante a replicação do DNA).

A PCR consiste basicamente de três etapas - desnaturação, anelamento e síntese - que formam um ciclo (veja representação na figura 2), repetido por muitas vezes (em geral de 25 a 35 vezes). A desnaturação é feita pela elevação da temperatura acima de 90°C. Portanto, a polimerase usada na PCR deve ser resistente a altas temperaturas. A seguir, a temperatura é reduzida para permitir o anelamento dos iniciadores nas sequências a que são complementares. No próximo ciclo a temperatura é elevada ao ponto ótimo para a atividade da polimerase, em geral 72°C. Após o período de síntese, o DNA é novamente desnaturado para início de um outro ciclo. Ao final de um ciclo o número de moléculas de DNA molde aumenta, pois as moléculas recém sintetizadas passam a ser usadas como molde também. Assim, a cada ciclo o número de cópias da região de DNA de interesse aumenta em progressão geométrica, como pode ser visualizado na figura 3. O produto final será formado preponderantemente por fragmentos de mesmo tamanho. Isso porque, como toda síntese é feita a partir da região onde se anelam os iniciadores, após os primeiros ciclos a maioria das moléculas de DNA a serem usadas como molde terminará no primeiro nucleotídeo do iniciador que deu início à sua síntese. Assim, as regiões complementares aos iniciadores são as extremidades dos fragmentos amplificados. Para compreender melhor, faça um esquema semelhante ao da figura 2 e prossiga representando os eventos que se seguem com o correr dos ciclos.

25 a 40 ciclos com 3 passos



**Figura 2.** Representação esquemática de PCR. Cada ciclo de síntese de DNA é constituído por 3 passos: desnaturação, anelamento de iniciadores e extensão das moléculas novas. Modificado de <http://users.ugent.be/~avierstr/principles/pcrsteps.gif>



**Figura 3.** Representação esquemática da amplificação do número de cópias de um segmento de DNA por PCR. Modificado de <http://users.ugent.be/~avierstr/principles/pcrcopies.gif>

## 4.2.2. TÉCNICAS PARA ESTUDO DE POLIMORFISMOS DE DNA

As técnicas básicas descritas acima são usadas como base para técnicas para detecção de polimorfismos de DNA. Veremos a descrição das técnicas mais comuns a seguir.

As enzimas de restrição são muito específicas, reconhecendo uma única sequência de nucleotídeos. Se uma mutação altera essa sequência a endonuclease

deixará de reconhecê-la como sítio de clivagem. A detecção dessa mudança de padrão de digestão é a base para uma técnica chamada de RFLP, a saber, polimorfismo de tamanho de fragmento de restrição (*restriction fragment length polymorphism*). A figura 4 esquematiza seu princípio:

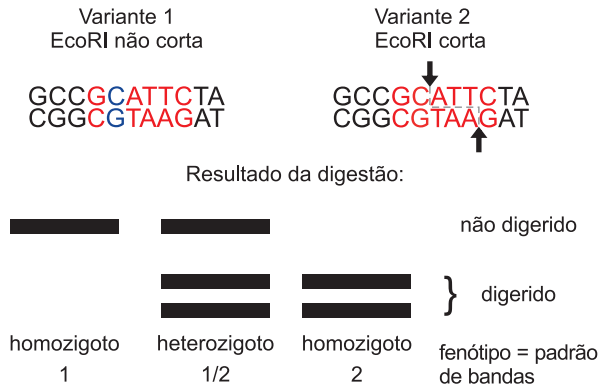


Figura 4. Representação esquemática do princípio do RFLP.

Um alelo (variante 1) não possui o sítio interno para a enzima EcoRI, produzindo apenas uma banda, o outro alelo (variante 2) contém um sítio interno, produzindo duas bandas. O heterozigoto apresentará três bandas. Modificado de <http://www.ucl.ac.uk/~ucbhjow/b241/images/rflp.gif>

### O mesmo segmento de DNA em três pessoas diferentes:

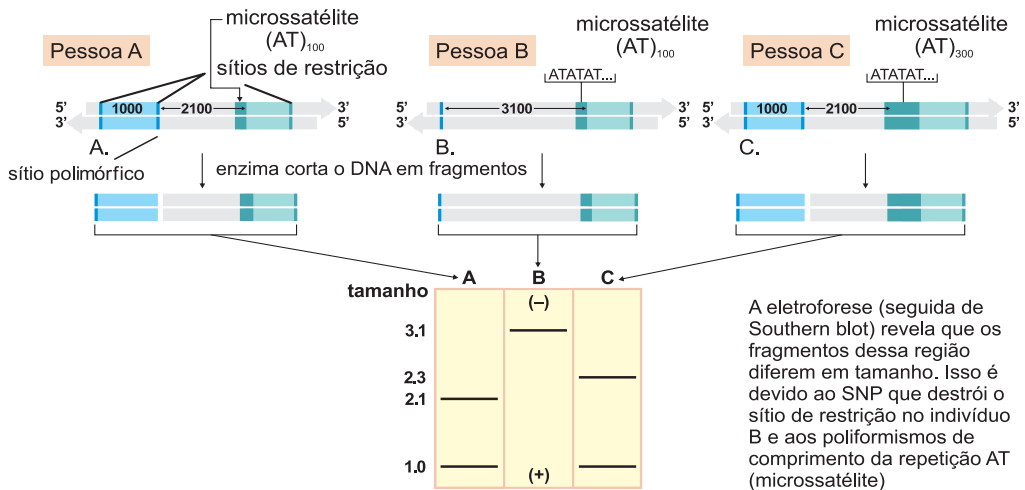


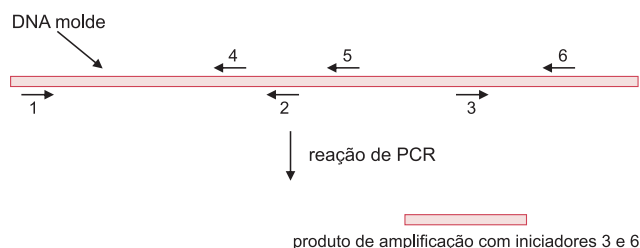
Figura 5. Representação esquemática de polimorfismo de microsatélites (amostras A e C) e RFLP (amostras A e B) e sua detecção por Southern blot. Modificado de <http://campus.queens.edu/faculty/jannr/molecular/DNAtech/RFLP.gif>

Em genomas grandes, o RFLP pode ser visto após *Southern blot*, com sonda que reconheça a região que exhibe o polimorfismo, ou por PCR, com iniciadores que amplifiquem o fragmento que traz o polimorfismo.

A PCR por si só pode ser usada para detectar polimorfismos de quantidade de repetições, como micro ou minissatélites. Nesse caso, os iniciadores utilizados devem ladear a região das repetições. Também *Southern blot* pode ser usado para a detecção de polimorfismos de repetições. Veja na figura 5 como um polimorfismo de microssatélite gerará bandas de tamanho diferente, que nesse exemplo foram detectadas por *Southern blot*. Esquematize agora a detecção desse mesmo polimorfismo usando PCR.

Outro tipo de PCR pode ser usado para detectar polimorfismos, embora sem que saibamos os loci desses polimorfismos no genoma. Essa técnica é a **amplificação aleatória de DNA polimórfico**, abreviada por **RAPD** (*Random Amplification of Polymorphic DNA*). Como o próprio nome sugere as sequências amplificadas não são conhecidas porque são aleatórias. Isso porque iniciadores de 10 pares de bases com sequências arbitrárias, ou seja, pode ser qualquer sequência, são usados. Como o tamanho dos iniciadores é pequeno existe a possibilidade de pareamento com diferentes pontos de um genoma, principalmente se esse genoma for grande, como o de animais ou plantas. A amplificação de um fragmento só pode ocorrer se iniciadores de orientação oposta parearem com sequências relativamente próximas no genoma. Veja a representação na figura 6.

**Figura 6.** Representação esquemática de RAPD. Note que embora muitos iniciadores tenham sido usados apenas foi possível a amplificação com o uso dos iniciadores 3 e 6. Isso devido aos demais terem orientações iguais (anelam na mesma fita) ou hibridizarem em sítios muito distantes entre si.



Modificado de <http://avery.rutgers.edu/WSSP/StudentScholars/project/archives/onions/rapd2.gif>

Os resultados de RFLP ou PCR de múltiplos loci podem resultar em um padrão complexo de bandas, que pode ser chamado de “impressão digital”

(*fingerprinting*) de DNA. Veremos a seguir como a detecção desses padrões de polimorfismos pode ser usada para diferentes fins.

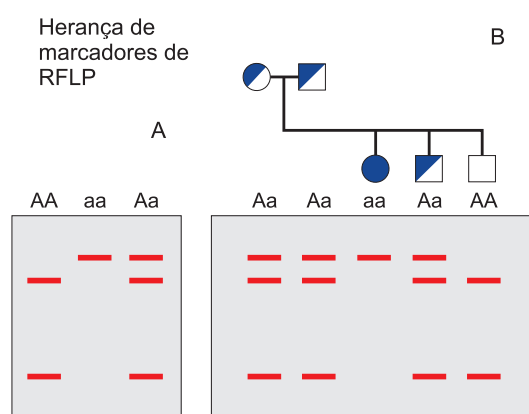
### 4.3. UTILIZAÇÃO DA CARACTERIZAÇÃO DE POLIMORFISMOS GENÉTICOS

Os polimorfismos genéticos, em seu sentido amplo, sempre foram usados para

estudo de transmissão de caracteres. Posteriormente, seu estudo, principalmente de tipos sanguíneos, foi reconhecido como útil na identificação de indivíduos em criminalística e no reconhecimento de paternidade. O desenvolvimento das técnicas de análise de DNA e detecção de um maior número de polimorfismos, aumentou significativamente seu uso. Exemplos de utilização de detecção de polimorfismos de DNA são amplamente divulgados pela televisão para o público em geral, seja em seriados policiais ou em programas sensacionalistas.

Isso porque uma característica chave dos polimorfismos genéticos é sua herança. A figura 7 esquematiza a herança de um RFLP.

**Figura 7.** Representação esquemática da herança de RFLP. Em (A) está representado o padrão de bandas após separação em gel de amostras de indivíduos com cada um dos três genótipos possíveis. Note que



o alelo *A* forma dois fragmentos após a digestão com a enzima de restrição, o que resulta em duas bandas em amostras de indivíduos *AA*. O alelo *a* gera apenas um fragmento nessa região, resultando em uma única banda no gel. Assim o heterocigoto *Aa* apresentará três bandas. Em (B) os fenótipos morfológicos estão representados no heredograma e abaixo estão apresentados os resultados moleculares, sendo que cada indivíduo do heredograma tem sua amostra separada na canaleta imediatamente abaixo de sua representação. Modificado de [http://www.ncbi.nlm.nih.gov/projects/genome/probe/IMG/RFLP\\_genotyping.gif](http://www.ncbi.nlm.nih.gov/projects/genome/probe/IMG/RFLP_genotyping.gif)

### 4.3.1. USO EM MAPEAMENTO GENÉTICO E/OU DIAGNÓSTICO DE DOENÇAS GENÉTICAS

Muitas vezes podem-se detectar polimorfismos moleculares que estão muito ligados, ou seja, com distâncias (genéticas) muito pequenas entre eles e um locus de doença genética. De fato, o mapeamento de genes envolvidos em doenças genéticas atualmente usa, de maneira intensa, a determinação de distâncias genéticas entre o locus de interesse e loci polimórficos do mesmo cromossomo. A figura 8 mostra a ligação entre o alelo mutante responsável por um fenótipo alterado e o alelo 2 de um locus polimórfico com oito alelos (que pode ser um locus de minissatélite). Essa ligação fica clara quando se observa que os filhos afetados herdaram da mãe afetada não apenas o alelo responsável por esse fenótipo, mas também o alelo 2, enquanto os filhos normais herdaram o alelo 7.

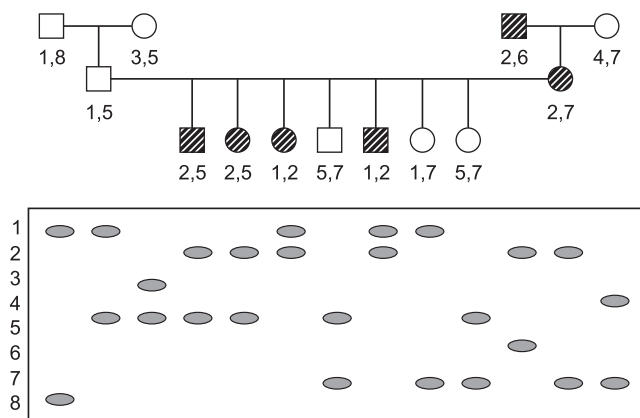


Figura 8. Representação esquemática da ligação de um alelo de fenótipo morfológico (representado por símbolos hachurados) e o alelo 2 de um polimorfismo molecular (por exemplo, de minissatélite) que possui oito alelos, cujas bandas correspondentes estão representadas no esquema abaixo do heredograma. Modificado de [http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/R/RFLPs\\_8.gif](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/R/RFLPs_8.gif)

Algumas vezes a própria mutação que causa a doença cria um RFLP, porque altera a sequência sujeita ao reconhecimento pela enzima. Essa mutação pode destruir um sítio de restrição presente no alelo selvagem ou criar um sítio que anteriormente não existia. Quando isso acontece, a detecção do alelo de RFLP mutante fornece diagnóstico da doença. A figura 9 mostra o exemplo da eliminação de um sítio reconhecido pela enzima *MstII* no gene da beta-globina que é característica do alelo da hemoglobina ciclêmica. Um indivíduo que seja homocigoto para esse alelo apresenta anemia falciforme. Assim, é possível a identificação de heterocigotos não sintomáticos ou de fetos homocigotos por um teste genético após digestão com *MstII* usando sonda específica em *Southern blot* ou PCR com iniciadores específicos.

RFLP na beta-globina associado à hemoglobina ciclêmica (falcêmica)

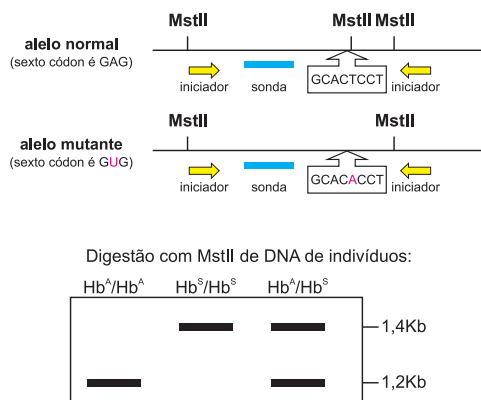


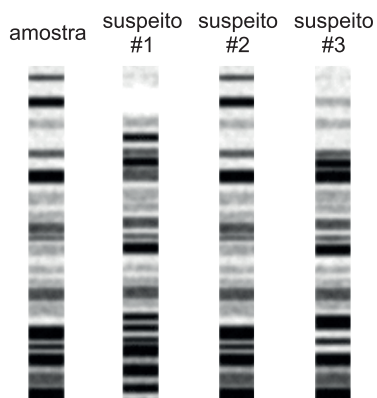
Figura 9. Representação esquemática de parte da região cromossômica do locus da beta-globina, salientando a localização de sítios da enzima *MstII* no alelo normal e no alelo ciclêmico. O esquema sugere a localização de iniciadores (em amarelo) para detecção do RFLP por PCR seguido de digestão e a posição de hibridização da sonda (em azul) se a detecção for feita por *Southern blot*. Na porção inferior estão representados os resultados esperados para indivíduos de diferentes genótipos.

### 4.3.2. USO DE POLIMORFISMOS GENÉTICOS NA IDENTIFICAÇÃO DE INDIVÍDUOS OU DEFINIÇÃO DE PATERNIDADE

Os polimorfismos genéticos podem ser usados na identificação de indivíduos porque podem fornecer padrões complexos se forem detectados alelos de

muitos sítios altamente polimórficos, o que geralmente acontece com micro e minissatélites. A existência de muitos alelos na população ajuda a configurar um indivíduo como único. Esses padrões complexos são chamados de *fingerprinting*

de DNA (seriam “impressões digitais” de DNA). Vejam um exemplo de identificação de indivíduo presente em cena de crime por *fingerprinting* de DNA, com alta probabilidade de certeza, no esquema da figura 10.



**Figura 10.** Representação esquemática de fingerprinting de DNA de amostra colhida em cena de crime e de três indivíduos suspeitos de participação no crime. O padrão de bandas indica que o suspeito de número 2 tenha deixado DNA na cena do crime, implicando em sua presença nesse lugar. Modificado de <http://evolution.berkeley.edu/evolibrary/images/news/dnafingerprints.gif>

Os testes de *fingerprinting* de DNA são também bastante usados para inclusão de paternidade, uma vez que um indivíduo herda alelos da mãe e do pai. Assim, se uma determinada banda presente no *fingerprint* de DNA de um indivíduo não existe nos resultados de sua mãe, foi muito provavelmente herdada de seu pai. Veja o exemplo na figura 11. Você é capaz de definir quem é o pai da criança?



**Figura 11.** Representação esquemática de fingerprinting de DNA de uma mãe, sua criança e dois possíveis pais. Este é um exemplo com um padrão muito simples de bandas, apenas para fins didáticos.

