

**Atena**  
Editora  
Ano 2021



# AVANÇOS METODOLÓGICOS EM BIOLOGIA MOLECULAR E BIOTECNOLOGIA

Anderson Barros Archanjo  
Aline Ribeiro Borçoi  
Suzanny Oliveira Mendes  
Adriana Madeira Álvares-da-Silva  
(Organizadores)

**Atena**  
Editora  
Ano 2021



# AVANÇOS METODOLÓGICOS EM BIOLOGIA MOLECULAR E BIOTECNOLOGIA

Anderson Barros Archanjo  
Aline Ribeiro Borçoi  
Suzanny Oliveira Mendes  
Adriana Madeira Álvares-da-Silva  
(Organizadores)

### **Editora Chefe**

Profª Drª Antonella Carvalho de Oliveira

### **Assistentes Editoriais**

Natalia Oliveira

Bruno Oliveira

Flávia Roberta Barão

### **Bibliotecária**

Janaina Ramos

### **Projeto Gráfico e Diagramação**

Natália Sandrini de Azevedo

Camila Alves de Cremo

Luiza Alves Batista

Maria Alice Pinheiro

### **Imagens da Capa**

Shutterstock

### **Edição de Arte**

Luiza Alves Batista

### **Revisão**

Os Autores

2021 by Atena Editora

Copyright © Atena Editora

Copyright do Texto © 2021 Os autores

Copyright da Edição © 2021 Atena Editora

Direitos para esta edição cedidos à Atena Editora pelos autores.



Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição *Creative Commons*. Atribuição-Não-Comercial-NãoDerivativos 4.0 Internacional (CC BY-NC-ND 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores, inclusive não representam necessariamente a posição oficial da Atena Editora. Permitido o *download* da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Todos os manuscritos foram previamente submetidos à avaliação cega pelos pares, membros do Conselho Editorial desta Editora, tendo sido aprovados para a publicação com base em critérios de neutralidade e imparcialidade acadêmica.

A Atena Editora é comprometida em garantir a integridade editorial em todas as etapas do processo de publicação, evitando plágio, dados ou resultados fraudulentos e impedindo que interesses financeiros comprometam os padrões éticos da publicação. Situações suspeitas de má conduta científica serão investigadas sob o mais alto padrão de rigor acadêmico e ético.

### **Conselho Editorial**

#### **Ciências Humanas e Sociais Aplicadas**

Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia do Paraná

Prof. Dr. Américo Junior Nunes da Silva – Universidade do Estado da Bahia

Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná

Prof. Dr. Antonio Gasparetto Júnior – Instituto Federal do Sudeste de Minas Gerais

Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília

Prof. Dr. Carlos Antonio de Souza Moraes – Universidade Federal Fluminense  
Prof. Dr. Crisóstomo Lima do Nascimento – Universidade Federal Fluminense  
Profª Drª Cristina Gaio – Universidade de Lisboa  
Prof. Dr. Daniel Richard Sant’Ana – Universidade de Brasília  
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia  
Profª Drª Dilma Antunes Silva – Universidade Federal de São Paulo  
Prof. Dr. Edvaldo Antunes de Farias – Universidade Estácio de Sá  
Prof. Dr. Elson Ferreira Costa – Universidade do Estado do Pará  
Prof. Dr. Eloi Martins Senhora – Universidade Federal de Roraima  
Prof. Dr. Gustavo Henrique Cepolini Ferreira – Universidade Estadual de Montes Claros  
Profª Drª Ivone Goulart Lopes – Istituto Internazionale delle Figlie de Maria Ausiliatrice  
Prof. Dr. Jadson Correia de Oliveira – Universidade Católica do Salvador  
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense  
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins  
Prof. Dr. Luis Ricardo Fernandes da Costa – Universidade Estadual de Montes Claros  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof. Dr. Marcelo Pereira da Silva – Pontifícia Universidade Católica de Campinas  
Profª Drª Maria Luzia da Silva Santana – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Pablo Ricardo de Lima Falcão – Universidade de Pernambuco  
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa  
Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa  
Prof. Dr. Rui Maia Diamantino – Universidade Salvador  
Prof. Dr. Saulo Cerqueira de Aguiar Soares – Universidade Federal do Piauí  
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Profª Drª Vanessa Ribeiro Simon Cavalcanti – Universidade Católica do Salvador  
Prof. Dr. William Cleber Domingues Silva – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

#### **Ciências Agrárias e Multidisciplinar**

Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano  
Prof. Dr. Arinaldo Pereira da Silva – Universidade Federal do Sul e Sudeste do Pará  
Prof. Dr. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás  
Profª Drª Carla Cristina Bauermann Brasil – Universidade Federal de Santa Maria  
Prof. Dr. Cleberton Correia Santos – Universidade Federal da Grande Dourados  
Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia  
Prof. Dr. Écio Souza Diniz – Universidade Federal de Viçosa  
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul  
Prof. Dr. Fágner Cavalcante Patrocínio dos Santos – Universidade Federal do Ceará  
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia  
Prof. Dr. Jael Soares Batista – Universidade Federal Rural do Semi-Árido  
Prof. Dr. Jayme Augusto Peres – Universidade Estadual do Centro-Oeste  
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro  
Profª Drª Lina Raquel Santos Araújo – Universidade Estadual do Ceará  
Prof. Dr. Pedro Manuel Villa – Universidade Federal de Viçosa  
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão  
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará  
Profª Drª Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

## **Ciências Biológicas e da Saúde**

Prof. Dr. André Ribeiro da Silva – Universidade de Brasília  
Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas  
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás  
Profª Drª Daniela Reis Joaquim de Freitas – Universidade Federal do Piauí  
Profª Drª Débora Luana Ribeiro Pessoa – Universidade Federal do Maranhão  
Prof. Dr. Douglas Siqueira de Almeida Chaves – Universidade Federal Rural do Rio de Janeiro  
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Profª Drª Elizabeth Cordeiro Fernandes – Faculdade Integrada Medicina  
Profª Drª Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília  
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina  
Profª Drª Eysler Gonçalves Maia Brasil – Universidade da Integração Internacional da Lusofonia Afro-Brasileira  
Prof. Dr. Fernando Lima Santos – Universidade Federal do Recôncavo da Bahia  
Profª Drª Fernanda Miguel de Andrade – Universidade Federal de Pernambuco  
Prof. Dr. Fernando Mendes – Instituto Politécnico de Coimbra – Escola Superior de Saúde de Coimbra  
Profª Drª Gabriela Vieira do Amaral – Universidade de Vassouras  
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria  
Prof. Dr. Helio Franklin Rodrigues de Almeida – Universidade Federal de Rondônia  
Profª Drª Iara Lúcia Tescarollo – Universidade São Francisco  
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande  
Prof. Dr. Jefferson Thiago Souza – Universidade Estadual do Ceará  
Prof. Dr. Jesus Rodrigues Lemos – Universidade Federal do Piauí  
Prof. Dr. Jônatas de França Barros – Universidade Federal do Rio Grande do Norte  
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará  
Prof. Dr. Luís Paulo Souza e Souza – Universidade Federal do Amazonas  
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande  
Prof. Dr. Marcus Fernando da Silva Praxedes – Universidade Federal do Recôncavo da Bahia  
Profª Drª Maria Tatiane Gonçalves Sá – Universidade do Estado do Pará  
Profª Drª Mylena Andréa Oliveira Torres – Universidade Ceuma  
Profª Drª Natiéli Piovesan – Instituto Federacl do Rio Grande do Norte  
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá  
Prof. Dr. Rafael Henrique Silva – Hospital Universitário da Universidade Federal da Grande Dourados  
Profª Drª Regiane Luz Carvalho – Centro Universitário das Faculdades Associadas de Ensino  
Profª Drª Renata Mendes de Freitas – Universidade Federal de Juiz de Fora  
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Profª Drª Welma Emidio da Silva – Universidade Federal Rural de Pernambuco

## **Ciências Exatas e da Terra e Engenharias**

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto  
Profª Drª Ana Grasielle Dionísio Corrêa – Universidade Presbiteriana Mackenzie  
Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás  
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná  
Prof. Dr. Cleiseano Emanuel da Silva Paniagua – Instituto Federal de Educação, Ciência e Tecnologia de Goiás  
Prof. Dr. Douglas Gonçalves da Silva – Universidade Estadual do Sudoeste da Bahia  
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná  
Profª Drª Érica de Melo Azevedo – Instituto Federal do Rio de Janeiro  
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará  
Profª Dra. Jéssica Verger Nardeli – Universidade Estadual Paulista Júlio de Mesquita Filho  
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande

Profª Drª Luciana do Nascimento Mendes – Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte  
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá  
Prof. Dr. Marco Aurélio Kistemann Junior – Universidade Federal de Juiz de Fora  
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Profª Drª Priscila Tessmer Scaglioni – Universidade Federal de Pelotas  
Prof. Dr. Sidney Gonçalves de Lima – Universidade Federal do Piauí  
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

#### **Linguística, Letras e Artes**

Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins  
Profª Drª Angeli Rose do Nascimento – Universidade Federal do Estado do Rio de Janeiro  
Profª Drª Carolina Fernandes da Silva Mandaji – Universidade Tecnológica Federal do Paraná  
Profª Drª Denise Rocha – Universidade Federal do Ceará  
Profª Drª Edna Alencar da Silva Rivera – Instituto Federal de São Paulo  
Profª Drª Fernanda Tonelli – Instituto Federal de São Paulo,  
Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões  
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná  
Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia do Paraná  
Profª Drª Miraniide Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará  
Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste  
Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia

#### **Conselho Técnico Científico**

Prof. Me. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo  
Prof. Me. Adalberto Zorzo – Centro Estadual de Educação Tecnológica Paula Souza  
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba  
Prof. Dr. Adilson Tadeu Basquerote Silva – Universidade para o Desenvolvimento do Alto Vale do Itajaí  
Profª Ma. Adriana Regina Vettorazzi Schmitt – Instituto Federal de Santa Catarina  
Prof. Dr. Alex Luis dos Santos – Universidade Federal de Minas Gerais  
Prof. Me. Alexsandro Teixeira Ribeiro – Centro Universitário Internacional  
Profª Ma. Aline Ferreira Antunes – Universidade Federal de Goiás  
Profª Drª Amanda Vasconcelos Guimarães – Universidade Federal de Lavras  
Prof. Me. André Flávio Gonçalves Silva – Universidade Federal do Maranhão  
Profª Ma. Andréa Cristina Marques de Araújo – Universidade Fernando Pessoa  
Profª Drª Andrezza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico  
Profª Drª Andrezza Miguel da Silva – Faculdade da Amazônia  
Profª Ma. Anelisa Mota Gregoleti – Universidade Estadual de Maringá  
Profª Ma. Anne Karynne da Silva Barbosa – Universidade Federal do Maranhão  
Prof. Dr. Antonio Hot Pereira de Faria – Polícia Militar de Minas Gerais  
Prof. Me. Armando Dias Duarte – Universidade Federal de Pernambuco  
Profª Ma. Bianca Camargo Martins – UniCesumar  
Profª Ma. Carolina Shimomura Nanya – Universidade Federal de São Carlos  
Prof. Me. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro  
Prof. Me. Carlos Augusto Zilli – Instituto Federal de Santa Catarina  
Prof. Me. Christopher Smith Bignardi Neves – Universidade Federal do Paraná  
Profª Drª Cláudia de Araújo Marques – Faculdade de Música do Espírito Santo  
Profª Drª Cláudia Taís Siqueira Cagliari – Centro Universitário Dinâmica das Cataratas  
Prof. Me. Clécio Danilo Dias da Silva – Universidade Federal do Rio Grande do Norte  
Prof. Me. Daniel da Silva Miranda – Universidade Federal do Pará  
Profª Ma. Daniela da Silva Rodrigues – Universidade de Brasília  
Profª Ma. Daniela Remião de Macedo – Universidade de Lisboa

Profª Ma. Dayane de Melo Barros – Universidade Federal de Pernambuco  
Prof. Me. Douglas Santos Mezacas – Universidade Estadual de Goiás  
Prof. Me. Edevaldo de Castro Monteiro – Embrapa Agrobiologia  
Prof. Me. Edson Ribeiro de Britto de Almeida Junior – Universidade Estadual de Maringá  
Prof. Me. Eduardo Gomes de Oliveira – Faculdades Unificadas Doctum de Cataguases  
Prof. Me. Eduardo Henrique Ferreira – Faculdade Pitágoras de Londrina  
Prof. Dr. Edwaldo Costa – Marinha do Brasil  
Prof. Me. Eliel Constantino da Silva – Universidade Estadual Paulista Júlio de Mesquita  
Prof. Me. Ernane Rosa Martins – Instituto Federal de Educação, Ciência e Tecnologia de Goiás  
Prof. Me. Euvaldo de Sousa Costa Junior – Prefeitura Municipal de São João do Piauí  
Prof. Dr. Everaldo dos Santos Mendes – Instituto Edith Theresa Hedwing Stein  
Prof. Me. Ezequiel Martins Ferreira – Universidade Federal de Goiás  
Profª Ma. Fabiana Coelho Couto Rocha Corrêa – Centro Universitário Estácio Juiz de Fora  
Prof. Me. Fabiano Eloy Atilio Batista – Universidade Federal de Viçosa  
Prof. Me. Felipe da Costa Negrão – Universidade Federal do Amazonas  
Prof. Me. Francisco Odécio Sales – Instituto Federal do Ceará  
Prof. Me. Francisco Sérgio Lopes Vasconcelos Filho – Universidade Federal do Cariri  
Profª Drª Germana Ponce de Leon Ramírez – Centro Universitário Adventista de São Paulo  
Prof. Me. Gevair Campos – Instituto Mineiro de Agropecuária  
Prof. Me. Givanildo de Oliveira Santos – Secretaria da Educação de Goiás  
Prof. Dr. Guilherme Renato Gomes – Universidade Norte do Paraná  
Prof. Me. Gustavo Krahl – Universidade do Oeste de Santa Catarina  
Prof. Me. Helton Rangel Coutinho Junior – Tribunal de Justiça do Estado do Rio de Janeiro  
Profª Ma. Isabelle Cerqueira Sousa – Universidade de Fortaleza  
Profª Ma. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia  
Prof. Me. Javier Antonio Albornoz – University of Miami and Miami Dade College  
Prof. Me. Jhonatan da Silva Lima – Universidade Federal do Pará  
Prof. Dr. José Carlos da Silva Mendes – Instituto de Psicologia Cognitiva, Desenvolvimento Humano e Social  
Prof. Me. Jose Elyton Batista dos Santos – Universidade Federal de Sergipe  
Prof. Me. José Luiz Leonardo de Araujo Pimenta – Instituto Nacional de Investigación Agropecuaria Uruguay  
Prof. Me. José Messias Ribeiro Júnior – Instituto Federal de Educação Tecnológica de Pernambuco  
Profª Drª Juliana Santana de Curcio – Universidade Federal de Goiás  
Profª Ma. Juliana Thaisa Rodrigues Pacheco – Universidade Estadual de Ponta Grossa  
Profª Drª Kamilly Souza do Vale – Núcleo de Pesquisas Fenomenológicas/UFGA  
Prof. Dr. Kárpio Márcio de Siqueira – Universidade do Estado da Bahia  
Profª Drª Karina de Araújo Dias – Prefeitura Municipal de Florianópolis  
Prof. Dr. Lázaro Castro Silva Nascimento – Laboratório de Fenomenologia & Subjetividade/UFPR  
Prof. Me. Leonardo Tullio – Universidade Estadual de Ponta Grossa  
Profª Ma. Lilian Coelho de Freitas – Instituto Federal do Pará  
Profª Ma. Lilian de Souza – Faculdade de Tecnologia de Itu  
Profª Ma. Liliani Aparecida Sereno Fontes de Medeiros – Consórcio CEDERJ  
Profª Drª Lúvia do Carmo Silva – Universidade Federal de Goiás  
Prof. Dr. Lucio Marques Vieira Souza – Secretaria de Estado da Educação, do Esporte e da Cultura de Sergipe  
Prof. Dr. Luan Vinicius Bernardelli – Universidade Estadual do Paraná  
Profª Ma. Luana Ferreira dos Santos – Universidade Estadual de Santa Cruz  
Profª Ma. Luana Vieira Toledo – Universidade Federal de Viçosa  
Prof. Me. Luis Henrique Almeida Castro – Universidade Federal da Grande Dourados  
Prof. Me. Luiz Renato da Silva Rocha – Faculdade de Música do Espírito Santo  
Profª Ma. Luma Sarai de Oliveira – Universidade Estadual de Campinas  
Prof. Dr. Michel da Costa – Universidade Metropolitana de Santos

Prof. Me. Marcelo da Fonseca Ferreira da Silva – Governo do Estado do Espírito Santo  
Prof. Dr. Marcelo Máximo Purificação – Fundação Integrada Municipal de Ensino Superior  
Prof. Me. Marcos Aurelio Alves e Silva – Instituto Federal de Educação, Ciência e Tecnologia de São Paulo  
Profª Ma. Maria Elanny Damasceno Silva – Universidade Federal do Ceará  
Profª Ma. Marileila Marques Toledo – Universidade Federal dos Vales do Jequitinhonha e Mucuri  
Prof. Dr. Pedro Henrique Abreu Moura – Empresa de Pesquisa Agropecuária de Minas Gerais  
Prof. Me. Pedro Panhoca da Silva – Universidade Presbiteriana Mackenzie  
Profª Drª Poliana Arruda Fajardo – Universidade Federal de São Carlos  
Prof. Me. Rafael Cunha Ferro – Universidade Anhembi Morumbi  
Prof. Me. Ricardo Sérgio da Silva – Universidade Federal de Pernambuco  
Prof. Me. Renan Monteiro do Nascimento – Universidade de Brasília  
Prof. Me. Renato Faria da Gama – Instituto Gama – Medicina Personalizada e Integrativa  
Profª Ma. Renata Luciane Polsaque Young Blood – UniSecal  
Prof. Me. Robson Lucas Soares da Silva – Universidade Federal da Paraíba  
Prof. Me. Sebastião André Barbosa Junior – Universidade Federal Rural de Pernambuco  
Profª Ma. Silene Ribeiro Miranda Barbosa – Consultoria Brasileira de Ensino, Pesquisa e Extensão  
Profª Ma. Solange Aparecida de Souza Monteiro – Instituto Federal de São Paulo  
Profª Ma. Taiane Aparecida Ribeiro Nepomoceno – Universidade Estadual do Oeste do Paraná  
Prof. Me. Tallys Newton Fernandes de Matos – Faculdade Regional Jaguaribana  
Profª Ma. Thatianny Jasmine Castro Martins de Carvalho – Universidade Federal do Piauí  
Prof. Me. Tiago Silvio Dedoné – Colégio ECEL Positivo  
Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista

## Avanços metodológicos em biologia molecular e biotecnologia

**Bibliotecária:** Janaina Ramos  
**Diagramação:** Natália Sandrini de Azevedo  
**Correção:** Giovanna Sandrini de Azevedo  
**Edição de Arte:** Luiza Alves Batista  
**Revisão:** Os Autores  
**Organizadores:** Anderson Barros Archanjo  
Aline Ribeiro Borçoi  
Suzanny Oliveira Mendes  
Adriana Madeira Álvares-da-Silva

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

A946 Avanços metodológicos em biologia molecular e biotecnologia / Organizadores Anderson Barros Archanjo, Aline Ribeiro Borçoi, Suzanny Oliveira Mendes, et al. – Ponta Grossa - PR: Atena, 2021.

Outra organizadora  
Adriana Madeira Álvares-da-Silva

Formato: PDF  
Requisitos de sistema: Adobe Acrobat Reader  
Modo de acesso: World Wide Web  
Inclui bibliografia  
ISBN 978-65-5983-175-3  
DOI 10.22533/at.ed.753210406

1. Biologia Molecular. 2. Biotecnologia. 3. DNA. I. Archanjo, Anderson Barros (Organizador). II. Borçoi, Aline Ribeiro (Organizadora). III. Mendes, Suzanny Oliveira (Organizadora). IV. Título.

CDD 572.8

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

**Atena Editora**

Ponta Grossa – Paraná – Brasil

Telefone: +55 (42) 3323-5493

[www.atenaeditora.com.br](http://www.atenaeditora.com.br)

contato@atenaeditora.com.br

## DECLARAÇÃO DOS AUTORES

Os autores desta obra: 1. Atestam não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao artigo científico publicado; 2. Declaram que participaram ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certificam que os artigos científicos publicados estão completamente isentos de dados e/ou resultados fraudulentos; 4. Confirmam a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhecem terem informado todas as fontes de financiamento recebidas para a consecução da pesquisa.

## APRESENTAÇÃO

O livro “Avanços Metodológicos em Biologia Molecular e Biotecnologia” é uma obra com o foco principal na valorização acerca do conhecimento aprofundado em Biologia Molecular e Biotecnologia, assim como do histórico e das técnicas recentes mais aplicadas para o estudo do tema.

A obra é fruto do trabalho dos professores e alunos da disciplina “Tópicos Especiais em Biotecnologia II: Avanços Metodológicos em Biologia Molecular e Biotecnologia (BIOTEC024)”, ofertada pelo Programa de Pós-Graduação em Biotecnologia da Universidade Federal do Espírito Santo.

Os diversos temas relacionados à Biologia Molecular e à Biotecnologia foram abordados tendo como foco os avanços nas metodologias utilizadas ao longo dos anos. A obra consta de sete capítulos, nos quais foram discutidos a evolução da biologia molecular e das ômicas, assim como os sequenciadores de próxima geração, os métodos de estudo em epigenética, a amplificação do DNA em single-cell, estudo de expressão gênica, ferramentas de edição gênica e expressão heteróloga de proteínas.

Deste modo a obra “Avanços Metodológicos em Biologia Molecular e Biotecnologia” apresenta uma ferramenta de estudo que proporciona aos leitores uma visão ampla das principais metodologias em Biologia Molecular e Biotecnologia, assim como as suas evoluções ao longo dos anos.

Desejamos uma excelente leitura a todos!

Anderson Barros Archanjo

Aline Ribeiro Borçoi

Suzanny Oliveira Mendes

Adriana Madeira Álvares-da-Silva

## SUMÁRIO

<b>LISTA DE ABREVIações E SIGLAS .....</b>	<b>1</b>
<b>CAPÍTULO 1.....</b>	<b>3</b>
A EVOLUÇÃO DA BIOLOGIA MOLECULAR E AS ÔMICAS	
Tamires dos Santos Vieira Suzanny Oliveira Mendes	
 <a href="https://doi.org/10.22533/at.ed.7532104061">https://doi.org/10.22533/at.ed.7532104061</a>	
<b>CAPÍTULO 2.....</b>	<b>14</b>
GENÔMICA E OS SEQUENCIADORES DE PRÓXIMA GERAÇÃO	
Paola Cerbino Doblas Iuri Drumond Louro	
 <a href="https://doi.org/10.22533/at.ed.7532104062">https://doi.org/10.22533/at.ed.7532104062</a>	
<b>CAPÍTULO 3.....</b>	<b>26</b>
EPIGENÉTICA E MÉTODOS DE ESTUDOS EPIGENÉTICOS	
Ivana Alece Arantes Moreno Suzanny Oliveira Mendes	
 <a href="https://doi.org/10.22533/at.ed.7532104063">https://doi.org/10.22533/at.ed.7532104063</a>	
<b>CAPÍTULO 4.....</b>	<b>35</b>
PRÍNCIPIOS BÁSICOS EM AMPLIFICAÇÃO DO DNA E GENÔMICA DE CÉLULAS ÚNICAS	
Bárbara Risse-Quaioto Anderson Barros Archanjo	
 <a href="https://doi.org/10.22533/at.ed.7532104064">https://doi.org/10.22533/at.ed.7532104064</a>	
<b>CAPÍTULO 5.....</b>	<b>46</b>
RT-qPCR: ESTUDOS DE EXPRESSÃO GÊNICA	
Mayara Mota de Oliveira Anderson Barros Archanjo	
 <a href="https://doi.org/10.22533/at.ed.7532104065">https://doi.org/10.22533/at.ed.7532104065</a>	
<b>CAPÍTULO 6.....</b>	<b>58</b>
FERRAMENTAS DE EDIÇÃO GÊNICA	
Joaquim Gasparini dos Santos Aline Ribeiro Borçoi	
 <a href="https://doi.org/10.22533/at.ed.7532104066">https://doi.org/10.22533/at.ed.7532104066</a>	
<b>CAPÍTULO 7.....</b>	<b>69</b>
EXPRESSÃO HETERÓLOGA DE PROTEÍNAS	
Karolinni Bianchi Britto	

Greiciane Gaburro Paneto

 <https://doi.org/10.22533/at.ed.7532104067>

<b>SOBRE OS ORGANIZADORES .....</b>	<b>85</b>
-------------------------------------	-----------

## LISTA DE ABREVIÇÕES E SIGLAS

2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
5hmC	5 hidroximetilcitosina
5mC	5-metilcitosina
AD	Domínio de Ativação
AMPK	Proteína Quinase Ativada por Adenosina Monofosfato
BACs	Cromossomos Artificiais Bacterianos
BiTS-ChiP	Isolamento Descontínuo da Cromatina Específica de Tecido para Imunoprecipitação
cDNA	DNA complementar
ChiP	Imunoprecipitação da Cromatina
ChIP-chip	Imunoprecipitação de Cromatina Ligada a Microarranjos
ChIP-Seq	Imunoprecipitação da Cromatina com Tecnologia de Sequenciamento de Nova Geração
CRISPR	Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespçadas (do inglês <i>Clustered Regularly Interspaced Short Palindromic Repeats</i> )
DCR	Doença Renal Crônica
DNA	Ácido Desoxirribonucleico
DNAr	DNA recombinante
DNTPs	Desoxirribonucleotídeos Fosfatados
DOP-PCR	Reação em Cadeia da Polimerase Iniciada por Oligonucleotídeo Degenerado
dPCR	PCR digital
DSB	Quebra Dupla da Fita de DNA (do inglês <i>Double-Stranded Break</i> )
ESI-MS	Espectrometria de Massa de Eletrospray
FACS	Classificação Celular Ativada por Fluorescência
FDA	Food and Drug Administration
<i>FOK I</i>	<i>Flavobacterium okeanokoites</i>
FRET	Transferência de Energia de Ressonância de Fluorescência
FTIR	Espectroscopia Infravermelha de Transformação Fourier
GAPDH	Gliceraldeído-3-fosfato Desidrogenase
GFP	Proteína Verde Fluorescente
GST	Glutathione S-transferase
HATs	Histonas Acetiltransferases
HDACs	Histonas Desacetiltransferases
HPLC	Cromatografia Líquida de Alta Resolução
HR	Recombinação Homóloga (do inglês <i>Homologous Recombination</i> )
INTACT	Isolamento de Núcleos Marcados em Tipo Celular Específico
LCM	Microdissecção de Captura a Laser
Lpp	Lipoproteína Mureína
MACS	Classificação de Células Ativadas Magneticamente
MALBAC	Múltiplos Ciclos de Amplificação Baseados em Anelamento e Loop
MBP	Proteína Ligante da Maltose
MDA	Amplificação de Deslocamento Múltiplo
miRNA	microRNA

NGS	Sequenciadores de Próxima Geração
NHEJ	Junções Terminais não Homólogas (do inglês <i>non-homologous end joining</i> )
NLS	Sinal de Localização Nuclear
NMR	Ressonância Magnética Nuclear
OGM	Organismos Geneticamente Modificados
OmpA	Proteína A da membrana externa
OmpT	Protease VII
PCR	Reação em Cadeia da Polimerase
PGM	Personal Genome Machine
PhoE	Proteína E dos poros da membrana externa
qPCR	Reação Quantitativa em Cadeia da Polimerase (qPCR)
RBS	Sequência Shine-Dalgarno (do inglês <i>Ribosome Binding Site</i> )
RFLP	Polimorfismo de Comprimento de Fragmento de Restrição (RFLP)
RNA	Ácido ribonucleico
RNA <sub>m</sub>	RNA mensageiro
RNA-seq	Sequenciamento de RNA
RT	Transcriptase reversa
RT-PCT	PCR de transcrição reversa
RT-qPCR	PCR de transcrição reversa quantitativa
RVS	Repetições Variáveis Diresíduos
Sars-cov-2	Severe acute respiratory syndrome coronavirus 2
SDS-PAGE de	Eletroforese em Gel de Poliacrilamida na presença de Dodecil Sulfato Sódio
SNP	Polimorfismo de Nucleotídeo Único
STR	Sequências de Repetições Curtas em Tandem
TALENs	Nucleases com Efetores do tipo Ativador Transcricional (do inglês <i>Transcription Activator-Like Effector Nucleases</i> )
TET	Translocação Metilcitosina Dioxigenase
TFIIIA	Fator de transcrição IIIA
T <sub>m</sub>	Temperatura de <i>melting</i>
TRX	Tiorredoxina
YACs	Cromossomos Artificiais de Leveduras
ZFN	Nucleases de Dedo de Zinco (do inglês <i>Zinc Finger Nucelases</i> )

## A EVOLUÇÃO DA BIOLOGIA MOLECULAR E AS ÔMICAS

*Data de aceite:* 10/05/2021

*Data de submissão:* 14/04/2021

### Tamires dos Santos Vieira

Universidade Federal do Espírito Santo, Vitória  
<http://lattes.cnpq.br/9014510901567961>  
<https://orcid.org/0000-0002-3899-3664>

### Suzanny Oliveira Mendes

Universidade Federal do Espírito Santo, Vitória  
<http://lattes.cnpq.br/5613486906366786>  
<https://orcid.org/0000-0001-8660-5139>

**RESUMO:** As ciências ômicas objetivam identificar, quantificar e caracterizar os processos celulares a fim de buscar entendimento das vias intracelulares. Por meio de técnicas que adotam uma visão abrangente da biologia molecular visando a identificação por meio da genômica, transcriptômica, proteômica e metabolômica. Essas diferentes técnicas que evoluíram ao longo dos anos permitiram que mais e mais conhecimento sobre essa maquinaria biológica fosse aos poucos decifrada, permitindo que esse sistema complexo pudesse ser compreendido. As tecnologias Ômicas podem ser aplicadas na busca por respostas dos diversos processos fisiológicos na saúde e na doença sendo para esse último capaz de desempenham o diagnóstico, prognóstico e compreensão na etiologia. Nesse presente capítulo foram compiladas informações sobre a evolução das ciências ômicas abrangendo as diversas áreas da Biologia, por meio de buscas em periódicos indexados. Deste modo, foi possível discutir, a integração de diversas ferramentas no entendimento da Biologia de Sistemas e definir os

mecanismos e funcionalidade dentro da biologia molecular.