### 4.3.3. OUTROS USOS DE DETECÇÃO DE POLIMORFISMOS GENÉTICOS

Polimorfismos genéticos moleculares são bastante úteis como marcadores para estudos populacionais, pois são em muitos casos codominantes e, assim, os heterozigotos são distinguíveis dos homozigotos. Outra finalidade em que o uso de polimorfismos moleculares tem se acentuado é o estudo de heranças complexas. Nesse caso, marcadores moleculares associados com fenótipos de herança multifatorial são chamados de “loci de características quantitativas”, cuja abreviatura é QTLs (*quantitative trait loci*).

## Autoavaliação

- a. O que são polimorfismos genéticos? Como são classificados?
- b. Quais são as técnicas para detectar polimorfismos de DNA? Em que técnicas básicas essas técnicas se baseiam?
- c. Para que a detecção de polimorfismos de DNA pode ser usada?

## Bibliografia indicada:

Entendendo a Biotecnologia. Borém, A. e Santos, F.R. 2008. Editora da Universidade Federal de Viçosa. Viçosa, MG.

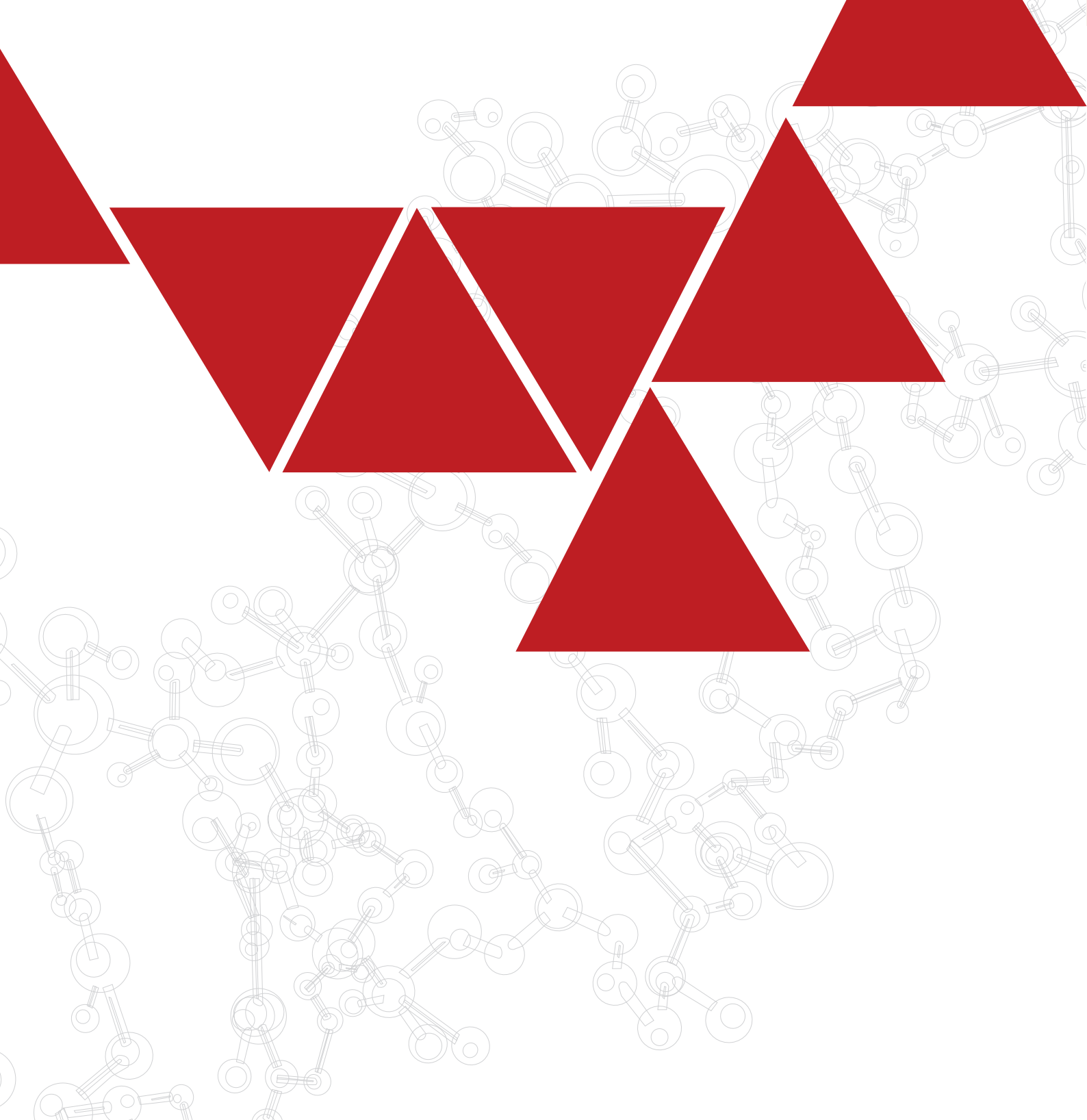
Genética – Um enfoque conceitual. Pierce, B.A. 2004. Ed. Guanabara Koogan, Rio de Janeiro, R.J.

Introdução à Genética. Griffiths, A.J.F; Wessler, S.R.; Lewontin, R.C. e Carroll, S.B.. Nona edição (2009). Ed. Guanabara Koogan, Rio de Janeiro, R.J.

Genes IX. Benjamin Lewis. Nona edição (2009). Ed. Artmed, Porto Alegre, RS.

Fundamentos de Genética. Snustad, D.P. e M.J. Simmons. Quarta edição (2008). E. Guanabara Koogan, Rio de Janeiro, R.J.





5

*Proteoma:  
as proteínas do momento*

## AULA 5: PROTEOMA: AS PROTEÍNAS DO MOMENTO

Ao final desta aula o aluno deverá:

- Entender o conceito de proteoma.
- Reconhecer e exemplificar as modificações pós-traducionais.
- Dominar o vocabulário associado a síntese protéica.

### 5.1. CONCEITO DE PROTEOMA E IMPORTÂNCIA DE SEU ESTUDO

Já sabemos que o conjunto de material genético de uma espécie é seu **genoma** e, portanto, estará presente em todos os indivíduos dessa espécie e em todas as células de cada indivíduo, se esses organismos forem multicelulares. Sabemos também que mesmo organismos unicelulares não expressam todos os seus genes ao mesmo tempo. Existe controle de expressão gênica, sendo bastante importante para bactérias a regulação do início de transcrição, embora a estabilidade dos mRNAs, o início de tradução, a estabilidade e a atividade das proteínas também possam ser reguladas (veja as aulas 8 e 9 de Bases Moleculares II). Em organismos eucariotos a transcrição não fornece diretamente o mRNA, sendo que os RNAs transcritos sofrem um processo de maturação antes de serem considerados mRNAs e exportados para o citoplasma, onde ocorre a tradução. O processamento para a produção de mRNAs também pode ser regulado e, especialmente, a retirada de íntrons pode variar, em um processo chamado de splicing alternativo. Se pensarmos em eucariotos multicelulares complexos, nos quais podemos encontrar células diferentes em um mesmo tecido, tecidos diferentes formando órgãos e sistemas compostos por diferentes órgãos, fica evidente a importância da produção de proteínas diferentes em células diferentes. Além disso, a função protéica pode ser dependente de modificações que ocorrem após a formação do polipeptídeo ou ser regulada por essas modificações.

Assim, ao contrário do genoma, o conjunto de proteínas de uma célula não é estático e pode se modificar dependendo das condições e estímulos, como modificações do meio ambiente ao qual o organismo está exposto e ação de drogas ou infecções. Para o conjunto de proteínas foi criada a denominação “proteoma”. A palavra proteoma e a sua derivada proteômica não se limita ao estudo das proteínas propriamente ditas, mas também a composição, estrutura, função e interação das proteínas responsáveis por toda a atividade celular, as mudanças que elas sofrem e os efeitos que elas têm dentro do organismo. Lembrando que apesar da complexidade de suas funções, as proteínas são relativamente simples: repetições de 20 unidades básicas, os aminoácidos. Em

função do número de genes humanos, 30 mil conhecidos, um gene é capaz de expressar uma ou mais proteínas, cada uma com uma função específica, uma célula pode expressar cerca de 20 mil diferentes proteínas, o que seriam cerca de um milhão de proteínas para serem estudadas apenas para a espécie humana. Isso torna complexo o seu estudo, mais complexo do que o estudo de ácidos nucléicos. Considerando que as proteínas como vimos, sofrem proteólise, se desestruturam e sofrem rearranjos diversos, uma proteína pode estar envolvida em mais de um processo, e funções similares podem ser executadas por diferentes proteínas, temos que concluir que o proteoma apresenta uma maior complexidade comparada ao genoma e, de maneira geral, pode-se afirmar que um genoma pode apresentar muitos proteomas.

Então, os desafios são enormes. A proteômica avança motivada por entender as vias metabólicas em diferentes situações celulares, o que significa:

- Identificar as proteínas que participam de um processo metabólico.
- Identificar possíveis sítios relacionados à ação de fármacos já conhecidos e de novos fármacos que atendam melhor determinadas respostas celulares.

Acompanhar o desenvolvimento de uma doença, desde a identificação, diagnóstico, evolução e tratamento para produção de fármacos alvo-específicos;

## 5.2. MODIFICAÇÕES PÓS-TRADUCIONAIS

Como vimos, a tradução da sequência de um mRNA em uma sequência de aminoácidos não é o final do processo de formação de uma proteína. A proteína funcional deve apresentar a conformação tridimensional adquirida através de interações não covalentes, modificações, ligações com co-fatores e associação a outras cadeias de aminoácidos. Podem sofrer também modificações mais drásticas, representadas por ligações covalentes exemplificadas por glicosilação, fosforilação e acetilação. Essas modificações são chamadas de modificações pós-traducionais, que resultam numa molécula com a menor energia livre associada a uma conformação e onde a informação necessária para ser produzida está contida na própria sequência de aminoácidos.

São exemplos de modificações que ocorrem na proteína pós traducionalmente: ligações dissulfeto; modificações individuais em aminoácidos; ligações com carboidratos; adição de grupos prostéticos; processamento proteolítico; perda de sequência sinal; adição de grupos isoprênicos, ubiquitinação, sulfatação ou acetilação.

Vejam o exemplo da transformação da pré-pró-insulina em insulina representada na figura 1, que mostra a modificação com redução no tamanho da proteína para torná-la um peptídeo funcional.

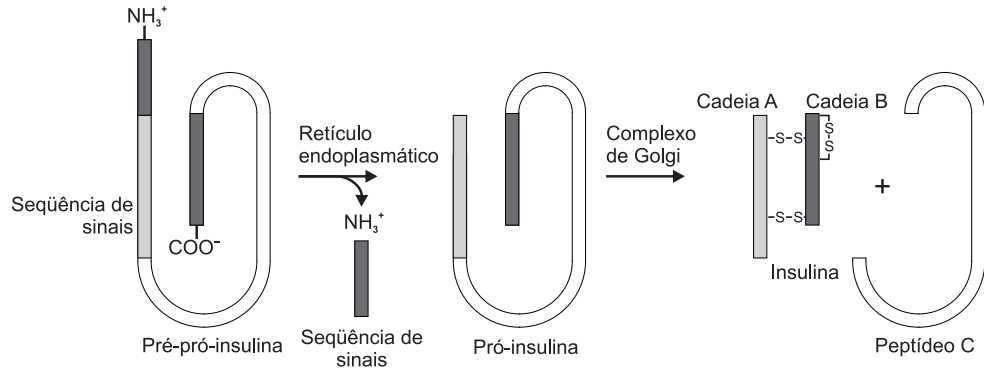


Figura 1: Transformações que produzem a insulina. O gene da insulina codifica a pré-pró-insulina. O peptídeo sinal é degradado durante a síntese da pró-insulina. Esta última é dobrada em uma configuração que permite que as ligações dissulfídicas entre as cadeias A e B se formem. Na sequência e em outro espaço celular, enzimas clivam a cadeia C, conhecida como peptídeo C, completando a síntese de insulina.

A modificação na estrutura covalente pode ser exemplificada pela fosforilação, ou seja, a ligação de grupamentos fosfato, provenientes do ATP, a resíduos de serina, treonina ou tirosina, que são aminoácidos que possuem o grupamento hidroxila. A ligação do grupamento fosfato à cadeia lateral do aminoácido tem natureza covalente e pode afetar a proteína por dois motivos: 1) a presença de cargas negativas altera as interações com novos ligantes, seja por repulsão ou atração e 2) este grupamento pode ser reconhecido por outros sítios de ligação e participar de um processo de junção de proteínas a peptídeos. A fosforilação é uma etapa importante de regulação (figura 2), dos processos relacionados ao crescimento e ciclo celular e ocorre após uma série de ativações denominadas **cascata de fosforilação** e em resposta a sinais que especificam alguma mudança no estado da célula. A fosforilação de resíduos específicos da tirosina de algumas proteínas tem sido associada a etapas de transformação de células normais em células cancerosas.

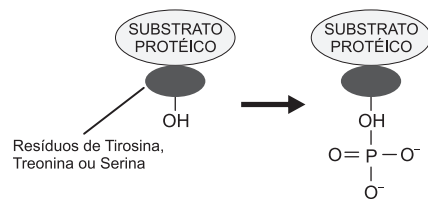


Figura 2: Esquema simples da fosforilação que ocorre na presença de ATP e de enzimas específicas em resíduos de tirosina, treonina ou serina. A fosforilação é um processo reversível promovido pelas quinases, enzimas que adicionam o grupamento fosfato nos resíduos e pelas fosfatases, que removem o grupamento.

A glicosilação (figura 3) é um processo de adição de cadeias curtas de açúcares (oligossacarídeos) em aminoácidos específicos, asparagina, serina ou treonina, presentes na cadeia protéica em posições definidas. É muito comum em proteínas que formam os componentes estruturais de membranas de células eucarióticas e em algumas proteínas secretadas. Essa modificação ocorre no retículo endoplasmático e no Golgi. A glicosilação está associada aos processos de reconhecimento celular, respostas imunes e a adesão. Os oligossacarídeos são ligados covalentemente às proteínas por ligações N-glicosídicas ou O-glicosídicas. No primeiro caso, a adição de um oligossacarídeo inteiro contendo 14 resíduos de açúcar é feita na cadeia aminolateral do aminoácido,

asparagina (Glicosilação N-ligada de proteína) na sequência Asn-X-Ser/Thr, onde X é qualquer aminoácido com exceção da prolina e do ácido aspártico. No segundo caso, e menos frequente, os oligossacarídeos são ligados ao grupamento hidroxila na cadeia lateral dos aminoácidos serina, treonina ou hidroxilisina.

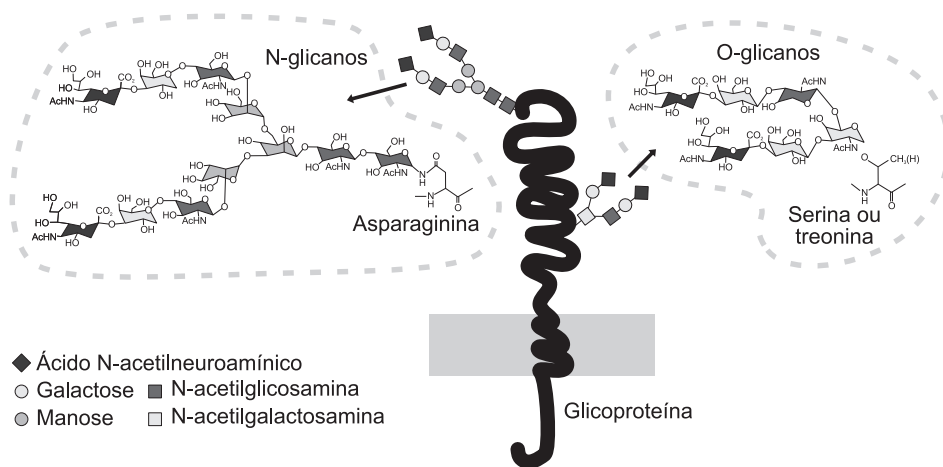


Figura 3: Modificação na cadeia lateral por adição de carboidratos, mostrando as cadeias adicionadas em detalhe.

Outra modificação importante é dada pela formação de ligações covalente do tipo pontes dissulfeto, unindo duas partes de uma mesma proteína ou entre proteínas diferentes. Essas ligações ajudam a proteger a conformação ativa e também dá estabilidade à proteína. A presença destas ligações cruzadas indica que essas proteínas devem ser exportadas das células eucarióticas.

Um sinal de que a proteína deve ser degradada, relacionada portanto com a “morte” da proteína, é dada pela ubiquitinação. Essa modificação consiste da ligação covalente de múltiplas moléculas de ubiquitina, uma proteína de 8 kDa, e que tem uma estrutura basicamente idêntica nas leveduras e nos seres humanos (um exemplo dos padrões repetitivos encontrados nos seres vivos apesar da aparente diversidade). Essa reação depende de ATP e permite as proteínas serem reconhecidas e degradadas em processos relacionados a diversos transtornos catabólicos como traumas, acidoses metabólicas, queimaduras, jejum e também na patogenia de doenças como atrofia e distrofia muscular, câncer, mal de Alzheimer.

A complexidade de um organismo é proporcional às modificações que acontecem na estrutura da proteína após a tradução e determinam entre outros fatores a regulação dos processos biológicos do organismo. As proteínas responsáveis pela fosforilação e defosforilação, as quinases e fosfatases, representam cerca de 2% dos genes tal a importância destes processos na função protéica.