**PALAVRAS-CHAVE:** Biologia de sistemas, biotecnologia, técnicas ômicas.

### THE EVOLUTION MOLECULAR BIOLOGY AND OMICS

**ABSTRACT:** Omic Sciences aim the to identificfy, quantify and characterize cellular processes, in order to understand to intracellular channels. By means of techniques that adopt a comprehensive vision of Molecular Biology aiming to identify by genomic, transcriptomic, proteomic and metabolomic means. These different techniques which evolved throughout the years allowed more and more knowledge about this biological machinery to be gradually deciphered. Omic science can be applied in the search for answers to several physiological processes of health and disease, for this last able to perform a diagnostic, prognostic in etiology. In presente chapter was compiled information about the evolution omic Sciences comprehensive to several áreas on biology, through searches in indexed journals. Thus, it was possible to discuss integration to several tools in understanding to Systems Biology and the mechanisms and functionality in side molecular biology.

**KEYWORDS:** Systems Biology, biotechnology, omic techniques.

### 1 | INTRODUÇÃO

As Ômicas representam análises sistêmicas das várias vias biológicas que objetivam identificar, quantificar e caracterizar os processos celulares a fim de buscar entendimento das vias intracelulares (SOUZA; RHODEN;

PAMPHILE, 2014). Logo as tecnologias que envolvem as Ciências Ômicas adotam uma visão abrangente da biologia molecular visando a identificação por meio da genômica, transcriptômica, proteômica e metabolômica, sendo a reunião dessas várias técnicas chamadas de biologia de sistemas (HORGAN; KENNY, 2011; ARIAS; FREIRE, 2008).

Mas antes de iniciarmos sobre a discussão acerca das ômicas é essencial relembremos fatos que foram importantes para que todos esses conjuntos de saberes fossem incorporados no nosso dia a dia. Desde a descoberta dos Ácidos Desoxirribonucleicos (DNA) até os dias de hoje muitos pesquisadores trabalharam, cada um em seu tempo e em seus laboratórios, para que esses conhecimentos fossem possíveis e disponibilizados nos dias de hoje.

Em 1944, Oswald Avery, Colin MacLeod e MaLyn McCarty em suas pesquisas permitiram-se conhecer a existência do DNA (AVERY; MACLEOD; MCCARTY, 1944). A partir deste ponto nos anos que seguiram, a busca dos cientistas em responder todas as perguntas aumentou. Em 1951, Rosalind Franklin aprimorou o uso da cristalografia de raios-X para criar imagens de matérias microscópicas, extraindo DNA e analisando com raios-X expondo assim que não havia apenas uma forma da molécula, e sim duas fitas em dupla hélice (ORTIZ; SILVA, 2016).

Em 1952 Alfred Hershey e Martha Chase demonstram que o DNA realmente era a molécula que transmitia as informações herdáveis. Mais tarde, em 1953, Watson e Crick baseando-se em estudos de outros pesquisadores como os dados cristalográficos de raios-X de Rosalind Franklin e Maurice Wilkins do King's College, descreveram a estrutura do DNA composta por duas cadeias complementares unificadas em direções inversas (FRIGOLET; GUTIÉRREZ-AGUILAR, 2017).

As descobertas envolvendo o DNA se sobrepuseram com a descoberta da sequência de aminoácidos que compõem as proteínas e do RNA e os conhecimentos foram assim sendo construídos. Logo os trabalhos posteriores objetivaram sequenciar o material genético, no entanto, esbararam em diversas limitações tecnológicas.

Fred Sanger realizou os primeiros sequenciamentos, não do DNA, mas sim de proteínas mostrando que proteínas tinham padrões definidos de resíduos de aminoácidos e que as proteínas formadas variavam entre as espécies e entre os indivíduos. No final dos anos 1960 muitas proteínas já haviam sido sequenciadas, bem como os estudos com RNAs, que também progrediram de 1968 a 1973. Contudo, em 1977 foi possível conhecer a ordem das bases de nucleotídeos, servindo de base para a análise e sequenciamento do genoma (SHENDURE et al., 2017). No entanto, com desenvolvimento tecnológico e científico alcançado nos anos seguintes, foi em 2001 que o genoma humano inteiro finalmente foi sequenciado (FRIGOLET; GUTIÉRREZ-AGUILAR, 2017).

À medida em que o conhecimento científico e tecnológico foi sendo desenvolvido e aprimorado tornou-se possível sequenciar o genoma inteiro e assim buscar entender cada vez mais sobre as doenças e como elas são distribuídas na população (BEDIA, 2018; SHENDURE et al., 2017).

Deste modo, as ciências ômicas permitiram que mais e mais conhecimento sobre essa maquinaria fosse aos poucos decifrada, permitindo que esse sistema complexo pudesse ser compreendido. As tecnologias Ômicas podem ser aplicadas na busca por

respostas dos diversos processos fisiológicos na saúde e na doença sendo para esse último capaz de desempenhar o diagnóstico, prognóstico e compreensão na etiologia (HORGAN; KENNY, 2011).

Assim iniciamos nosso capítulo discorrendo sobre as ciências ômicas divididas entre as suas partes, genômica, transcriptômica, proteômica e metabolômica. Além disso, ao final discorreremos sobre as novas subáreas de conhecimento que surgiram e foram incorporadas nessas fontes de saberes.

## 2 | GENÔMICA

A genômica compreende o estudo da sequência de DNA e segundo Horgan e Kenny (2011) os genes sempre foram analisados individualmente e o advento das técnicas envolvendo microarranjos permitiu-se viabilizar o avanço desta ômica possibilitando medir tanto as diferenças na sequência de DNA entre os indivíduos quanto a expressão de milhares de genes simultaneamente.

Logo a disponibilidade de diversas tecnologias de sequenciamento contribuiu para um enorme aumento no número de sequências do genoma disponíveis publicamente. Principalmente no campo da “genética populacional” preocupada com a variação genética dentro e entre as populações, relacionada à migração e demografia (BRAGG et al., 2015). Os dados gerados permitiram conhecer as características de diversas populações e inferir sobre aparecimento de doenças e rastreamento de algumas espécies.

Do mesmo modo possibilitou conhecer diversos tipos de variações que ocorrem normalmente e que podem trazer malefícios ou não ao indivíduo. Dentre essas variações que podem ocorrer na natureza encontra-se: variações de nucleotídeo único, com pequenas inserções/deleções, variações estruturais que incluem indels grandes; variantes de número de cópias e inversões que podem afetar a expressão e a sequência de proteínas.

Essas variações de nucleotídeos únicos também conhecidas como polimorfismos de nucleotídeo único (SNPs) podem ocorrer com frequência maiores que 1% na população (MANZONI et al., 2018). E ocorrem quando um nucleotídeo é substituído por outro e essa substituição pode significar alteração de um códon para um aminoácido diferente e assim a proteína formada pode ser inativada ou funcionar erroneamente.

Existe um interesse particular quando associados a doenças com uma determinação genética. O perfil único de polimorfismo de nucleotídeos também tem um papel na farmacogenômica e nutrigenômica na exploração de respostas individuais de pacientes a medicamentos e dietas (HORGAN; KENNY, 2011).

As técnicas empregadas na genômica são também utilizadas em outras ômicas como a transcriptômica, a qual tornou-se a camada da biologia de sistema mais analítica. Sendo assim as técnicas compreendem a reação quantitativa em cadeia da polimerase (qPCR) e arranjos de DNA e sequenciamento de próxima geração (NGS) (BEDIA, 2018).

As técnicas e conhecimentos da genômica podem ser aplicadas em diversas áreas de conhecimentos desde a seleção de medicamentos a definição de melhor tratamento e prognóstico. Na área de medicamentos para transfusões a testagem do DNA é aplicado

à tipagem tornando assim a disponibilidade de doadores apropriados e viáveis (CHOU; WESTHOFF, 2017).

Outras vias de aplicação desta ômica ocorre em doenças de tratamentos mais complexos como mieloma múltiplo, doença com características heterogênicas devido às aberrações genéticas que desempenham papéis diretos no prognóstico e na resposta ao tratamento (SOEKOJO et al., 2018), sendo assim torna-se possível prevenir complicações da doença ao se conhecer mais sobre as aberrações encontradas no genoma. Outro exemplo de aplicação da genômica envolvem as doenças inflamatórias intestinais poligênicas, cujo objetivo é diagnosticar pacientes que apresentem fenótipos extremos e assim pensar em terapias cada vez mais específicas e personalizadas (UHLIG; MUISE, 2017).

O risco do aparecimento de doenças é multifatorial e a compreensão dos determinantes genéticos de risco entre os indivíduos de diferentes ascendências oferece novas oportunidades para diferentes terapias, assim como melhor compreensão do risco de doenças na população geral (GURDASANI et al., 2019).

Logo, o sequenciamento do genoma humano permitiu estudar os diversos mecanismos biológicos humanos em relação a fatores ambientais como drogas, poluentes e a dietas e ainda estudar sobre a prevenção das doenças (ARIAS; FREIRE, 2008) e vislumbrar uma medicina mais individualizada e mais eficaz. Apesar disso, a genômica não fornece informações sobre os mecanismos biológicos e não reflete à regulação da expressão genica e o efeito ambiental na célula (BEDIA, 2018), para isso, outras técnicas precisaram ser desenvolvidas para que cada vez mais se aprofunde o conhecimento acerca da biologia de sistemas.

### 3 | TRANSCRIPTÔMICA

Antes de iniciarmos discutindo sobre a transcriptômica precisamos voltar ao que se entende por Dogma Central da Biologia Celular. De acordo com algumas pressuposições a partir da sequência codificadora de nucleotídeos o RNA mensageiro seria gerado, e posteriormente traduzido em uma sequência de aminoácidos que gerariam proteínas, definindo assim o caminho de informações em uma célula do DNA para o RNA e posteriormente para a proteína.

No entanto, a esse respeito diversos pesquisadores discordam ao afirmarem, como é dissertado por Camacho e colaboradores (2019) em sua revisão, que não é possível compreender totalmente o universo de reações que ocorrem desde o DNA até a expressão da proteína por sua alta complexidade. Parte desse enredamento seria a existência de RNA's não codificantes que poderiam exercer ou não funções intracelulares bem como sofrerem modificações pós transcricionais e traducionais como os *Splicings* alternativos e modificações de proteínas, respectivamente.

Sendo assim, na transcriptômica o objetivo é identificar e quantificar os diferentes transcritos de genes representando uma camada muito dinâmica devido aos processos de transcrição específicos, contínuos e reflexos da atividade das células e suas respostas a estímulos externos (BEDIA, 2018). Compreendendo um elemento essencial para estudar mecanismos funcionais do genoma e desencadeadores de doenças.

O sequenciamento de RNA (RNA-seq) considerado padrão ouro em plataformas de sequenciamento de próxima geração (NGS) possibilitou a análise de milhões de transcritos de amostras biológicas. Podendo ser empregados em diversos estudos de transcriptômica nas mais diversas áreas como em animais de produção (MACHADO et al., 2018), no auxílio do diagnóstico de cânceres e outras doenças (LOMAS-SORIA et al., 2018; TIXIER et al., 2020). Gerando tantos transcritos que assim como na genômica se faz necessário o aprofundamento em estudos de softwares e bioinformática para compreender melhor essas relações existentes.

A partir do estudo da expressão dos RNAs transcritos, torna-se possível conhecer e compreender melhor a biologia de sistemas e como o meio ambiente pode inter-relacionar-se com o interior celular por meio da modificação da expressão gênica (ESPINDOLA et al., 2010).

Em seu estudo Zang et al. (2019) realizaram o sequenciamento de RNA usando tecidos hepáticos de peixe-zebra adultos e descobriram que extrato de chá verde pôde melhorar os fenótipos obesos por meio da ativação da sinalização da via Wnt/ $\beta$ -catenina e proteína quinase ativada por adenosina monofosfato (AMPK). Além disso, a análise comparativa do transcriptoma revelou que peixes-zebra e mamíferos podem compartilhar uma resposta molecular comum.

Mais recentemente com a busca por métodos eficazes de diagnóstico do novo coronavírus (Sars-cov-2) as técnicas de transcriptômica foram sendo implementadas a fim de conhecer e diagnosticar a doença em casos assintomáticos e até ambientalmente. De modo que a aplicação dessas técnicas possibilitaria implementar protocolos de rastreio a fim de traçar casos e fornecer uma visão abrangente da biologia viral e do hospedeiro podendo auxiliar na mitigação de danos médicos e socioeconômicos, bem como estabelecer redes de vigilância protetora que podem ajudar a se defender contra futuras pandemias (BUTLER et al., 2020). Nessa doença pesquisadores estudaram três módulos de expressão gênica representando os principais sinalizadores de processos inflamatórios que podem ser alvo de medicamentos aprovados (ISLAM1; FISCHER, 2020).

Na área do câncer os estudos de transcriptômica buscam associar a expressão gênica a exames de imagens a fim de fechar um diagnóstico melhor e prever um prognóstico mais certo a fim de personalizar os tratamentos por meio de processos moleculares específicos de tumores (TIXIER et al., 2020). Na obesidade essas técnicas também têm gerado trabalhos interessantes que relacionam a obesidade materna à predisposição da prole à obesidade, resistência insulínica, doença hepática gordurosa não alcoólica a partir dos estudos do transcriptoma hepático usando RNA-seq (LOMAS-SORIA et al., 2018). Esses estudos direcionam cada vez mais a ciência a desvendar os possíveis desencadeadores de doenças.

## 4 | PROTEÔMICA

Proteômica surge como uma tecnologia indispensável para interpretar a informação codificada nos genomas, sendo uma técnica empregada para explicar as classes proteicas contidas em uma célula ou tecido, usada para desvendar mudanças de proteínas entre os

estados saudável e doente (ARIAS; FREIRE, 2008).

O interesse no estudo proteômico parte de buscar entendimento sobre como, onde, quando e por que centenas de proteínas são produzidas e sobre quais mecanismos estão envolvidos nesses processos. Logo a proteômica é considerada uma metodologia que parte de uma análise integrativa e multidisciplinar desde os genes até o fenótipo expresso nas proteínas. Estas análises podem envolver tecnologias “ômicas” e estudos a partir de banco de dados (ESPINDOLA et al., 2010).

Um único gene pode originar vários tipos diferentes de proteínas graças às modificações pós-transcricionais e até mesmo pós traducionais que podem modificar, ativar ou inativar proteínas.

As modificações que as proteínas traduzidas podem sofrer segundo Frigolet e Gutiérrez-Aguilar (2017) compreendem cortes, fosforilação, glicosilação, sumoolização, que alteram a estrutura, controlam a formação de complexos funcionais de proteínas, regulam a atividade das proteínas as tornando-as ativas ou inativas.

A metodologia para o estudo da proteômica consiste principalmente em: separação de proteínas por técnicas cromatográficas (líquidas, gasosas, etc.) ou eletroforéticas (2D-PAGE); digestão proteica; detecção de fragmentos peptídicos por espectrometria de massa; identificação de proteínas. Ambas as técnica consistem em posteriormente realizar leitura por meio de um software (BARBOSA et al., 2012).

Em estudos direcionados ao câncer de mama, por exemplo, a proteômica pode ser empregada para avaliar os efeitos de mutações genéticas a fim de elucidar as consequências funcionais de mutações somáticas e identifica possíveis alvos terapêuticos (MERTINS et al., 2016).

Logo a aplicação da proteômica à clínica permite encontrar biomarcadores capazes de diagnosticar doenças, possibilitando oferecer tratamento mais adequado no futuro a partir da identificação de proteínas e identificação de um perfil proteico e sua interação em uma rede funcional (FRIGOLET; GUTIÉRREZ-AGUILAR, 2017).

## 5 | METABOLÔMICA

O termo metabolômica foi criado e tem sido usado para abranger o estudo do metabolismo sob perturbações ambientais e genéticas. Logo pode ser definida como o estudo da influência da expressão gênica nos metabólitos das células, tecido ou órgãos atuando em várias funções celulares. Podendo sofrer conversões enzimáticas e químicas, compondo macromoléculas ou mesmo reservas energéticas (ARIAS; FREIRE, 2008). E devido a isto a identificação, quantificação e reações dos metabólitos são importantes no contexto da biologia de sistemas (CANUTO et al., 2017).

A metabolômica compreende a ômica que mais necessita de estudos computacionais devido à variabilidade de resultado com difícil interpretação. Geralmente são ricos em dados, sendo necessário o uso de ferramentas estatísticas e de bioinformática para avaliação e sistematização dos dados, em que propriedades bioquímicas e relações celulares podem ser mapeadas em plataformas de software que podem reforçar a interpretabilidade dos

dados (ESPINDOLA et al., 2010).

Os métodos para obtenção dos metabólitos utilizam técnicas de separação cromatográficas clássicas que podem compreender Espectroscopia Infravermelha de Transformação Fourier (FTIR), Espectrometria de Massa de Eletrospray (ESI-MS) e Espectroscopia de Ressonância Magnética Nuclear (NMR) (FRIGOLET; GUTIÉRREZ-AGUILAR, 2017). Devido à facilidade de preparação da amostra e a capacidade de quantificar os diferentes níveis de metabólitos produzidos no interior da célula e presente no organismo NMR torna-se uma plataforma preferida para estudos metabólicos clínicos de longo prazo ou em larga escala (EMWAS et al., 2019).

A metabolômica pode ser empregada em diversos estudos inclusive combinada com outras ômicas em doenças como doença renal crônica (DRC), obesidade, diabetes e câncer. Essas doenças afligem muitas pessoas e há necessidade de melhores biomarcadores para diagnóstico, monitoramento da doença e tratamento fatores relevantes para a tomada de decisões terapêuticas individualizadas (HOCHER; ADAMSKI, 2017).

No diagnóstico do câncer o estudo da metabolômica pode direcionar ao diagnóstico do tipo de tumor sendo ele benigno ou maligno a partir da distinção dos metabólitos em casos recursivos e que responderam positivamente a terapia. No câncer de próstata alterações específicas da carcinogênese e progressão podem ajudar na representação de potenciais biomarcadores metabólicos (KELLY et al., 2016).

Em relação aos estudos envolvendo à dieta a metabolômica também é empregada para investigar a associação entre nutrição e estado de saúde e sua aplicação em estudos epidemiológicos que permitem a identificação de metabólitos purificados de boa qualidade no sangue e no plasma (HOCHER; ADAMSKI, 2017).

A compreensão e o estudo dessa ômicas aliada as demais aqui citadas permitem o completo entendimento da biologia de sistemas, as quais podem ser empregadas nas mais diferentes áreas.

## **6 I OUTRAS DISCIPLINAS DERIVADAS DAS ÔMICAS**

Ao final desse capítulo discorreremos conceitualmente sobre as ômicas sobre a farmacogenética e nutrigenética que são ômicas que derivaram dos estudos das genômicas e que são empregadas em conjuntos com uma ou mais ômicas. Ambas são poderosas ferramentas da biologia molecular e impulsionadas por estudos de diversas áreas.

### **6.1 Farmacogenética**

Essa área de estudo permite a partir da utilização de técnicas genômicas identificar bases genéticas da variação interindividual na eficácia, metabolismo e transporte de fármacos (FRIGOLET; GUTIÉRREZ-AGUILAR, 2017), técnicas genômicas estas que compreendem: sequenciamento de DNA, mapeamento genético e a bioinformática.

O estudo da variabilidade individual na resposta a medicamentos pode ser usado para individualizar e potencializar o tratamento medicamentoso sendo importante na oncologia e em outras doenças que visam um tratamento mais eficiente e preciso (HORGAN; KENNY, 2011).

O conhecimento sobre a assertividade individualizada de fármacos em abordagens sistêmicas a condições como o câncer e doenças cardiovasculares e a obesidade compreendem excelentes oportunidade de facilitar muito o sucesso da seleção de novos alvos para tratamentos e desenvolvimento de medicamentos e conseqüentemente o sucesso no tratamento.

Um exemplo dessa aplicação corresponde à terapia para o tratamento com antidepressivos e dor crônica em que encontrar uma variante relacionada à metabolização tornou-se essencial para facilitar o atendimento clínico adequado (SMITH et al., 2018). Sendo assim, o estudo farmacogenético direciona a intervenção para a seleção dos alvos mais precisos a fim de priorizar a ligação do fármaco ao alvo e assim resultar em melhor efeito a uma doença específica possibilitando tanto o diagnóstico quanto as decisões terapêuticas (LOPEZ, 2018).

## 6.2 Nutrigenética

Na nutrição, os saberes que surgiram após o projeto genoma humano possibilitaram o desenvolvimento de dois conceitos de estudo, a nutrigenética e nutrigenômica. Embora esses dois conceitos sejam vistos em conjunto, o termo nutrigenética refere-se às interações entre hábitos dietéticos e o perfil genético de cada indivíduo. A nutrigenética é baseada em observações das respostas individuais à determinada modificação na dieta e em que estas diferentes respostas sejam associadas à presença ou ausência de marcadores biológicos como os polimorfismos genéticos capazes de prever a resposta individual à dieta (KAUR; ALLAHBADIA, 2018).

O termo nutrigenômica refere-se às influências de fatores dietéticos sobre o epigenoma, essa combinação de nutrição e genômica, biologia computacional e bioinformática pode auxiliar no combate a doenças crônicas contribuindo para diminuir o risco de vida relacionado (KAUR; ALLAHBADIA, 2018).

Essas áreas visam buscar entendimento sobre a resposta interindividual de intervenções dietéticas ou a hábitos alimentares (SCHUCH et al., 2010). Segundo Fournier, Poulain e Jacob (2019) a nutrigenética e nutrigenômica referem-se, respectivamente ao estudo da influência dos genes sobre as respostas às práticas alimentares e a influência das práticas alimentares sobre a expressão dos genes.

É fundamental a aplicação das ômicas na área de nutrição podendo ser em cultura de células, modelos animais, estudos pré-clínicos e clínicos. Possibilitando a identificação de biomarcadores que respondam especificamente a um determinado nutriente ou composto bioativo oriundo dos alimentos a fim de refinar as recomendações dietéticas individuais para redução de doenças crônicas não transmissíveis e garantindo a promoção da saúde (FIALHO; MORENO; ONG, 2008).

## 7 | CONSIDERAÇÕES FINAIS

O presente capítulo abrangeu as diversas áreas da Biologia que visam integrar, visualizar e modelar as informações das células a fim de gerar um conhecimento amplo possibilitando novas descobertas, tratamentos e diagnósticos. Deste modo, as ciências

Ômicas e as tecnologias que as envolvem como pudemos discutir, integram as ferramentas para o entendimento da Biologia de Sistemas, contribuindo assim para o entendimento necessário para definir os mecanismos e funcionalidade dentro da biologia molecular.

## REFERÊNCIAS

ARIAS, A. R. L.; FREIRE, M. M. **Contribution of the “omics” sciences to human health: advances and challenges.** Caderno de Saúde Coletiva, v. 16, n. 4, p. 717–732, 2008.

AVERY, O. T.; MACLEOD, C. M.; MCCARTY, M. **Studies on the chemical nature of the substance inducing transformation of Pneumococcal Types.** The Journal of Experimental Medicine, v. 79, n. 2, p. 137–158, 1944.

BARBOSA, E. B. *et al.* **Proteômica: metodologias e aplicações no estudo de doenças humanas.** Revista da Associação Médica Brasileira, v. 58, n. 3, p. 366–375, 2012.

BEDIA, C. **Experimental Approaches in Omic Sciences.** In: Comprehensive Analytical Chemistry. 1. ed. Elsevier . v. 82, p. 13–36, 2018.

BRAGG, J. G. *et al.* **Genomic variation across landscapes: insights and applications.** New Phytologist, v. 207, n. 4, p. 953–967, 2015.

BUTLER, D. J. *et al.* **Shotgun Transcriptome and Isothermal Profiling of SARS CoV-2 Infection Reveals Unique Host Responses, Viral Diversification, and Drug Interactions.** bioRxiv, 2020.04.20.048066, 2020.

CAMACHO, M. P. **The Central Dogma Is Empirically Inadequate... No Matter How We Slice It.** Philosophy, Theory, and Practice in Biology, v. 11, n. 20, p. 1–15, 2019.

CANUTO, G. *et al.* **Metabolômica: definições, estado-da-arte e aplicações representativas.** Química Nova, v. 41, n. 1, p. 75–91, 2017.

CHOU, S. T.; WESTHOFF, C. M. **Application of genomics for transfusion therapy in sickle cell anemia.** Blood Cells, Molecules, and Diseases, v. 67, p. 148–154, 2017.

EMWAS, A.-H. *et al.* **NMR Spectroscopy for Metabolomics Research.** Metabolites, v. 9, n. 7, p. 123, 2019.

ESPINDOLA, F. S. *et al.* **Bioinformatic resources applied on the omic sciences as genomic, transcriptomic, proteomic, interatomic and metabolomic.** Bioscience Journal, v. 26, p. 463–477, 2010.

FIALHO, E.; MORENO, F. S.; ONG, T. P. **Nutrição no pós-genoma: fundamentos e aplicações de ferramentas ômicas.** Revista de Nutrição, v. 21, n. 6, p. 757–766, 2008.

FOURNIER, T.; POULAIN, J.; JACOB, M. **Genômica nutricional: (re)considerando as relações alimentação-saúde via religação das ciências sociais, biomédicas e da vida.** INTER-LEGERE, v. 2, n. 25, p. 1–21, 2019.

FRIGOLET, M. E.; GUTIÉRREZ-AGUILAR, R. **Ciências “ômicas”, ¿como ayudan a las ciencias de la salud?** Revista Digital Universitaria, v. 18, n. 7, p. 0–15, 2017.