## Autoavaliação

1. Listar os vocábulos próprios desta aula e dar o significado de cada um.
2. Fazer um fluxograma da informação entre DNA e proteína, assinalando as possibilidades de regulação.
3. A obtenção da insulina é um exemplo de modificações pós-traducional. Por que este evento acontece?
4. Exemplifique outras modificações pós-traducionais.

### Para você lembrar de outra forma, acesse os vídeos:

Transcrição - Síntese Protéica

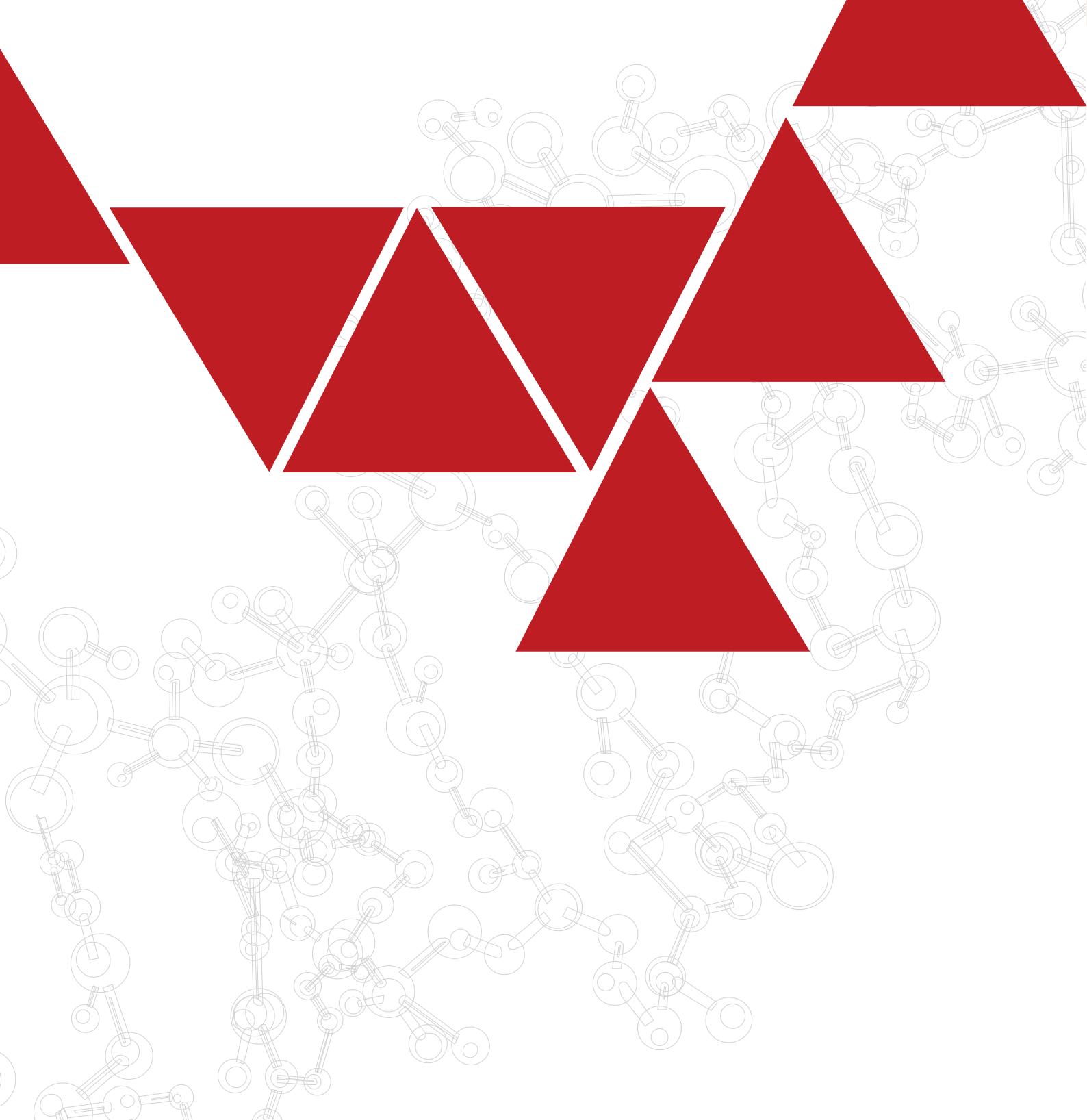
<http://www.youtube.com/watch?v=nNjbYfhfglo>

Transcription: Advanced Look

<http://vcell.ndsu.edu/animations/transcription/dna-tfid.htm>

Tradução - Síntese Protéica

<http://www.youtube.com/watch?v=DcCnmPeutP4o>



# 6

## *Como estudar o proteoma: Cromatografia*

## AULA 6: COMO ESTUDAR O PROTEOMA: CROMATOGRAFIA

Ao final desta aula o aluno deverá:

- Dominar o vocabulário relacionado à cromatografia;
- Diferenciar os tipos de cromatografia por adsorção e por partição;
- Diferenciar os princípios de separação das cromatografias em papel, em coluna, em camada fina e gás-líquido;
- Definir e ilustrar os conceitos e parâmetros cromatográficos.

### 6.1. O QUE É A CROMATOGRAFIA?

É uma das técnicas mais antigas de separação e identificação de componentes de um mistura, inicialmente utilizada para separação dos pigmentos encontrados em vegetais. Uma gota de extrato de vegetal, colocada sobre um papel e este tendo uma das suas extremidades em contato com um solvente, deixava um rastro, uma marca colorida no papel, enquanto o solvente arrastava esta gota até a outra extremidade. Os pigmentos marcavam o papel, semelhante a uma “escrita”, o que deu origem à palavra cromatografia; ou seja, a escrita em cores derivadas dos diferentes pigmentos vegetais.

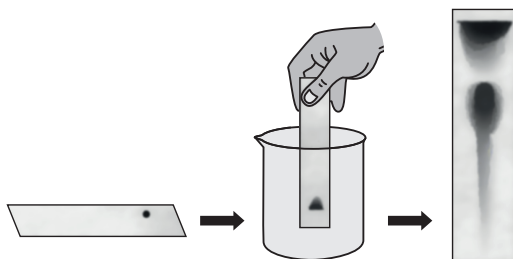


Figura 1: Cromatografia em papel

O papel usado como suporte na cromatografia inicial, nada mais é do que uma matriz de celulose, unidimensional, não totalmente compactada. A partir dessa ideia original foi possível desenvolver colunas de

celulose, bidimensionais, nas quais a matriz de celulose pode ser manuseada de tal forma a controlar os poros. Você pode verificar estes poros se colocar uma folha de papel sob a luz e perceber porções mais claras e mais escuras, proporcionais a quantidade de celulose ali presente. A coluna de cromatografia é um aperfeiçoamento da técnica, na qual a celulose, ou outro material, é colocado em um tubo de vidro, a coluna. É um dos métodos mais comuns de purificação de proteínas.

São utilizadas diferentes substâncias para preencher a coluna, denominada de fase estacionária. Essas substâncias podem apresentar a consistência de gel, a exemplo da acrilamida, também utilizada na eletroforese; pode ser a celulose modificada, ou mesmo resina de diferentes naturezas e composições, como a agarose, que é um polímero de galactose. O processo cromatográfico

de separação das proteínas presentes na fase móvel, a amostra contendo as proteínas, será de acordo com as propriedades do material da fase estacionária. A corrida cromatográfica consiste na passagem da proteína através de uma coluna (fase estacionária), que é desenhada para reter ou diminuir a velocidade da passagem do solvente, também chamado de eluente ou fase móvel, que irá arrastar a amostra onde estão as macromoléculas (proteínas). A corrida cromatográfica obedece a uma ou mais propriedades, como o tamanho, a massa molecular, a polaridade, a hidrofobicidade, a carga ou a afinidade química entre as fases. O processo de separação recebe a sua denominação de acordo com a propriedade: se separadas de acordo com a carga, cromatografia de troca iônica, por hidrofobicidade, cromatografia de interação hidrofóbica, por tamanho, filtração em gel ou pela sua capacidade de se ligar a grupos químicos particulares, afinidade química ou especificidade. Resumindo, a separação ocorre devido às diferentes distribuições entre as fases, sendo duas as forças atuantes: a força impulsora, dada pela fase móvel, e a força retardatória, dada pela fase estacionária.

A amostra contendo a mistura de proteínas é dissolvida em solução aquosa colocada em contato com a fase estacionária e eluída através deste suporte com uma solução aquosa (fase móvel).

## 6.2. TIPOS DE CROMATOGRAFIA

Há vários tipos de cromatografia se classificadas a partir dos processos de Adsorção e Partição.

### 6.2.1. Partição

A cromatografia de partição tem como princípio que se duas fases estão em contato uma com a outra e se uma ou ambas as fases contem um soluto, o soluto se distribuirá entre as duas fases. O coeficiente de partição nos dá a proporção na qual o soluto se distribui:

O coeficiente de partição de 0,8 indica que 20% do soluto se encontram na fase líquida, livre, e 80% na fase sólida, portanto ainda retido.

Como exemplos de cromatografia de partição podemos citar a cromatografia de papel, de camada fina, gás-líquida e de filtração em gel.

$$\text{Coeficiente de Partição} = \frac{\text{Quantidade de soluto na Fase 1}}{\text{Quantidade de soluto na Fase 2}}$$

### 6.2.2. Adsorção

Na cromatografia de adsorção, a separação ocorre devido às interações eletrostáticas e forças de Van Der Waals entre a fase estacionária (sólido) e os componentes a separar da fase móvel (líquido ou gás). A natureza da fase

estacionária pode ser muito diversa, sendo as mais comuns a sílica gel, a alumina ou a celulose.

A adsorção é a fixação de moléculas de uma substância, o adsorvato, na superfície de outra, o adsorvente. Uma maneira simples de explicar este processo é a escrita em papel: a tinta (adsorvato) é adsorvida pelo papel (adsorvente) onde o número de moléculas adsorvidas depende da concentração. Se uma dada concentração de moléculas é aplicada a uma superfície ou suporte e deixarmos fluir um solvente, uma quantidade fixa é ligada e o remanescente moverá ao longo da coluna. O material que avança será retardado por sua adsorção ao material da coluna. A fração ligada não é constante, mas diminui com a diminuição da concentração.

A cromatografia de adsorção combinada com substâncias específicas dá origem a outros tipos de cromatografia: troca iônica, afinidade e imunoafinidade.

A cromatografia convencional ou de fase normal é aquela na qual a fase estacionária é mais polar que a fase móvel (solvente). Se a polaridade da fase estacionária for menor do que a polaridade do solvente, a cromatografia recebe a denominação de cromatografia de fase reversa.

### 6.3. PROCEDIMENTOS EXPERIMENTAIS

A cromatografia pode ser aplicada experimentalmente de quatro maneiras diferentes: em papel; em coluna; em camada fina e gás-líquido.

#### 6.3.1. Cromatografia em papel

Neste procedimento, a celulose em forma de tiras de papel é utilizada como suporte. A separação ocorre entre a água ligada ao suporte e um solvente escolhido. A posição relativa no papel após a movimentação do solvente, denominada de front é utilizada como método para identificação dos componentes da amostra.

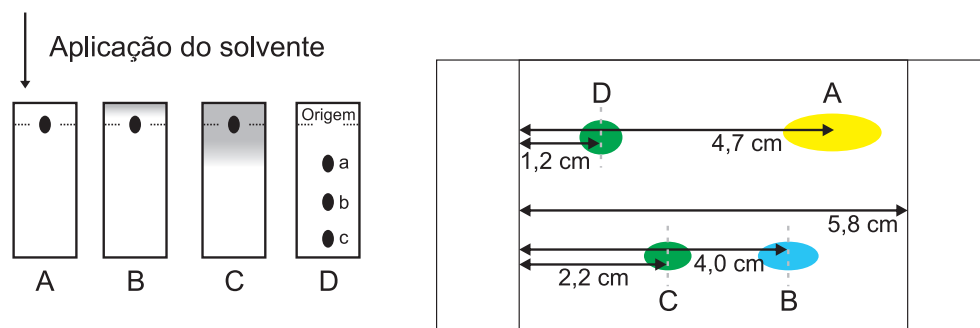


Figura 2: A direção do solvente, de cima para baixo, arrasta a amostra. A eluição identificou três componentes com fronts diferentes - a, b, c - em relação à distância total, percorrida pelo solvente. A comparação entre estas medidas permite identificar os componentes da amostra. Essa medida é conhecida como fator de

retenção ou  $R_f$  ( $R_f = \text{distância percorrida pela amostra} / \text{distância máxima que poderia percorrer}$ ). Ao lado, um Cromatograma com indicação do fator de retenção para uma amostra com quatro componentes. O valor de **A** (5,8 cm) representa o valor máximo que poderia ter sido percorrido. Os demais valores indicam a distância percorrida,  $R_f$ , para os quatro componentes da amostra.

É possível obter uma corrida bidimensional se submetermos a cromatografia com um solvente numa direção e, a seguir, recromatografar em direção perpendicular utilizando um segundo solvente. Componentes que falham na separação em um solvente podem ser separados com o segundo solvente.

As características químicas das cadeias laterais dos aminoácidos, polar ou apolar, ácidas, básicas ou neutras, podem ser úteis no processo de separação de uma mistura complexa de aminoácidos, possibilitando identificar cada um deles, o que pode ser conseguido pela cromatografia em papel. Os aminoácidos precisam ser corados para a identificação, uma vez que eles não podem ser vistos pelo olho humano. O corante ninhidrina reage com grupamentos amino livres produzindo um composto de cor púrpura. A reação da ninhidrina com prolina gera um composto amarelo e já a reação com a tirosina produz uma coloração azul metálica.

A detecção poderá ser feita utilizando reações químicas, como ninidrina, descrito anteriormente, que identifica aminoácidos pela cor, por fluorescência ou radioatividade. É necessário aplicar um padrão com distância de eluição conhecida ( $R_f$ ) para comparação com a amostra.

### **6.3.2. Cromatografia em camada**

É utilizada para a separação de pequenas moléculas e utiliza a superfície de plástico ou de vidro para espalhar, em camadas finas de 0,5mm, a fase estacionária; daí o nome de cromatografia em camada. A fase estacionária pode ser a sílica, óxido de alumínio, silicatos, poliamida, polietileno, entre outras. É mais sensível, tem fácil detecção, maior poder de resolução e maior velocidade de separação do que a cromatografia em papel.

### **6.3.3. Cromatografia em coluna**

Nesta cromatografia, uma coluna de vidro é utilizada como recipiente para o material da fase estacionária. O solvente flui ao longo da coluna e este processo é denominado de eluição. O material que sai da coluna quando um determinado volume de solvente já tiver passado é denominado de eluente e será coletado em pequenas frações, ou alíquotas, que serão identificadas por diferentes métodos. A eluição pode ocorrer utilizando um ou mais solventes, se aplicarmos um solvente e a seguir aplicarmos outro, este processo é denominado de *step-wise*. Quando alteramos a proporção entre dois ou mais solventes, aumentando a concentração de um componente (por exemplo a força iônica, sal ou pH), este processo é denominado de eluição com gradiente.

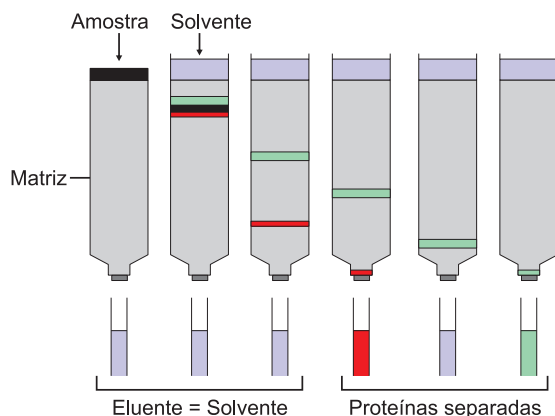


Figura 3: Separação em cromatografia em coluna.

### 6.3.4. Cromatografia de filtração em gel, exclusão molecular ou peneira molecular

A cromatografia de filtração em gel separa as proteínas com base no seu tamanho. A matriz é um composto inerte, que contém pequenas esferas porosas empacotadas. Se uma

amostra contendo moléculas de vários tamanhos for aplicada, as moléculas maiores que os poros moverão somente nos espaços entre as esferas e por este motivo não serão retardadas pelo material da matriz. As moléculas menores que os poros, difundem-se para dentro e para fora dos poros e serão assim retardadas no seu movimento ao longo da coluna. A probabilidade de penetração é o principal fator determinante na velocidade. As moléculas serão eluídas em ordem decrescente de tamanho ou, se a forma for constante, decrescente na massa molecular.

Ao fazer passar a solução de proteínas, por exemplo, pela matriz da coluna, as moléculas pequenas entram nos poros das esferas demorando a atravessá-las, enquanto que as grandes passam entre as esferas sendo separadas primeiro. As dimensões dos poros da matriz determinam a faixa de massa molecular da amostra que pode ser separada na cromatografia.

A matriz é uma rede molecular tridimensional, com polímeros inertes, como exemplo a dextrana, a agarose e a poliacrilamida, encontradas comercialmente com poros de tamanhos variados, o que permite determinar a massa molecular de proteínas e purificar amostras que contenham componentes com massas moleculares diferentes.