- GURDASANI, D. *et al.* **Genomics of disease risk in globally diverse populations.** *Nature Reviews Genetics*, v. 20, n. 9, p. 520–535, 2019.
- HOCHER, B.; ADAMSKI, J. **Metabolomics for clinical use and research in chronic kidney disease.** *Nature Reviews Nephrology*, v. 13, n. 5, p. 269–284, 2017.
- HORGAN, R. P.; KENNY, L. C. **‘Omic’ technologies: genomics, transcriptomics, proteomics and metabolomics.** *The Obstetrician & Gynaecologist*, v. 13, n. 3, p. 189–195, 2011.
- ISLAM1, R.; FISCHE, A. **A transcriptome analysis identifies potential preventive and therapeutic 1 approaches towards COVID-19.** *Rezaul Islam*. p. 1–21, 2020.
- KAUR, K. K.; ALLAHBADIA, G. **Impact of Nutrigenomics on Various Metabolic Disorders in Relation to Life Style Alteration.** *J Nutri Food Sci*. v. 6, n. 1, p. 1–10, 2018.
- KELLY, R. S. *et al.* **Metabolomic Biomarkers of Prostate Cancer: Prediction, Diagnosis, Progression, Prognosis, and Recurrence.** *Cancer Epidemiology Biomarkers & Prevention*, v. 25, n. 6, p. 887–906, 2016.
- LOMAS-SORIA, C. *et al.* **Maternal obesity has sex-dependent effects on insulin, glucose and lipid metabolism and the liver transcriptome in young adult rat offspring.** *The Journal of Physiology*, v. 596, n. 19, p. 4611–4628, 2018.
- LOPEZ, D. **Pharmacogenetics: An Important Part of Drug Development with A Focus on Its Application.** *International Journal of Biomedical Investigation*, v. 1, n. 2, p. 1–16, 2018.
- MACHADO, A. *et al.* **“Out of the Can”: A Draft Genome Assembly, Liver Transcriptome, and Nutrigenomics of the European Sardine, *Sardina pilchardus*.** *Genes*, v. 9, n. 10, p. 485, 2018.
- MANZONI, C. *et al.* **Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences.** *Briefings in Bioinformatics*, v. 19, n. 2, p. 286–302, 2018.
- MERTINS, P. *et al.* **Proteogenomics connects somatic mutations to signalling in breast cancer.** *Nature*, v. 534, n. 7605, p. 55–62, 2016.
- ORTIZ, E.; SILVA, M. R. D. **O uso de abordagens da história da ciência no ensino de biologia: uma proposta para trabalhar a participação da cientista Rosalind Franklin na construção do modelo da dupla hélice do DNA.** *Investigações em Ensino de Ciências*, v. 21, n. 1, p. 106, 2016.
- SCHUCH, J. B. *et al.* **Nutrigenética: a interação entre hábitos alimentares e o perfil genético individual.** *Revista Brasileira de Biociências*, v. 4849, n. 2755, p. 73–84, 2010.
- SHENDURE, J. *et al.* **DNA sequencing at 40: past, present and future.** *Nature*, v. 550, n. 7676, p. 345–353, 2017.
- SMITH, D. M. *et al.* **Clinical application of pharmacogenetics in pain management.** *Personalized Medicine*, v. 15, n. 2, p. 117–126, 2018.
- SOEKOJO, C. *et al.* **Potential Clinical Application of Genomics in Multiple Myeloma.** *International Journal of Molecular Sciences*, v. 19, n. 6, p. 1721, 2018.
- SOUZA, L. D. L.; RHODEN, S. A.; PAMPHILE, J. A. **The Importance of Omics As a Tool for the Study of Exploration of Microorganisms: Prospects and Challenges.** *Uningá Review*, v. 18, n. 2, p. 16–21, 2014.

TIXIER, F. *et al.* **Transcriptomics in cancer revealed by Positron Emission Tomography radiomics.** Scientific Reports, v. 10, n. 1, p. 5660, 2020.

UHLIG, H. H.; MUISE, A. M. **Clinical Genomics in Inflammatory Bowel Disease.** Trends in Genetics, v. 33, n. 9, p. 629–641, 2017.

ZANG, L. *et al.* **RNA-seq based transcriptome analysis of the anti-obesity effect of green tea extract using zebrafish obesity models.** Molecules, v. 24, n. 18, p. 3256-3267, 2019.

## GENÔMICA E OS SEQUENCIADORES DE PRÓXIMA GERAÇÃO

*Data de aceite: 10/05/2021*

*Data de submissão: 15/04/2021*

### **Paola Cerbino Doblaz**

Licenciada em Ciências Biológicas e Mestranda em Biotecnologia pela Universidade Federal do Espírito Santo (UFES).

<http://lattes.cnpq.br/4830901683783673>

### **Iuri Drumond Louro**

Graduado em Medicina pela UFES, PhD em Biochemistry and Molecular Genetics e Pós-Doutorado em Cancer Genetics pela (UAB), Professor Titular da UFES e do Programa de Pós-graduação em Biotecnologia, UFES.

<http://lattes.cnpq.br/3817361438227180>

**RESUMO:** A genômica é um campo que surgiu para estudar a estrutura, a função e a expressão dos genes nos genomas dos organismos. Contudo, para estudar os genes é preciso ferramentas adequadas, como os sequenciadores de próxima geração (NGS). Estes passaram por diversas melhorias ao longo da história, possibilitando um sequenciamento mais preciso, mais automatizado, e mais rápido. Com isso, esta revisão tem como objetivo auxiliar os pesquisadores na escolha da plataforma adequada para sua pesquisa, demonstrando as diferentes características entre os sequenciadores de próxima geração, e suas principais vantagens e desvantagens. Com os NGSs é possível detectar variantes gênicas e, em alguns casos, variações no número de cópias. Através dos NGS, é possível sequenciar o DNA de qualquer organismo, tendo atualmente uso na prática médica, forense, estudos evolutivos e

em biotecnologia. O sequenciamento por meio dos NGSs necessita de quatro ou cinco etapas dependendo da plataforma, sendo elas, a etapa de extração do DNA, construção de uma biblioteca de fragmentos de DNA, amplificação dos fragmentos, leitura das sequências, e análise dos dados. Nesta revisão são citadas as plataformas: Roche, Illumina, Solid, IonTorrent, PacBio e Nanopore. A escolha da plataforma adequada depende dos objetivos e recursos de cada usuário.

**PALAVRAS-CHAVE:** NGS, sequenciamento, DNA.

### GENOMICS AND NEXT GENERATION SEQUENCERS

**ABSTRACT:** Genomics is a field created to study the DNA structure, function and gene expression of any organism. However, when studying genes, one must have access to the appropriate tools, such as Next Generation Sequencers. These have undergone several improvements throughout history, enabling a more accurate, automated and faster sequencing. Thus, this review aims to assist researchers in choosing the appropriate platform for their research, showing different characteristics among next-generation sequencers along with their main advantages and disadvantages. With NGSs, it is possible to observe genetic variants and, in some cases, copy number variations. NGS are capable of sequencing DNA from any organism. They are currently used in clinical practice, forensic, evolutionary studies and biotechnology. NGS sequencing requires four or five steps, depending on the platform: DNA extraction, DNA fragment library construction, amplification, sequencing and data analysis. The platforms cited in this review are: Roche Illumina, Solid, IonTorrent, PacBio and Nanopore. Choosing the right platform depends on

the purpose and resources of each user.

**KEYWORDS:** NGS, sequencing, DNA.

## 1 | A HISTÓRIA DO SEQUENCIAMENTO DO DNA

Com a descoberta da estrutura tridimensional do DNA, em 1953, os pesquisadores começaram a despertar ainda mais interesse por esta molécula, surgindo assim um novo campo na ciência, a genômica.

A genômica é a ciência que estuda a estrutura, a função e a expressão dos genes nos genomas dos organismos (HORGAN; LOUISE, 2011; SALZANO, 2014). Embora os pesquisadores conhecessem as estruturas presentes no DNA, sequenciá-lo ainda era uma tarefa difícil. Somente na década de 70, novas técnicas para o sequenciamento do DNA começaram a ser desenvolvidas (FIETTO; MACIEL, 2015; HEATHER; BENJAMIN, 2016).

Na época, dois grupos trabalhavam em desenvolver técnicas de sequenciamento: Alan Maxam e Walter Gilbert, do qual o princípio da técnica era marcar o DNA de interesse e depois quebrá-lo, com reagentes químicos, em pequenos fragmentos que podem ser lidos em gel de poliacrilamida; Frederick Sanger e Alan Coulson, utilizaram a técnica baseada em reações enzimáticas, como a da DNA polimerase. Ambos faziam uso de marcadores radioativos. Em 1977, Sanger aperfeiçoou sua técnica utilizando didesoxinucleotídeos, e assim se deu início ao sequenciamento de primeira geração (HEATHER; BENJAMIN, 2016; SALZANO, 2014). Ao longo do tempo, as tecnologias avançaram e com isso a técnica de Sanger foi aprimorada e automatizada, até surgirem novas formas de sequenciamento.

Na década de 90, os sequenciadores trouxeram grandes feitos para a ciência, marcando a história dos sequenciadores. Em 1990, foi iniciado o Projeto Genoma Humano, que teve como objetivo obter toda a sequência do genoma humano em 15 anos, e em 2003 o projeto foi concluído (COLLINS; MORGAN; PATRINOS, 2003). Diversos fatores influenciaram para a que a leitura fosse realizada em menos tempo, como a automatização das técnicas de PCR e de Sanger, grandes avanços da bioinformática, além da melhoria do sequenciamento pela técnica de *Shotgun Sequencing*.

Em 1995, a técnica de *Shotgun Sequencing* possibilitou o sequenciamento completo dos genomas das espécies: *Hameophilus influenzae* e *Mycoplasma genitalium*. Ela acabou sendo revolucionária para o projeto genoma humano, pelo fato de permitir o sequenciamento em paralelo, e sem precisar do conhecimento prévio das sequências adjacentes. A técnica consiste em fragmentar o DNA aleatoriamente com enzimas de restrição, construir uma biblioteca de fragmentos em vetores de sequenciamento, e posterior alinhamento das sequências. (LOMAN; PALLEEN, 2015; SALZANO, 2014; VENTER et al., 1998). Ao longo do tempo as tecnologias permitiram o surgimento de novos sequenciadores, conhecidos como Sequenciadores de Próxima Geração (NGSs).

A técnica de Sanger é utilizada até hoje e tem grande importância, principalmente para confirmar as variantes encontradas pelos NGSs, isso ocorre, pois há limitações em suas técnicas. Contudo, os NGSs permitem uma maior agilidade no sequenciamento, e geram maiores informações. Os NGSs, apesar de serem bem diferentes entre si em relação a leitura da sequência do DNA, possuem características semelhantes, como a necessidade

da construção de uma biblioteca de fragmentos, leitura em paralelo e repetitiva e a ausência do uso de gel. Esta revisão tem como objetivo auxiliar os pesquisadores na escolha da plataforma adequada para suas pesquisas, demonstrando as principais diferenças entre os sequenciadores de próxima geração, e suas principais vantagens e desvantagens.

## 2 | APLICAÇÕES DOS NGSS

Os NGSS são eficazes para observar variantes gênicas, como mutações e polimorfismos, inserções e deleções pequenas (*indels*), e dependendo do pipeline utilizado, podem detectar variações no número de cópias. São eficientes para estudos genômicos, metagenômico e epigenômicos. Estes sequenciadores são capazes de sequenciar qualquer DNA, como o de cloroplastos, mitocôndrias ou DNA nuclear de células germinativas ou somáticas, e de quaisquer organismos, como vírus, bactérias e humanos (VOELKERDING; DAMES; DURTSCHI, 2009; YOHE; THYAGARAJAN, 2017).

Devido à sua ampla capacidade de leitura e geração de dados, os NGSS são usados na prática médica. Com a leitura gênica é possível identificar alterações genéticas hereditárias ou não, como deficiências, síndromes, cegueira, surdez, entre diversos distúrbios renais, neurológicos, do tecido conjuntivo, entre outros. Na oncologia, o sequenciamento de determinados genes pode ajudar a entender se o câncer é hereditário ou somático, assim como definir opções de tratamento. É possível também verificar predisposição ao câncer por meio de painéis genéticos (BEHJATI; TARPEY, 2013; YOHE; THYAGARAJAN, 2017). Estes painéis também podem ser estabelecidos para outras doenças genéticas, como doenças cardiovasculares, neurológicas, hematológicas, etc. Os testes genéticos podem ajudar a tomar medidas preventivas (JÚNIOR et al., 2019; KALAYINIA et al., 2017).

Como os NGSS são eficientes para quaisquer organismos, na medicina, eles ajudam a identificar patógenos como vírus, bactérias e outros. O sequenciamento dos genomas desses organismos facilita o diagnóstico, e ajudam no desenvolvimento de vacinas e remédios.

Em meio a pandemia do coronavírus em 2020, o sequenciamento rápido do vírus SarsCov-2 permitiu verificar as variações genéticas importantes para o desenvolvimento de possíveis vacinas e drogas. Conhecer a sequência gênica também facilitou para o desenvolvimento de diagnósticos por RT-PCR ou sorologia, ajudou a identificar outros microorganismos associados, além de possibilitar a identificação de cepas variantes (DIAS et al., 2020; TANG et al., 2020).

Na área forense, as sequências de repetições curtas em tandem (STR) são as mais usuais para identificação do DNA coletado. Os NGSS permitem a amplificação de amostras degradadas e com pouca quantidade de DNA (ALVAREZ-CUBERO et al., 2017).

Os novos sequenciadores também ajudam a reconstruir as filogenias e com isso fazer inferências evolutivas. Nos diferentes organismos há muitas regiões homólogas e regiões com polimorfismo ou mutações, com os NGSS estas regiões e as variações gênicas puderam ser melhor observadas (KADLEC, 2017).

### 3.1 SEQUENCIADORES DE PRÓXIMA GERAÇÃO (NGS)

Os sequenciadores de próxima geração, também são conhecidos como sequenciadores massivamente paralelos. Receberam esse nome pelo salto na sua capacidade de sequenciamento quando comparados a tecnologias previamente existentes, ou seja, são capazes de sequenciar um número ainda maior de fragmentos de DNA simultaneamente (GOODWIN; MCPHERSON; MCCOMBIE, 2016; NARCISO; BEVITORI, 2014; YOHE; THYAGARAJAN, 2017).

Serão apresentados neste capítulo as seguintes plataformas: Roche, Illumina, Solid e IonTorrent PacBio e Nanopore

A leitura por meio dos NGSs necessita de quatro ou cinco etapas: 1-Extração do DNA de interesse; 2- construção de uma biblioteca de sequenciamento; 3- amplificação dos fragmentos a serem sequenciados; 4- leitura das sequências; 5- Análise dos dados (YOHE; THYAGARAJAN, 2017).

#### 3.1 Extração do DNA

Existem diferentes protocolos para extração de DNA, da qual a escolha depende do tipo de material biológico e da pesquisa. Para o sequenciamento, o método de extração não terá influência, desde que a extração apresente uma boa qualidade e uma quantidade razoável de DNA. Como os NGS são bem sensíveis, não necessitam de grandes quantidades. A verificação da qualidade e quantidade do DNA podem ser medidas pelos equipamentos Nanodrop ou gel de agarose e Qiubit ou picogreen (SIMBOLO et al., 2013; YOHE; THYAGARAJAN, 2017).

Existem protocolos de extração mais baratos como os realizados por fenol e/ou clorofórmio ou etanol, porém são substâncias que podem agir como inibidores, por este motivo necessitam de várias etapas de lavagem, tornando a técnica mais demorada e mais suscetível a erros devido as diversas manipulações. Um alternativa seriam kits comerciais. Algumas empresas, como a Thermo Fisher Scientific, Kasvi, Qiagen ou Promega disponibilizam kits específicos para cada material biológico, com diferentes rendimentos dependendo do valor (SCHIEBELHUT et al., 2016).

#### 3.2 Construção de uma biblioteca de sequenciamento

Há diversos protocolos e kits que permitem a construção da biblioteca, da qual se resume em fragmentar o DNA e adicionar adaptadores nas extremidades dos fragmentos (DIJK; JASZCZYSZYN; THERMES, 2014; FIETTO; MACIEL, 2015; YOHE; THYAGARAJAN, 2017).

Como o NGS sequencia fragmentos menores que os de Sanger, o DNA é quebrado em fragmentos menores (*reads*). A quebra do DNA pode ser efetuada por métodos químicos, como o uso das enzimas de restrição, ou métodos físicos, como a sonicação (VAN DIJK; JASZCZYSZYN; THERMES, 2014; YOHE; THYAGARAJAN, 2017).

Em seguida os adaptadores, que são oligonucleotídeos, são adicionados às extremidades dos fragmentos pelas enzimas ligases. Estes podem ter a função de ancorar o fragmento de DNA à superfície da plataforma em *beads* ou *flowcell*, complementar o

primer para anelamento, ou identificar os fragmentos por barcode ou índice em plataformas que permitam a multiplexação (GOETZE; PASSAIA; SPERB-LUDWIG, 2017; YOHE; THYAGARAJAN, 2017).

Nas tecnologias mais atuais, adaptadores tipo *hairpin*, possuem a função de conectar um filamento de DNA a outro (Figura 1). A tecnologia Pacbio utiliza o adaptador *hairpin* em ambas as extremidades, enquanto a de nanopore, usa o adaptador *hairpin* em apenas uma das extremidades (GROHME et al., 2013).

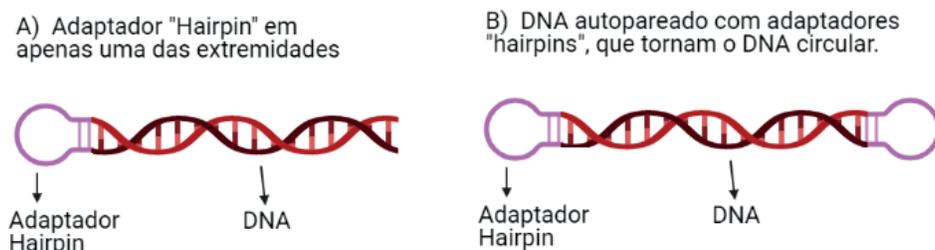


Figura 1. Adaptadores tipo *hairpin*. A) Adaptador *hairpin* em apenas uma extremidade, comum na tecnologia de nanopore. B) Adaptadores *hairpin* em ambas as extremidades, tornando o DNA um único filamento circular autopareado, usado na plataforma Pacbio. Fonte: Próprio autor.

Dependendo da tecnologia são adicionados adaptadores complementares para anelamento do primer somente em uma ou nas duas extremidades. Quando o sequenciamento ocorre em apenas um sentido, sequenciamento *single end*, é adicionado o adaptador (*forward*) em apenas uma extremidade (Figura 2). A maioria dos NGSs usam essa técnica, como a Roche, SOLID e Ion Torrent. O sequenciamento *paired end* é quando há o sequenciamento em ambos os sentidos do fragmento, sendo adicionado adaptadores em ambas as extremidades (*forward* e *reverse*), como ocorre na plataforma Illumina (ESCALONA; ROCHA; POSADA, 2016; GREH, 2018).

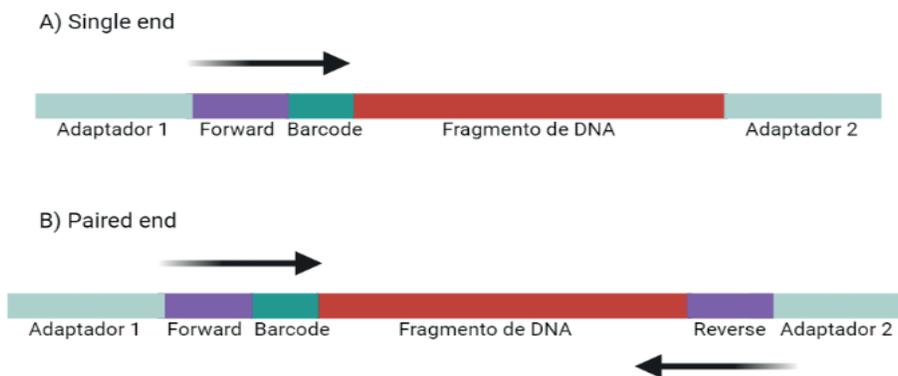


Figura 2. Adição dos adaptadores aos Fragmentos. As setas indicam o sentido da leitura. a) *Single end*: Leitura em apenas um sentido. b) *Paired end*: Leitura em ambos os sentidos. Fonte: Próprio autor.

O sequenciamento *mate-pair* são sequências mais longas (*scaffold*), que também permitem a leitura em ambos os sentidos. Sua construção é realizada adicionando biotina nas extremidades dos *scaffolds*. As biotinas se ligam formando um DNA circular, para que depois sejam fragmentados. Os fragmentos com biotina são selecionados e então são adicionados adaptadores nas extremidades (Figura 3) (NIEUWERBURGH et al., 2011). As plataformas como Roche e Solid são capazes de realizar este tipo de leitura (BERGLUND; KIIALAINEN; SYVÄNEN, 2011).

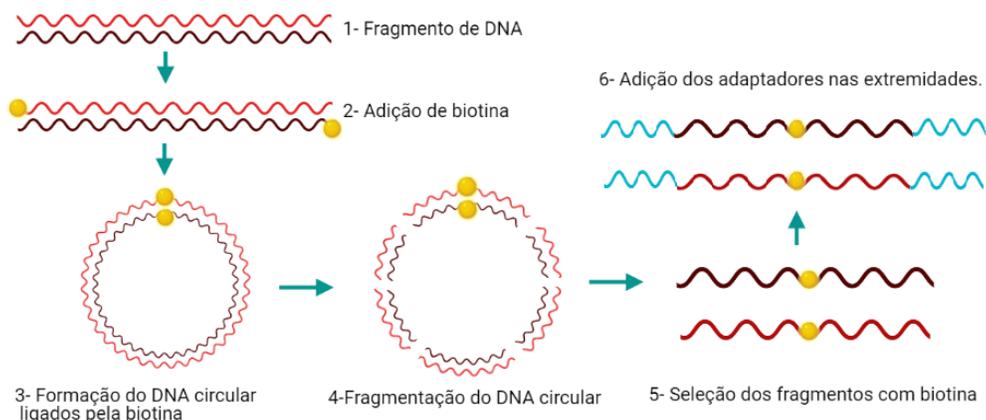


Figura 3. Construção da biblioteca de sequenciamento a partir de *scaffolds*, para sequenciamento *mate-pair*. Fonte: Próprio autor

Após a construção da biblioteca de sequenciamento o fragmento está disponível para a próxima etapa, podendo ser amplificado ou sequenciado.

### 3.3 Amplificação do fragmento

É muito comum que haja amplificação dos fragmentos selecionados, porém alguns equipamentos atuais como a da tecnologia de PacBio ou Nanopore não necessitam dessa etapa. A amplificação também é conhecida como enriquecimento, pois gera uma maior quantidade de DNA adaptados. A amplificação pode ser feita por PCR, na qual os fragmentos podem ser ancorados em microesferas (*beads*) ou em placa de vidro (*flow cell*) (VAN DIJK; JASZCZYSSZYN; THERMES, 2014; YOHE; THYAGARAJAN, 2017).

A amplificação em *beads* ocorre em uma solução aquosa com óleo, conhecido como PCR em Emulsão. Da qual cada microgota aquosa com um *bead* fica imersa em uma solução oleosa. Para a multiplicação dos fragmentos são necessários os reagentes de PCR e os fragmentos adaptados. As microesferas possuem adaptadores por toda sua superfície, que se ligam aos fragmentos em uma de suas extremidades, fixando-os na superfície. E assim o *primer* pode se anelar à outra extremidade do fragmento para que ocorra a amplificação. As plataformas Roche 454, Solid e Ion torrent utilizam esta amplificação (GOODWIN; MCPHERSON; MCCOMBIE, 2016; FIETTO; MACIEL, 2015).

A plataforma Illumina faz uso da amplificação por pontes em *flowcell* que são cobertas por primers *forward* e *reverse*. O fragmento de interesse se liga, em uma de suas extremidades ao primer complementar presente na placa (*forward*), enquanto a

outra extremidade se liga ao primer complementar mais próximo na placa (*reverse*). Os nucleotídeos são adicionados formando uma dupla fita em forma de ponte. Quando há o processo de desnaturação, as fitas se separam e apenas uma das extremidades ficará fixada e a outra ficará livre, permitindo que no próximo ciclo haja o anelamento com outros primers mais próximos na *flowcell*, havendo um aumento do número de fragmentos (FIETTO; MACIEL, 2015; KNETSCH et al., 2019).

### 3.4 Sequenciamento

O sequenciamento do DNA depende de cada equipamento. A leitura pode ocorrer em poços como os equipamentos da Roche 454 e Ion Torrent ou em placas como da Illumina e a SOLID (GOODWIN; MCPHERSON; MCCOMBIE, 2016; YOHE; THYAGARAJAN, 2017). Nos equipamentos atuais costuma-se utilizar a nanotecnologia como nanoporos em células de fluxo na tecnologia *Oxford Nanopore Technologies* ou matriz nanofotônica ZMW (*zero-mode waveguide*) utilizada pela Pacbio (FENG et al, 2015; HEATHER; BENJAMIN, 2016).

A maioria das plataformas, para fazer a leitura, utilizam a enzima DNA polimerase, da qual é adicionado nucleotídeos à cadeia (FIETTO; MACIEL, 2015). A tecnologia Pacbio, apesar de fazer uso da DNA polimerase, sua enzima é específica e necessita de ficar imobilizada na superfície próxima ao detector ZMW (EID et al., 2009). A tecnologia SOLID, por sua vez, não utiliza a DNA polimerase, pois faz uso de sondas que são ligadas pela enzima DNA ligase (FIETTO; MACIEL, 2015). A tecnologia da *Oxford Nanopore* não necessita da construção de uma nova fita para detecção, dessa forma, a leitura ocorre quando o fragmento de interesse passa pelo nanoporo (FENG et al, 2015).

Para detectar qual nucleotídeo ou sonda está sendo incorporado as tecnologias podem fazer uso de fluorescência, pirofosfato ou voltagem. Na tecnologia da Roche 454 a detecção da incorporação nucleotídica ocorre pela liberação de um pirofosfato. A plataforma Illumina é semelhante à de Sanger, do qual a detecção nucleotídica ocorre por fluorescência emitida a cada nucleotídeo terminal reversível incorporado.

Outras plataformas como a SOLID e a Pacbio também fazem uso de fluorescência. As tecnologias que utilizam a voltagem são: Ion torrent e *Oxford Nanopore*. A tecnologia Ion Torrent faz detecção pela mudança de pH gerada pela incorporação de cada nucleotídeo (CHRISTOFF, 2017), enquanto a da *Oxford Nanopore* à medida que cada nucleotídeo passa pelo nanoporo, há uma mudança na corrente elétrica, gerando uma voltagem diferente respectiva a cada base (LEGGETT; CLARK, 2017).

### 3.5 Análise dos dados

Cada plataforma irá gerar um gráfico segundo sua tecnologia de detecção. A plataforma Roche 454 e de Ion Torrent, por exemplo, devido a dispensação nucleotídica ser de apenas um tipo de base nitrogenada, conhecida pelo programa, cada nucleotídeo incorporado gera um pico de acordo com a quantidade incorporada, formando um gráfico (pirograma) no caso da Roche 454 ou um gráfico de voltagem como no equipamento Ion Torrent (FIETTO; MACIEL, 2015; GOODWIN; MCPHERSON; MCCOMBIE, 2016).

As tecnologias que utilizam a fluorescência também geram um tipo de gráfico, comumente chamado de cromatograma, pois cada base nitrogenada possui uma cor específica. Dessa forma, o programa consegue visualizar qual nucleotídeo foi incorporado

(CHRISTOFF, 2017). O mesmo ocorre com a tecnologia de Nanopore, que apesar de não ser por fluorescência, cada nucleotídeo que passa pelo nanoporo irá gerar uma voltagem específica, formando um gráfico.

Para que seja possível compreender os resultados apresentados nos gráficos, e interpretar a grande quantidade de dados gerados, os softwares realizam uma análise, traduzindo os gráficos em palavras (*string*) para que assim seja possível obter a sequência. Os resultados passam por uma série de processos, chamados de *pipelines* de análises genômicas, que são programas sequenciais dependentes uns dos outros (MELO, 2009; YOHE; THYAGARAJAN, 2017).

Algumas plataformas por permitirem a multiplexação, ou seja, a mistura de amostras, necessitam da etapa de desmultiplexação, que é quando o programa identifica e separa as amostras (ALEXANDER et al, 2012; YOHE; THYAGARAJAN, 2017). A sequência interpretada pelos softwares precisa ser salva na forma de texto ou binária em formatos que variam entre plataformas, como FASTA, FASTQ, SFF, BAM ou SAM, entre outros (ESCALONA; ROCHA; POSADA, 2016; YOHE; THYAGARAJAN, 2017). Assim, obtendo-se a sequência é possível verificar variantes e realizar inferências genômicas.

## 3.6 Vantagens e desvantagens de cada técnica

### 3.6.1 Roche 454

A grande vantagem de utilizar esta tecnologia é a redução do tempo e a capacidade de fazer leitura de *reads* mais longas com 600pb. Contudo existe uma maior chance de gerar *indels* (inserção e deleção), isso ocorre principalmente nas regiões homopolímeras, que são sequências de um mesmo nucleotídeo. Outro fator importante é o custo mais elevado devido a grande quantidade de reagentes para o pirosequenciamento. Tendo em vista suas vantagens e desvantagens ele é mais usual para sequenciamento de genomas de organismos menores como vírus e bactérias e ideal para verificar mutações pontuais, porém sua tecnologia não tem sido mais comercializada (FIETTO; MACIEL, 2015; LEDERGERBER; DESSIMOZ, 2011; CHRISTOFF, 2017; ZHANG et al., 2011; ZHOU et al., 2010).

### 3.6.2 Illumina/Solexa

A vantagem dessa técnica é o alto rendimento e possibilidades de multiplex. Além disso, devido ao seu bloqueio para adicionar mais de um nucleotídeo por vez, não há dificuldade em sequenciar regiões homopolímeras. Por outro lado, o equipamento é extremamente caro e pode haver erros de incorporação das bases. São ideais para aplicações genômicas e metagenômicas, RNAs, detecção de mutações somáticas, testes pré-natais e aplicações forenses (CHRISTOFF, 2017; SHENDURE; JI, 2008).

### 3.6.3 Solid

A principal vantagem dessa técnica é a qualidade dos dados. Entretanto, sua técnica é mais lenta que as demais. São bastante usuados para genomas complexos, detecção de

mutações e polimorfismos, além de permitir multiplex (CHRISTOFF, 2017).

### 3.6.4 *Ion Torrent*

Devido à semelhança com a plataforma Roche, as vantagens e as desvantagens da plataforma Ion costumam ser as mesmas. Seus equipamentos são mais baratos, capazes de ler *reads* mais longos e o sequenciamento é mais rápido se comparado à outras plataformas. Porém tem problemas com as regiões homopoliméricas, e o sequenciamento de um genoma completo pode ainda ser um problema. São ideais para detectar variações como polimorfismos e mutações pontuais. A plataforma tem se voltado muito para área clínica, com equipamentos voltados para diagnóstico como a Ion Personal Genome Machine (PGM) e Ion S5 (CHRISTOFF, 2017; GOODWIN; MCPHERSON; MCCOMBIE, 2016).

### 3.6.5 *Pacific Biosciences*

A plataforma possui vantagens como por exemplo o sequenciamento ser mais rápido que outros métodos, isso é possível pois não há processo de amplificação e nem etapas de lavagens. Outra vantagem é o comprimento das *reads*, que podem chegar até 60kb. O problema está no custo que ainda é elevado e a taxa de erros que é considerada alta. A leitura por CCS diminui a taxa de erro devido a quantidade de vezes que o fragmento é sequenciado gerando uma maior precisão. A tecnologia é ideal para sequenciamento de genomas, transcriptomas e até mesmo epigenomas de organismos complexos (GOODWIN; MCPHERSON; MCCOMBIE, 2016; YOHE; THYAGARAJAN, 2017; ZOLET et al., 2017).

### 3.6.6 *Nanopore*

As principais vantagens do nanopore é leitura de longas sequências em menos tempo, detecções de variantes, e a possibilidade de transportar o equipamento. Ideais para estudos genômicos, metagenômicos e epigenômicos, tem sido amplamente utilizado para detecção de patógenos. Apesar de apresentar altas taxas de erro, a plataforma tem investido em diminuí-los, aumentando cada vez mais a precisão do sequenciamento, porém ainda não é o ideal para leitura de sequenciamento de novos genomas (CHRISTOFF, 2017; JAIN et al., 2016).

## 4 | CONSIDERAÇÕES FINAIS

Sequenciar o DNA tem grande importância para as práticas na área da saúde, na biotecnologia, forense, agropecuária entre outras. Os Sequenciadores novos podem ajudar e muito a otimizar o trabalho. É preciso analisar as vantagens e desvantagens de cada técnica. Desta forma, a escolha da plataforma vai depender do objetivo e dos recursos de cada usuário.

## REFERÊNCIAS

- ALEXANDER, C. *et al.* **HighSSR: High throughput SSR characterization and locus development from next gen sequencing data.** Oxford University Press, 2012. Disponível em: [https://www.researchgate.net/profile/Peter\\_Houde/publication/230810081\\_HighSSR\\_High-throughput\\_SSR\\_characterization\\_and\\_locus\\_development\\_from\\_next-gen\\_sequencing\\_data/links/560d4fb508ae6cf68153e93c.pdf](https://www.researchgate.net/profile/Peter_Houde/publication/230810081_HighSSR_High-throughput_SSR_characterization_and_locus_development_from_next-gen_sequencing_data/links/560d4fb508ae6cf68153e93c.pdf). Acesso em: 07 de jul de 2020.
- ALVAREZ-CUBERO, M. J. *et al.* **Next generation sequencing: an application in forensic sciences.** *Annals of human biology*, v. 44, n. 7, p. 581-592, 2017.
- BEHJATI, S.; TARPEY, P. S. **What is next generation sequencing?** *Archives of Disease in Childhood-Education and Practice*, v. 98, n. 6, p. 236-238, 2013.
- BERGLUND, E.C.; KIIALAINEN, A.; SYVÄNEN, A.C. **Next-generation sequencing technologies and applications for human genetic history and forensics.** *Investig Genet.* v.2, n.23, p.2-15, 2011.
- CHRISTOFF, A.P. **Genômica e sequenciamento de nova geração.** In TURCHETTO-ZOLET, *et al.* **Marcadores Moleculares na Era Genômica: Metodologias e Aplicações.** São Paulo: Sociedade Brasileira de Genética, 2017. p.21-50. Disponível em: [https://www.sbg.org.br/sites/default/files/e\\_book\\_marcadores\\_moleculares\\_sbg\\_2017\\_final.pdf#page=22](https://www.sbg.org.br/sites/default/files/e_book_marcadores_moleculares_sbg_2017_final.pdf#page=22). Acesso em: 5 de jul de 2020.
- COLLINS, F. S.; MORGAN, M.; PATRINOS, A. **The Human Genome Project: lessons from large-scale biology.** *Science*, v. 300, n. 5617, p. 286-290, 2003.
- DIAS, V.M.C.H *et al.* **Orientações sobre Diagnóstico, Tratamento e Isolamento de Pacientes com COVID-19.** *J Infect Control*, v. 9, n. 2, p. 1-20, 2020.
- EID, J. *et al.* **Real-time DNA sequencing from single polymerase molecules.** *Science*, v. 323, n. 5910, p. 133-138, 2009.
- ESCALONA, M.; ROCHA, S.; POSADA, D. **A comparison of tools for the simulation of genomic next-generation sequencing data.** *Nature Reviews Genetics*, v. 17, n. 8, p. 459, 2016.
- FENG, *et al.* **Nanopore-based Fourth-generation DNA Sequencing Technology.** *Genomics, Proteomics & Bioinformatics.* v. 13, n.1, p.4-16, 2015.
- FIETTO, J. L. R.; MACIEL, T. E. F. Sequenciando genomas. In: MOREIRA, L.M. **Ciências genômicas: fundamentos e aplicações.** Ribeirão Preto, SP: Sociedade Brasileira de Genética, 2015. v. 1, p. 27–64. Disponível em: <http://professor.pucgoias.edu.br/SiteDocente/admin/arquivosUpload/18497/material/Seq%C3%AAnciamdo%20genomas.pdf>. Acesso em: 5 de jul de 2020.
- GOETZE, M.; PASSAIA, G.; SPERB-LUDWIG, F. Marcadores moleculares baseados em restrição: AFLP e suas variações. In TURCHETTO-ZOLET, *et al.* **Marcadores Moleculares na Era Genômica: Metodologias e Aplicações.** São Paulo: Sociedade Brasileira de Genética, 2017. p.60-76.
- GOODWIN, S.; MCPHERSON, J. D.; MCCOMBIE, W. R. **Coming of age: ten years of next-generation sequencing technologies.** *Nature Reviews Genetics*, v. 17, n. 6, p. 333, 2016.
- GREHL, C. *et al.* **How to design a whole-genome Bisulfite sequencing experiment.** *Epigenomes*, v. 2, n. 4, p. 21, 2018.
- GROHME, M. A. *et al.* **Microsatellite marker discovery using single molecule real-time circular consensus sequencing on the Pacific Biosciences RS.** *BioTechniques*, v. 55, n. 5, p. 253-256, 2013.

HEATHER, J. M.; CHAIN, B. **The sequence of sequencers: The history of sequencing DNA.** *Genomics*, v. 107, n. 1, p. 1-8, 2016.

HORGAN, R. P.; KENNY, L. C. **'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics.** *The Obstetrician & Gynaecologist*, v. 13, n. 3, p. 189-195, 2011.

JAIN, M. *et al.* **The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community.** *Genome Biology*, v. 17, n. 239, p. 1-11, 2016.

LAMOUNIER JÚNIOR, A. *et al.* **Importância do Teste Genético na Miocardiopatia Dilatada: Aplicações e Desafios na Prática Clínica.** *Arquivos brasileiros de cardiologia*, v. 113, n. 2, p. 274-281, 2019.

KADLEC, M. *et al.* **Targeted NGS for species level phylogenomics: "made to measure" or "one size fits all"?** *PeerJ*, v. 5, 2017.

KALAYINIA, S. *et al.* **Next generation sequencing applications for cardiovascular disease.** *Annals of medicine*, v. 50, n. 2, p. 91-109, 2018.

KNETSCH, C. W. *et al.* DNA sequencing. In: VAN PELT-VERKUIL, E.; VAN LEEUWEN, W.B.; TE WITT, R. **Molecular Diagnostics.** Singapore: Springer, 2019. p. 339-360. Disponível em: [https://link.springer.com/chapter/10.1007/978-981-13-1604-3\\_8](https://link.springer.com/chapter/10.1007/978-981-13-1604-3_8). Acesso em: 07 de jul de 2020.

LEDERGERBER, C.; DESSIMOZ, C. **Base-calling for next-generation sequencing platforms.** *Briefings in bioinformatics*, v. 12, n. 5, p. 489-497, 2011.

LEGGETT, R. M.; CLARK, M. D. **A world of opportunities with nanopore sequencing.** *Journal of Experimental Botany*, v. 68, n. 20, p. 5419-5429, 2017.

LOMAN, N. J.; PALLEN, M.J. **Twenty years of bacterial genome sequencing.** *Nature Reviews Microbiology*, v. 13, n. 12, p. 787-794, 2015.

MELO, H. V. F. *et al.* **Desenvolvimento de um pipeline para análise genômica e transcriptômica com base em web services.** 108f. 2009. Dissertação (Mestrado em Biotecnologia) - Universidade Federal de São Carlos, São Paulo, 2009.

NARCISO, M. G.; BEVITORI, R. **Utilização de ferramentas computacionais para análise de expressão de genes.** Embrapa (INFOTECA-E), p.8, 2014. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1003470/1/ct219.pdf>. Acesso: 07 de jul de 2020.

NIEUWERBURGH, F.P. *et al.* **Illumina mate-paired DNA sequencing-library preparation using Cre-Lox recombination.** *Nucleic Acids Research*, v.4, n3, p. 24, 2011.

VENTER, J.C. *et al.* **Shotgun Sequencing of the Human Genome.** *Science*, v. 280, p. 1540-1542, 1998.

VOELKERDING, K. V.; DAMES, S. A.; DURTSCHI, J. D. **Next-generation sequencing: from basic research to diagnostics.** *Clinical chemistry*, v. 55, n. 4, p. 641-658, 2009.

SALZANO, F.M. **Genômica e Evolução: moléculas, organismos e sociedade.** São Paulo. Oficina de Textos, 2012. p.272.

SCHIEBELHUT, L. M. *et al.* **A comparison of DNA extraction methods for high-throughput DNA analyses.** *Molecular ecology resources*, v.17, n.4, p.721-729, 2016.

SHENDURE, J.; JI, H. **Next-generation DNA sequencing**. Nature biotechnology, v. 26, n. 10, p. 1135-1145, 2008.

SIMBOLO, M. *et al.* **DNA qualification workflow for next generation sequencing of histopathological samples**. PloS one, v. 8, n. 6, 2013.

STRÖHER, P. R. **Sequenciamento de nova geração e entomologia: Novas perspectivas para antigos questionamentos**. Revista da Biologia. v. 18, n.1, p.6-16, 2018.

TANG, X. *et al.* **On the origin and continuing evolution of SARS-CoV-2**. National Science Review . v.7, ed. 6, p. 1012-1023, 2020.

VAN DIJK, E. L.; JASZCZYSZYN, Y.; THERMES, C. **Library preparation methods for next-generation sequencing: tone down the bias**. Experimental cell research, v. 322, n. 1, p. 12-20, 2014.

YOHE, S.; THYAGARAJAN, B. **Review of clinical next-generation sequencing**. Archives of pathology & laboratory medicine, v. 141, n. 11, p. 1544-1557, 2017.

ZHANG, J. *et al.* **The impact of next-generation sequencing on genomics**. Journal of genetics and genomics, v. 38, n. 3, p. 95-109, 2011.

ZHOU, X. *et al.* **The next-generation sequencing technology: A technology review and future perspective**. Ciência China Life Sciences , v. 53, n. 1, p. 44-57, 2010.

## EPIGENÉTICA E MÉTODOS DE ESTUDOS EPIGENÉTICOS

Data de aceite: 10/05/2021

Data de submissão: 15/04/2021

### Ivana Alece Arantes Moreno

Bacharel em Ciências Biológicas, Mestra e  
Doutoranda em Biotecnologia pela Universidade  
Federal do Espírito Santo (UFES)  
Alegre - Espírito Santo  
<http://lattes.cnpq.br/4011255253507606>

### Suzanny Oliveira Mendes

Bacharel em Ciências Biológicas, Mestra,  
Doutora e Pós-Doutoranda em Biotecnologia  
pela Universidade Federal do Espírito Santo/  
RENORBIO  
Alegre - Espírito Santo  
<http://lattes.cnpq.br/5613486906366786>

**RESUMO:** A epigenética é o estudo das mudanças na expressão gênica sem alterar a sequência primária do DNA podendo ser herdáveis ou não herdáveis e reversíveis. Os principais mecanismos de regulação epigenética incluem a metilação do DNA, modificações pós-traducionais das histonas e regulação por miRNAs. Compreender como os mecanismos epigenéticos interagem no desenvolvimento de doenças complexas é um grande desafio. Além disso, avanços recentes também apontam novas possibilidades terapêuticas. Aqui, revisamos e discutimos os recentes conceitos dos mecanismos epigenéticos e destacamos os principais métodos tecnológicos de detecção neste campo.

**PALAVRAS-CHAVE:** metilação de DNA, modificação das histonas, mRNA, métodos, epigenética.

### EPIGENETICS AND METHODS OF EPIGENETIC STUDIES

**ABSTRACT:** Epigenetics is the study of changes in gene expression without altering the primary DNA sequence and can be either reversible and inheritable or non-inheritable. The main mechanisms of epigenetic regulation include DNA methylation, post-translational modifications of histones and regulation by miRNAs. Understanding how epigenetic mechanisms interact in the development of complex diseases is a major challenge. Moreover, recent advances also point to new therapeutic possibilities. Here, we review and discuss the recent concepts of epigenetic mechanisms and highlight the main technological methods of detection in this field.

**KEYWORDS:** DNA methylation, histones modifications, miRNAs, methods, epigenetics.

### 1 | INTRODUÇÃO

Cada indivíduo possui o seu próprio genoma que é encontrado em quase todas as células do seu corpo (DOR; CEDAR, 2018). O genoma de um organismo abrange toda a informação genética contida em sequências de nucleotídeos que compõem o DNA do indivíduo. Embora o genoma contenha estas informações, somente isso não revela como as suas funções nos organismos são reguladas. Diante disto, o estudo da epigenética contribuiu para elucidar os mecanismos de funcionamento dos genes (DAVID SWEATT, 2019). Desta forma, a regulação epigenética é um dos mecanismos que modula a expressão dos genes e seu estudo contribui para o melhor entendimento sobre o epigenoma e a influência do ambiente na

regulação da expressão gênica.

A epigenética começou a ser estudada pelo embriologista e geneticista Conrad Waddington, definindo-a como “o ramo da biologia que estuda as interações casuais entre os genes e seus produtos que dão origem ao fenótipo” (WADDINGTON, 1939). De acordo com os novos conhecimentos sobre o tema, os conceitos foram atualizados. Após, Arthur Riggs e colaboradores a definiram como “o estudo de alterações mitoticamente e/ou meioticamente herdáveis na função do gene que não podem ser explicadas por alterações na sequência do DNA” (RUSSO; MARTIENSSEN; RIGGS, 1996).

Entretanto, esse foi um dos paradigmas da epigenética tradicional de que as marcas epigenéticas eram imutáveis a partir do momento em que elas ocorriam, sugerindo que os mecanismos epigenéticos eram estáticos e de acordo com descobertas recentes, esses conceitos foram ampliados e atualizados (BIRD, 2007; DAVID SWEATT, 2019). Atualmente a epigenética refere-se ao estudo dos mecanismos que alteram a expressão gênica, sem modificar a sequência primária do DNA. Como também, os mecanismos epigenéticos podem ser herdáveis ou não herdáveis e reversíveis (DEANS; MAGGERT, 2015; HIRST; MARRA, 2009; TRONICK; HUNTER, 2016).

Os mecanismos epigenéticos incluem processos moleculares como as alterações estruturais do DNA e suas proteínas relacionadas e incorporações de RNA não codificantes (DAVID SWEATT, 2019). Estas modificações podem ser induzidas por estímulos ambientais, como dieta, estresse, hormônios e medicamentos (DOR; CEDAR, 2018). Além disso, essas modificações podem ser de natureza transitória e hereditária.

As modificações epigenéticas contribuem para o entendimento de diversos processos celulares, participando em mecanismos biológicos cruciais como diferenciação, regulação gênica, desenvolvimento e o envelhecimento. Estudos sugerem que alterações no equilíbrio epigenético podem desempenhar um papel importante na susceptibilidade e progressão de diversas doenças, como por exemplo, o câncer (BENNETT; LICHT, 2018; HIRST; MARRA, 2009; MASUI et al., 2020). Desta forma, avanços recentes de estudos do epigenoma vêm sendo promissores no desenvolvimento de novas estratégias de tratamento de doenças.

A ação conjunta dos mecanismos epigenéticos forma uma rede interativa complexa com a atividade de outros mecanismos, contribuindo para a mesma regulação transcricional e pós-transcricional (DEANS; MAGGERT, 2015; KLENGEL; BINDER, 2015). Esses processos contribuem para facilitar ou dificultar o acesso de fatores de transcrição, regulando a leitura dos genes.

O foco deste capítulo é fornecer os conceitos e as metodologias que foram desenvolvidas para detectar os mecanismos epigenéticos. Vale ressaltar que a escolha da técnica adequada permite a obtenção de uma resposta compatível para responder à pergunta biológica do pesquisador e outros fatores precisam ser levados em consideração como custos, sensibilidade e especificidade, disponibilidade de equipamentos e quantidade de amostra do DNA (KURDYUKOV; BULLOCK, 2016).

Existem métodos disponíveis podem ser úteis em aplicações específicas ou são técnicas que atualmente não são empregadas, logo, não serão discutidas aqui. Sendo assim, vamos nos concentrar nos métodos que consideramos com alto rendimento, mais robustos e prontamente disponíveis para a comunidade científica.

## 2 | METILAÇÃO DO DNA

A metilação do DNA é o mecanismo epigenético mais abundante e caracterizado em modelos animais e humanos (GOUIL; KENIRY, 2019; HIRST; MARRA, 2009; O'DONNELL; MEANEY, 2020). O processo consiste em uma modificação covalente do DNA com a adição um grupo metil ( $\text{CH}_3$ ) no carbono 5' que ocorre exclusivamente nos nucleotídeos citosina, convertendo-a em 5-metilcitosina (5mC), catalisada pela enzima metiltransferase de DNA (DNMT) (KOHLI; ZHANG, 2013; LI; ZHANG, 2014).

Outra marca epigenética, a 5 hidroximetilcitosina (5hmC) produzida através da oxidação da 5mC catalisada pela enzima de translocação metilcitosina dioxigenase (TET) encontrada em mamíferos e está associada ao processo de desmetilação do DNA (KOHLI; ZHANG, 2013; LI; ZHANG, 2014). Na maioria dos casos, o processo de metilação do DNA ocorre em dinucleotídeos seguidos por uma guanina (5'-CpG-3'), que são regiões ricas em CpGs, contexto chamado de "ilhas CpG" (FEDOTOVA; ILLARIOSHKIN, 2019; VIALOU et al., 2013). Em todo genoma humano, encontra-se cerca de 28 milhões de sítios CpG e destes, aproximadamente 70 a 80% são metilados em células somáticas normais (EHRlich et al., 1982; SKVORTSOVA; STIRZAKER; TABERLAY, 2019).

A metilação do DNA em regiões promotoras dos genes é um mecanismo que promove a repressão da transcrição gênica e ao silenciamento de genes. A repressão gênica ocorre porque um complexo de proteínas que reconhecem os grupos metil é recrutado, promovendo a remodelação da cromatina, formando uma estrutura de cromatina mais condensada e evitando a transcrição gênica (DOMCKE et al., 2015). Além do mais, entre os mecanismos epigenéticos descritos, a metilação do DNA é considerada a mais estável, podendo se manter até a idade adulta do indivíduo e no decorrer das gerações celulares (DOR; CEDAR, 2018; KOHLI; ZHANG, 2013). Em adição, a metilação do DNA é específica do tecido, variando em padrões de metilação em comparação com outros tipos celulares de um mesmo organismo (FENG; LOU, 2019).

Estudos atuais têm se concentrado em evidências de que os mecanismos epigenéticos desempenham um papel fundamental no processo de envelhecimento biológico, na busca de marcadores para doenças complexas e nos efeitos transgeracionais em pais que foram expostos em determinadas condições ambientais (BORÇOI et al., 2020; LOCKE et al., 2019; RYAN et al., 2019; UNNIKRIISHNAN et al., 2019).

As metodologias empregadas possuem como objetivo determinar o status de metilação de amostras de DNA. Aqui, será discutido os métodos para distinguir a 5-metilcitosina (5mC) da citosina e detecção da 5-hidroximetilcitosina. Além disso, serão discutidos os métodos que são utilizados para avaliação da metilação do DNA de todo o genoma ou em determinadas regiões, que por sua vez, podem ser regiões reguladoras de genes de interesse. Atualmente, os avanços nas tecnologias de sequenciamento de nova geração (NGS) possibilitaram a criação de perfis e mapas genômicos de metilação do DNA. Vale ressaltar que os padrões de metilação do DNA não são perdidas após o processo de extração de DNA (GOUIL; KENIRY, 2019).

O método padrão ouro para análise de metilação do DNA na detecção da 5mC é baseado no tratamento do DNA genômico com bissulfito de sódio. A técnica consiste na conversão de resíduos de conversão de resíduos da citosina não metilados em uracila,

através do processo de desaminação e as citosinas metiladas permanecem intactas (GOUIL; KENIRY, 2019). Dessa forma, a metilação do DNA pode ser interpretada por sequenciamento de Sanger, Pirosequenciamento ou Illumina. Entretanto, essa técnica não permite a distinção de outras modificações na citosina, como por exemplo, a 5hm (GOUIL; KENIRY, 2019; LIU; LIAO; LU, 2015). Existem métodos complementares que permitem a identificação das formas oxidadas da citosina, um desses métodos é o sequenciamento de bissulfito oxidativo (OxBS-seq) específico para a detecção da 5hmC (BOOTH et al., 2012).

No caso de baixas quantidades de amostras de DNA, recomenda-se a utilização da tecnologia MethyLight que é uma técnica quantitativa, sensível e dependente do tratamento de bissulfito de sódio (LIU; LIAO; LU, 2015; ZHAN; LUO, 2019). A execução deste método requer iniciadores específicos e sondas fluorescentes específicas para metilação (LIU; LIAO; LU, 2015).

### 3 | MODIFICAÇÃO DAS HISTONAS

No núcleo, o DNA genômico é envolvido por proteínas histônicas gerando um complexo chamado cromatina. A unidade básica da cromatina é chamada de nucleossomo, sendo formada por octamero de histonas, que constituída pelas histonas H2A, H2B, H3 e H4, em eucariotos (COBOS; BENNETT; TORRENTE, 2019). As modificações pós-transcricionais das histonas vem sendo consideradas os componentes essenciais da regulação da atividade gênica, ao marcar as regiões ativas e inativas da cromatina através dos níveis de compactação do DNA (LENNARTSSON; EK WALL, 2009; WANG et al., 2015).

As histonas são covalentemente modificadas de várias maneiras, incluindo fosforilação, ubiquitinação, acetilação e metilação (LENNARTSSON; EK WALL, 2009). Devido ao impacto na ativação e repressão da expressão gênica, as modificações mais estudadas são a acetilação e metilação, que ocorrem em resíduos de lisina ou arginina conservados nos domínios da cauda-N-terminal de cada proteína (EGGER et al., 2004).

O mecanismo de acetilação das histonas está relacionado com a ativação transcricional. As enzimas responsáveis por esse mecanismo são as histonas acetiltransferases (HATs) e histonas desacetiltransferases (HDACs), que adicionam e retiram grupos acetil ( $\text{COCH}_3$ ) das caudas das histonas (PENNER-GOEKE; BINDER, 2019).

Já para o mecanismo de metilação das histonas, os grupos metil são adicionados pelas metiltransferases de histonas e são retirados pelas metiltransferases, que podem contribuir para a repressão ou ativação da expressão gênica dependendo em qual resíduo de lisina ou arginina a metilação ocorre (PARK; HAN, 2019; VERMEULEN et al., 2010).

Progressos recentes no desenvolvimento de tecnologia tornaram possível a realização de estudos sistemáticos em todo o genoma da influência das modificações de histonas na etiologia de determinadas doenças, como por exemplo, a influência de modificações pós-transcricionais aberrantes de histonas relacionadas à neurodegeneração (COBOS et al., 2019). Estudos recentes também encontraram a relação entre mudanças nos padrões de modificações pós-transcricionais das histonas e sua relação com diversos tipos de cânceres (LI; LI; ZHU, 2019; PARK; HAN, 2019; QIN et al., 2019; ZHAO; SHILATIFARD, 2019).

A técnica amplamente utilizada para avaliar a status de modificação das histonas é a imunoprecipitação da cromatina (ChIP) com algumas variações de acordo com o tipo de estudo (LENNARTSSON; EKWALL, 2009). Neste método, o ensaio de ChIP opera como uma técnica versátil e de alto rendimento usada para sondar interações entre proteínas específicas ou formas modificadas de proteínas e uma região de DNA genômico (LIU; LIAO; LU, 2015). A técnica consiste, de modo geral, na utilização de anticorpos que interagem especificamente em determinados tipos de histonas de interesse (FUREY, 2012; HUEBERT et al., 2006).

Nesta técnica, a ligação marcada entre DNA-histonas é revertida em DNA para verificação em análise de sequenciamento ChIP-Seq (imunoprecipitação da cromatina com tecnologia de sequenciamento de nova geração) ou por hibridação com microarranjos ChIP-chip (imunoprecipitação de cromatina ligada a microarranjos), que são duas técnicas acopladas ao ChIP que são amplamente empregadas. Além disso, são utilizadas para obtenção de uma imagem mais detalhada da distribuição de metilação da histona ao longo dos genes. Ho e colaboradores (2011) destacaram as principais diferenças entre os ensaios e relataram que a técnica de ChIP-seq se demonstra mais vantajosa devido às suas vantagens práticas e análise de dados.

#### 4 | ANÁLISES DE miRNA

O estudo dos RNAs não codificantes tem se demonstrado relevantes na regulação do transcriptoma. Recentemente, muitos estudos buscaram avanços na compreensão da relação entre a parte não codificante do genoma e os mecanismos biológicos subjacentes a doenças (MUÑOZ-CULLA et al., 2016).

Os microRNAs (miRNA) fazem parte do grupo de RNAs não codificadores que contém cerca de 19 a 25 nucleotídeos e conservados entre as espécies. Por meio do emparelhamento de bases com o mRNA, regulam a expressão gênica induzindo a degradação e repressão transcricional do mRNA ou inibição do processo da tradução exercido (BARTEL, 2009). Os miRNAs atuam como reguladores cruciais em várias vias de sinalização biológica, incluindo o desenvolvimento, proliferação celular, diferenciação, resposta imune, angiogênese, apoptose e no ciclo celular (PAN et al., 2019).

Diversos estudos mostraram que a desregulação da expressão de miRNAs tem sido associada à tumorigênese (FABBRI; CALIN, 2010; HUMPHRIES; WANG; YANG, 2019; YANG et al., 2020). Mais especificamente, na relação entre a regulação negativa de genes supressores de tumores e a progressão do câncer (BANDRES et al., 2009). Além disso, novos estudos também buscam entendimento da interação epigenética de miRNA na obtenção de novos tratamentos contra o câncer (MIRANDA FURTADO et al., 2019; SERMER et al., 2019). Como resultado, ressaltando a importância de miRNAs como possível biomarcador para o câncer. Estudos também tem sugerido que outros mecanismos epigenéticos, como a metilação do DNA e a modificação das histonas são as principais causas de desregulação de miRNAs no câncer (HAN et al., 2007; SUZUKI et al., 2013).

A pesquisa por meio do miRNA podem seguir os seguintes passos: primeiramente, ocorre a detecção do miRNA, em seguida é feita a predição dos alvos e por último os alvos são validados. As metodologias de detecção dos miRNAs podem ser estudadas usando

NGS ou RT-PCR. Os alvos de miRNA podem ser obtidos por previsões usando softwares, como por exemplo o TargetScan. É válido ressaltar que existem técnicas de validação de alvo e análises funcionais disponíveis (LIU et al., 2015).

Atualmente, informações sobre o entendimento das funções dos miRNAs circulantes para uma melhor compreensão do seu papel no desenvolvimento e na comunicação do câncer ainda são limitadas. Entretanto, as pesquisas sobre essa relação estão em constante crescimento.