### 6.3.5. Cromatografia de troca iônica

O trocador de íons é um sólido que tem quimicamente ligado a ele grupos carregados aos quais íons estão eletrostaticamente associados. Por competição, podem trocar estes íons por íons da solução aquosa. A separação é feita de acordo com a densidade e distribuição da carga. A separação se dá em duas etapas: as substâncias a serem separadas são ligadas ao trocador e a seguir a coluna é eluída com tampões de diferentes pH, com composição e força iônica diferentes. Os componentes do tampão competirão pelos sítios de ligação na matriz. Esta matriz é também uma rede tridimensional que contém grupos carregados ligados covalentemente. Se o grupo é carregado negativamente, ele poderá trocar íons positivos e por este motivo esta coluna é uma trocadora catiônica, possuindo, por exemplo, íons sulfatos imobilizados na fase estacionária. Se o grupo for carregado positivamente, íons amônio, por exemplo, irá trocar anions

e, portanto, será um trocador aniônico.

Depois de fazer passar a solução pela coluna recorre-se ao procedimento *salting out* para separar a amostra, com carga positiva, por exemplo, da coluna com carga negativa. A técnica do *salting out* utiliza uma solução salina de elevada concentração a qual compete com a amostra na ligação à coluna. Como a matriz tem maior afinidade para a carga dos sais do que para a das proteínas presentes na amostra, estas acabam por ser liberadas, mantendo-se os sais ligados à coluna. As proteínas com fracas interações iônicas serão liberadas com uma concentração baixa de sal. A adição de sal deverá ser feita de uma forma gradual e, no final, para ter a certeza de que todas as proteínas foram liberadas da coluna, deve ser colocado nesta uma solução de concentração extrema de sal ( 2-3mol/L).

Cuidados devem ser tomados quando se usa proteínas como amostra: as mudanças no pH alteram a carga das proteínas, por isso é necessário conhecer o seu ponto isoelétrico (pH na qual a carga da proteína é 0) e certificar-se de que o pH do sistema está convenientemente ajustado e tamponado. Ao final da eluição, o sal deverá ser retirado das proteínas eluídas.

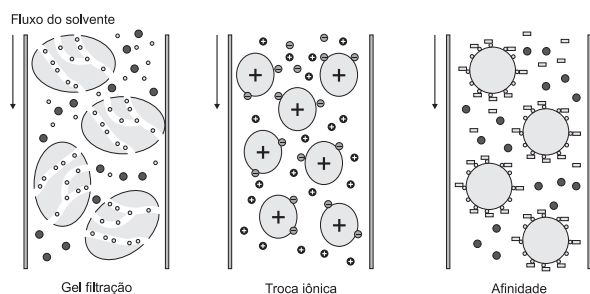
### 6.3.6. Cromatografia por afinidade ou interação bio-específica

A afinidade específica é usada entre uma substância a ser isolada e uma outra molécula, o ligante, que estará imobilizada.

O ligante deve ser acoplado sem que haja alteração nas suas propriedades e a interação deste ligante com a molécula de interesse deve ser forte, mas deve permitir a eluição da amostra, sem contudo destruí-la ou alterar suas propriedades funcionais. Estas interações são do tipo interações eletrostáticas, efeitos hidrofóbicos, ponte de hidrogênio. O tipo mais comum de cromatografia por afinidade envolve uma pequena biomolécula específica, por exemplo, anticorpo, que é imobilizado e ligado a uma matriz da coluna, como a celulose ou poliacrilamida. A proteína a ser analisada passa através da coluna e fica retida pelo seu ligante, enquanto que outras proteínas serão liberadas. A separação da proteína alvo é normalmente feita passando através da coluna uma solução

que contenha uma elevada concentração de ligante livre.

É um método muito eficiente de purificação, uma vez que se baseia numa especificidade biológica da proteína que se pretende analisar, tal como a afinidade de uma



enzima a um substrato ou a ligação entre antígeno e anticorpo.

Figura 4: Exemplos de cromatografia em coluna utilizando propriedades diferentes da fase estacionária para o processo de separação.

### 6.3.7. Cromatografia gás-líquido

Nesta modalidade a fase móvel é um gás, como o hélio, argônio ou nitrogênio. A fase estacionária é um líquido adsorvido numa superfície de uma coluna ou em um suporte sólido, como o teflon. A amostra deve ser passível de volatilizar para poder ser aplicada. A amostra líquida é adicionada ao gás, submetida a aquecimentos e aplicada na coluna, que tem como característica ser extremamente longa, podendo atingir 20 metros e um pequeno (0,5 cm) diâmetro.

A amostra sofre então contínua distribuição entre a fase aquosa e a fase líquida, de acordo com a afinidade entre estas fases, o que chamamos de coeficiente de partição. Após a separação dos componentes da mistura na coluna, e após a sua eluição, estes são conduzidos para o detector onde são identificados e quantificados. Vários são os detectores utilizados: espectrofotômetros que utilizam a luz ultravioleta; detectores de fluorescência, de índice de refração e eletroquímicos, entre outros. Detectores de polarimetria diferenciam compostos quirais, através da rotação de seus estereoisômeros frente à luz plano polarizada.

O sinal transmitido pelo detector é enviado para um integrador, onde são processados os dados e obtido o resultado na forma de um cromatograma. A sensibilidade desta cromatografia permite detectar amostra com  $10^{-12}$  g em tempos bem reduzidos (menos de um minuto).

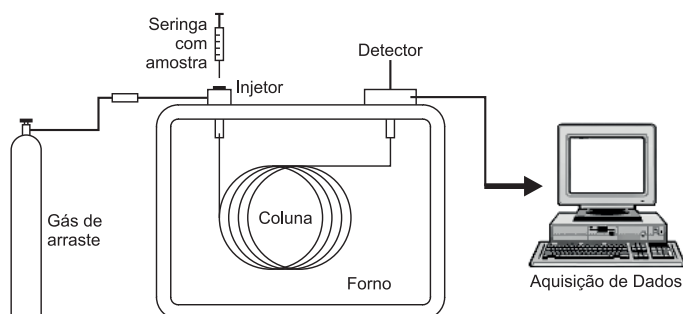
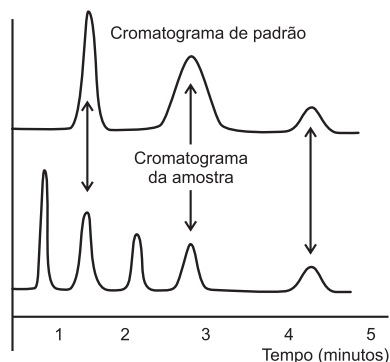


Figura 5: Sistema de cromatografia gás-líquido e o cromatograma obtido para o padrão conhecido e para a amostra.



### 6.4. HPLC

A cromatografia convencional oferece inúmeras vantagens, mas está limitada pela não homogeneidade nas matrizes que causa uma fluidez desigual do solvente através da coluna. A

possibilidade técnica de sintetizar partículas muito pequenas, menores do que  $10 \mu\text{m}$ , o que permitiu o empacotamento denso e uniformidade das partículas deu origem a cromatografia de alta eficiência, o HPLC (*High Performance Liquid*

*Cromatography*). É necessária que uma força mais intensa, gerada por uma bomba, impulse de modo a superar a resistência da coluna ao escoamento da fase móvel. A HPLC é uma cromatografia de alta pressão, também chamada cromatografia de alta precisão ou eficiência. Nesta modalidade, utiliza-se um pequeno volume da amostra, 10  $\mu\text{L}$ , que será submetido a uma alta pressão para poder passar pelo empacotamento denso das partículas da matriz, que é mantida numa coluna de aço inoxidável para suportar a alta pressão utilizada. A coluna apresenta diâmetro de 3 a 6 milímetros, com 10 a 20 centímetros de comprimento e deve suportar pressões de 5 a 20  $\text{kg}/\text{cm}^2$ . A separação é rápida e com excelente resolução.

### 6.5. Outros conceitos e parâmetros cromatográficos

Visto que o objetivo da cromatografia é “resolver”, ou seja, separar os componentes de uma mistura, o cromatograma registra a sucessão de picos obtidos após a eluição. Alguns termos que caracterizam uma análise cromatográfica são mostrados na Figura 6.

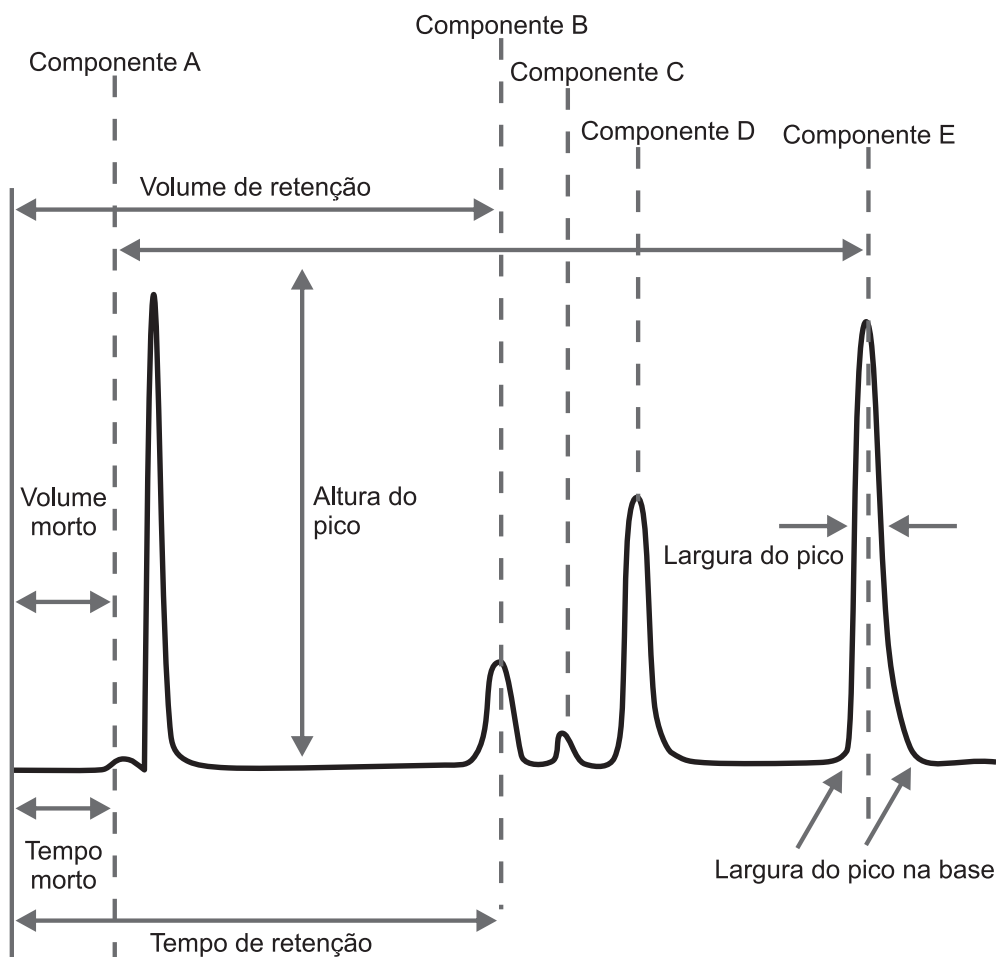
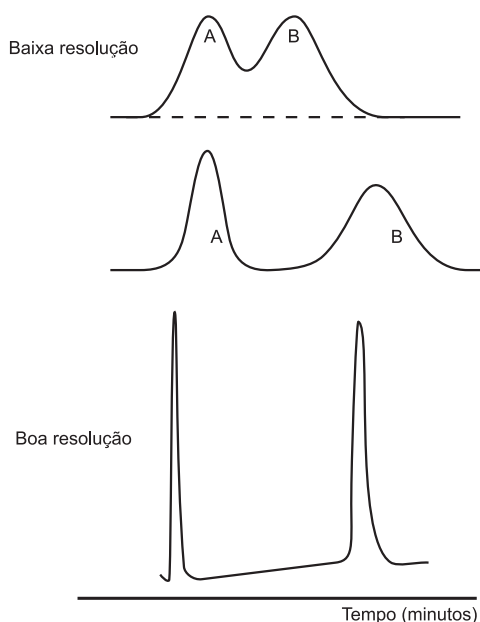


Figura 6: Alguns parâmetros que caracterizam um cromatograma.

O volume total de uma coluna é dado pelo volume ocupado pelo sólido e pelo líquido, ou seja, da fase estacionária e da fase móvel. O volume da fase móvel é também denominado de void,  $V_0$  ou volume de retenção. O volume de eluição é a quantidade de líquido que deve ser adicionado para obter um pico do efluente. As moléculas do soluto transportadas pela fase móvel têm tempos de retenção diferentes de acordo com os respectivos graus de adsorção na fase estacionária. O tempo de retenção ou tempo estequiométrico,  $t_r$ , é o tempo que decorre entre o ponto de injeção até o máximo do pico. Vários são os fatores que interferem no tempo de retenção: a velocidade com que um constituinte migra ao longo da coluna; o coeficiente de partição que, como visto anteriormente, é a relação entre a concentração do soluto na fase estacionária dividido pela concentração do soluto na fase móvel; os espaços vazios da coluna; o

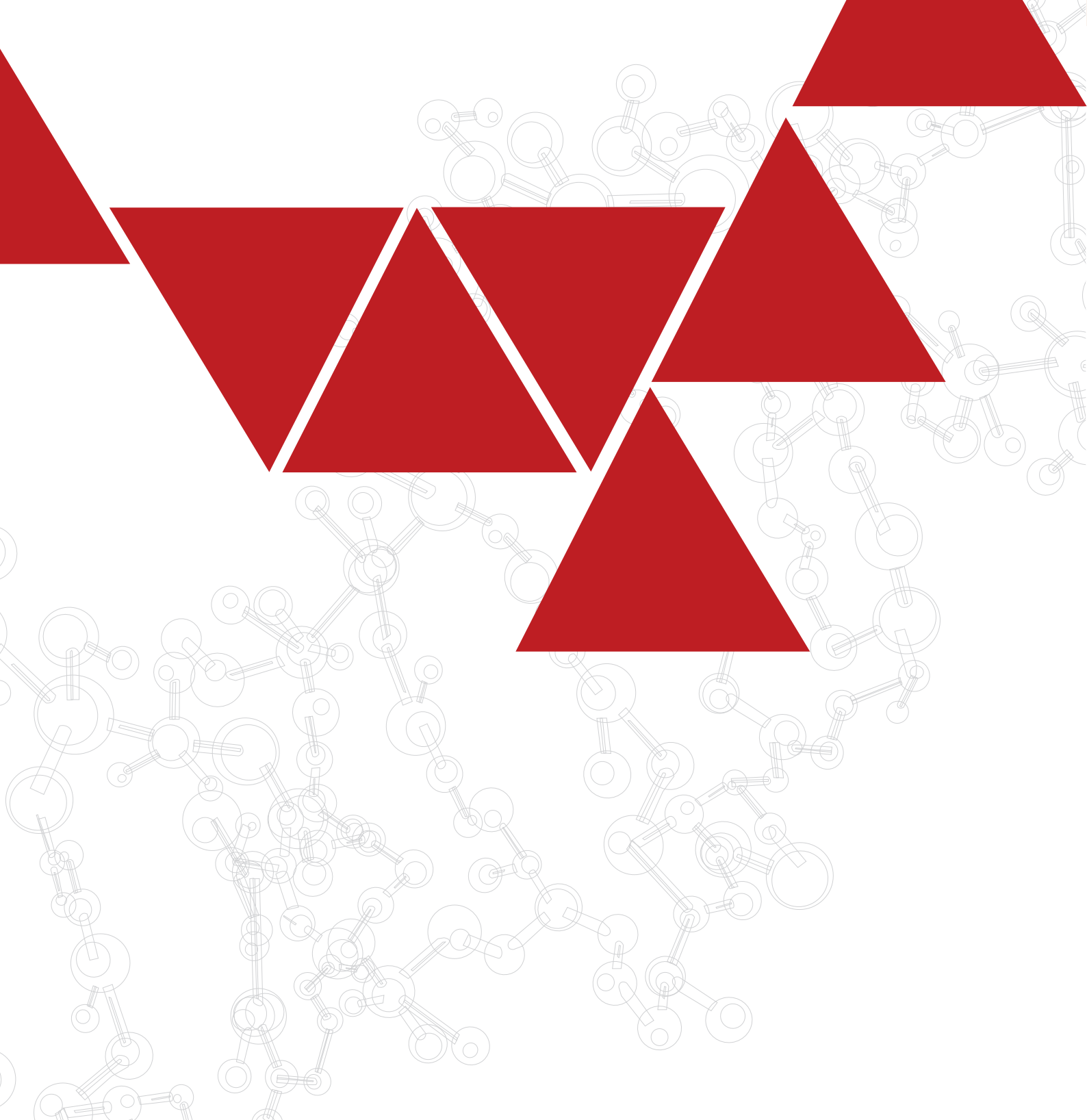


volume de retenção, que é o volume de eluente, dentre outros fatores. A resolução dos picos cromatográficos, ou seja, a separação real entre dois picos depende de dois parâmetros operatórios: a seletividade, que é o fator de separação e a eficiência da coluna. Aumentando a seletividade da coluna, os picos afastam-se até atingir uma distância máxima entre dois picos. Uma melhor eficiência significa maior altura e menor largura do pico. Uma boa resolução implica em picos bem definidos, finos e separados, como ilustrados na figura 7.