## 5 | CONSIDERAÇÕES FINAIS

A compreensão e o entendimento da relação entre as modificações epigenéticas e distúrbios que dão origem a várias doenças humanas está aumentando. Embora agora tenhamos uma compreensão geral do papel de determinados tipos de mecanismos epigenéticos, como por exemplo, a metilação do DNA e a modificação das histonas, ainda existem lacunas para muitas áreas que são de atual relevância.

Assim, foram descritas aqui as principais metodologias e protocolos que foram desenvolvidos para detectar os mecanismos epigenéticos. Além disso, fornecemos uma visão geral dos recentes avanços dos papéis que a modificação do DNA e das histonas, e a ação de miRNA desempenham na medicina.

## REFERÊNCIAS

BANDRES, E. *et al.* **Epigenetic regulation of microRNA expression in colorectal cancer.** *International Journal of Cancer*, v. 125, n. 11, p. 2737–2743, 2009.

BARTEL, D. P. **MicroRNAs: target recognition and regulatory functions.** *Cell*, v. 136, n. 2, p. 215–233, 2009.

BENNETT, R. L.; LICHT, J. D. **Targeting Epigenetics in Cancer.** *Annual Review of Pharmacology and Toxicology*, v. 58, n. 1, p. 187–207, 2018.

BIRD, A. **Perceptions of epigenetics.** *Nature*, v. 447, n. 7143, p. 396, 2007.

BOOTH, M. J. *et al.* **Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution.** *Science*, v. 336, n. 6083, p. 934–937, 2012.

BORÇOI, A. R. *et al.* **Risk factors for depression in adults: NR3C1 DNA methylation and lifestyle association.** *Journal of Psychiatric Research*, v. 121, p. 24–30, 2020.

COBOS, S. N.; BENNETT, S. A.; TORRENTE, M. P. **The impact of histone post-translational modifications in neurodegenerative diseases.** *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, v. 1865, n. 8, p. 1982–1991, 2019.

DAVID-SWEATT, J. D. **The epigenetic basis of individuality.** *Current Opinion in Behavioral Sciences*, v. 25, p. 51–56, 2019.

DEANS, C.; MAGGERT, K. A. **What do you mean, “epigenetic”?** *Genetics*, v. 199, n. 4, p. 887–896, 2015.

DOMCKE, S. *et al.* **Competition between DNA methylation and transcription factors determines binding of NRF1.** *Nature*, v. 528, n. 7583, p. 575–579, 2015.

DOR, Y.; CEDAR, H. **Principles of DNA methylation and their implications for biology and medicine.** *The Lancet*, v. 392, n. 10149, p. 777–786, 2018

EGGER, G. *et al.* **Epigenetics in human disease and prospects for epigenetic therapy.** *Nature*, v. 429, n. 6990, p. 457, 2004.

EHRlich, M. *et al.* **Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells.** *Nucleic acids research*, v. 10, n. 8, p. 2709–2721, 1982.

FABBRI, M.; CALIN, G. A. **Epigenetics and miRNAs in Human Cancer.** *In: HERCEG, Zdenko; USHIJIMA, Toshikazu B. T. Advances in Genetics (org.). Epigenetics and Cancer, Part A.* Academic Press, v. 70p. 87–99, 2010.

FEDOTOVA, E. Y.; ILLARIOSHKIN, S. N. **DNA Methylation in Neurodegenerative Diseases.** *Russian Journal of Genetics*, v. 55, n. 3, p. 271–277, 2019.

FENG, L.; LOU, J. **DNA Methylation Analysis.** *Methods in molecular biology (Clifton, N.J.)*, EUA, v. 1894, p. 181–227, 2019.

FUREY, T. S. **ChIP-seq and beyond: new and improved methodologies to detect and characterize protein–DNA interactions.** *Nature Reviews Genetics*, v. 13, n. 12, p. 840–852, 2012.

GOUIL, Q.; KENIRY, A. **Latest techniques to study DNA methylation.** *Essays in biochemistry*, v. 63, n. 6, p. 639–648, 2019.

HAN, L. *et al.* **DNA methylation regulates microRNA expression.** *Cancer Biology & Therapy*, v. 6, n. 8, p. 1290–1294, 2007.

HIRST, M.; MARRA, M. A. **Epigenetics and human disease.** *The International Journal of Biochemistry & Cell Biology*, v. 41, n. 1, p. 136–146, 2009.

HO, J. W. K. *et al.* **ChIP-chip versus ChIP-seq: lessons for experimental design and data analysis.** *BMC genomics*, v. 12, p. 134, 2011.

HUEBERT, D. J. *et al.* **Genome-wide analysis of histone modifications by ChIP-on-chip.** *Methods*, v. 40, n. 4, p. 365–369, 2006.

HUMPHRIES, B.; WANG, Z.; YANG, C. **MicroRNA regulation of epigenetic modifiers in breast Cancer.** *Cancers*, v. 11, n. 7, p. 897, 2019.

KLENGEL, T.; BINDER, E. B. **Epigenetics of stress-related psychiatric disorders and gene× environment interactions.** *Neuron*, v. 86, n. 6, p. 1343–1357, 2015.

KOHLI, R. M.; ZHANG, Y. **TET enzymes, TDG and the dynamics of DNA demethylation.** *Nature*, v. 502, n. 7472, p. 472–479, 2013.

KURDYUKOV, S.; BULLOCK, M. **DNA methylation analysis: choosing the right method.** *Biology*, v. 5, n. 1, p. 3, 2016.

LENNARTSSON, A.; EKWALL, K. **Histone modification patterns and epigenetic codes.** *Biochimica et Biophysica Acta (BBA) - General Subjects*, v. 1790, n. 9, p. 863–868, 2009.

LI, E.; ZHANG, Y. **DNA methylation in mammals.** *Cold Spring Harbor perspectives in biology*, v. 6, n. 5, p. a019133, 2014.

LI, Y.; LI, Z.; ZHU, W. **Molecular mechanisms of epigenetic regulators as activatable targets in cancer theranostics.** *Current medicinal chemistry*, v. 26, n. 8, p. 1328–1350, 2019.

LIU, Y.; LIAO, J.; LU, Q. **Chapter 2. Laboratory Methods in Epigenetics.** *Epigenetics and Dermatology*, p. 7–35, 2015.

LOCKE, W. J. *et al.* **DNA Methylation Cancer Biomarkers: Translation to the Clinic.** *Frontiers in genetics*, v. 10, p. 1150, 2019.

MASUI, K. *et al.* **Codependency of Metabolism and Epigenetics Drives Cancer Progression: A Review.** *ACTA HISTOCHEMICA ET CYTOCHEMICA*, v. 53, n. 1, p. 1–10, 2020.

MIRANDA FURTADO, C. L. *et al.* **Epidrugs: targeting epigenetic marks in cancer treatment.** *Epigenetics*, v. 14, n. 12, p. 1164–1176, 2019. DOI: 10.1080/15592294.2019.1640546.

MUÑOZ-CULLA, M. *et al.* **SncRNA (microRNA & snoRNA) opposite expression pattern found in multiple sclerosis relapse and remission is sex dependent.** *Scientific Reports*, v. 6, n. 1, p. 20126, 2016.

O'DONNELL, K. J.; MEANEY, M. J. **Epigenetics, Development, and Psychopathology.** *Annual Review of Clinical Psychology*, v. 16, n. 1, p. 327–350, 2020.

PAN, L. *et al.* **Genome-Wide miRNA Analysis Identifies Potential Biomarkers in Distinguishing Tuberculous and Viral Meningitis.** *Frontiers in cellular and infection microbiology*, v. 9, p. 323, 2019.

PARK, J. W.; HAN, J. **Targeting epigenetics for cancer therapy.** *Archives of Pharmacal Research*, v. 42, n. 2, p. 159–170, 2019.

QIN, J. *et al.* **Histone Modifications and their Role in Colorectal Cancer (Review).** *Pathology & Oncology Research*, 2019.

RUSSO, V. E. A.; MARTIENSSEN, R. A.; RIGGS, A. D. **Epigenetic mechanisms of gene regulation.** Cold Spring Harbor Laboratory Press, 1996.

RYAN, J. *et al.* **Systematic Review and Meta-analysis of Environmental, Lifestyle, and Health Factors Associated With DNA Methylation Age.** *The Journals of Gerontology: Series A*, v. 75, n. 3, p. 481–494, 2019.

SERMER, D. *et al.* **Emerging epigenetic-modulating therapies in lymphoma.** *Nature reviews. Clinical oncology*, v. 16, n. 8, p. 494–507, 2019.

SKVORTSOVA, K.; STIRZAKER, C.; TABERLAY, P. **The DNA methylation landscape in cancer.** *Essays in biochemistry*, v. 63, n. 6, p. 797–811, 2019.

SUZUKI, H. *et al.* **Epigenetic alteration and microRNA dysregulation in cancer.** *Frontiers in Genetics*, 2013.

TRONICK, E.; HUNTER, R. G. **Waddington, dynamic systems, and epigenetics.** *Frontiers in behavioral neuroscience*, v. 10, p. 107, 2016.

UNNIKRISHNAN, A. *et al.* **The role of DNA methylation in epigenetics of aging.** *Pharmacology & Therapeutics*, v. 195, p. 172–185, 2019.

VERMEULEN, M. *et al.* **Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers.** *Cell*, v. 142, n. 6, p. 967–980, 2010.

VIALOU, V. *et al.* **Epigenetic mechanisms of depression and antidepressant action.** *Annual review of pharmacology and toxicology*, v. 53, p. 59–87, 2013.

WADDINGTON, C. H. **Preliminary notes on the development of the wings in normal and mutant strains of *Drosophila*.** *Proceedings of the National Academy of Sciences of the United States of America*, v. 25, n. 7, p. 299, 1939.

WANG, Y. *et al.* **Analytical strategies used to identify the readers of histone modifications: A review.** *Analytica Chimica Acta*, v. 891, p. 32–42, 2015.

YANG, X. *et al.* **Epigenetic modulations of noncoding RNA: a novel dimension of Cancer biology.** *Molecular Cancer*, v. 19, n. 1, p. 64, 2020.

ZHAN, Y.; LUO, G. **DNA methylation detection methods used in colorectal cancer.** *World journal of clinical cases*, v. 7, n. 19, p. 2916–2929, 2019.

ZHAO, Z.; SHILATIFARD, A. **Epigenetic modifications of histones in cancer.** *Genome Biology*, v. 20, n. 1, p. 245, 2019. DOI: 10.1186/s13059-019-1870-5.

## PRÍNCIPIOS BÁSICOS EM AMPLIFICAÇÃO DO DNA E GENÔMICA DE CÉLULAS ÚNICAS

*Data de aceite: 10/05/2021*

*Data de submissão: 14/04/2021*

### **Bárbara Risse-Quaioto**

Licenciada em Ciências Biológicas pela Universidade São Camilo-ES, Bacharela em Ciências Biológicas pelo Instituto Federal do Espírito Santo e Mestranda em Biotecnologia pela Universidade Federal do Espírito Santo  
Vitória - ES  
<http://lattes.cnpq.br/7948833367111238>

### **Anderson Barros Archanjo**

Farmacêutico, Doutor e Pós-Doutorando em Biotecnologia pela Universidade Federal do Espírito Santo/RENORBIO  
Vitória - ES  
<http://lattes.cnpq.br/5529149503714764>

**RESUMO:** A compreensão da heterogeneidade celular tem motivado o desenvolvimento tecnológico, resultando em diversos métodos poderosos para analisar o genoma de células únicas. O sequenciamento de célula única consiste em isolar uma única célula, extrair e amplificar seu DNA ou RNA para análises posteriores, envolvendo principalmente sequenciamento de nova geração. Dessa forma, objetivou-se descrever as principais técnicas utilizadas no método single-cell genomics, além de identificar os principais avanços e contribuições da técnica para estudos em biologia molecular. Os principais métodos para isolamento de células únicas incluem: Classificação Celular Ativada por Fluorescência (FACS); Classificação de Células Ativadas Magneticamente (MACS); Microdissecação de

Captura a Laser (LCM); Coleta Manual de Células ou Micromanipulação; e Microfluídicos. Além dos métodos de isolamento celular, existem métodos para isolamento direto do núcleo: Isolamento de Núcleos Marcados em Tipo Celular Específico (INTACT) e; Isolamento Descontínuo da Cromatina Específica de Tecido para Imunoprecipitação (BiTS-ChIP). Métodos de amplificação de genoma inteiro foram desenvolvidos para gerar quantidade de DNA suficiente para o posterior sequenciamento, sendo os três principais: Reação em Cadeia da Polimerase Iniciada por Oligonucleotídeo Degenerado (DOP-PCR); Amplificação de Deslocamento Múltiplo (MDA); e Múltiplos Ciclos de Amplificação Baseados em Anelamento e Loop (MALBAC). Os avanços nas análises de células únicas possibilitaram novas descobertas na biologia molecular, bem como na medicina. O sequenciamento genômico de células únicas ampliará cada vez mais o conhecimento a cerca da complexidade e heterogeneidade celular em organismos normais, bem como no desenvolvimento dos seres vivos e das doenças.

**PALAVRAS-CHAVE:** Biologia Molecular; Biotecnologia; Genômica.

### BASIC PRINCIPLES IN DNA AMPLIFICATION AND SINGLE CELL GENOMICS

**ABSTRACT:** The understanding of cellular heterogeneity has motivated technological development, resulting in several powerful methods to analyze the genome of single cells. Single-cell genome sequencing consists of isolating a single cell, extracting and amplifying its DNA or RNA for further analysis, mainly involving new generation sequencing. Thus, the objective was to describe the main techniques used in the single-cell genomics method, in addition to identifying

the main advances and contributions of the technique to studies in molecular biology. The main methods for isolating single cells include: Fluorescence-activated cell sorting (FACS); Magnetic activated cell sorting (MACS); Laser capture microdissection (LCM); Manual cell collection or Micromanipulation; e Microfluids. In addition to cell isolation methods, there are methods for direct isolation of the nucleus: Isolation of nuclei in tagged cell types (INTACT); Isolation of tissue-specific chromatin for immunoprecipitation (BiTS-ChIP). Methods of whole genome amplification were developed to generate enough DNA for subsequent sequencing, the main three being: Degenerate oligonucleotide primed-polymerase chain reaction (DOP-PCR); Multiple Displacement Amplification (MDA); e Multiple annealing and looping based amplification cycles (MALBAC). The advances of single-cell analysis enabled new discoveries in molecular biology as well as in medicine. Genomic sequencing of single cells will increasingly expand knowledge about the complexity and cellular heterogeneity in normal organisms, as well as in the development of living beings and diseases.

**KEYWORDS:** Molecular biology; Biotechnology; Genomics.

## 1 | INTRODUÇÃO

Quase todo o conhecimento sobre o genoma e sua regulação foi obtido através de estudos ao nível de população celular, geralmente analisando uma mistura de milhares ou milhões de células. Isso porque os métodos tradicionais de sequenciamento podem obter apenas a média de muitas células, o que gera a perda de informações de heterogeneidade celular (MACAULAY; VOET, 2014). Então, é importante considerarmos que os tecidos raramente são homogêneos, onde o genoma está sujeito a adquirir mutações genéticas e a variabilidade da expressão gênica está presente em todas as células, mesmo dentro de uma população fenotipicamente homogênea (HUANG, 2009).

Na última década, a compreensão da heterogeneidade celular tem motivado o desenvolvimento tecnológico, resultando em diversos métodos poderosos para analisar o genoma de células únicas (KALISKY; BLAINEY; QUAKE, 2011). A tecnologia de sequenciamento de DNA vem melhorando constantemente desde o seu surgimento, na década de 1970. Inicialmente a tecnologia de sequenciamento de células únicas teve seu uso restrito, devido ao alto custo. Porém, conforme o passar do tempo, novos métodos foram criados, permitindo a ampliação da utilização, sendo atualmente cada vez mais utilizada em vários campos (SHAPIRO; BIEZUNER; LINNARSSON, 2013). Em 2013, a técnica foi escolhida pela Nature Methods como a tecnologia do ano.

O sequenciamento de célula única consiste em isolar, em uma única célula, extrair e amplificar seu DNA ou RNA (dependendo do foco do estudo) para análises posteriores, envolvendo principalmente sequenciamento de nova geração.

Os desafios nos estudos com base em sequenciamento de célula única incluem o isolamento eficiente das células, a amplificação do genoma com obtenção de material suficiente para as posteriores análises, sequenciamento econômico para identificar variações, bem como a interpretação dos dados, levando em consideração possíveis erros durante essas etapas (GAWAD; KOH; QUAKE, 2016). A seguir, descreveremos as principais técnicas envolvidas na realização do isolamento de células únicas, bem como amplificação do seu genoma.

## 2 I MÉTODOS DE ISOLAMENTO DE CÉLULA ÚNICA

A primeira etapa do estudo com células únicas requer o isolamento de células individuais. Para tanto, foram desenvolvidas diferentes tecnologias para isolamento celular, sendo as principais descritas mais a diante neste capítulo.

Os métodos de isolamento de célula única podem seguir dois tipos de amostragem, a imparcial ou a tendenciosa. A amostragem imparcial é realizada de forma aleatória, refletindo melhor a composição do tecido estudado. Já a amostragem tendenciosa é feita de maneira direcionada, onde é possível isolar tipos celulares raros ou específicos de interesse (SHAPIRO; BIEZUNER; LINNARSSON, 2013).

Os principais métodos para isolamento de células únicas incluem: 1) Classificação Celular Ativada por Fluorescência (FACS); 2) Classificação de Células Ativadas Magneticamente (MACS); 3) Microdissecção de Captura a Laser (LCM); 4) Coleta Manual de Células ou Micromanipulação; e 5) Microfluídicos (Figura 1).

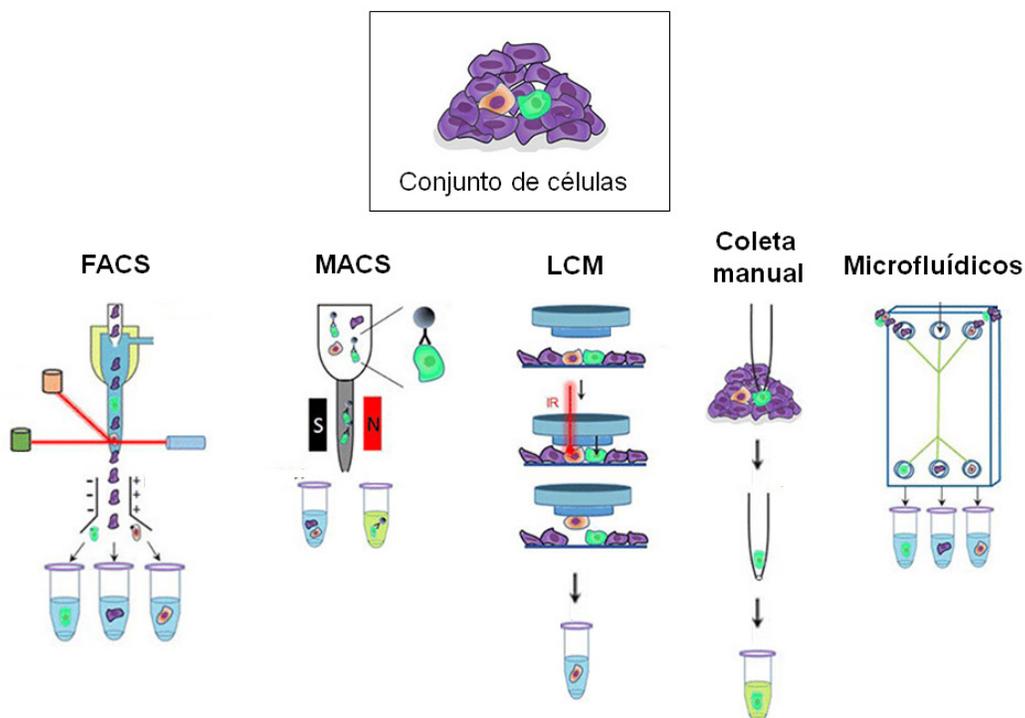


Figura 1. Visão geral das técnicas de isolamento celular. Fonte: Adaptado de Hu et al. (2016).

### **1) Classificação Celular Ativada por Fluorescência (FACS):**

Este método é baseado em citometria de fluxo que utiliza a marcação de células alvo com de sondas fluorescentes, como por exemplo, anticorpos monoclonais, que reconhecem marcadores específicos nas células. Conforme as células marcadas em

suspensão percorrem a citometria, um laser é emitido pelo aparelho, e dessa forma, as células podem ser separadas conforme características específicas (HERZENBERG; SWEET; HERZENBERG, 1976).

## **2) Classificação de Células Ativadas Magneticamente (MACS):**

A separação de células através deste método é baseada na separação por carga, utilizando anticorpos, enzimas, lectinas ou estreptavidinas associadas a esferas magnéticas para ligar proteínas específicas nas células-alvo. Posteriormente a esta marcação, as células são colocadas em um campo magnético externo e as esferas magnéticas ativadas. Consequentemente, células marcadas com as esferas magnéticas polarizam e as demais são lavadas (MILTENYI et al., 1990).

## **3) Microdissecção de Captura a Laser (LCM):**

O método LCM permite a separação de células a partir de amostras de tecido principalmente sólidas em uma lâmina de microscópio. É utilizado um microscópio invertido com um sistema de captura a laser infravermelho ou ultravioleta. A partir da visualização e reconhecimento da célula alvo, o laser é emitido a fim de derreter um fino filme termoplástico (polímero de acetato de etileno e vinil) acima da célula, fundindo-o a ela. A remoção do filme permite também a remoção da célula, enquanto o restante do tecido permanece na lâmina (EMMERT-BUCK et al., 1996).

## **4) Coleta Manual de Células ou Micromanipulação:**

Este método é frequentemente utilizado, em virtude do baixo custo e do simples processo. A técnica consiste na utilização de um microscópio invertido combinado com micropipetas, onde o pesquisador visualiza o conjunto de células e isola-as de forma imparcial.

## **5) Microfluídicos:**

Existem diferentes categorias dentro deste método, no qual a técnica baseada em cromatografia de afinidade celular é a mais utilizada. Utiliza-se um chip modificado com anticorpos específicos, capazes de ligar-se ao antígeno na superfície celular das células de interesse. À medida que a amostra flui através de microcanais, as células alvo são ligadas ao chip através da interação antígeno-anticorpo e são imobilizadas, enquanto as outras células fluem normalmente. Este método é geralmente associado à FACS ou MACS, por exemplo.

Além dos métodos de isolamento celular descritos acima, existem métodos para isolamento direto do núcleo. A seguir iremos descrever rapidamente os dois principais métodos: 1) Isolamento de Núcleos Marcados em Tipo Celular Específico (INTACT) e; 2) Isolamento Descontínuo da Cromatina Específica de Tecido para Imunoprecipitação (BiTS-ChIP) (Figura 2).

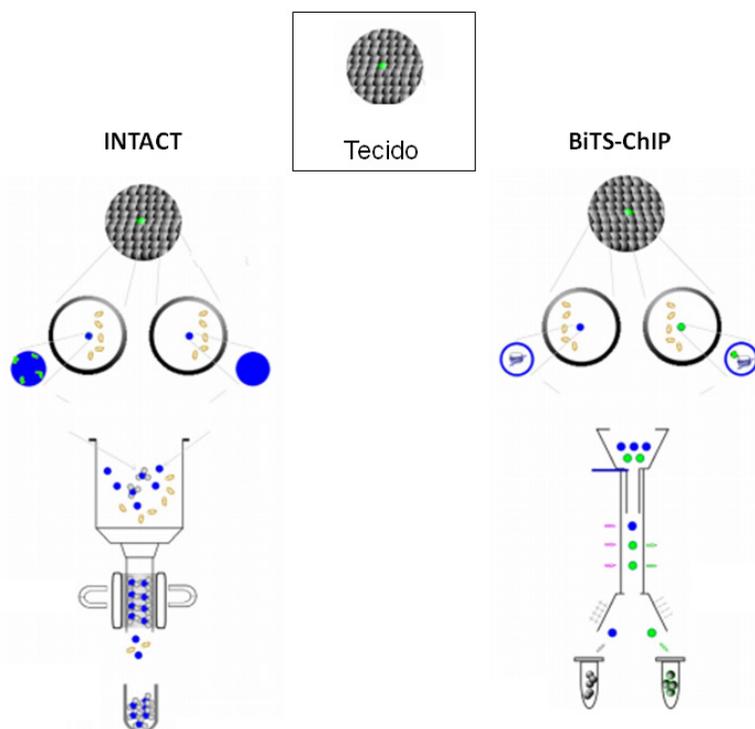


Figura 2. Visão geral das técnicas de isolamento de núcleos. Fonte: Adaptado de Zeb et al. (2019).

### **1) Isolamento de Núcleos Marcados em Tipo Celular Específico (INTACT):**

INTACT é um método para isolar núcleos de células específicas, baseado em afinidade. O método utiliza marcação de núcleos com biotina a partir de uma proteína de fusão direcionada ao envelope nuclear externo. A purificação dos núcleos marcados é realizada com o auxílio de esferas magnéticas revestidas com estreptavidina (DEAL; HENIKOFF, 2010).

### **2) Isolamento Descontínuo da Cromatina Específica de Tecido para Imunoprecipitação (BiTS-ChIP):**

Esta técnica é realizada marcando com fluorescência uma proteína nuclear expressa no tipo celular de interesse. Os núcleos fluorescentes são classificados através de FACS e posteriormente a cromatina é cortada e extraída. O corte da cromatina possibilita a imunoprecipitação do DNA associado à proteína de interesse (BONN et al., 2012).

### 3 | AMPLIFICAÇÃO DO DNA E SINGLE-CELL GENOMICS

Quando trabalhamos com metodologias de amplificação do DNA, seja por técnicas convencionais ou por métodos de célula única, a qualidade e a quantidade do DNA genômico extraído são fundamentais. Uma célula humana diplóide, por exemplo, contém aproximadamente 7 pg de DNA genômico e para realizarmos as diversas análises é necessária à sua amplificação (MACAULAY; VOET, 2014).

Dessa forma, muitos métodos de amplificação de genoma inteiro foram desenvolvidos para gerar quantidade de DNA suficiente para o posterior sequenciamento. Os três principais métodos de amplificação do genoma incluem: Reação em Cadeia da Polimerase Iniciada por Oligonucleotídeo Degenerado (DOP-PCR); Amplificação de Deslocamento Múltiplo (MDA); e Múltiplos Ciclos de Amplificação Baseados em Anelamento e Loop (MALBAC).

Os métodos possuem diferentes vantagens e desvantagens, as quais devem ser levadas em consideração na hora do pesquisador escolher o melhor método para sua pesquisa. Bourcy e colaboradores (2014) comparando quantitativamente o desempenho dos métodos MDA e MALBAC na amplificação de todo o genoma concluíram que nenhum método apresenta total desempenho em todos os critérios (uniformidade da cobertura, taxas de erro e nível de contaminação). Dessa forma, a escolha do método deve obedecer ao objetivo do experimento.

#### **1) Reação em Cadeira da Polimerase Iniciada por Oligonucleotídeo Degenerado (DOP-PCR):**

A PCR é uma técnica bastante utilizada na biologia molecular com a finalidade de amplificar uma região do DNA, sendo capaz de gerar milhares a milhões de cópias após vários ciclos da reação. Este método é dependente da enzima Taq DNA polimerase e de primers específicos para a região de interesse.

O método DOP-PCR foi proposto em 1992 por Telenius e colaboradores, utilizando primers degenerados (Figura 3). Sabe-se que o código genético é degenerado, ou seja, um mesmo aminoácido pode ser codificado por diferentes códons. Considerando isso, para que todas as possibilidades de códons sejam expressas, primers degenerados não são formados por uma sequência específica e única, mas sim por um conjunto de sequências, de acordo com as possíveis bases em determinadas posições.

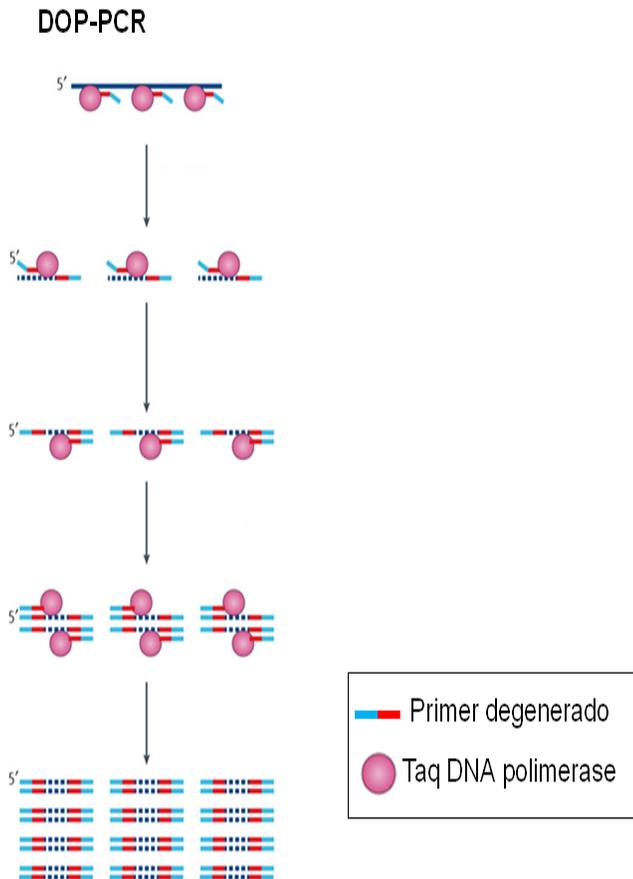


Figura 3. Visão geral da técnica DOP-PCR. A dupla fita de DNA é desnaturada para permitir o anelamento dos primers degenerados às fitas simples. Posteriormente, a Taq DNA polimerase inicia a amplificação da fita simples de DNA. Depois de repetidos ciclos, existem várias cópias do mesmo DNA, e as fitas simples se unem por complementariedade. Fonte: Adaptado de Gaward, Koh e Quake (2016).

O princípio do método DOP-PCR é a utilização do primer degenerado. A reação inicia-se com o processo de pré-amplificação, onde ocorre a desnaturação da dupla fita de DNA, seguida do anelamento do primer a uma temperatura inicial baixa. A pré-amplificação é seguida pela amplificação por PCR das fitas de DNA. Após vários ciclos, o DNA molde é então amplificado em várias cópias.

Para uma amplificação eficiente, os autores indicam dois pontos fundamentais, que incluem uma baixa temperatura inicial, a fim de permitir ao primer iniciar o processo a partir de sequências alvo curtas, e utilização de primer degenerado. Com relação ao rendimento, fatores limitantes incluem a concentração do primer utilizado e da Taq polimerase, pois, diferente da PCR convencional, a DOP-PCR busca amplificar o maior número possível de locais (TELENIUS et al., 1992).

## 2) Amplificação de Deslocamento Múltiplo - MDA

Diferente dos métodos baseados em PCR que utilizam a Taq DNA polimerase para a polimerização de novas fitas de DNA, Dean e colaboradores descreveram o método MDA em 2002, utilizando a enzima polimerase replicativa do fago *Bacillus subtilis* phi29 (phi29 DNA polimerase).

A phi29 DNA polimerase além de realizar o processo de polimerização, é capaz de produzir deslocamento da fita recém-sintetizada (BLANCO et al., 1989). Dessa forma, polimerizações secundárias são iniciadas a partir de produtos primários (Figura 4).

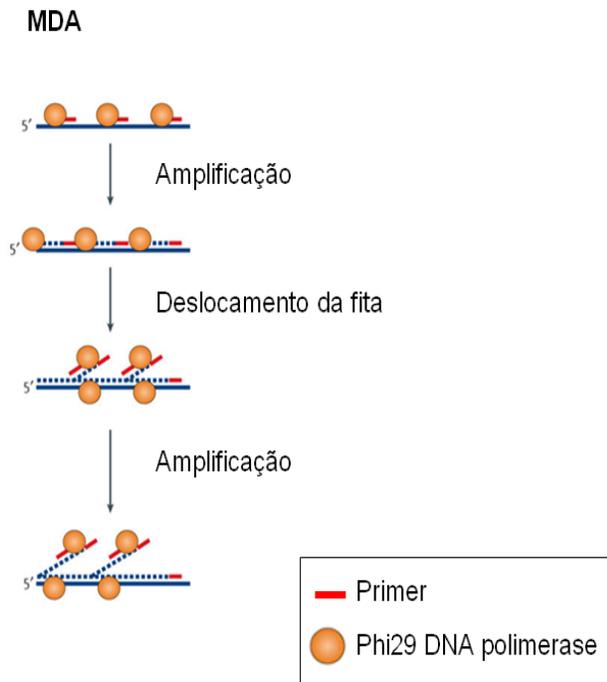


Figura 4. Visão geral da técnica de MDA. Após anelamento do primer, a Phi29 DNA polimerase inicia a amplificação do DNA. A fita recém-sintetizada é deslocada ainda pela Phi29 DNA polimerase, permitindo novos anelamentos de primers e amplificações secundárias. Fonte: Adaptado de Gaward, Koh e Quake (2016).

O processo do método é semelhante à PCR, ocorrendo desnaturação da dupla fita de DNA, anelamento do primer e sintetização de novas fitas. Porém, como a polimerase em questão possibilita além a amplificação, o deslocamento da fita recém-sintetizada, amplificações secundárias são realizadas nessas fitas, gerando maior rendimento.

MDA é uma técnica simples, a qual pode ser feita com uma variedade de amostras biológicas e seu produto de DNA gerado é maior que 10 kb. O método é eficiente para diversas aplicações, como análise de polimorfismo de nucleotídeo único (SNP), polimorfismo de comprimento de fragmento de restrição (RFLP) e hibridação comparativa de genoma. Além disso, o MDA não depende de elevada temperatura para a desnaturação da dupla fita de DNA, dessa forma, as chances de degradação do DNA são reduzidas, e a especificidade

da amplificação da sequência melhora (DEAN et al., 2002).

### 3) Múltiplos Ciclos de Amplificação Baseados em Anelamento e Loop – MALBAC

Em 2012, Zong e colaboradores descreveram um novo método de amplificação, denominado MALBAC. Este é um método híbrido, onde inicialmente ocorre amplificação com a phi29 DNA polimerase, seguida de amplificação por PCR com a Taq DNA polimerase. (Figura 6).

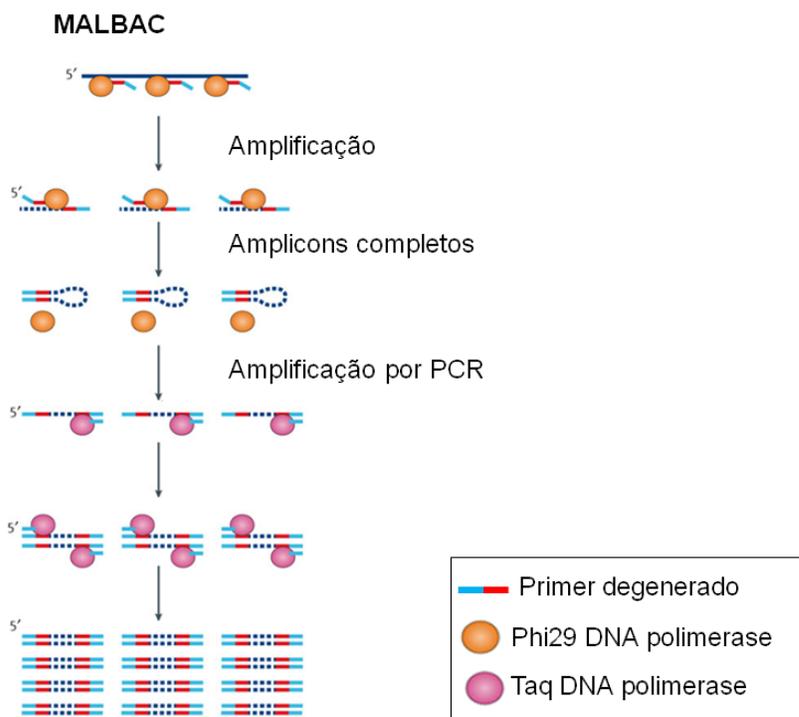


Figura 6. Visão geral da técnica MALBAC. O método híbrido inicia com anelamento do primer degenerado e amplificação e deslocamento da fita pela Phi29 DNA polimerase. Essa primeira amplificação dá origem a semiamplicons. Ocorre outra amplificação originando amplicons completos que possuem a extremidade 3' complementar à extremidade 5'. Essas extremidades complementares se unem, gerando um DNA em loop. Os amplicons completos são amplificados pela Taq DNA polimerase em uma reação de PCR, permitindo a formação de diversas cópias do mesmo DNA. Fonte: Adaptado de Gaward, Koh e Quake (2016).

A amplificação pela Phi29 DNA polimerase gera semiamplicons. Posteriormente, ocorre uma segunda amplificação nestes semiamplicons, gerando amplicons denominados completos, que por possuírem sua extremidade 3' complementar à sequência na extremidade 5', formam uma estrutura de DNA em loop. Após repetidos ciclos de amplificação e deslocamento, os amplicons completos são encaminhados para a amplificação por PCR pela Taq DNA polimerase.

O método garante maior uniformidade na cobertura do genoma, chegando a uma cobertura de até 93% do genoma, bem como detecção mais precisa de variações no número de cópias e polimorfismos de nucleotídeo único (ZONG et al., 2012).

## 4 | APLICAÇÕES E CONSIDERAÇÕES FINAIS

Os avanços nas tecnologias de análises de células únicas possibilitaram novas descobertas na biologia molecular, bem como na medicina. Tecnologias unicelulares podem direcionar o desenvolvimento de novos medicamentos, principalmente imunoterápicos contra o câncer (YOFE; DAHAN; AMIT, 2020).

As pesquisas com câncer foram muito beneficiadas com os avanços nas análises de célula única, uma vez que tumores são complexos e apresentam grande heterogeneidade celular. Entender o ambiente intratumoral possibilita melhores respostas aos tratamentos propostos. Além disso, ressalta-se o grande potencial de tecnologias unicelulares na detecção de lesões pré-neoplásicas, contribuindo no diagnóstico da doença (BASLAN; HICKS, 2017).

O sequenciamento genômico de célula única também contribuiu para o avanço em diversas outras áreas da biologia molecular, incluindo a identificação de variações cromossômicas, como número de cópias e variações de nucleotídeo único, evolução de tumores, gênese dos gametas, mosaïcismo somático, o que se reflete na heterogeneidade genômica entre uma população de células (HU et al., 2016).

Percebe-se o crescente progresso nos últimos tempos nos métodos de análise de células únicas. Ainda existem limitações a serem superadas, como o aumento na taxa de rendimento do isolamento celular, bem como melhorias na amplificação do DNA. Fato é que o sequenciamento genômico de células únicas ampliará cada vez mais o nosso conhecimento acerca da complexidade e heterogeneidade celular em organismos normais, bem como no desenvolvimento dos indivíduos e das doenças.

## REFERÊNCIAS

BASLAN, T.; HICKS, J. **Unravelling biology and shifting paradigms in cancer with single-cell sequencing.** *Nature Reviews Cancer.* v. 17, p. 557–569. 2017.

BOURCY, C. F. *et al.* **A quantitative comparison of single-cell whole genome amplification methods.** *PloS one.* v. 9, n. 8, p 1-9. 2014.

BLANCO, L. *et al.* **Highly efficient DNA synthesis by the phage phi29 DNA polymerase.** *The journal of biological chemistry.* v. 264, p. 8935-8940. 1989.

BONN, S. *et al.* **Cell type-specific chromatin immunoprecipitation from multicellular complex samples using BiTS-ChIP.** *Nature Protocols.* v. 7, n. 5, p. 978-994. 2012.

DEAL, R. B.; STEVEN, H. **A simple method for gene expression and chromatin profiling of individual cell types within a tissue.** *Developmental Cell.* v. 18, p. 1030–1040. 2010.

DEAN, F. B. *et al.* **Comprehensive human genome amplification using multiple displacement amplification.** Proceedings of the National Academy of Sciences. v. 99, n. 8, p. 5261–5266. 2002.

EMMERT-BUCK, M. R. *et al.* **Laser capture microdissection.** Science. v. 274, p. 998-1001. 1996.

GAWARD, C.; KOH, W.; QUAKE, S. R. **Single-cell genome sequencing: current state of the science.** Nature Reviews Genetics. v. 17, p.175–188. 2016.

GROSS, A. *et al.* **Technologies for Single-Cell Isolation.** International journal of molecular sciences. v. 16, n. 8, p. 16897–16919. 2015.

HERZENBERG, L. A.; SWEET, R. G.; HERZENBERG, L. A. **Fluorescence-Activated Cell Sorting.** Scientific American. v. 234, n. 3, p. 108–117. 1976.

HUANG, L. *et al.* **Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications.** Annual Review of Genomics and Human Genetics. v. 16, p. 79-102. 2015.

HUANG, S. **Non-genetic heterogeneity of cells in development: more than just noise.** Development. v. 136, p. 3853–3862. 2009.

HU, P. *et al.* **Single Cell Isolation and Analysis.** Frontiers in cell and developmental biology. v. 4, p. 1-12. 2016.

KALISKY, T.; BLAINEY, P.; QUAKE, S. R. **Genomic analysis at the single-cell level.** Annual review of genetics. v. 45, p. 431–445. 2011.

MACAULAY, I.C.; VOET, T. **Single cell genomics: advances and future perspectives.** PLoS Genetics. v. 30, p. 1-9. 2014.

MILTENYI, S. *et al.* **High gradient magnetic cell separation with MACS.** Cytometry. v. 11, p. 231-238. 1990.

SHAPIRO, E.; BIEZUNER, T.; LINNARSSON, S. **Single-cell sequencing-based technologies will revolutionize whole-organism science.** Nature Reviews Genetics. v. 14, p. 618–630. 2013.

TELENIUS, H. *et al.* **Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer.** Genomics. v. 13, p. 718-725. 1992.

**Method of the Year 2013.** Nature Methods. v.11, n. 1, p. 1. 2014.

YOFE, I.; DAHAN, R.; AMIT, I. **Single-cell genomic approaches for developing the next generation of immunotherapies.** Nature Medicine. v. 26, p. 171–177. 2020.

ZEB, Q. *et al.* An Overview of Single-Cell Isolation Techniques. In: BARH, D.; AZEVEDO, V. **Single-Cell Omics: Technological Advances and Applications.** 1. ed. Cambridge: Academic Press, 2019. cap. 6, p. 101-135.

ZONG, C. *et al.* **Genome-wide detection of single-nucleotide and copy-number variations of a single human cell.** Science. v. 338, p. 1622–1626. 2012.

*Data de aceite: 10/05/2021*

*Data de submissão: 08/04/2021*

### Mayara Mota de Oliveira

Programa de Pós-Graduação em Biotecnologia/  
RENORBIO, Universidade Federal do Espírito  
Santo, Vitória, Espírito Santo, Brasil.  
<http://lattes.cnpq.br/5495329160438980>

### Anderson Barros Archanjo

Programa de Pós-Graduação em Biotecnologia  
RENORBIO, Universidade Federal do Espírito  
Santo, Vitória, Espírito Santo, Brasil.  
<http://lattes.cnpq.br/5529149503714764>

**RESUMO:** O entendimento sobre o transcriptoma trouxe novas perspectivas sobre a biologia da célula, no qual o estudo do RNA mensageiro (RNAm) e da expressão gênica poderiam auxiliar no entendimento sobre as respostas obtidas à medicamentos, tratamentos, e comportamentos biológicos associados. O PCR de transcrição reversa quantitativa (RT-qPCR) é resultado da evolução da técnica amplamente empregada na biologia molecular de reação em cadeia da polimerase (PCR) e as plataformas de quantificação. A técnica baseia-se na formação de uma fita de DNA complementar (cDNA) a partir de RNA (RNAm), uma enzima transcriptase reversa e a extensão de um ou mais oligonucleotídeo. Para além da pesquisa a ser aplicada, o RT-qPCR ganhou um conjunto de estratégias metodológicas que podem variar nos métodos de detecção, na escolha sobre os iniciadores da transcrição reversa, nas etapas de execução e outras. Este

capítulo reúne os processos de evolução da técnica e apresenta algumas das estratégias amplamente empregadas na RT-qPCR.

**PALAVRAS-CHAVE:** RT-qPCR; Expressão Gênica; Biologia Molecular

### RT-qPCR: GENE EXPRESSION STUDIES

**ABSTRACT:** The understanding of the transcriptome brought new perspectives on the cell's biology. Thus, the study of messenger RNA (mRNA) and gene expression could assist in understanding the responses obtained to drugs, treatments, and associated biological behaviors. The quantitative reverse transcription PCR (RT-qPCR) is the result of the evolution of the technique widely used in molecular biology of polymerase chain reaction (PCR) and the quantification platforms. The technique is based on the formation of a complementary DNA strand (cDNA) from RNA (mRNA), a reverse transcriptase enzyme and the extension of one or more oligonucleotides. The RT-qPCR gained a set of methodological strategies that can vary in the detection methods, in the choice of the initiators of reverse transcription, in the stages of execution, and others. This study brings together the processes of evolution of the technique and presents some of the strategies widely used in RT-qPCR.

**KEYWORDS:** RT-qPCR; Gene expression; Molecular biology

## 1 | INTRODUÇÃO

O percurso evolutivo das tecnologias de diagnóstico avançou ao passo da criação de novas máquinas, insumos e tecnologias. Seguindo esse percurso a reação em cadeia da

polimerase (PCR<sup>1</sup>), usada em larga escala para pesquisas biológicas, no campo clínico e de diagnóstico e forense, alcançou novos alvos genômicos, deixando de ser restrito ao DNA, e aplicando-se ao RNA e suas variantes mi-, pi-, sno-, si-, sr-, U-RNA (HOLLAND et al., 1991; KUBISTA et al., 2006).

O entendimento sobre o transcriptoma trouxe novas perspectivas sobre a biologia da célula, e as especificações e desdobramentos relacionados ao RNA mensageiro (RNAm) e a regulação gênica. Entender o padrão de expressão para diferentes tecidos e as correlações químicas, bioquímicas e respostas fisiológicas poderiam auxiliar no entendimento sobre as respostas obtidas à medicamentos, tratamentos, os comportamentos associados e até sobre os mecanismos iniciadores (ROCHA; SANTOS; PACHECO, 2015; SANDERS et al., 2018).

Para se alcançar o que conhecemos hoje como PCR de transcrição reversa quantitativa (RT-qPCR<sup>2</sup>), foi necessário o aprimoramento e derivações da técnica, na qual o PCR digital (dPCR), PCR quantitativo em tempo real (qPCR) e PCR de transcrição reversa (RT-PCR), somam-se para solucionar a necessidade de uma análise qualitativa e quantitativa dos ácidos nucleicos nos diferentes sistemas biológicos (TAYLOR; MRKUSICH, 2014) (Tabela 1).

Para melhor entendimento, as técnicas citadas têm como pilar à detecção dos produtos do PCR com o auxílio de uma Taq polimerase estável ao calor, Desoxirribonucleotídeos Fosfatados (dNTPs) e as sondas iniciadoras (oligonucleotídeo). Mediante a evolução do termocicladores, acrescentou-se uma molécula repórter florescente que auxilia a identificação e quantificação dos produtos em análise produzida em cada ciclo (HOLLAND et al., 1991; GIBSON; HEID; WILLIAMS, 1996).

	<i>PCR</i>	<i>qPCR</i>	<i>RT-qPCR</i>
<b>Princípio da Técnica</b>	Baseia-se na presença ou ausência da amplificação do material genético.	A marcação fluorescente permite a detecção da amplificação (dados) de acordo com os ciclos da PCR.	Reação direcionada ao entendimento sobre o RNA e expressão gênica.
<b>Método</b>	<ul style="list-style-type: none"> <li>· Amplificação de uma Região gênica;</li> <li>· Combinação de DNA polimerase, cofatores, nucleotídeos, e sequencia direcionadora (<i>Primer</i>);</li> <li>· Precisa fazer eletroforese em gel para ver seu produto de PCR amplificado.</li> </ul>	<ul style="list-style-type: none"> <li>· Uso de marcador fluorescente, somado aos constituintes da reação de PCR;</li> <li>· Resultado é atribuído ao <i>Ct</i> de amplificação em resultado digital.</li> </ul>	<ul style="list-style-type: none"> <li>· Etapa antecessora de conversação do RNA em cDNA;</li> <li>· <i>RT: Transcriptase reversa;</i></li> <li>· Segue os parâmetros da qPCR para análise dos transcritos.</li> </ul>

1. Dados Históricos: O PCR foi inventado por Kary B. Mullis e Michael Smith no início dos anos 80. Ambos ganharam o Nobel de Química de 1993 pela invenção.

2. Nos anos 90 foi considerado como o “padrão ouro” na análise quantitativa de ácido nucleico, devido à sua alta sensibilidade, boa reprodutibilidade, ampla faixa dinâmica de quantificação, fácil uso e razoável custo-benefício.

<b>Aplicações Comuns</b>	<ul style="list-style-type: none"> <li>· Sequenciamento;</li> <li>· Clonagem;</li> <li>· Detecção de Mutações;</li> <li>· Validação de OGM;</li> <li>· Detecção de patógenos;</li> <li>· Testes forenses.</li> </ul>	<ul style="list-style-type: none"> <li>· Genotipagem;</li> <li>· Análise de Mutações;</li> <li>· Análise de Metilação do DNA;</li> <li>· Detecção de patógenos;</li> <li>· Quantificação de carga viral.</li> </ul>	<ul style="list-style-type: none"> <li>· Análise de perfis de expressão gênica;</li> <li>· Aplica-se ao RNA e suas variantes.</li> </ul>
--------------------------	--	---	--

Tabela 1. Classificação e diferenciação das técnicas de PCR, qPCR e RT-qPCR

## 2 | A TRANSCRIÇÃO REVERSA

Para o estudo do RNA, de forma que este seja passível à técnica de PCR necessitou de uma nova etapa que antecederesse a técnica e denominada transcrição reversa (RT da sigla). Esta etapa baseia-se na formação de uma fita de DNA complementar (cDNA) a partir de RNA (RNAm), uma enzima transcriptase reversa e a extensão de um ou mais oligonucleotídeo iniciadores (GIBSON; HEID; WILLIAMS, 1996; BUSTIN, 2000).

Inicialmente observado no comportamento dos retrovírus nos processos infecciosos, para a aplicação molecular o processo de iniciação da reação de transcrição reversa pode usar três abordagens, que são: *Random Primers*, *Oligo(dT)s* e os *Primers* de sequência específica (*Sequence Specific Primers*). Os *Random Primers*, amplamente empregados atualmente, baseiam-se em sondas variadas que se anelam e pontos variados do RNA como ponto de partida. Os *Oligo(dT)s* baseiam-se na estrutura do RNAm, na qual o *primer* se âncora a extremidade 5' da cauda poli (A). Já o *Primers* de sequência específica baseia-se em *Primers* personalizados que visam a sequência específica de um mRNA, garantindo que a síntese do cDNA seja específica a um gene de interesse (STANGEGAARD; DUFVA; DUFVA, 2006; LI; BROWNLEY, 2010; GREEN; SAMBROOK, 2019) (Figura 1).

O processo de síntese deste cDNA é dependente das sequencias de anelamento destes iniciadores de reação, polimerização e inativação da atividade da enzimaDNA polimerase RNA-dependente (*Transcriptase reversa*), em um processo dependente do aquecimento similar ao PCR (BUSTIN, 2000; NAKAMURA et al., 2003).

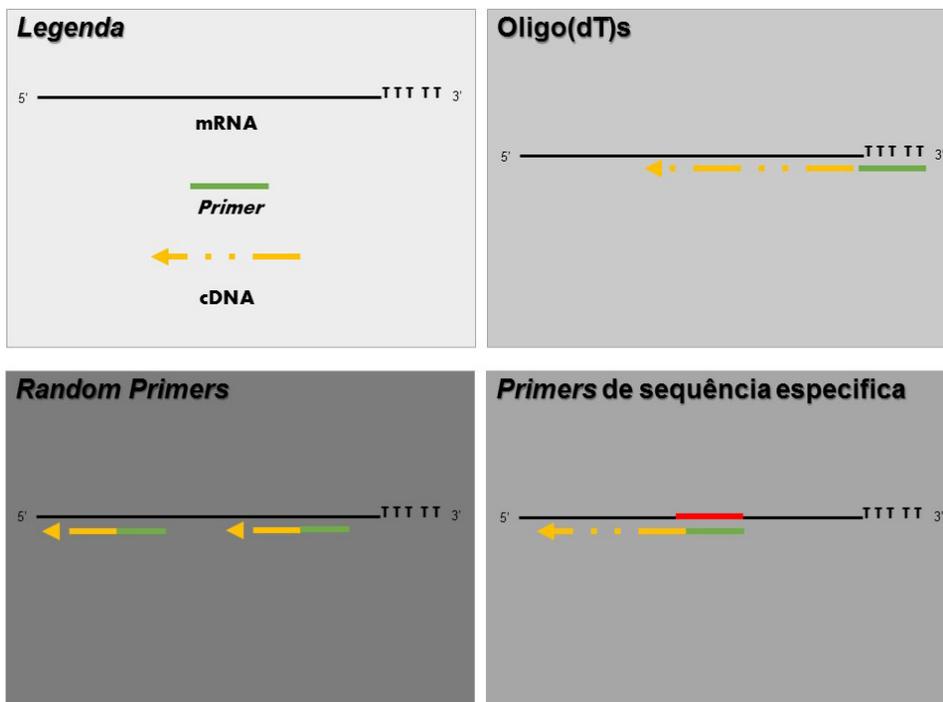


Figura 1. Iniciadores da Transcrição Reversa. Os iniciadores *Oligo(dT)s* se ligam a calda poli(A) do RNA mensageiro. Os iniciadores *RandomPrimers* se ligam a pontos diversos do RNA mensageiro. Os iniciadores *Primers* de sequência específica se ligam à uma região específica do RNA mensageiro.

### 3 | ONE-STEP X TWO-STEP

Entre as estratégias para realização das reações do RT-qPCR está a escolha de kits a serem realizados em uma ou duas etapas. Essas estratégias baseiam-se na combinação ou não da etapa da transcrição reversa e PCR em tempo real, modificando-se assim as estratégias de iniciação, os processos de otimização dos *primers*, e as condições de reação (WONG; MEDRANO, 2005).

Entre as vantagens e desvantagens relatadas pelas marcas disponíveis no mercado demonstra-se que a escolha por fazer as reações em um único tubo garante uma menor variação experimental e uma maior reprodutibilidade do experimento, no entanto, questões como otimização das reações, baixa sensibilidade e a detecção de menos alvos por amostra são pontos negativos para a escolha desta abordagem experimental (Figura 2) (DE PAULA et al., 2004; RASHED-UL ISLAM; JAHAN; TABASSUM, 2015).

Na abordagem mais comum, no qual as etapas são feitas de forma separada, a vantagem de maior destaque é a possibilidade de armazenamento do cDNA por longos períodos, e a disponibilidade deste para novas reações, além da possibilidade de otimização e flexibilização das reações. No entanto, o processo em duas etapas segue sendo mais oneroso em tempo e no uso de materiais plásticos.

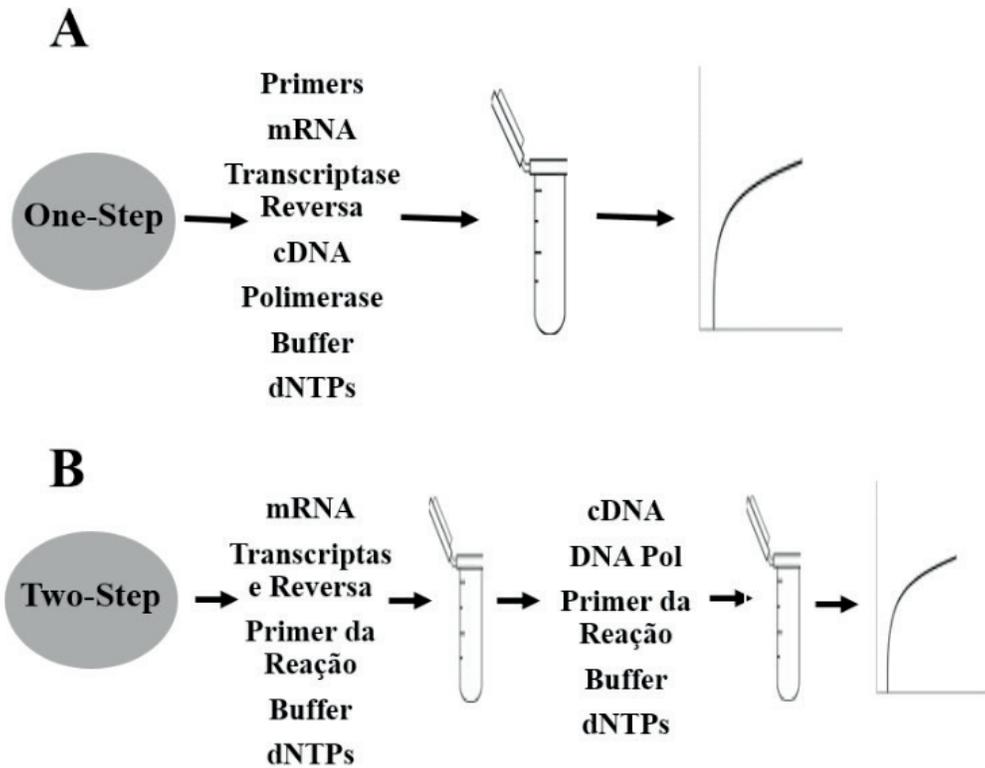


Figura 2. Estratégias para realização das reações do RT-qPCR. **A.** Reação estabelecida em uma única etapa, na qual os reagentes para a transcrição reversa e para o qPCR participam da mesma reação. **B.** Reação estabelecida em 2 etapas. A primeira o referente a transcrição reversa e a segunda referente ao qPCR.