Figura 7: A resolução do cromatograma é função de seletividade e eficácia da coluna.

### Autoavaliação

1. Elaborar um glossário com os termos associados à cromatografia.
2. A cromatografia em fase gasosa acoplada a diferentes detectores tem demonstrado ser prática, rápida e eficiente no controle da dopagem no esporte. Pesquise e leia artigos relacionados a este tema.
3. O processo de separação cromatográfico pode ser definido a partir do conceito de pratos teóricos. Pesquise, defina e discuta o processo de separação a partir do conceito de pratos teóricos.
4. Pesquise sobre o conceito de coeficiente de partição e compare com aquele de pratos teóricos.



**7**

*Como estudar o proteoma:  
A eletroforese*

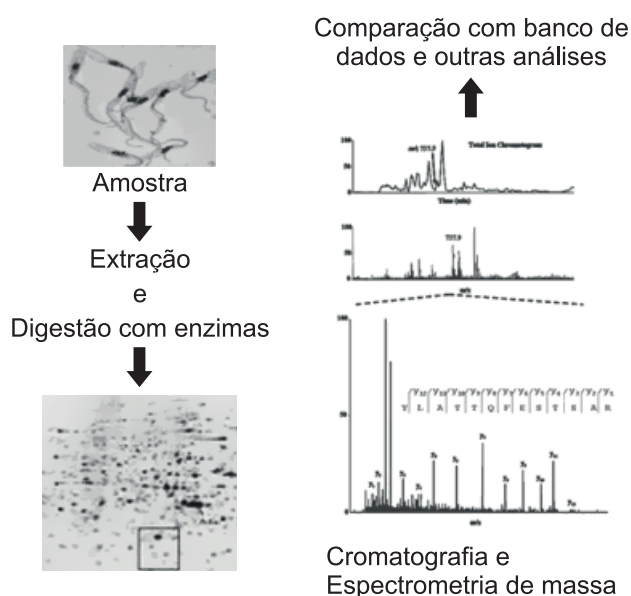
## AULA 7: COMO ESTUDAR O PROTEOMA: A ELETROFORESE

Ao final desta aula o aluno deverá ser capaz de:

- Citar técnicas que podem anteceder a eletroforese.
- Discutir os princípios da eletroforese.
- Escrever as equações de força aplicadas à eletroforese, explicando o significado de cada variável.
- Citar os meios ou suportes mais comuns utilizados na eletroforese, descrevendo suas características.
- Descrever as etapas de uma corrida eletroforética.
- Descrever as etapas da eletroforese bidimensional e o princípio de separação utilizado em cada uma.
- Descrever as etapas da técnica de *Western blotting*.
- Reconhecer as limitações da eletroforese.

### 7.1. INTRODUÇÃO

Como estudar o proteoma? De acordo com as palavras do grupo de pesquisas em proteomas da Fiocruz: “A combinação da eletroforese em gel de poliacrilamida bidimensional com espectrometria de massa é uma das estratégias para o estudo do proteoma. A eletroforese permite a separação de proteínas de um dado sistema biológico com alta resolução e reprodutibilidade, enquanto que a



espectrometria de massa permite a identificação de proteínas e peptídeos. Abordagens que usam cromatografia líquida de alta resolução para a purificação e fracionamento de proteínas e de peptídeos de uma amostra também têm sido desenvolvidas e aplicadas com sucesso.”

Figura 1: Exemplo de estratégias para a análise proteômica. Retirado de <http://www.fiocruz.br/chagas/>.

Vamos então descrever a técnica da eletroforese. Inicialmente é necessário que se diga que há diversos tipos de eletroforese: a eletroforese livre (frente móvel); a eletroforese de zona; a eletroforese em papel; a eletroforese em acetato de celulose; a eletroforese em gel (SDS-PAGE); a focalização isoelétrica e a eletroforese bidimensional.

## 7.2. O princípio da técnica eletroforese

A eletroforese é uma técnica de separação de moléculas carregadas que se baseia em conceitos da física elétrica, na qual partículas carregadas movem-se quando colocadas sobre a influência de um campo elétrico (daí a denominação que tem origem no grego forese, que significa transporte).

É um método poderoso de separação de proteínas e outras macromoléculas. A velocidade de migração depende da força do campo aplicado, do tamanho, da forma e da carga elétrica da molécula em estudo, da viscosidade e temperatura do meio onde ocorre a separação. A força do campo aplicado se opõe à resistência do meio e quando estas forças se igualam as moléculas migram com velocidade constante. A voltagem aplicada na extremidade inferior de um gel gera um campo elétrico com uma intensidade definida pelo comprimento do gel e pela diferença de potencial (V/cm). A equação 1 define a velocidade de migração ( $V$ ) de uma molécula em um campo elétrico, onde  $E$  representa a intensidade de campo,  $Z$  a carga global da molécula e  $f$  é o coeficiente de atrito. Este último depende da massa, da forma da molécula e da viscosidade do meio  $\eta$ . Considerando a molécula como sendo uma esfera de raio  $r$ , o coeficiente de atrito  $f$  pode ser expresso pela equação 2.

$$V = E \frac{Z}{f} \quad \text{Equação 1} \quad \text{e} \quad f = 6\pi \eta r \quad \text{Equação 2}$$

A força elétrica  $E \times Z$  que impulsiona a molécula é equilibrada pela resistência da viscosidade resultante do atrito entre as moléculas em movimento e o meio.

Os meios ou suportes mais comuns utilizados na eletroforese são os géis ao invés de soluções livres, pois os géis suprimem as correntes de convecção produzidas por pequenos gradientes de temperatura, o que é muito importante para que haja uma boa separação entre as moléculas. Usualmente o gel fornece por si mesmo quase toda a resistência no circuito. Para uma dada corrente, a diminuição da malha do gel ou da força iônica da solução aumenta a resistência e, conseqüentemente, aumenta o gradiente de voltagem através do gel e a mobilidade eletroforética da amostra. Se houver resistência e corrente, haverá manifestação da potência na forma de calor e alguns géis podem dissipar somente uma quantidade particular de energia sem aumentar a temperatura do gel. Acima desse ponto, pequenos aumentos na voltagem podem causar significantes e potencialmente desastrosos aumentos na temperatura do gel. É muito importante conhecer quanta potência um gel em particular pode dissipar facilmente e monitorar cuidadosamente a temperatura dos géis durante a corrida eletroforética.

Outra particularidade dos géis é o fato deles atuarem como peneiras moleculares que torna mais evidente a separação. Assim, os géis são os preferidos para atuarem como suportes para a separação de moléculas. Os géis mais usados são os de acrilamida, podendo também ser usados os de agarose. O gel de acrilamida, PAGE, é quimicamente inerte e facilmente formado pela polimerização da acrilamida. Variações nas concentrações de acrilamida e de um reagente que favorece a interligação, que é o etileno- bisacrilamida, permite controlar o tamanho da rede, ou seja, o tamanho dos poros feitos na malha do gel. Poros grandes deixam passar através deles fragmentos maiores. Quanto menor o poro, maior o grau de separação. É possível também preparar géis que apresentam variação do tamanho dos poros ao longo do gel, ou variação de pH ao longo da aplicação, que é denominado de gradiente.

O gel é aplicado em uma superfície de vidro, na forma de uma lâmina fina vertical denominada por placa. As moléculas são separadas de acordo com o seu tamanho, pois as de menor massa irão migrar mais rapidamente que as de maior massa. Em alguns casos, o formato das moléculas também influi, pois algumas terão maior facilidade para migrar pelo gel. As proteínas podem ser separadas principalmente pela massa molecular através da eletroforese em gel de poliacrilamida em condições desnaturantes. Para isso são adicionados alguns reagentes à mistura de proteínas, o que chamamos de amostra. Um destes agentes desnaturantes é o detergente aniônico dodecilsulfato ou lauril sulfato de sódio, SDS, que é utilizado tanto para dissolver a amostra, rompendo quase todas as interações não-covalentes, quanto para conferir aos componentes da amostra a mesma carga negativa e garantir que a velocidade de migração seja, principalmente, em função da massa molecular: quanto maior, mais lenta a migração. Quando a voltagem for aplicada a amostra migrará para o pólo positivo. Resumidamente, o objetivo do SDS é desnaturar a proteína, isto é, converter a proteína numa estrutura linear (a sua forma nativa é, geralmente, globular) e conferir-lhe densidade de carga uniforme, de forma a serem separadas por eletroforese somente em função do tamanho (massa molecular). O SDS tem uma alta carga negativa pH 7 e tem uma cauda hidrofóbica que interage com as cadeias das proteínas. O número de moléculas de SDS que se ligam às proteínas é proporcional ao número de aminoácidos que constituem as mesmas, pois o detergente liga-se na razão de uma molécula por ligação peptídica. Outros agentes desnaturantes que podem ser adicionados são o  $\beta$ -Mercaptoetanol ou o Ditioneitol, utilizados para promover a redução das pontes dissulfetos.

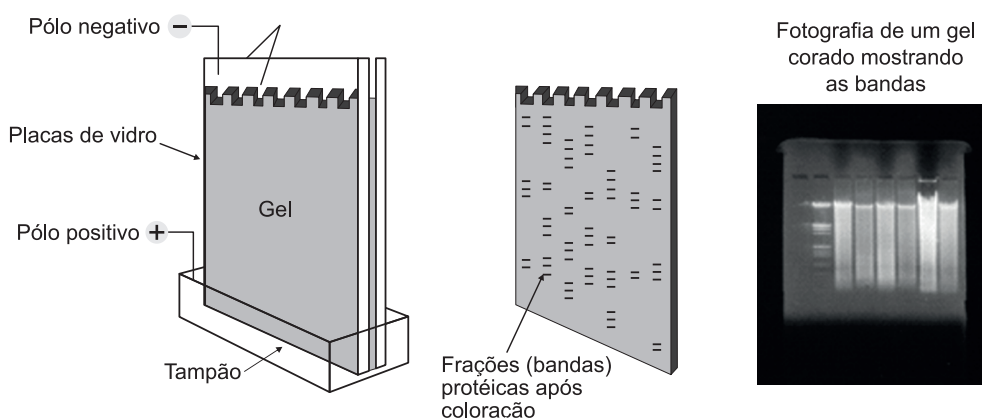
A superfície do gel é mantida em uma solução tampão e contem canaletas onde a amostra modificada é aplicada suavemente com uma pipeta e é então submetida ao campo elétrico, como ilustrado na figura 2. A amostra é empurrada por esta força e caminha ao longo do gel. O sentido da eletroforese é de cima para baixo, sendo o pólo negativo superior e o pólo positivo inferior. Partículas pequenas migram com maior velocidade para parte inferior do gel e as partículas maiores ficam retidas superiormente perto do ponto de aplicação.

Após as proteínas terem “corrido” o gel, é preciso fixá-las, para evitar que difundam quando se procede à coloração do mesmo. O gel é tipicamente corado com azul de Coomassie ou pelo método da prata, este último mais

sensível detectando menores quantidades da amostra. A fixação e coloração podem ser preparadas na mesma solução, usando como solvente o metanol ou sais de prata e depois descorado. O Coomassie se liga inespecificamente a praticamente todas as proteínas. É menos sensível que corar com a prata, mas efetivo também, além de ser fácil de corar. A coloração pela prata é 50 vezes mais sensível do que aquela obtida com Coomassie. A maior sensibilidade oferece vantagens como a necessidade de se utilizar menor quantidade e menor concentração da amostra no gel.

O aumento na sensibilidade também pode ser alcançado se utilizarmos a autorradiografia, ou seja, quando a amostra da proteína for previamente marcada por um radioisótopo. A energia do radioisótopo poderá ser captada por filmes apropriados e assim identificar a amostra. A fluorescência também é utilizada.

Figura 2: A eletroforese ocorre dentro de uma matriz ou gel. A placa montada permanece mergulhada em



solução tampão que também está em contato com os eletrodos e mantém o pH constante. Quando as proteínas aplicadas são submetidas ao campo elétrico podem migrar para o pólo positivo ou negativo, em função da carga líquida que apresenta e da massa molecular.

A eletroforese de amostras de sangue pode detectar, por exemplo, mutações em indivíduos com suspeitas de doença falciforme (HbS). É utilizado um tampão com pH entre 8 e 9, uma vez que nessa faixa de pH a mutação que deu origem à hemoglobina promove uma mudança de carga elétrica da molécula de HbS, tornando-a menos negativa em relação à HbA. Assim, quando amostras de sangue com diferentes genótipos são submetidas à eletroforese, a HbS se move eletricamente de forma mais lenta que a HbA.

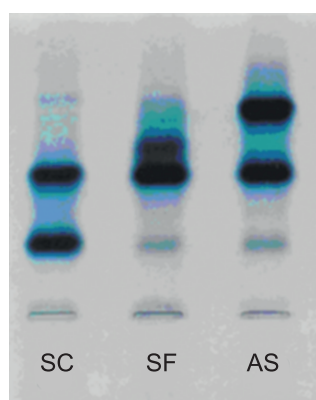
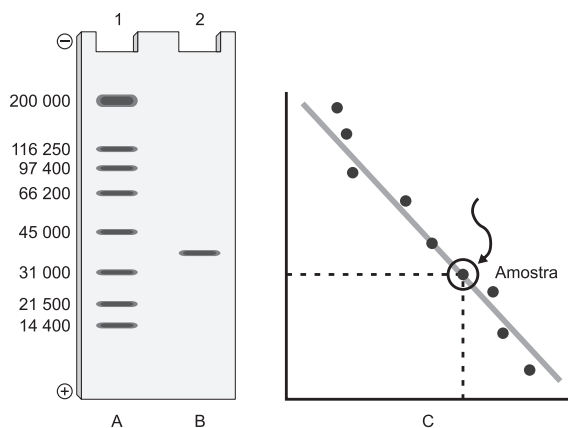


Figura 3: eletroforese alcalina de hemoglobinas em gel de agarose. Diferenciação da mobilidade eletroforética dos genótipos SC, SF e AS. Há traços de Hb A nos genótipos SC e SF por provável contaminação da amostra. Fonte: [www.hemoglobinopatias.com.br](http://www.hemoglobinopatias.com.br)

A distância percorrida pela amostra a partir do ponto de aplicação é comparada com a distância que outras proteínas de massa molecular conhecida percorrem no mesmo gel. Essas proteínas são os marcadores ou padrões de massa molecular que são aplicados em uma canaleta do gel no início do processo. A partir da medida dessas migrações,



um gráfico é construído e por extrapolação à reta obtida encontra-se valor estimado da massa molecular da amostra, como ilustrado na figura 4.

Figura 4: A aplicação de padrões de massas moleculares conhecidas (A) e da amostra (B) permite a determinação da massa molecular da proteína desconhecida através da interpolação no gráfico construído a partir dos padrões.

### 7.3. VARIAÇÕES NA TÉCNICA DE ELETROFORESE: A ELETROFORESE BIDIMENSIONAL

A eletroforese bidimensional é o método mais eficiente de separação de proteínas, mas por ser uma etapa menos automatizada exige maior habilidade do pesquisador. Através desta técnica as proteínas são submetidas a dois processos consecutivos em diferentes dimensões de separação baseados em duas propriedades diferentes das proteínas:

#### 7.3.1. Focalização isoeétrica

As proteínas migram em um gel contendo gradiente de pH até atingirem uma posição estacionária onde possuem carga zero (ponto isoeletrico). As proteínas com carga zero não podem se mover e, então, diz-se que estão focalizadas. Como cada proteína tem seu ponto isoeletrico específico, o gel possuirá manchas em diferentes posições ao final da corrida, cada uma correspondendo a um tipo de proteína.

#### 7.3.2. Massa molecular

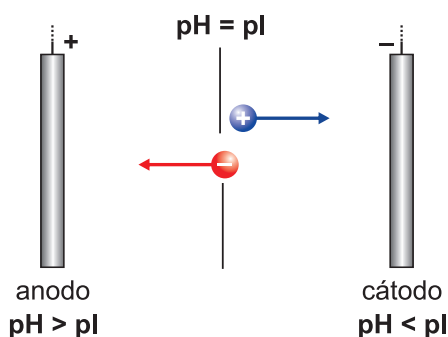
O gel da primeira dimensão é submetido a um segundo tipo de eletroforese, no qual as proteínas focalizadas serão separadas pelo seu tamanho ou massa molecular. As proteínas presentes no gel podem então ser coradas e visualizadas, permitindo que, ao final do processo, milhares de manchas de proteínas possam ser observadas em um único gel bidimensional.

Podemos afirmar que a eletroforese bidimensional separa as proteínas de massa

molecular idêntica, que diferem em seu  $pI$ , ou proteínas com  $pI$  similar, porém com massas moleculares diferentes. A sequência da eletroforese bidimensional é primeiramente a focalização isoelétrica, no sentido vertical (coluna), e depois a eletroforese com SDS no sentido horizontal (placa de gel); assim cada sentido indica uma diferença no  $pI$  ou na massa molecular.

A eletroforese **bidimensional** é fundamentada nas diferentes propriedades das proteínas. A carga das proteínas é determinada pelo pH do meio onde estão suspensas. As proteínas possuem uma carga nativa que pode ser positiva, negativa ou nula, dependendo do pH do meio, e para cada proteína existe um

valor de pH do meio para o qual a carga da mesma é 0 ( $pI$ , ponto isoelétrico).