### 3.1 Instrumentação e princípios de detecção da técnica

O princípio das análises de PCR em “tempo real” pauta-se na identificação dos fluorocoromo “*Reporter*” associados ao produto do PCR durante os ciclos de reação (HEID et al., 1996). Nesse sentido, o entendimento sobre esta técnica baseia-se na escolha dos sistemas de detecção, e seus derivados como: módulo de reação, sistema de detecção óptica, e os softwares (BUSTIN, 2000).

O módulo de reação consiste no processo de aquecimento e resfriamento do equipamento em uso, o entendimento deste processo relaciona-se as etapas de separação das fitas duplas, hibridização dos *primers* e a polimerização. Ligado a esta etapa também está a escolha correta das placas ou tubos usados para perfeito aproveitamento da reação. Em relação ao sistema de detecção é importante entender a interação entre fluorescência e filtros que acompanham o equipamento e o fluorocoromo “*Reporter*” à ser usado. O software em questão é a forma de interação com a reação e necessária para o processo de interpretação dos resultados. O entendimento destas etapas garante que erros sejam evitados na escolha dos kits, e o processo de otimização e configurações ganhem mais celeridade (BUSTIN, 2000).

Entre os métodos de detecção disponíveis no mercado, os sistemas mais comuns são o de detecção específica (Sistema TaqMan®) ou inespecífica (Sistema SYBR Green I). No primeiro, utiliza-se uma sonda fluorescente ancorada na extremidade 5' por um fluorocromo "Reporter", que tem como papel emitir a fluorescência, e em uma molécula "Quencher" na sua extremidade 3', que atua na absorção dessa fluorescência enquanto a sonda está estável, em um sistema denominado transferência de energia de ressonância de fluorescência (FRET). O processo de detecção é dependente da clivagem da sonda pela enzima Taq DNA polimerase, que libera o fluorocromo "Reporter" que fica livre para ser excitado e produzir fluorescência (HEID et al., 1996) (Figura 3). Nesse sentido, a fluorescência produzida é proporcional aos alvos recém-sintetizados a cada ciclo.

Já o sistema SYBR Green I, utiliza fluorocoromos com afinidade por moléculas de DNA dupla-fita, de forma que com o decorrer dos ciclos as moléculas se intercalamos produtos do PCR. A etapa de detecção ocorre no processo de polimerização da dupla fita neste sistema, já que a fluorescência é reduzida no processo de desnaturação (MORRISON; WEIS; WITTEWER, 1998). Por ser um método mais barato, a técnica é amplamente utilizada pelos pesquisadores, no entanto, a inespecificidade da fluorescência aos alvos de escolha pode-se tornar um problema na amplificação de produtos inespecíficos ou na ocorrência de dímeros de *primers* (Figura 3).

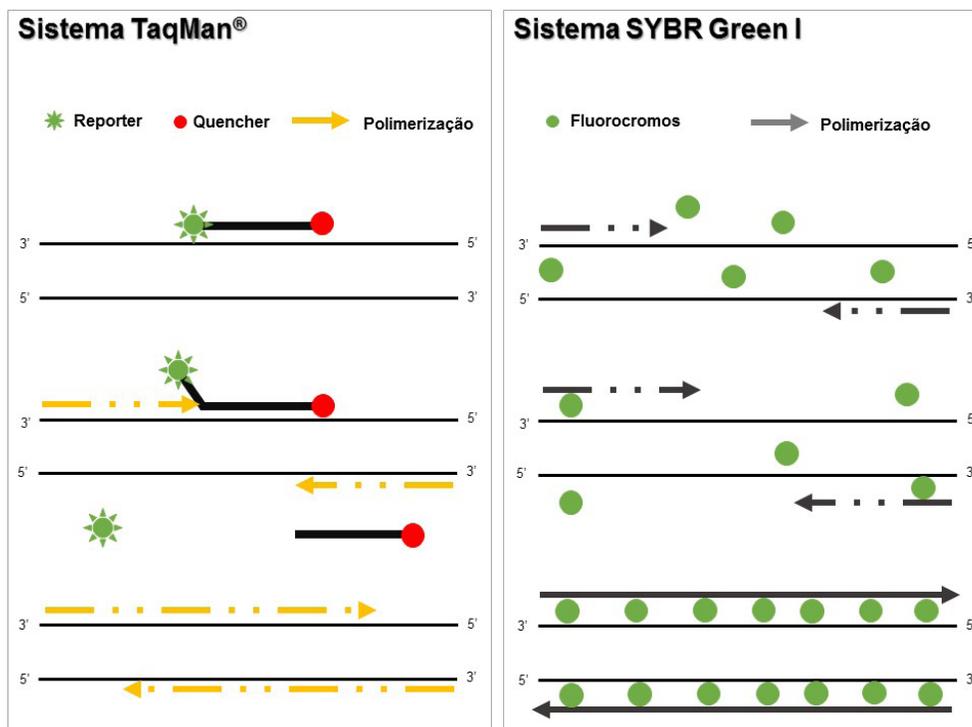


Figura 3. Sistema de detecção no qPCR. O Sistema TaqMan® baseia-se na desestabilização do sistema transferência de energia de ressonância de fluorescência (FRET) entre fluorocromo "Reporter" e a molécula "Quencher". O rompimento da ligação libera o fluorocromo "Reporter" para ser excitado e produzir fluorescência. O sistema SYBR Green I baseia-se na ligação do fluorocromo as moléculas de DNA dupla-fita.

Com o entendimento dos métodos de detecção a quantificação da técnica baseia-se na fluorescência obtida em relação aos ciclos de PCR. Nesse sentido, a representação gráfica da curva de amplificação pode ser dividida em fase de iniciação, fase exponencial e fase platô (HEID et al., 1996; DING; CANTOR, 2004).

Os primeiros ciclos caracterizam-se por uma baixa detecção da fluorescência, e com o passar dos ciclos e com um acúmulo exponencial dos produtos do PCR, o sinal de fluorescência torna-se detectável nomeando-se esta etapa como ciclo de quantificação. O valor basal de detecção é a base da quantificação, demonstrando-se que em reações que você entra com a mesma concentração de material biológico consegue-se comparar a presença gene em estudo (DING; CANTOR, 2004; TAYLOR et al., 2010).

Este conceito também torna essencial as etapas de padronização das reações, já que a quantidade de material biológico presente no início da reação é um fator determinante para que o ciclo de quantificação seja identificado nos ciclos iniciais ou tardios da reação de qPCR (BUSTIN, 2000).

O ciclo de quantificação também pode ser entendido pelos termos apresentados pelo software, entre estes estão (Figura 4):

- $\Delta RN$ : é o sinal da fluorescência calculado pela subtração da fluorescência emitida pelo produto da reação e emissão basal detectada, e os valores são apresentados com base nos ciclos;
- THRESHOLD: Com base no limite de detecção da fluorescência o próprio software determina limiar baseado sinal de amplificação real para reação. Seu cálculo é baseado no desvio padrão do sinal médio da fluorescência entre os 15 primeiros ciclos da reação;
- $Ct$ : é definida como o valor em que o  $\Delta RN$  de uma amostra ultrapassa o limiar mínimo de detecção. Este valor compõem a base cálculo das metodologias de quantificação.

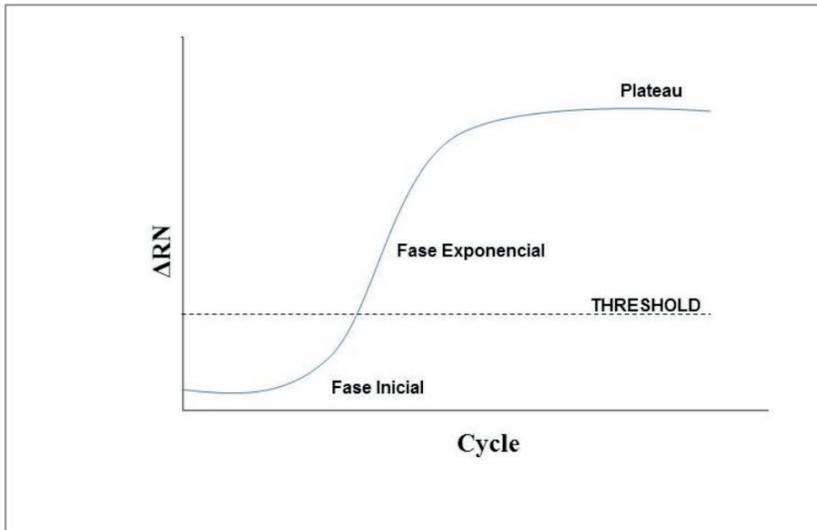


Figura 4. Exemplificação gráfica dos conceitos relacionados a curva de amplificação e definição do  $C_t$  da reação.

#### 4 | TEMPERATURA DE *melting* ( $T_m$ ):

Um dos princípios básicos para quem trabalha com PCR é a temperatura na qual a metade do produto de PCR está dissociado e a outra metade hibridado. A temperatura de *melting* está diretamente relacionada ao tamanho do fragmento em estudo e a presença de guanina e citosina neste (Figura 5).

Os *amplicons* produzidos na reação tem um pico de fusão característico à temperatura, nesse sentido, no RT-qPCR a análise sobre o *melting* contribui sobre a identificação e entendimento de agentes contaminantes ou da inespecificidade do *primer* (NAKAMURA et al., 2003).

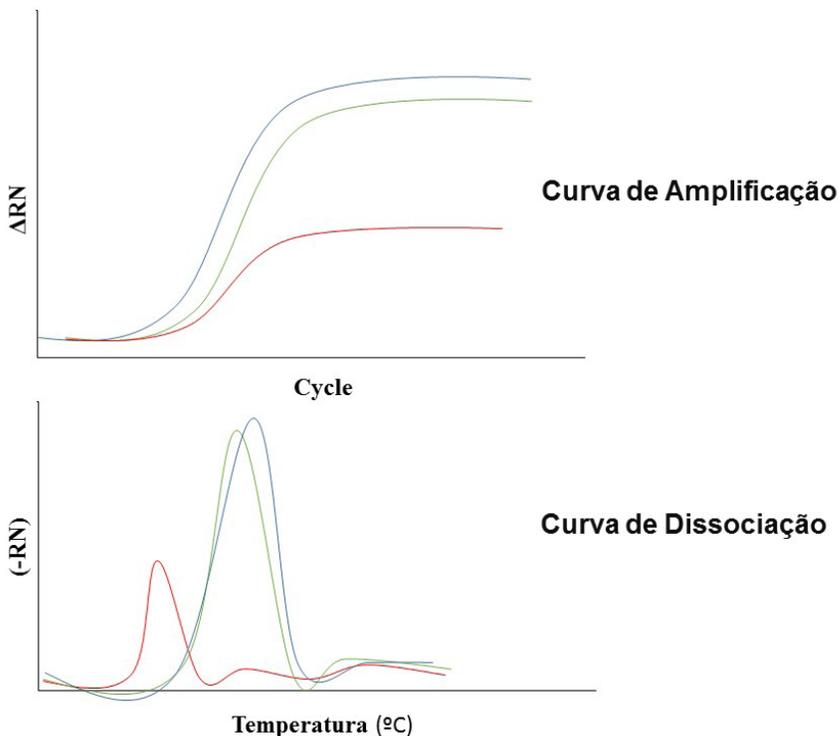


Figura 5. Exemplificação das Curva de Amplificação e de dissociação. A curva de amplificação demonstra a detecção de produtos amplificados em uma amostra. A Curva de dissociação que na reação foi encontrado diferentes temperaturas de *melting* para o produto da amplificação.

## 5 | ANÁLISE QUANTITATIVA

Como a grande parte dos métodos de quantificação em análises experimentais, para o RT-qPCR a escolha das abordagens de quantificação também pode ser absoluta ou relativa. O método de quantificação absoluta, está diretamente relacionado ao número de cópias ou produtos da PCR detectados, baseados em uma correlação com a curva padrão desenhada na reação (ROCHA; SANTOS; PACHECO, 2015).

A definição de curva padrão baseia-se na relação absorbância/ concentração. A aplicação no RT-qPCR o *Ct* torna-se o valor base de cálculo, porém é dependente de um conhecimento prévio dos padrões em uso que devem ser identificados por métodos independentes. Desta forma, o *Ct* “desconhecido” é comparado ao padrão para determinar o número de cópias iniciais da reação, que podem variar entre amostras de plasmídeo para o gene de interesse ou oligonucleotídeo sintético de cadeia simples (LIVAK; SCHMITTGEN, 2001).

Em relação a análise relativa, ela tem como objetivo avaliar o padrão de expressão gênica em comparação à uma amostra e a um gene referência, sendo assim denominado Método CT comparativo. A escolha por esta metodologia de análise é iniciada a partir da escolha adequada e validação dos genes endógenos a serem usados como controle de reação. Questões como tratamento e as vias que podem integrar a resposta devem ser

priorizadas para que este gene controle não tenha a sua expressão afetada pelo tratamento. A próxima etapa para à análise baseia-se no controle do experimento, que muitas vezes mimetiza situações em que a expressão do gene em estudo tenha um acréscimo de destaque (APPLIED BIOSYSTEMS, 2004).

Passados estes passos a quantificação relativa baseia-se no método de limiar comparativo (método 2-Ct), na qual o nível de expressão do gene alvo ( $\Delta Ct$ ) é calibrado pelo controle endógeno Ct da “amostra” -Ct do gene calibrador, na qual o gene calibrador exerce o papel de “limpeza” da reação retirando o que seria esperado como uma expressão basal. Exemplos como Gliceraldeído-3-fosfato desidrogenase (GAPDH), beta-actina e Proteína Ribossômica S18, são exemplos de controles endógenos bem estabelecidos e citados nos trabalhos (LIVAK; SCHMITTGEN, 2001).

Passados esta etapa, a expressão do gene de interesse é calibrada com o controle da reação experimental obtendo-se assim o  $\Delta\Delta Ct$ , seguido do cálculo  $2^{-\Delta\Delta Ct}$  (Tabela 2).

	Gene Calibrador	Gene Alvo	$\Delta Ct$ <i>Ct Alvo – Ct Calibrador</i>	$\Delta Ct$ Médio	$\Delta\Delta Ct$ <i><math>\Delta Ct</math> Alvo – <math>\Delta Ct</math> Calibrador</i>	$2^{-\Delta\Delta Ct}$
<b>Controle da reação no experimento</b>	17.93	22.42	4.49	4.39 ± 0.09	0 ± 0.09	1 (0.9 – 1.1)
	17.97	22.34	4.37			
	17.95	22.26	4.31			
	20.87	19.48	-1.39			
<b>Experimento Teste</b>	20.91	19.44	-1.47	-1.36 ± 0,13	-5,75 ± 0,13	53.82 (49.18 – 58.89)
	20.85	19.63	-1.22			

Tabela 2. Explicação metodologia para quantificação relativa baseada no 2-  $\Delta\Delta Ct$

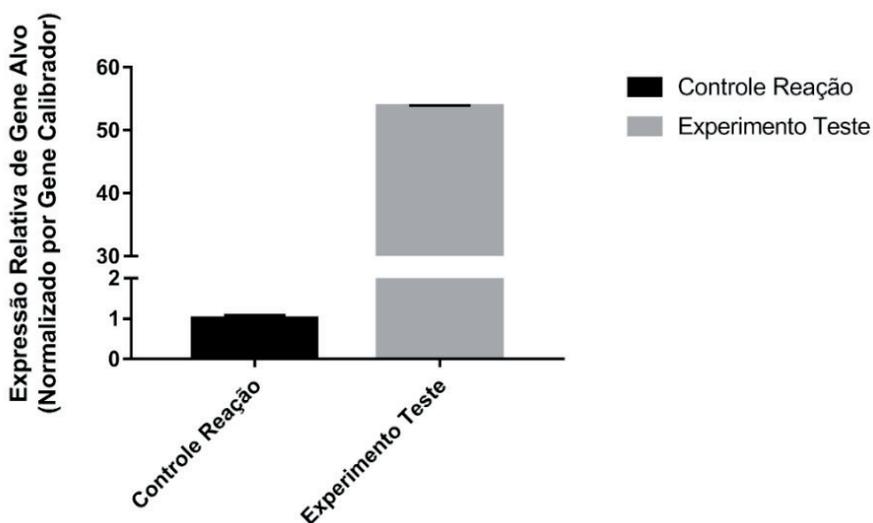


Figura 6. Representação gráfica dos resultados apresentados na Tabela 2.

## 6 | CONCLUSÃO

O entendimento sobre o DNA e RNA bem como as correlações baseadas no controle da expressão gênica permanecem entre os principais alvos de estudo na clínica médica e áreas correlatas. O aprofundamento sobre a genética molecular trouxe avanços no entendimento sobre síndromes, mutações, rastreamento genético, desenvolvimento de doenças, patógenos etc. As metodologias de estudo baseadas na PCR permanecem sendo essenciais como ferramenta de estudo e de diagnóstico. Nesse sentido, o RT-qPCR possibilitou a abordagem quantitativa dos dados, por meio da inserção dos fluorocromos, e tornou os experimentos e testes mais reprodutíveis, com acréscimos metodológicos na sensibilidade, precisão e acurácia dos testes.

## REFERÊNCIAS

APPLIED BIOSYSTEMS. **Guide To Performing Relative Quantitation of Genes Expression**. Using Real-Time Quantitative PCR, Foster City, California, 2004.

BUSTIN, S.A. **Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays**. J Mol Endocrinol., v.25, n.2, p.169-193, 2000.

DE PAULA, S.O.; LIMA, C.M.; TORRES, M.P.; et al. **One-Step RT-PCR protocols improve the rate of dengue diagnosis compared to Two-Step RT-PCR approaches**. J Clin Virol., v.30, n.4, p. 297-301, 2004

DING, C.; CANTOR, C.R. **Quantitative analysis of nucleic acids - the last few year sof progress**. J Biochem Mol Biol., v.37, n.1, p.1-10, 2004.

GIBSON, U.E.; HEID, C.A.; WILLIAMS, P.M. **A novel method for real time quantitative RT-PCR**. Genome Res., v.6, n.10, p. 995-1001, 1996.

- GREEN, M.R.; SAMBROOK, J. **Synthesis of cDNA Probes from mRNA Using Random Oligonucleotide Primers.** Cold Spring Harb Protoc., v.2019, n.7, 2019.
- HEID, C. A.; STEVENS, J.; LIVAK, K.J.; et al. **Real time quantitative PCR.** Genome Res., v.6, n.10, p.986-993, 1996.
- HOLLAND, P.M.; ABRAMSON, R.D.; WATSON, R.; et al. **Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of *Thermus aquaticus* DNA polymerase.** Proc Natl Acad Sci USA, v.88, n.16, p. 7276-7280, 1991.
- KUBISTA, M.; ANDRADE, J.M.; BENGTSSON, M.; et al. **The real-time polymerase chain reaction.** Mol Aspects Med. v.27, n.2-3, p.95-125, 2006.
- LI, K.; BROWNLEY, A. **Primer design for RT-PCR.** Methods Mol Biol., v. 630, p. 271-299, 2010.
- LIVAK, K.J.; SCHMITTGEN, T.D. **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)).** Method. Methods., v.25, n.4, p. 402-408, 2001.
- MORRISON, T.B.; WEIS, J.J.; WITWER, C.T. **Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification.** BioTechniques., v.24, n.6, p.954-962, 1998.
- NAKAMURA, T.; SCORILAS, A.; STEPHAN, C.; et al. **Quantitative analysis of macrophage inhibitory cytokine-1 (MIC-1) gene expression in human prostatic tissues.** Br J Cancer., v.88, n.7, 2003.
- RASHED-UL ISLAM, S.M.; JAHAN, M.; TABASSUM, S. **Evaluation of a Rapid One-step Real-time PCR Method as a High-through put Screening for Quantification of Hepatitis B Virus DNA in a Resource-limited Setting.** Euroasian J Hepatogastroenterol., v.5, n.1, p.11-15, 2015.
- ROCHA, D.J.; SANTOS, C.S.; PACHECO, L.G. **Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis.** Meta-Analysis. Antonie Van Leeuwenhoek., v.108, n.3, p.685-693, 2015.
- SANDERS, R.; BUSTIN, S.; HUGGETT, J.; MASON, D. **Improving the standardization of mRNA measurement by RT-qPCR.** Biomol Detect Quantif., v.15, p. 13-17, 2018.
- STANGEGAARD, M.; DUFVA, I.H.; DUFVA, M. **Reverse transcription using random pentadecamer primers in creases yield and quality of resulting cDNA.** Biotechniques, v.40, n.5, p.649-657, 2006.
- TAYLOR, S.; WAKEM, M.; DIJKMAN, G.; et al. **A practical approach to RT-qPCR - Publishing data that conform to the MIQE guidelines.** Methods., v.50, n.4, p.S1-S5, 2010.
- TAYLOR, S.C.; MRKUSICH, E.M. **The state of RT-quantitative PCR: first hand observations of implementation of minimum information for the publication of quantitative real time PCR experiments (MIQE).** J Mol Microbiol Biotechnol., v.24, n.1, p.46-52, 2014.
- WONG, M.L.; MEDRANO, J.F. **Real-time PCR for mRNA quantitation.** Biotechniques, v.39, n.1, p.75-85, 2005.

# CAPÍTULO 6

## FERRAMENTAS DE EDIÇÃO GÊNICA

Data de aceite: 10/05/2021

### Joaquim Gasparini dos Santos

Bacharel em Ciências Biológicas, Mestre e Doutor em Biotecnologia pela Universidade Federal do Espírito Santo/RENORBIO  
<http://lattes.cnpq.br/2629284844712364>  
<https://orcid.org/0000-0001-6096-116X>  
joaquimgasparini@gmail.com, Universidade Federal do Espírito Santo

### Aline Ribeiro Borçoi

Farmacêutica pela UFJF, Mestre em Farmacologia pela UNIFESP, Doutora e Pós-Doutorado em Biotecnologia pela Universidade Federal do Espírito Santo/RENORBIO.  
<http://lattes.cnpq.br/0558531160628107>  
<https://orcid.org/0000-0003-4594-9888>  
alineborcoi@gmail.com, Universidade Federal do Espírito Santo

**RESUMO:** A surpreendente capacidade do DNA em armazenar e transmitir informações necessárias para o desenvolvimento e manutenção da vida possibilitou a criação de um novo campo de pesquisa denominado de engenharia genética. Este segmento é totalmente voltado para o desenvolvimento de tecnologias capazes de manipular e modificar de maneira precisa e pontual as informações contidas no DNA de diversos tipos de organismos. A partir da descoberta da atividade das enzimas de restrição e de outras ferramentas como ZFN, TALEN e CRISPR, abriu-se um mundo de possibilidades para se realizar uma edição genômica cada vez mais precisa.

**PALAVRAS-CHAVE:** CRISPR-Cas9; ZFN; TALEN; Edição gênica.

### GENOME EDITING TOOLS

**ABSTRACT:** The surprising ability of DNA to store and transmit information necessary for the development and maintenance of life, made it possible to create a new field of research called genetic engineering. This segment is totally focused on the development of technologies capable of manipulating and modifying in a precise and punctual way the information contained in the DNA of different types of organisms. From the discovery of the activity of restriction enzymes and other tools such as ZFN, TALEN and CRISPR, opening up a world of possibilities to carry out an increasingly accurate genomic edition.

**KEYWORDS:** CRISPR-Cas9; ZFN, TALEN, Genic edition

## 1 | INTRODUÇÃO

Em meados de 1970, com a descoberta das enzimas de restrição – as quais possuem a capacidade de clivar regiões específicas do DNA de invasores em bactérias - e de sua junção com outras técnicas, forneceu aos cientistas a capacidade de manipulação e de transferência de fragmentos de DNA para diversos organismos, dando se início a tecnologia do DNA recombinante (GAJ; GERSBACH; BARBAS, 2013a).

Apartir desse primeiro passo, começaram a emergir outras técnicas de manipulação que possibilitaram aos cientistas uma maior precisão, dentre as quais podemos citar as nucleases de dedo de zinco (do inglês Zinc Finger Nucleases – ZFN), nucleases com efetores do tipo ativador transcricional (do inglês transcription activator-like effector nucleases – TALENs) e repetições palindrômicas curtas agrupadas e regularmente

interespaçadas (do inglês Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR) conjugadas com uma enzima Cas9, permitindo aos cientistas: realizarem com maior precisão alterações em regiões responsáveis pelo desenvolvimento de diversos tipos de doenças; desenvolvimento de fármacos mais precisos e eficazes; criação de cultivares mais eficientes na produção de alimentos; desenvolvimento de novos biocombustíveis e outras diversas aplicações (GUPTA; MUSUNURU, 2014).

A base de aplicação das técnicas de edição gênica tem como princípio gerar uma quebra na dupla da fita de DNA em uma região de interesse. Em células de eucariotos essa quebra pode tomar dois caminhos para ser reparadas que são: reparo por junções terminais não homólogas (do inglês non-homologous end joining - NHEJ) ou reparo por recombinação homóloga (do inglês Homologous Recombination - HR). O tipo de reparo que será realizado pela célula após a quebra da fita dupla de DNA vai depender muito da fase do ciclo celular que a célula está. O reparo por NHEJ tem seu ponto de ocorrência durante a fase G1 do ciclo celular, já o reparo por HR tem seu ponto de ocorrência no final da fase S até o final da fase G2 do ciclo celular, pois nesse ponto temos uma maior disponibilidade de fita homóloga para que seja realizado esse reparo (GUPTA; MUSUNURU, 2014; NEMUDRYI et al., 2014).

O reparo por NHEJ (Figura 1) ocorre de forma aleatória e tem uma eficiente junção da quebra realizada. Entretanto, esse reparo muitas vezes pode gerar perda de informação naquela região devido a frequentes ocorrências de inserção ou deleções (INDEL) de nucleotídeos. Desta maneira, pode causar a inativação por completo de um gene devido a formação de um mRNA que será degradado ou de uma proteína truncada causada por uma tradução de um mRNA sem sentido. Esse tipo de reparo muito vezes é desejado para gerar um gene *knockout* ou o silenciamento da região específica do genoma que se pretende estudar (JOUNG; SANDER, 2013).

O reparo por HR (Figura 1) é um tipo mais complexo para se reparar a quebra na fita dupla de DNA, pois ele só será realizado mediante a presença de uma fita homóloga. Para induzir esse tipo de reparo é inserido na célula, em conjunto com a técnica selecionada para edição gênica, um DNA doador ao qual contém a região de interesse que o pesquisador queira inserir no genoma do receptor. Esse tipo de abordagem muitas vezes é utilizado para se corrigir uma região com erros na sequência do DNA causadoras de doenças, corrigindo desta forma a função normal do gene. Também pode ser utilizado para inserir um gene de interesse para que a célula exerça uma nova função pretendida, como no desenvolvimento de novos cultivares, onde características de interesse podem ser inseridas no genoma da planta (MA et al., 2017).

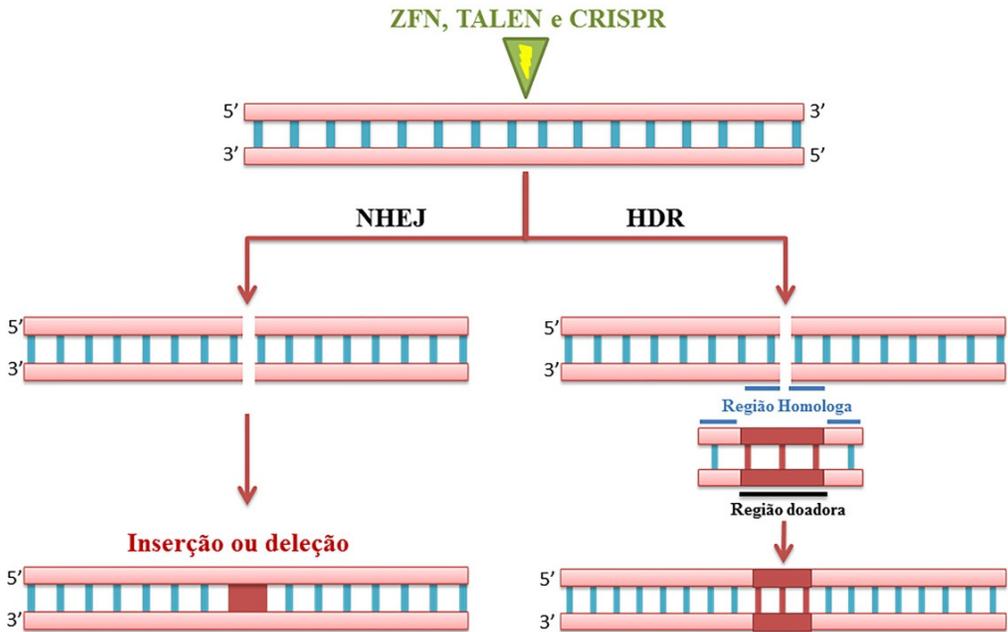


Figura 1. Duas possíveis vias de reparo de uma clivagem genômica da dupla fita de DNA, realizado por meio do uso de técnicas de edição. Alternativamente, a ruptura da fita dupla de DNA pode ser reparada por junções terminais não homogêneas (NHEJ), ao qual leva a mutações no local da clivagem através de inserção ou deleção de pares de bases presente na fita dupla do DNA. Quando fornecido a célula um DNA doador homólogo, o reparo poderá prosseguir por recombinação homóloga (HDR), onde será utilizado o doador como modelo.