Numa solução com pH acima do ponto isoelétrico, a proteína está negativamente carregada e migra para o ânodo do campo elétrico. Abaixo do ponto isoelétrico, a proteína é positivamente carregada e migra para o cátodo, como ilustrado na figura 5.

Figura 5: O ponto isoelétrico é aquele onde a carga da proteína é zero e se aplicarmos um campo elétrico

não haverá migração.

Utilizando-se deste comportamento das proteínas, a primeira etapa da eletroforese bidimensional, ou primeira dimensão, consiste em uma focalização isoelétrica (IEF), na qual as proteínas são separadas pela sua carga elétrica em um gel de poliacrilamida com um gradiente de pH (Figura 6).

Quando uma proteína é colocada num tubo estreito contendo gel de poliacrilamida, no qual o gradiente de pH é estabelecido pela mistura de soluções tampão especiais e sujeita a um campo elétrico, a amostra irá inicialmente mover-se em direção ao eletrodo com carga oposta. Durante a migração através do gradiente de pH, a proteína irá captar ou perder prótons. Enquanto a proteína migra a sua velocidade vai diminuindo até chegar ao ponto no qual o valor de pH será igual ao seu  $pI$ . Neste ponto a proteína terá carga total neutra e como consequência deixa de migrar. Se a proteína se difundir para uma região fora do seu  $pI$ , irá adquirir carga e, conseqüentemente, move-se novamente para a posição onde é globalmente neutra. Pode ter grupos ionizados, mas globalmente é neutra.

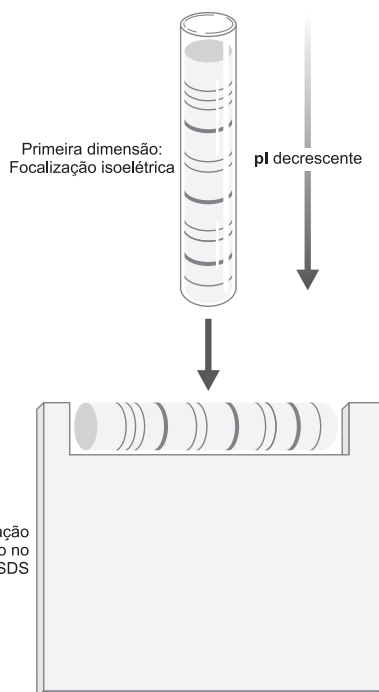


Figura 6: A aplicação das duas dimensões; inicialmente a focalização seguida de PAGE-SDS.

Na segunda dimensão, as proteínas previamente separadas pela IEF são submetidas a uma eletroforese desnaturante em gel de poliacrilamida, onde são separadas de acordo com suas massas moleculares. Como os parâmetros usados na primeira dimensão (ponto isoelétrico) e na segunda dimensão (massa molecular) são independentes, a separação atinge um alto nível de resolução, o que permite a visualização de centenas de proteínas diferentes ao mesmo tempo. Após a eletroforese bidimensional, as proteínas são visualizadas diretamente no gel, através de métodos de coloração com azul de Coomassie ou por coloração com prata, resultando num perfil bidimensional de pontos denominados

*spots*, cada um contendo múltiplas cópias de uma proteína.

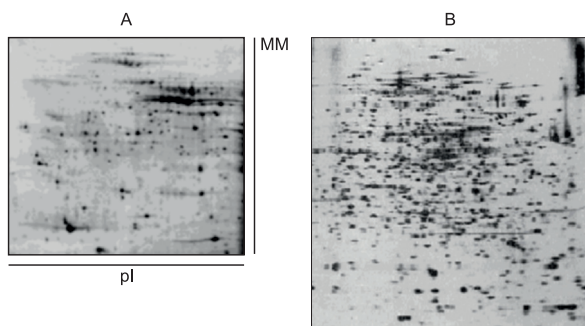


Figura 7: Eletroforese bidimensional mostrando em A as propriedades pI e massa molecular utilizadas para a separação e em B a separação para a *E. coli*, mostrando mais de 2000 proteínas.

Como as bandas protéicas tendem a sobrepor-se, os métodos unidimensionais de separação, como a eletroforese em gel de poliacrilamida, podem apenas separar um número relativamente pequeno de proteínas (geralmente menos de 50), enquanto a eletroforese de um gel bidimensional, ao combinar dois processos de separação distintos, pode ser usada para separar mais de 1000 proteínas. Entretanto, o número de proteínas focalizadas no gel e o número de identificações obtidas representam uma

pequena parte de todo o proteoma. De forma complementar, uma análise quantitativa das proteínas, empregando marcadores isotópicos seguido por cromatografia líquida acoplada à espectrometria de massa, poderá auxiliar na identificação de proteínas expressas. As únicas proteínas que não são separadas serão aquelas que terão idêntico tamanho e ponto isoelétrico, uma situação relativamente rara. O poder de separação é tão grande que duas proteínas que diferem em apenas um aminoácido carregado podem ser prontamente distinguidas.

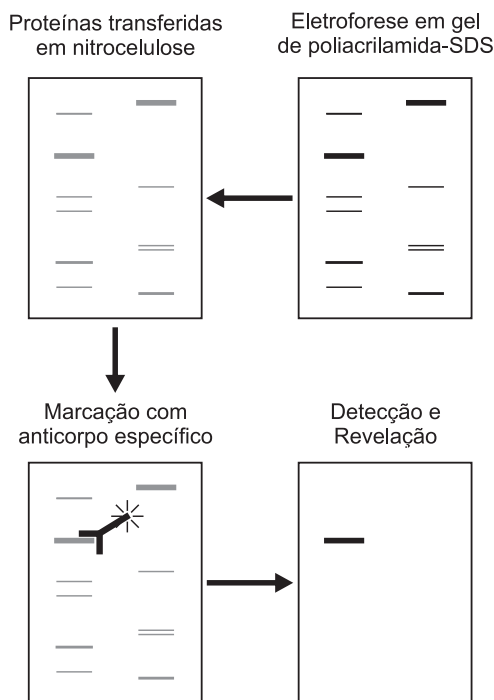


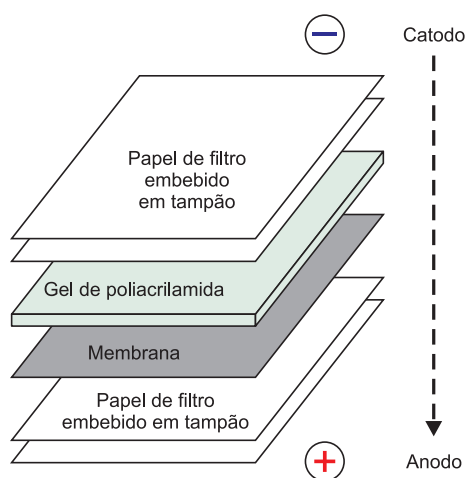
Figura 8: As etapas do *Western blotting*.

Outra modificação na técnica de eletroforese é denominada de *western blotting*. Trata-se de método semiquantitativo altamente específico para proteínas. Consiste

de quatro etapas após a extração das proteínas: separação em condições desnaturantes; eletrotransferência; reações imunoquímicas com anticorpos marcados; e detecção e revelação a partir, por exemplo, da presença de moléculas quimioluminescentes.

As proteínas são aplicadas em um gel de poliacrilamida, em condições desnaturantes (SDS, betamercaptoetanol) submetidas a separação pela ação do campo elétrico. Após este fracionamento é feita a transferência das proteínas para uma membrana de nitrocelulose, sobreposta a papéis de filtro de boa qualidade. O processo de montagem deve ser feito cuidadosamente a fim de evitar a formação de bolhas de ar que entre o gel de acrilamida e a membrana. O sistema é então montado em placas, fazendo contato entre o anodo e o catodo, como mostrado na figura 9.

A transferência das proteínas inicia-se quando aplicada uma corrente constante ao sistema montado. A corrente aplicada, em geral, não deve exceder  $0,8 \text{ mA/cm}^2$  de área de gel, usualmente por 1h, que é o tempo suficiente para a transferência ocorrer.



Pode ser feita uma coloração do gel com Comassie para verificar se a transferência foi completa. A seguir, as proteínas-alvo são identificadas por anticorpos específicos. A revelação do sistema é feita com outro anticorpo (secundário) fluorescente ou radioativo.

Figura 9: Montagem do sistema de transferência.

## 7.4. LIMITAÇÕES DA ELETROFORESE

A eletroforese bidimensional é útil mas sofre limitações quando utilizada para separação de misturas complexas. A técnica da eletroforese poderá ser utilizada mesmo não tendo nenhum conhecimento do proteoma a ser analisado. A eletroforese é uma técnica de baixa complexidade mas um pouco frágil devido tanto à própria natureza do gel, que pode se quebrar facilmente durante a sua manipulação, quanto ao fato de ainda não ser automatizada. Outra desvantagem da técnica é a necessidade de preparo da amostra, às vezes com várias etapas (centrifugação fracionada, extração de lipídios, desalinização, redução das pontes dissulfetos, entre outras). Algumas desvantagens da eletroforese estão relacionadas à quantidade de proteínas encontradas na amostra. Em um gel não é possível determinar todas as proteínas, pois a diversidade e quantidade impedem a visualização individual. Uma proteína que se encontra em quantidade relativamente menor do que outra da amostra não será visualizada, pois as que se encontram em grande quantidade dificultam a visualização. Daí, a necessidade de se encontrar uma técnica que permita o preparo da amostra de maneira a analisar qualquer proteína presente,

independente da sua contribuição relativa no proteoma.

É preciso também lembrar que a eletroforese é feita em condições desnaturantes o que provoca o rompimento de arranjos internos da proteína, arranjos estes com ação funcional. Os arranjos protéicos com outras moléculas dificultam a solubilização. Vimos que proteínas podem ter natureza hidrofóbica ou hidrofílica, o que exige métodos diferentes de solubilização; pH e tamanho da proteína também interferem na solubilização e na eletroforese, visto que proteínas muito grandes terão sua mobilidade eletroforética dificultada e proteínas muito pequenas podem formar uma única banda na eletroforese por superposição no mesmo local do gel. Técnicas como a espectrometria de massa e a cromatografia líquida têm sido aliadas à eletroforese para concentrar e separar os componentes específicos de uma amostra.

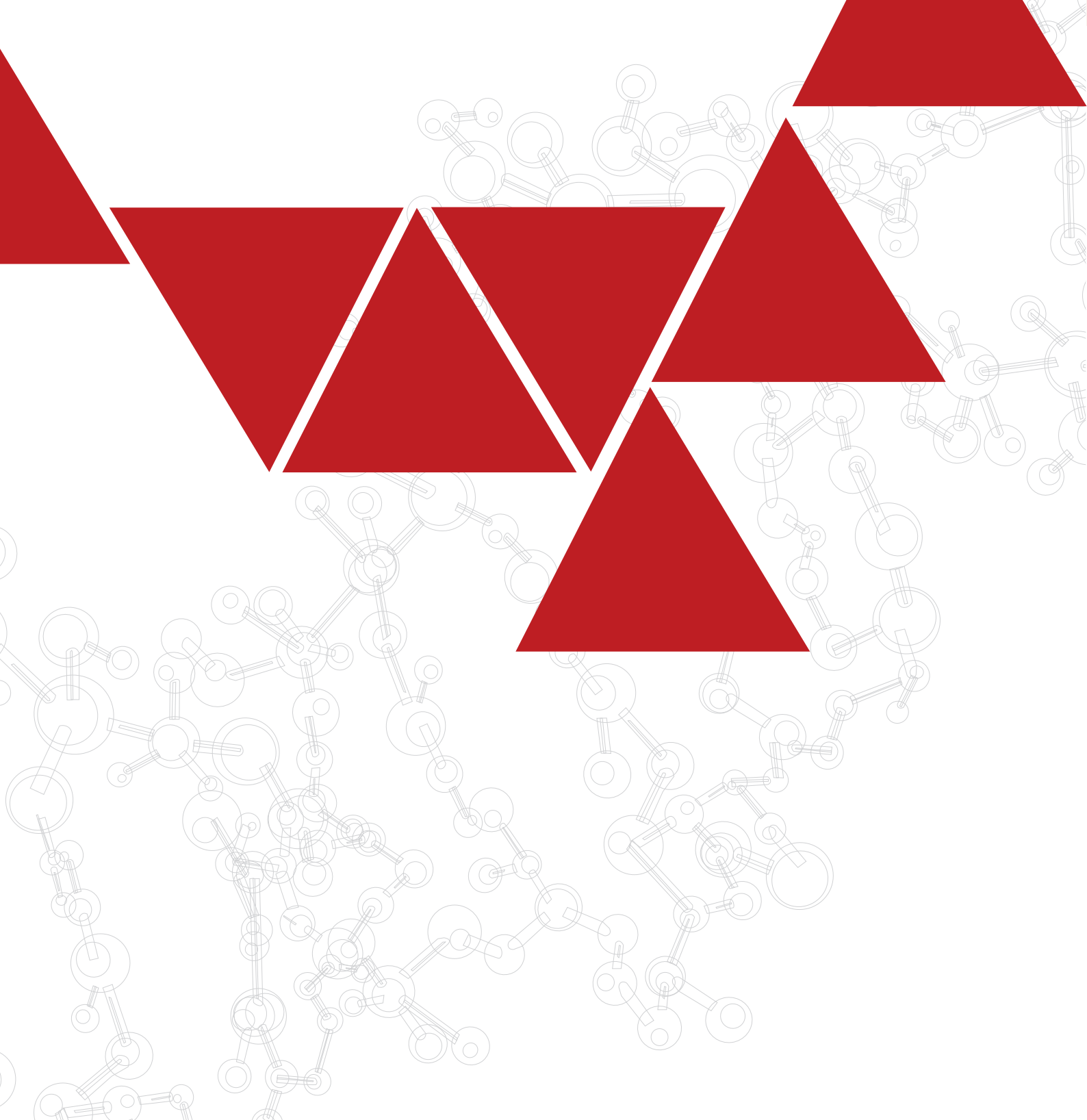
Outras técnicas podem anteceder a eletroforese. A centrifugação, por exemplo, é usada especialmente se o que quisermos for identificar a localização das proteínas, ou seja, se elas estão inseridas na membrana ou se são encontradas no citoplasma. Já a imunoprecipitação, auxilia na definição da presença ou não de associações covalentes e não-covalentes de proteínas.

É bom lembrar que a identificação da seqüência de aminoácidos de uma nova proteína não garante que a sua função também foi identificada, a menos que a proteína pertença a uma família com funções já conhecidas. Como afirmado anteriormente, ensaios *in vitro* e/ou *in vivo* serão necessários para testar e comprovar a função da proteína analisada.

Num complexo multiprotéico são encontradas várias proteínas e um dos objetivos dos projetos de proteoma é separar e visualizar as proteínas individualmente. A eletroforese bidimensional feita em gel de poli(acrilamida), a 2D PAGE, é uma técnica eficiente e poderosa para separar, purificar e identificar proteínas em amostras que contém pequenas quantidades. Essa técnica fornece informação de quais e quantas proteínas são expressas em um organismo ou tecido. Modificações pós-traducionais também podem ser identificadas. Para isso, as amostras de proteínas devem ser previamente incubadas com fósforo radioativo  $^{32}\text{P}$  ou hidrogênio radioativo, trítio, associado a manose,  $[^3\text{H}]$ -manose, evidenciando os grupamentos fosfato ou manose-glicanas. Anticorpos específicos para fosfotreonina, fosfoserina e fosfotirosina podem ser aplicados no gel para localizar proteínas fosforiladas.

### Autoavaliação

- a. A técnica denominada de *western blotting* permite detectar uma proteína numa mistura de proteínas dando também informação a cerca do seu tamanho. Detalhe o princípio desta técnica e compare-a com a eletroforese bidimensional.
- b. Afirmamos que é possível controlar o tamanho dos poros do gel de acrilamida de maneira a formar uma peneira molecular. Explique como isso é feito na prática. Utilize as palavras acrilamida, metileno bis-acrilamida e poli(acrilamida).
- c. É sabido que os sais interferem diretamente no processo eletroforético. Explique como se dá esta interferência e indique metodologias utilizadas para eliminar a presença de sais na amostra.



8

*Espectrofotometria e  
Espectrometria de massa  
e o estudo do proteoma*

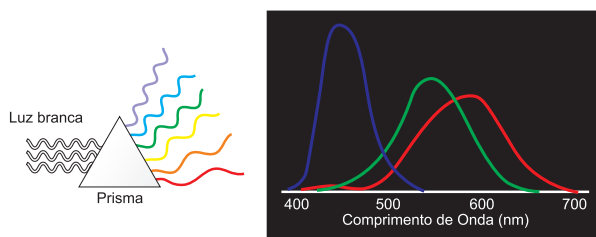
## AULA 8: ESPECTROFOTOMETRIA E ESPECTROMETRIA DE MASSA E O ESTUDO DO PROTEOMA

Ao final desta aula o aluno deverá:

- a. Conceituar qual é quantitativamente o espectro eletromagnético;
- b. Enunciar a Lei de Planck associada à energia da radiação;
- c. Exemplificar cromóforos encontrados em moléculas biológicas;
- d. Usar a Lei de Lambert-Beer associando variáveis aos fenômenos observados de transmitância e absorvância;
- e. Descrever métodos de determinações quantitativas de soluções coloridas;
- f. Descrever os princípios da análise por espectrometria de massa;
- g. Descrever os princípios de funcionamento para um espectrômetro de massa, a saber, spray eletrostático e o laser;
- h. Citar as técnicas utilizadas para o estudo do proteoma;
- i. Enumerar as dificuldades iniciais para o preparo de uma amostra;
- j. Explicar o uso de proteases no estudo proteômico;
- k. Descrever sucintamente como é obtido um banco de dados de proteínas;
- l. Citar como as modificações pós-traducionais podem ser identificadas.

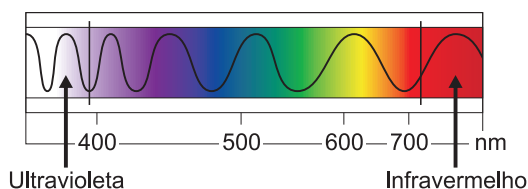
### 8.1. O ESPECTRO ELETROMAGNÉTICO

Luz, radiação e energia são sinônimos? Se interpretarmos a luz como um fenômeno energético, podemos atribuir a ela além de energia um determinado comprimento de onda ou frequência. A luz branca é capaz de refletir todas as cores, ou seja, ela contém todos os comprimentos de onda. Ao incidirmos sobre um prisma, esta luz branca será decomposta nos seus comprimentos de onda associados, o que denominamos de espectro eletromagnético, onde podem ser associados energia, frequência e comprimento de onda. A região do visível é a única sensível ao olho humano. Graças a ela é possível a percepção das cores provocadas pela absorção seletiva de certos comprimentos de onda por um objeto corado. A faixa do visível vai de 400 a 750 nm, única percebida pelos seres humanos na forma de sensações visuais denominadas cores. Abaixo de 400 nm e acima de 750 nm, essas sensações não são percebidas. Assim, cor é sensação associada a um comprimento de onda. O preto é ausência de todas as cores. Devemos fazer a distinção entre cor e pigmento. Chama-se cor ao conjunto



de comprimentos de onda ou espectro, e pigmento a qualquer substância ou matéria que dá a sensação de cor. Um feixe de luz vermelha é cor, uma mancha de tinta azul é pigmento.

Figura 1: A luz branca pode ser decomposta em seus diferentes comprimentos de onda



A espectrofotometria baseia-se na absorção da radiação nos comprimentos de onda entre o ultravioleta e o infravermelho, que delimitam a faixa acima e abaixo da luz visível.

Faixa	Comprimento de onda (nm)
Ultravioleta	<400
Violeta	390 - 455
Azul	455 - 492
Verde	492 - 577
Amarelo	577 - 597
Laranja	597 - 622
Vermelho	622 - 780
Infravermelho	>760

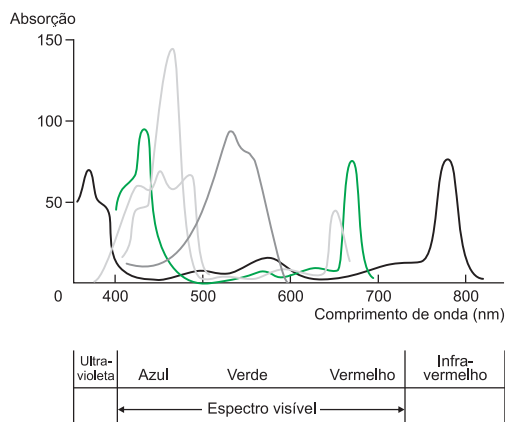
Figura 2. O espectro eletromagnético aplicado à espectrofotometria utiliza a radiação compreendida entre o ultravioleta e o infravermelho.

A Energia (E) associada ao comprimento de onda ( $\lambda$ ), esta medida em nanômetros (1 nm pode ser escrito como  $10^{-9}$  m) pode ser descrito pela Lei de Planck:

$E = h \cdot f$ <p>ou</p> $E = h \cdot c / \lambda$	<p>Onde:</p> <p><math>c = 3 \cdot 10^{10}</math> cm/s, a velocidade da luz no vácuo</p> <p><math>h = 6,6260693(11) \times 10^{-34}</math> J . s, a constante de Planck</p> <p>f é a frequência da radiação</p>
-----------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 8.2. ESPECTROS DE ABSORÇÃO

Quando uma solução ou composto absorve vários comprimentos de onda, a este conjunto denominamos **espectro de absorção** e cada molécula, composto ou substância, possui o seu próprio espectro. De posse de uma amostra, é possível obter o seu espectro de absorção incidindo sobre ela vários comprimentos de onda, medindo a energia absorvida e detectando qual o comprimento de onda que provoca o máximo de absorção. A espectrofotometria utiliza dos padrões de absorção para identificar e quantificar substâncias. A identificação é feita pelo espectro de absorção e a quantificação é feita sabendo que a quantidade



de luz absorvida está relacionada com a concentração da substância seguindo a lei de Lambert – Beer, como veremos a seguir.

Figura 3. Registro das absorções de cinco substâncias diferentes. A substância 1, por exemplo, absorve nas regiões do ultravioleta e do infravermelho, mas absorve muito pouco na região do visível. A substância 4 absorve na faixa verde do espectro e será vista pelo olho humano pela cor transmitida, a cor complementar, que é o vermelho.

Nas moléculas existem grupamentos químicos

responsáveis pela absorção da energia proveniente do comprimento de onda incidido sobre ela. A este grupamento é dado o nome de cromóforo. Uma molécula só absorve energia em determinado comprimento de onda se existir uma transição de energia associada à radiação eletromagnética naquela frequência. Quando o elétron volta ao seu estado fundamental, emite energia no mesmo comprimento de onda em que foi excitado. Essa energia será detectada em aparelho próprio, o espectrofotômetro.

As moléculas orgânicas responsáveis pela absorção luminosa são os aminoácidos, ácidos nucleicos e outros grupamentos cromóforos. Os aminoácidos têm uma significativa absorção no ultravioleta. Os sistemas vivos também apresentam biomoléculas que absorvem luz na faixa visível. A maioria dos cromóforos é derivada do grupo das porfirinas, da qual faz parte a hemoglobina. Quando ligada ao oxigênio (oxihemoglobina), a hemoglobina tem diferentes características de absorção. A oxihemoglobina absorve mais luz infravermelha, enquanto a deoxihemoglobina (hemoglobina desoxigenada) absorve mais luz vermelha. A figura 4 ilustra a absorção diferenciada destas duas espécies: observe a faixa de luz vermelha, no comprimento de onda entre 600-750nm, e a faixa de luz infravermelha de 850-1000nm. Esta absorção diferenciada permite a verificação instantânea do estado de oxigenação arterial de um indivíduo se acoplada a um espectrofotômetro específico nomeado oxímetro.

Outro exemplo de cromóforo é a melanina, que é um pigmento encontrado

na pele e que possui também um importante papel na conversão da energia em outra forma de energia, como por exemplo, em calor (efeito fototérmico). Alguns cromóforos são mostrados no Quadro 1.

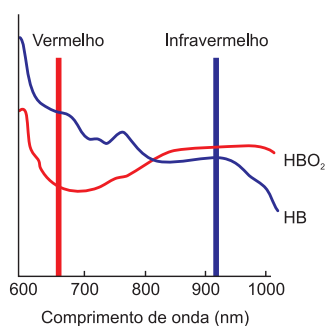


Figura 4: Espectro de absorção da hemoglobina (Hb) e da oxihemoglobina (HbO<sub>2</sub>).

Cromóforo	Exemplo	$\lambda_{\text{máx}}$ (nm)
Carbonila	$\begin{array}{c} \text{O} \\    \\ \text{CH}_3\text{CCH}_3 \end{array}$	186 280
	$\begin{array}{c} \text{O} \\    \\ \text{CH}_3\text{C H} \end{array}$	180 293
Carboxila	$\begin{array}{c} \text{O} \\    \\ \text{CH}_3\text{C OH} \end{array}$	204
Amida	$\begin{array}{c} \text{O} \\    \\ \text{CH}_3\text{C NH}_2 \end{array}$	214
Azo	$\text{CH}_3\text{CNH}_2$	339
Nitro	$\text{CH}_3\text{NO}_2$	280
Nitroso	$\text{C}_4\text{H}_9\text{NO}$	300 665
Nitrato	$\text{C}_2\text{H}_5\text{ONO}_2$	270

Quadro 1: Exemplos de grupos cromóforos e o comprimento de onda associado

### 8.3. ESPECTROFOTOMETRIA E QUANTIFICAÇÃO: A LEI DE LAMBERT-BEER

O aparelho usado para determinar a concentração de soluções coloridas através da passagem de um feixe de luz por um prisma, que atua dispersando esta luz em um espectro, é o espectrofotômetro. A luz, agora monocromática, passa por uma fenda e atravessa o tubo contendo uma solução colorida. Uma porção de luz é absorvida e outra passa pela solução. A luz que passa pela solução é detectada pela célula fotoelétrica, que converte em corrente elétrica, a qual é registrada por um galvanômetro. O espectrofotômetro permite saber a quantidade de luz absorvida a cada comprimento de onda.

Todo espectrofotômetro possui os seguintes componentes: fontes de radiação eletromagnética; componentes ópticos; detectores; obturadores eletrônicos. As fontes de radiação são as lâmpadas de deutério para excitação na região do ultravioleta (comprimentos de onda entre 180 a 370nm) e as lâmpadas de tungstênio ou tungstênio-halogênio ( $\lambda > 350\text{nm}$ ) para excitação na região do visível

Utilizamos um recipiente de vidro ou de quartzo, na forma cúbica, a cubeta, para fazer medidas relativas à quantidade de luz. Quando a luz incide sobre um meio homogêneo, uma parcela da luz incidente é refletida, outra parcela é absorvida pelo meio e o restante é transmitido. Se a intensidade da radiação da luz incidente for representada por  $I_o$  e da luz absorvida por  $I_a$ , a da transmitida por  $I_t$  e a da refletida por  $I_r$ , então:

$$I_o = I_a + I_t + I_r$$

Ou seja, a luz incidente terá uma parte absorvida, uma parte refletida e outra parte transmitida. Para diminuir o número dessas variáveis, a parcela  $I_r$  pode ser eliminada se previamente conhecida pelas propriedades da cubeta, onde se pode admitir que cerca de 4% a 2% da luz incidente sejam refletidos. A expressão anterior fica:

$$I_o = I_a + I_t$$

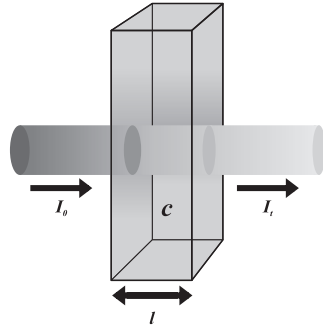
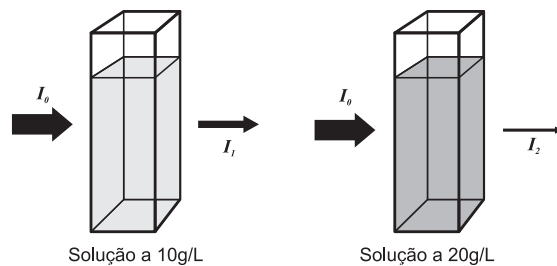


Figura 5: A cubeta de vidro ou quartzo, com caminho óptico  $l$ , contendo solução de concentração  $c$  sofrendo a incidência de radiação  $I_0$  e transmitindo radiação  $I_t$ .

Variando a concentração ou variando a espessura da solução, no caso da cubeta a ser atravessada pela luz, o que estamos promovendo é a interação entre a solução e a radiação. Os princípios que definem a absorção são conhecidos pela Lei de Lambert – Beer:

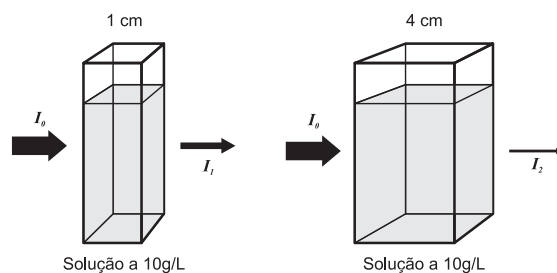
- A intensidade de luz monocromática, ou seja, aquela que apresenta apenas um comprimento de onda, transmitida por um corpo homogêneo é proporcional à intensidade de luz incidente, isto é,  $I_t = k I_0$ .
- A intensidade de luz monocromática transmitida decresce exponencialmente com o aumento da espessura da camada que sofre a incidência de luz.

Vejam alguns exemplos de absorção:



1. A absorção da luz é tanto maior quanto **mais concentrada** for a solução.

Figura 6: Relação entre a concentração de uma solução e a luz absorvida. A solução a 20g/L absorve **o dobro** da solução a 10g/L. A luz transmitida  $I_2$  é a metade de  $I_1$ .



2. A absorção da luz é tanto maior quanto maior for a distância percorrida (caminho óptico) pelo feixe luminoso através da amostra:

Figura 7: Relação entre a distância percorrida pelo feixe luminoso e a luz absorvida por uma solução. A solução contida na cubeta com  $l$  de 4 cm absorve 4 vezes mais luz que a contida na cubeta de  $l$  igual a 1 cm.

Matematicamente, a lei de Lambert – Beer pode ser expressa como:

$$A_{\lambda} = \epsilon_{\lambda} \cdot c \cdot l$$

Onde os símbolos representam:

$A_{\lambda}$  é a absorvância obtida num comprimento de onda,  $\lambda$ , fixo;

$\epsilon_{\lambda}$  é o coeficiente de extinção ou absortividade molar ao comprimento de onda  $\lambda$ , que é constante para um  $\lambda$  fixo, mede a probabilidade de um único comprimento de onda ser absorvido por uma determinada solução ou a sua capacidade em absorver energia;

$C$  é a concentração molar da solução absorvente;

$l$  é a distância percorrida pelo feixe luminoso através da amostra (caminho óptico).

O valor de  $\epsilon$  varia segundo os materiais absorventes e com o comprimento de onda para cada material em particular e deve ser determinado experimentalmente. Para quantificar espectrofotometricamente uma substância é necessário, obviamente, saber o valor de  $\epsilon$ .

Se a concentração for expressa em gramas por litro e usarmos as unidades de intensidade de luz incidente, como descrito anteriormente, a expressão da lei torna-se:

$$I_t = I_0 \cdot 10^{-acl}$$

Onde a constante  $a$  é a mesma absortividade molar  $\epsilon$  e é mais usada quando não se conhece a natureza do material absorvente (portanto a sua massa molar) e a concentração não está expressa em molar. Já a absortividade molar ( $\epsilon$ ) é preferível quando se deseja comparar quantitativamente a absorção de várias substâncias. Para uma mesma espessura do absorvedor (caminho óptico), quanto maior o valor de  $\epsilon$  maior a sensibilidade do método.

Trabalhando a equação  $I_t = I_0 \cdot 10^{-acl}$ , podemos chegar a forma .

$$\log_{10} \frac{I_0}{I_t} = -\log_{10} T = \epsilon C l$$

A relação  $\log_{10} \frac{I_0}{I_t}$  é chamada de absorvância ou absorbância:

Absorbância	Transmitância	Relação A e T
$A = \log (I_0 / I_t)$ ;	$T = I/I_0$	$A = \log (1/T) = -\log T$

Quadro 2: Relação entre absorbância e transmitância

Resumindo, a equação de Lambert- Beer pode ser expressa da seguinte forma:  $A = \epsilon l c$  ou no caso de não se utilizar a concentração em mol por litro  $A = a l c$ .

Veja que há uma relação linear entre a absorvância,  $A$ , e a concentração  $c$ . Se o caminho óptico for mantido constante ( $l$ ), pode-se determinar a concentração de uma espécie na amostra através da medida de absorvância. Na prática, uma curva de calibração (absorvância versus concentração) da amostra de interesse é construída e a concentração é determinada através dela. Na prática, observa-se que para concentrações muito elevadas a linearidade não é obedecida, tornando-se necessário diluir a amostra para que se obedeça à lei.

Vimos que os termos absorção e transmissão são utilizados quando incidimos luz sobre uma solução. Absorção e transmissão são complementares, ou seja, ao incidirmos luz sobre uma amostra, se a absorção for de 90%, a transmissão será de 10%.

## 8.4. MÉTODOS COLORIMÉTRICOS

Se a amostra de nosso interesse não absorver significativamente a luz em nenhum comprimento de onda, é possível promover uma reação da amostra com um reagente específico, de maneira a termos uma cor cuja intensidade é diretamente proporcional à concentração da substância na mistura original. Esta técnica é conhecida como método colorimétrico. Para as proteínas utilizamos a absorvância a 280 nm, sendo esta proporcional à concentração de proteína; entretanto os ácidos nucleicos também absorvem neste comprimento de onda. Se as proteínas de nosso interesse encontram-se num complexo, misturadas a ácidos nucleicos, por exemplo, escolhemos a reação do biureto ou de Lowry para formar um complexo reagente-proteína colorido. O complexo biureto-proteínas absorve em 540nm e proteína-lowry em 660 nm.

Na figura 8, vemos que, independente da concentração, há um valor que é mantido constante, que é o comprimento de onda onde a absorção é máxima. O conhecimento do comprimento de onda onde há maior absorção é muito importante para a determinação das concentrações. Para isso é necessário preparar uma série de soluções do composto a quantificar, de concentração conhecida, fazê-las entrar

em contato com o reagente e medir as absorvâncias ao comprimento de onda adequado, como ilustrado na figura 9.

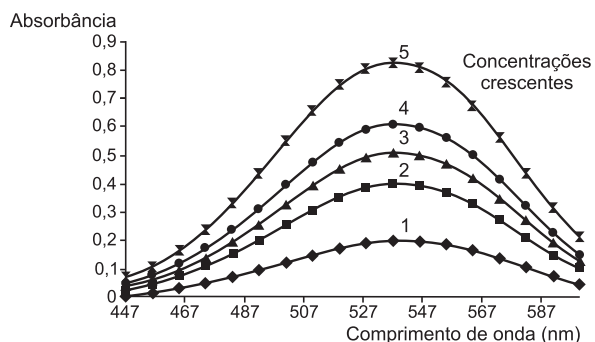


Figura 8: Espectro de absorção em função da concentração obtido na faixa de 440 a 600nm.

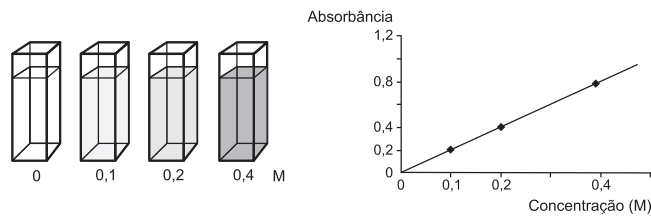


Figura 9: Calibração de um método colorimétrico. A absorbância ao comprimento de onda escolhido é diretamente proporcional à concentração do composto na solução.

No exemplo da Figura 9, há uma relação linear perfeita entre a concentração molar da substância e a absorbância ao comprimento de onda  $\lambda$  fixo. A reta obtida expressa a relação  $A_{\lambda} = \epsilon_{\lambda} \cdot c \cdot l$  definida anteriormente. Sabendo esta relação, podemos fazer corresponder uma absorbância medida, a uma concentração de substância na solução a analisar, por interpolação à reta obtida.

Assim, podemos fazer determinações quantitativas de soluções coloridas através do percentual de absorbância. Para isso, determina-se qual o comprimento de onda máximo na qual uma solução absorva mais intensamente (como ilustrado na Figura 8), faz-se uma série de medidas com soluções-padrões (como ilustrado na Figura 9), ou seja, de concentração conhecida, acham-se os valores da absorbância de cada uma das amostras e constrói-se um gráfico do tipo **Absorbância x Concentração**. Os valores da concentração estarão no eixo das abcissas, enquanto que os valores da absorbância estarão no eixo das ordenadas. Depois, é só fazer a leitura da absorbância da solução-problema e comparar o resultado com os pontos marcados no gráfico e determinar a sua concentração.

## 8.5. O EQUIPAMENTO DE ESPECTROSCOPIA NO ULTRAVIOLETA- VISÍVEL

A figura 10 mostra um diagrama do espectrofotômetro onde são mostradas as partes básicas. A faixa de trabalho do aparelho é definida pelas lâmpadas que são as fontes de radiação. Para a região do ultravioleta, na faixa de 200 nm – 400 nm, a lâmpada é de hidrogênio; já para a região visível, 400 nm – 760 nm, usa-se uma lâmpada de filamento de tungstênio. A absorção no UV-visível requer a presença de cromóforos como mencionamos anteriormente. Há um espelho que se movimenta automaticamente com a variação do comprimento de onda do feixe. Faz parte do equipamento um detector, que funciona como um fotomultiplicador que transmite ao registrador a razão  $I_0 / I$ , ao qual está acoplado um registrador ou computador para registrar os valores de absorbância. No diagrama, mostramos um espectrofotômetro de duplo feixe, nas quais a fonte primária de luz é incidida em dois feixes, um dos quais passa através de uma célula contendo a solução da amostra e o outro passa através da célula contendo o solvente de referência. Assim, a absorção da luz pelo solvente é subtraída eletronicamente da absorção da solução da amostra, retirando os efeitos do solvente.

## Espectrofotômetro de feixe duplo

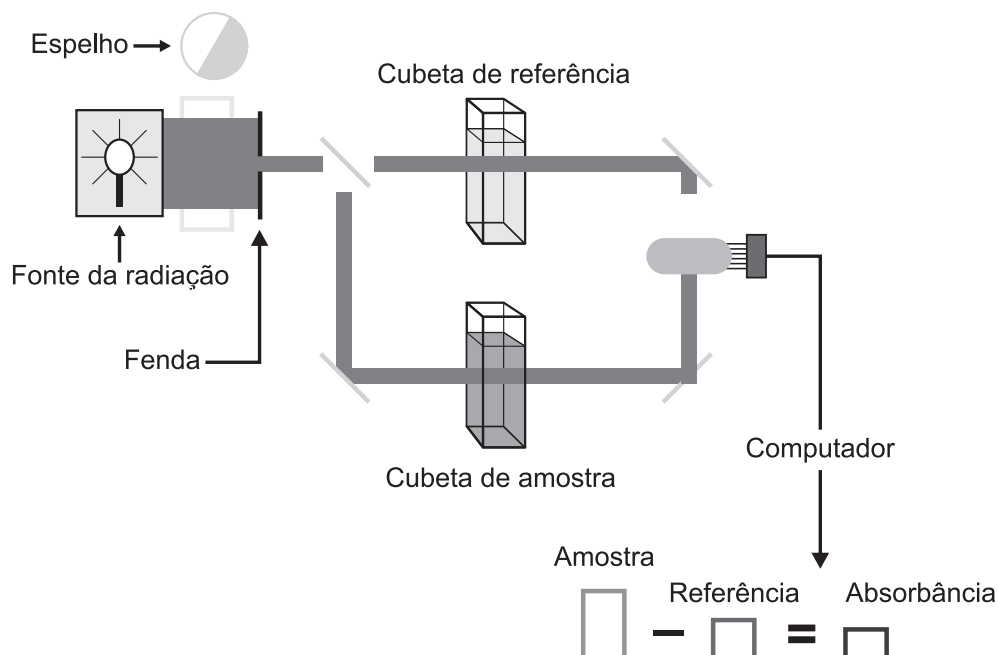


Figura 10: Diagrama de um espectrofotômetro de feixe duplo

## 8.6. ESPECTROMETRIA DE MASSA

A maioria dos espectrômetros de massa possui: uma fonte de ionização, para produzir os íons; um ou mais analisadores de massa, para separar os íons e um detector para quantificar os íons. A amostra ionizada é separada de acordo com a proporção entre a massa e a carga quando submetida a condições específicas de um campo elétrico e/ou magnético.

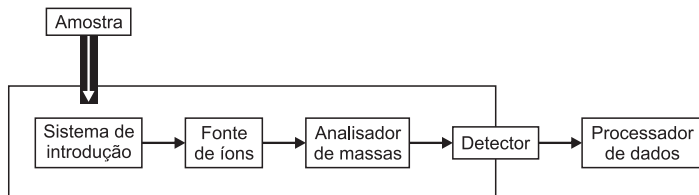


Figura 11: Diagrama de um espectrofotômetro de massa.

Definido assim, percebe-se que para a análise de uma amostra ser feita pelo

espectrômetro de massa é preciso inicialmente gaseificar, ou seja, vaporizar e ionizar a amostra. O tempo que cada íon leva para atingir o detector depende das suas massas moleculares ( $m$ ) e carga ( $z$ ), sendo proporcional à razão  $m/z$ ; portanto, é possível determinar a massa molecular de uma amostra a partir de um gráfico da intensidade relativa *versus* tempo de retenção, sendo que a

análise dos picos é feita por programas de computador, uma vez que a presença de uma grande quantidade de cargas em diferentes moléculas de um mesmo tipo gera diversos picos.

Há diferentes princípios de funcionamento para um espectrômetro de massa, mas, na sua maioria, a informação obtida em um espectro complementa a informação obtida no outro. Um dos princípios de funcionamento é a aplicação de uma descarga elétrica nas proteínas que provocará as fragmentações em partes eletricamente carregadas, as quais serão então identificadas de acordo com suas massas. Essa é a técnica de ionização por *spray* eletrostático, empregada no estudo de moléculas solúveis em água, como a hemoglobina, peptídeos,

proteínas, polissacarídeos e ácidos nucleicos. Outro princípio de funcionamento é o que utiliza um laser disparado sobre as proteínas depositadas em uma matriz ou cristal, a dessorção branda por laser, tornando as proteínas eletricamente carregadas. Por meio dessa técnica, pode-se avaliar as estruturas de proteínas encontradas nas membranas das células. A figura 12 mostra esses dois princípios de funcionamento. Na figura também são mostrados gráficos de espectro de massa onde aparecem picos de intensidades variáveis, cada pico correspondendo a íons com uma razão  $m/z$ .

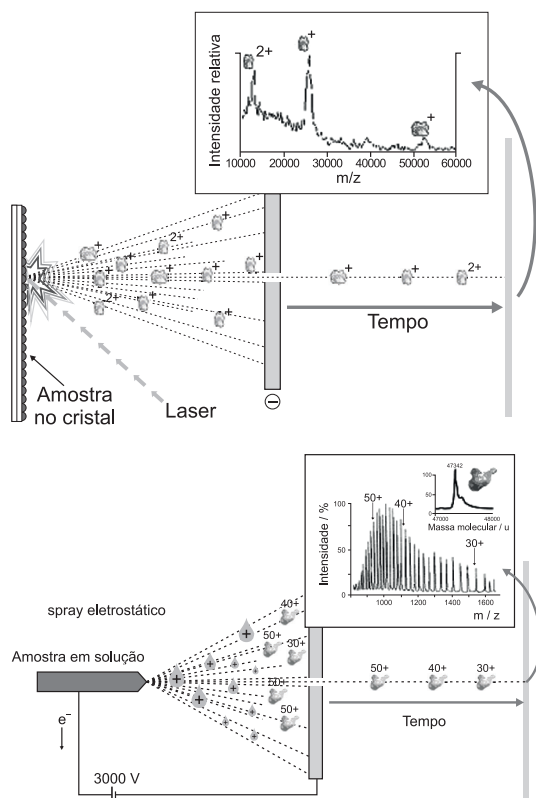


Figura 12: Diagrama de um espectrômetro de massa com ionização por laser e por *spray* eletrostático.

A técnica do *spray* eletrostático é uma técnica de ionização que permite a preservação das interações não covalentes entre moléculas em solução na fase gasosa, o que a torna ideal para as moléculas biológicas e para a análise proteômica.

A espectrometria de massas permite a identificação das proteínas presentes nos perfis de eletroforese bidimensional ou eluídas das colunas de cromatografia, que tratamos nas aulas anteriores. É exigido que a proteína seja submetida a algum tipo de fracionamento como aquele obtido pela ação de proteases, a tripsina, por exemplo.

## 8.7. O ESTUDO DO PROTEOMA

Técnicas já bem conhecidas, como a eletroforese, ganharam impulso quando associadas a novas técnicas, como a difração de Raios-X e a espectrometria de massa. A espectrometria de massas é uma técnica bastante utilizada para moléculas inorgânicas e recentemente foi aperfeiçoada para o estudo de proteínas. Tecidos de animais, cultura de células, plantas, bactérias, fungos são utilizados como amostras. Há uma sequência de passos que vai desde o preparo da amostra, a separação das proteínas, a identificação, a determinação da sequência dos aminoácidos e os cruzamentos das informações obtidas com aquelas presentes em um banco de dados. Algumas amostras possuem estruturas complexas e/ou às vezes a fração obtida é muita pequena e pouco concentrada, tornando-se necessário realizar um pré-fracionamento ou a concentração da amostra para se ter a quantidade suficiente para realizar as análises com precisão e qualidade.

Ao começarmos uma análise, o primeiro passo é definir o que queremos saber e onde está o nosso objeto de análise. Determinar um constituinte protéico encontrado em líquidos orgânicos, como a urina, por exemplo, requer a coleta do líquido seguida de concentração do material sólido ali presente. Para o veneno de serpentes, esta etapa não é necessária, pois o complexo inicial já se encontra extremamente concentrado, mas em quantidade às vezes não suficiente para a análise de seus componentes. O procedimento de isolamento de uma amostra deve garantir também a sua integridade funcional.

Para a separação dos componentes de uma amostra são empregadas técnicas como a eletroforese e a cromatografia e suas derivadas, como a eletroforese bidimensional em gel de poliacrilamida e a cromatografia líquida multidimensional, que são técnicas que separam a amostra em suas diferentes partes, fracionando-a de acordo com as suas propriedades físico-químicas. O fracionamento também pode ser obtido com o uso de enzimas, por exemplo, através da incubação (ou digestão) com a tripsina. A tripsina reconhece sítios para a quebra dentro da sequência específica da proteína e, assim, é possível obter um padrão de clivagem, único para cada amostra analisada, o que permite identificar a sequência protéica quando comparada a um banco de dados.

O complexo protéico que foi submetido ao processo de fracionamento em diferentes componentes também pode ser identificado através da sua massa molecular e pela sua atividade e funcionalidade. Ensaio clínico e *in vitro* permitem identificar qual a fração protéica tem atividade funcional. Às vezes, em uma amostra de veneno, várias são as frações obtidas, mas apenas uma responde pela ação neurotóxica; outra corresponde à função hemolítica e precisam ser caracterizadas quimicamente e bioquimicamente.

É possível caracterizar a massa molecular pela técnica de espectrometria de massa, que se tornou mais eficiente e sensível com os avanços da automatização. Proteínas têm sido identificadas através da comparação da composição de aminoácidos, ponto isoelétrico, massa molecular, com aquelas previamente inseridas em bancos de dados, o qual é disponibilizado na internet.

Sequenciando diretamente os aminoácidos da proteína, obtêm-se a composição de uma proteína. Pode-se também sequenciar o DNA e, a partir daí, deduzir a sequência de aminoácidos, utilizando o código genético. A quebra teórica de proteínas por proteases é feita por programas que utilizam as mais de um milhão de sequências de proteínas depositadas nos bancos de dados e, desta maneira, são obtidos um conjunto de números que correspondem às massas dos peptídeos produzidos teoricamente. Cada um desses conjuntos é comparado com os dados experimentais e, se a sequência da proteína-problema estiver no banco de dados, ela é imediatamente identificada através dessa comparação. Com os dados de identificação de proteínas são preparados então os chamados mapas de proteoma.

O mapa conceitual mostrado na figura 13 resume o estudo do proteoma:

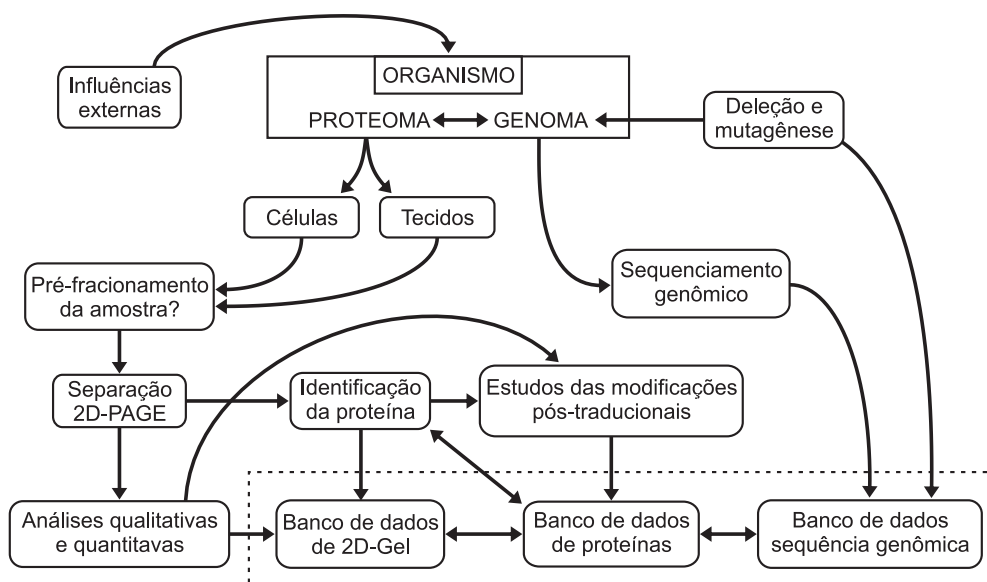


Figura 13: Como estudar um organismo. Adaptado de Marc R. Wilkins and Denis F. Hochstrasser em "Thinking bigproteome studies in a post-genome era".

### Autoavaliação

- a. Da equação que define a radiação em frequência, deduza qual é a relação que há entre frequência e comprimento de onda. Faça um escala de comprimento de onda e frequência na região do visível.
- b. Explique por que a afirmativa está correta: a cor complementar ao vermelho é o verde.
- c. Diferenciar matematicamente a absorbância da transmitância.
- d. A clorofila é um cromóforo? Justifique sua resposta e dê exemplos de cromóforos.
- e. Pesquise e resuma os métodos colorimétricos para dosagem de proteínas: método do biureto e método de Lowry.
- f. Quais são as limitações do uso da espectrometria de massa na análise proteômica?
- g. Faça o seu mapa conceitual de como estudar o proteoma.





Composto em caracteres Aller, Arial, Calibri, PT Sans e Times New Roman.  
Editorado pelo Centro de Apoio à Educação a Distância da UFMG (CAED-UFMG).  
Impresso pela Adescryn Gráfica Editora LTDA.  
Capa em Supremo, 250g, 4X0 cores - Miolo Off Set 120g, 4X4 cores.  
Julho - 2011

ISBN 978-85-64724-01-3



9 788564 724013