## 2 | NUCLEASE DE DEDO DE ZINCO (ZFN)

Nuclease de dedo de zinco (ZFN) está presente em 3% do genoma humano e foi descoberto em 1985 por Miller, enquanto realizava estudos com o Fator de transcrição IIIA (TFIIIA) em oocistos imaturos de *Xenopus*. Os ZFNs são proteínas de fatores de transcrição formadas por um arranjo de motivos em tandem, onde cada motivo do dedo de zinco contém aproximadamente 30 aminoácidos arranjados em uma estrutura beta-beta-alfa. Cada folha  $\beta$  anti-paralelas contém dois resíduos de Cisteínas e duas Histidinas na subunidade alfa-hélice, ligadas e estabilizadas por um zinco (Cys2 - Zn - His2) (Figura 2). A região alfa-hélice de cada dedo de zinco possui sequências específicas de aminoácidos capazes de reconhecer e se ligar com fidelidade a três pares de bases do DNA (URNOV et al., 2010; CARROLL, 2011).

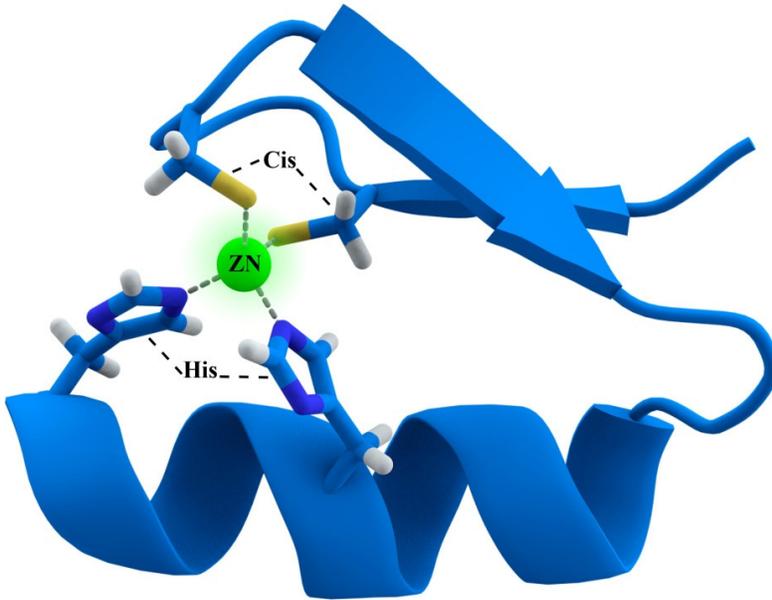


Figure 2. Representação esquemática da proteína de dedo de zinco (ZFP) (Modificado de Thomas Splettstoesse. Wikipedia ).

A primeira aplicação prática utilizando dedo zinco foi realizada em 1994, a partir de um estudo realizado *in vitro* por Cho e colaboradores, capaz de reconhecer e inibir a transcrição do oncogene *bcr-abl*. Esse oncogene é denominado de cromossomo Filadélfia, formado a partir da translocação entre o cromossomo 9 e 22, capaz de levar a uma divisão desordenada das células, causando a leucemia mielóide crônica. Esse importante feito abriu caminho para a criação das primeiras empresas de biotecnologia - a Sangamo BioSciences na Califórnia e Gendaq Ltda no Reino Unido com intuito de explorar e desenvolver nova tecnologia para o uso terapêutico (JAMIESON; MILLER; PABO, 2003; CHANDRASEGARAN, 2017).

Em 1996 Chandrasegaran junto com sua equipe conseguiram desenvolver à primeira endonuclease de restrição quimérica com capacidade de gerar uma quebra dupla da fita de DNA (do inglês Double-Stranded Break – DSB) em um local específico *in vitro*, a partir da fusão da proteína de dedo de zinco com um domínio de clivagem não específico *FOK I* (domínio catalítico presente na *Flavobacterium okeanoikoites*) (Figura 3). O desenvolvimento da ZFN foi uma divisora água para os futuros trabalhos de engenharia genética, pois facilitou a criação de modelos *knockout* para estudos de doença; possibilitou a inserção de gene doadores ao genoma do receptor, como a inserção de uma proteína marcadora que emite fluorescência verde, reparo de regiões responsáveis pelo desenvolvimento e doenças ou até mesmo regular a expressão de genes específicos; além de possibilitar a produção de novas cultivares pela eliminação ou adição de novos genes para aumento da produção (CHANDRASEGARAN, 1996; KLUG, 2010).

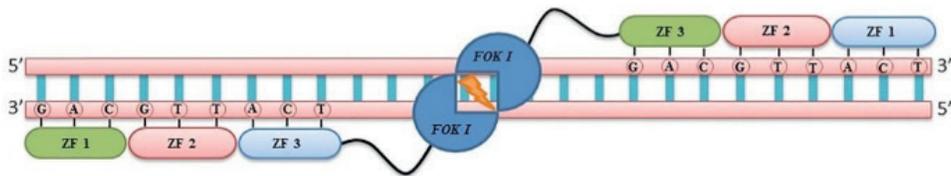


Figura 3. Cada nuclease de dedo de zinco (ZFN) consiste de domínios protéicos que reconhecem sequências específicas de três nucleotídeos do DNA. Na ponta de cada dedo de zinco encontra-se um domínio de clivagem composta por uma nuclease FOK I que quando unidas causa o corte abrupto na fita dupla de DNA.

Um dos mais importantes estudos conduzindo ZFN com intuito terapêutico aplicável em humanos foi a deleção do gene CCR5 em linfócitos T de portadores do vírus HIV. Sabe-se que o gene CCR5 é responsável por gerar um co-receptor que facilita a infecção do vírus HIV em linfócitos T. A partir desses dados, os cientistas realizaram os primeiros estudos com linfócitos CD4+ T *in vitro*, ao qual inativaram o gene CCR5 através do desenvolvimento de um ZFN específico e, após a modificação, inseriram as células em camundongos infectados com vírus do HIV. Foi observado que os camundongos que receberam as células modificadas possuíam uma baixa carga viral quando comparado com os selvagens. Esse estudo foi de início a um teste clínico em humanos e para o desenvolvimento de uma nova terapia para o vírus causador da Síndrome da Imunodeficiência Adquirida (AIDS) (PEREZ et al., 2008).

O processo de edição gênica realizada por ZFN são realizados através da introdução de duas nucleases, cada uma compostas de 3 a 9 dedos de zinco conjugadas com um domínio de clivagem *FOK I*, cada uma ZFN com sua capacidade de reconhecimento de 9 a 18 pares de base presentes no DNA (PEREZ et al., 2012).

Outros tipos de proteínas de dedo de zinco com diferentes domínios funcionais foram desenvolvidas para abordagens mais específicas. Essas ZFN são conjugadas com domínios que não tem atividade de gerar quebra na fita dupla de DNA, mas sim com fatores que causam a repressão de promotores gênicos, realizam a modificação da cromatina, realizam o controle da expressão gênica por meio de ligantes específicos ou ativam constantemente a expressão do gene de interesse ou podem realizar a metilação de regiões específicas do DNA (URNOV et al., 2010; PEREZ et al., 2012).

Um ponto importante em relação ao uso da técnica de edição gênica por ZFN, são as vantagens e a limitação para seu uso. Como vantagens pode se destacar que é uma técnica que não tem risco de ser incorporada no genoma, devido a utilizar uma proteína durante a fase de edição. Outro ponto é o baixo ponto de off-targets, ou seja, pontos de ação da nuclease que estão fora do alvo gênico pretendido, já que possui uma alta especificidade de ligação a região ao qual foi desenhada. Além disso essa técnica de edição genica pode se trabalhada em qualquer tipo de célula. Como pontos negativos pode se ressaltar principalmente a dificuldade de desenvolvimento, pois o processo de construção demanda tempo uma complexa engenharia na hora de desenhar. Além disso você só pode atingir um alvo de interesse de cada vez, não podendo realizar a modificação em vários alvos ao mesmo tempo, por fim ela é uma técnica que ainda possui um baixo

grau de eficiência (GAJ; GERSBACH; BARBAS, 2013a).

### 3 I NUCLEASES COM EFETORES DO TIPO ATIVADOR TRANSCRICIONAL (TALEN)

A técnica de edição gênica baseada em TALEN surgiu como uma opção alternativa ao uso de ZFN, apesar das duas técnicas compartilharem similaridades, como ser conjugada com uma nuclease não específica a *FOK I*. Os efetores do tipo ativador transcricional (TALEs), são proteínas secretadas pela *Xanthomona spp.*, uma bactéria fitopatogênica causadora da ferrugem bacteriana, que invade células de plantas, alterando a expressão gênica da mesma. Os TALEs são compostos por uma região C-terminal onde encontramos uma porção ao qual é um sinal de localização nuclear (NLS). Em conjunto com essa porção encontra-se um domínio de ativação (AD), ao qual tem como função ativar constitutivamente a transcrição da célula hospedeira. A parte mais complexa de TALE possui uma porção modular composta por 10 a 30 repetições, cada uma com aproximadamente 33 a 35 repetições de aminoácidos, denominado de repetições variáveis diresíduos (RVS). Na posição 12-13 das RVS encontram-se os aminoácidos variáveis responsáveis por realizar o reconhecimento e a ligação aos nucleotídeos do DNA do hospedeiro, ou seja, sequências de aminoácidos presentes nessa região como Asn e Ile (NI) se ligam a base adenina (A), Asn e Gly (NG) se ligam a base guanina (G), duas Asn (NN) irá se ligar a guanina (G) ou adenina (A) e His and Asp (HD) irão se ligar a citosina (C) (Figura 4) (JOUNG; SANDER, 2013; NEMUDRYI et al., 2014)

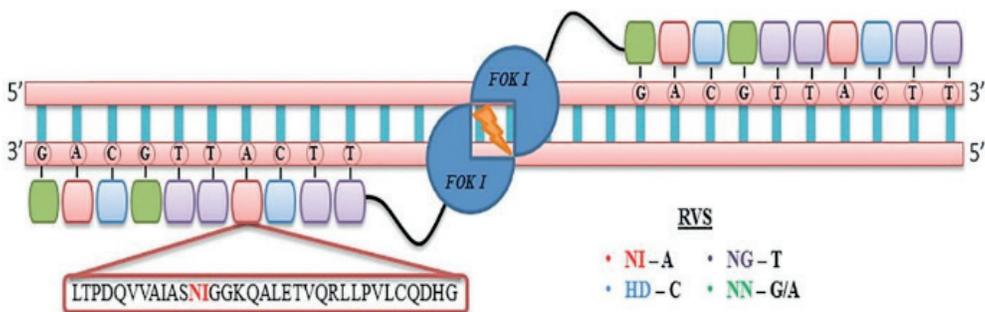


Figura 4. Estrutura esquemática de TALEN com seus domínios de repetição e o domínio de clivagem composto pela nuclease FOK I. Cada repetição TALE é constituída por 34 aminoácidos, onde os aminoácidos presentes nas posições 12 e 13, denominados de repetições variáveis diresíduos (RVDs), determinam a especificidade de ligação a bases presentes no DNA.

A técnica de TALEN vem sendo amplamente utilizada nas áreas de desenvolvimento de terapias humanas e agrária. No desenvolvimento de terapias humanas foi desenvolvido com TALEN a inativação do gene CCR5, o mesmo objetivo citado anterior realizado com ZFN. Outro interessante projeto é o desenvolvimento de células de defesa UCARTs (Universal Chimeric Antigen Receptor T-cells products), que são células T de defesa provenientes de um doador saudável modificadas *in vitro* e inserida no paciente refratários aos medicamentos disponíveis contra o câncer. A modificação realizada nessas células

provocam capacidade de reconhecimento e eliminação das células do câncer (KUSANO et al., 2016).

Na área agrária podemos citar o desenvolvimento de uma cultivar de arroz resistente a ataque de *Xanthomona* spp., onde TALEN foi utilizado para inativar a região gênica de ligação de TALE, inibindo dessa forma as perdas de produção causada pela ferrugem bacteriana (NEMUDRYI et al., 2014).

Outro projeto interessante utilizando a tecnologia de TALEN foi desenvolvido pela empresa Recombinectis (Minnesota – EUA), em que dois machos mestiços de raça leiteiros mochos não desenvolviam chifres ao longo do seu crescimento, evitando dessa forma um processo doloroso para o gado de descorna. O processo de edição gênica se deu a partir de uma célula doadora de um macho de vaca leiteira com presença de chifres e um alelo doador céltico da raça Angus, ao qual não apresenta chifre. Dessa forma, utilizou-se a tecnologia de TALEN em que a região do genoma responsável por produzir os chifres foi cortada e inserida no lugar o alelo céltico dominante. Por fim, foram selecionadas as células que sofreram a modificação desejada e clonadas a partir da transferência de núcleo, para assim, criar o embrião que seria transplantado na vaca aceptora, gerando com sucesso bezerros sem chifres. Entretanto, o projeto não teve sucesso. A agência regulatória de alimentos e medicamentos (FDA) dos Estados Unidos da América considerou o animal transgênico e teve sua comercialização proibida em 2019, já que parte do plasmídeo utilizado para entregar o alelo céltico foi integrada ao genoma dos bezerros (ZHANG et al., 2013; CHAUDHARY; PRATAP; SHARMA, 2016; LI et al., 2020).

O processo de edição gênica realizada por TALEN é bem parecido com o de ZFN, devido introdução de duas quimeras nucleases proteicas na célula alvo. TALEN pode ser desenvolvida para reconhecer de 12 a 20 pares de bases. Além disso pode ser substituído a nuclease *FOK I* por outros domínios funcionais como as citadas em ZFN (JOUNG; SANDER, 2013).

As vantagens de utilizar a técnica de edição gênica por TALEN está na forma mais fácil de desenhar e projetar, além de apresentar um baixo efeito de off-target e citotoxicidade quando comparado com a técnica de ZFN. Além disso, TALEN possui uma taxa de 96% de afinidade com a região ao qual foi projetada para se ligar. As desvantagens estão na baixa eficiência em regiões de alta metilação do DNA, um custo elevado para se projetar e requer um processo metodológico de clonagem complexo (CHAUDHARY; PRATAP; SHARMA, 2016).

## 4 | CRISPR-CAS9

Observado pela primeira vez no genoma de *E. coli* por Ishino em 1987, batizada em 2002 por Mujica com o nome de repetições palindrômicas curtas agrupadas e regularmente interespçadas (do inglês Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR), só em 2007 foi descoberta a sua função como um sistema de defesa de bactérias contra vírus invasores por Philippe Horvath e colaboradores. A infecção viral em bactéria muitas vezes leva a morte do hospedeiro, mas algumas bactérias que conseguem sobreviver a esse ataque integram ao seu DNA denominada de “CRISPR locus” o material

genético do invasor. Assim, quando essa bactéria é novamente infectada pelo mesmo vírus, é ativada a transcrição da região genômica responsável pela formação do complexo entre o RNA guia que irá reconhecer o DNA invasor e endonuclease Cas9 ao qual tem a capacidade de clivar o DNA invasor, evitando que ele infecte a bactéria e cause à morte ou danos à mesma (BARRANGOU; HORVATH, 2017).

No ano de 2012 as pesquisadoras Emmanuelle Charpentier e Jennifer Doudna, conseguem criar um RNA guia sintético capaz de guiar a Cas9 para regiões desejadas, conseguindo aperfeiçoar e simplificar o método de produção de CRISPR/Cas9. Nesse mesmo ano foi registrado o pedido de patente nos EUA para CRISPR/Cas9 como nova forma de edição gênica. Quase ao mesmo tempo o pesquisador Feng Zhang (Instituto Broad) e George Church (Universidade de Harvard), também registraram pedido de patente, onde o processo de autoria e patente se arrasta até hoje na justiça (JINEK et al., 2012; PICKAR-OLIVER; GERSBACH, 2019).

Um fato interessante de se observar em relação à produção científica sobre CRISPR/Cas9, é que de 2002 - ano em que Mujica deu nome à essa região - até a o ano de 2013 - após início da briga pela patente – apenas 29 estudos foram publicados. De 2013 a 2019 o número de publicações saltou para 3.221 publicações, com a China liderando as pesquisas com relação a utilização dessa técnica na edição gênica (PICKAR-OLIVER; GERSBACH, 2019).

Estudos voltados para edição gênica de mamíferos tiveram adaptações e testes em células humanas foram realizadas por Feng Zhang em 2013. A partir desse estudo, Jiahao Sha em conjunto com outros pesquisadores na universidade da China, conseguiram em 2014 gerar macacos com modificações de genes gerados por CRISPR/Cas9. Um grande avanço na área médica foi possível e revelado em 2015, quando Church conseguiu eliminar do genoma de suínos 62 retrovirus (PERVS) com potencial patogênico a humanos. Esse estudo foi recebido com bastante impacto pelo meio científico, pois abriu as portas para aprimorar o xenotransplante de órgão em humanos, com grande possibilidade de diminuir a espera por um órgão na fila de transplante.

No ano de 2016, You Lu da universidade Schian na China tiveram sucesso a desativação do gene PD-1 de células T, estimulando as mesmas a combaterem células de câncer de pulmão de pacientes refratários aos tratamentos disponíveis. Em 2017, Ha Youn Shin, em conjunto com outros pesquisadores do Instituto Nacional de Saúde dos Estados Unidos, conseguiram identificar que a técnica de CRISPR/Cas9 pode causar efeitos off-targets fora do alvo de estudo. No mesmo ano um feito inédito realizado por Shoukhrat Mitalipov em conjunto com sua equipe na Universidade de Saúde e Ciência do Oregon, conseguiram reparar com sucesso uma mutação do gene *MYBPC3* causador da cardiomiopatia hipertrófica em embriões inviáveis, abrindo as portas para novos estudos com doenças geneticamente herdáveis (MA et al., 2017; SCHAEFER et al., 2017).

Em resumo, muitos avanços na pesquisa utilizaram como técnica CRISPR/Cas9. Entretanto, pouco se discutia sobre a ética da utilização das técnicas de edição gênica em embriões viáveis e até que ponto a ciência pode interferir no processo natural da vida ou na geração de indivíduos dotados de capacidades ou habilidades por meio de modificações genéticas.

Em novembro de 2018 o pesquisador He Jiankui da Universidade de Ciência e tecnologia do Sul da China veio a público anunciar que tinha gerado gêmeas geneticamente modificadas por CRISP/Cas9. A modificação realizada por ele foi o silenciamento do gene CCR5 nos embriões, ao qual gera um receptor responsável por facilitar a infecção pelo vírus HIV. He Jiankui foi julgado e condenado a três anos de prisão por exercício ilegal da medicina na China, a pagar multa de US\$430,000, além de ter sido banido da comunidade científica (LU et al., ; THURTLE-SCHMIDT; LO, 2018; PICKAR-OLIVER; GERSBACH, 2019).

A técnica de CRISPR/Cas9 é vista como revolucionária no cenário da edição gênica devido ao seu baixo custo, sua versatilidade e facilidade de desenvolvimento quando comparada com ZFN e TALEN. CRISPR/Ca9 é composta por um RNA guia de aproximadamente 23 nucleotídeos, dentre eles três nucleotídeos iniciais que se denomina motivo PAM (Protospacer Adaptor Motif), responsáveis por reconhecer a região de interesse do DNA e permitir o anelamento dos outros 20 nucleotídeos presentes no RNA guia. Em conjunto com RNA guia há uma Cas9 com duas subunidades catalíticas (RuvC1 e HNH), responsável por realizar o corte na fita dupla do DNA (Figura 5).

Após a descoberta de CRISPR/Cas9, diversas outras variantes foram desenvolvidas para finalidades específicas de edição gênica, repressão ou ativação de regiões do genoma, em que cada modelo vai depender do objetivo do seu estudo (LU et al., ; GAJ; GERSBACH; BARBAS, 2013b; BARRANGOU; HORVATH, 2017).

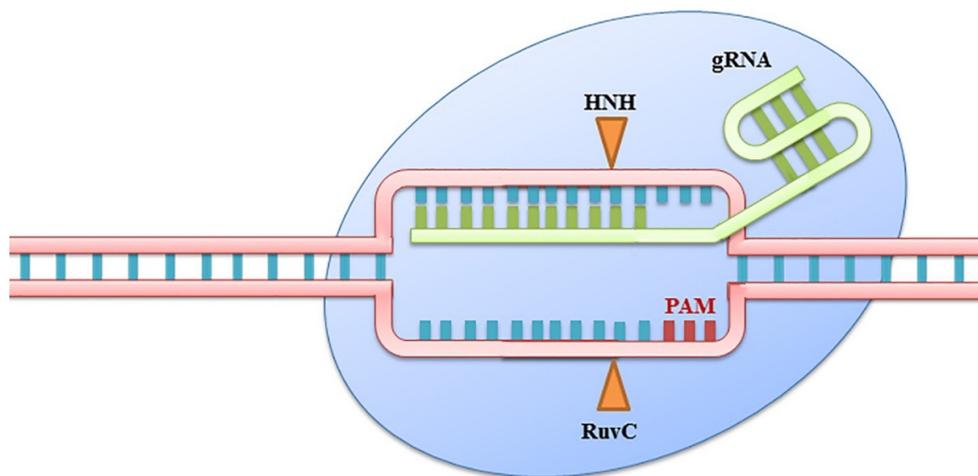


Figura 5. Estrutura esquemática de CRISPR-Cas9 ao qual é orientado por uma fita guia de RNA (gRNA) e um protospacer de motivo adjacente (PAM), responsável por direcionar sua especificidade, devido sua complementariedade e emparelhamento complementar com uma sequência alvo presente no DNA. Ao realizar o emparelhamento dois domínios catalíticos, RuvC e HNH, realizam a quebra da fita dupla de DNA.

As vantagens de usar CRISPR/Cas9 ou qualquer outra variante estão no seu baixo custo para desenvolvimento, alta eficiência para realização de edições gênicas,

possibilidade de se realizar diversas modificações gênicas ao mesmo tempo, o que não é possível utilizando ZFN e TALEN, além de ter sido testado com sucesso em diversos tipos de células. Como desvantagens estão na limitação na disponibilidade para região de reconhecimento de PAM (ADLI, 2018).

## 5 | CONSIDERAÇÕES FINAIS

O emprego de diferentes técnicas de edição gênica vem gerando um avanço no meio científico, seja na possibilidade de desenvolvimento de novas terapias para doenças que afetam os seres humanos, seja no desenvolvimento de bases mais eficientes na agricultura, com novas cultivares mais eficientes na produção de alimentos, resistências a pragas e na diminuição de uso de defensivos agrícolas, além de possibilitar a criação de animais com características desejadas, não precisando passar pelo demorado processo de seleção clássica das características desejadas através de diversos cruzamento.

Além disso, levanta uma discussão ética de até onde a pesquisa pode prosseguir, sem ser utilizada para caráter extremamente exclusiva de interesse de apenas uma minoria, como a realização de melhoramento genético humano, seleção de características vistas como padrões aceitáveis, ou até mesmo para criação de verdadeiras máquinas de guerra ou novas armas químicas. Toda nova tecnologia sempre será bem-vinda, quando for para alimentar o bem comum de todos e sempre respeitando a liberdade de escolha de cada indivíduo.

## REFERÊNCIAS

ADLI, M. **The CRISPR tool kit for genome editing and beyond.** Nature Communications, v. 9, n. 1, 2018.

BARRANGOU, R.; HORVATH, P. **A decade of discovery : CRISPR functions and applications.** Nature Publishing Group, v. 2, n. June, p. 1–9, 2017.

CARROLL, D. **Genome Engineering With Zinc-Finger Nucleases.** Genetics, v. 188, n. August, p. 773–782, 2011.

CHANDRASEGARAN, S. **Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain.** Proceedings of the National Academy of Sciences of the United States of America, v. 93, n. February, p. 1156–1160, 1996.

CHANDRASEGARAN, S. **Recent advances in the use of ZFN-mediated gene editing for human gene therapy.** Cell and Gene Therapy Insights, v. 3, n. 1, p. 33–41, 2017.

CHAUDHARY, K.; PRATAP, D.; SHARMA, P. K. **Transcription activator-like effector nucleases (TALENs ): An efficient tool for plant genome editing.** Engineering in Life Sciences, p. 330–337, 2016.

GAJ, T.; GERSBACH, C. A.; BARBAS, C. F. **ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.** Trends in Biotechnology, v. 31, n. 7, p. 397–405, 2013a.

GAJ, T.; GERSBACH, C. A.; BARBAS, C. F. **ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.** Trends in Biotechnology, v. 31, n. 7, p. 397–405, jul. 2013b..

GUPTA, R. M.; MUSUNURU, K. **Expanding the genetic editing tool kit : ZFNs , TALENs , and CRISPR-Cas9.** Journal of Clinical Investigation, v. 124, n. 10, 2014.

JAMIESON, A. C.; MILLER, J. C.; PABO, C. O. **Drug discovery with engineered zinc-finger proteins.** Nature Reviews Drug Discovery, v. 2, p. 14–20, 2003.

JINEK, M. *et al.* **A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity.** Science, v. 337, n. 6096, p. 816–822, 2012.

JOUNG, J. K.; SANDER, J. D. **TALENs: a widely applicable technology for targeted genome editing.** Nature Reviews Molecular Cell Biology, v. 14, n. 1, p. 49–55, 2013..

KLUG, A. **The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation.** Quarterly Reviews of Biophysics, v. 43, n. 1, p. 1-21, 2010.

KUSANO, H. *et al.* **A simple Gateway-assisted construction system of TALEN genes for plant genome editing.** Nature Publishing Group, n. July, p. 1–7, 2016.

LI, H. *et al.* **Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects.** Signal Transduction and Targeted Therapy, v. 5, 2020.

LU, Y. *et al.* **Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer.** Nature Medicine, v. 26, p. 732-740, 2020.

MA, H. *et al.* **Article Correction of a pathogenic gene mutation in human embryos.** Nature, v. 548, p. 413-419, 2017.

NEMUDRYI, A. A. *et al.* **TALEN and CRISPR / Cas Genome Editing Systems: Tools of Discovery.** Acta Naturae, v. 3, n. 22, p. 19–40, 2014.

PEREZ, E. E. *et al.* **Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases.** Nature Biotechnology, v. 26, n. 7, p. 808–816, 29 jul. 2008.

PEREZ, E. E. *et al.* **Editing Using Zinc-Finger Nucleases.** Nat Biotechnol, v. 26, n. 7, p. 808–816, 2012.

PICKAR-OLIVER, A.; GERSBACH, C. A. **The next generation of CRISPR–Cas technologies and applications.** Nature Reviews Molecular Cell Biology, v. 20, n. 8, p. 490–507, 2019.

SCHAEFER, K. A. *et al.* **Unexpected mutations after CRISPR-Cas9 editing *in vivo*.** Nature Methods, v. 14, p. 547-548, 2017.

THURTLIE-SCHMIDT, D. M.; LO, T. W. **Molecular biology at the cutting edge: A review on CRISPR/ CAS9 gene editing for undergraduates.** Biochemistry and Molecular Biology Education, v. 46, n. 2, p. 195–205, 2018.

URNOV, F. D. *et al.* **Genome editing with engineered zinc finger nucleases.** Nature, v. 11, n. 9, p. 636–646, 2010.

Data de aceite: 10/05/2021

Data de submissão: 13/04/2021

### **Karolinni Bianchi Britto**

Universidade Federal do Espírito Santo  
Vitória - ES  
<http://lattes.cnpq.br/3604508015056223>

### **Greiciane Gaburro Paneto**

Universidade Federal do Espírito Santo  
Alegre - ES  
<http://lattes.cnpq.br/8176374147579841>

**RESUMO:** Proteínas são macromoléculas importantes para o funcionamento dos organismos pois desempenham inúmeros papéis centrais, além de serem amplamente utilizadas como reagentes na pesquisa, sendo aplicadas em diversas áreas como indústria farmacêutica, agricultura e diagnóstico de doenças. A expressão heteróloga de proteínas, possível através da tecnologia do DNA recombinante, se destaca por permitir que organismos que não expressam a proteína normalmente possam sintetizá-la em escalas necessárias ao uso que se propõem. Para tal, são necessários vetores de expressão com características adequadas e que possam direcionar corretamente a transcrição e tradução do gene de interesse na célula hospedeira selecionada. Diferentes sistemas de expressão heteróloga de proteínas são conhecidos, podendo-se destacar as bactérias, especialmente *Escherichia coli*, leveduras, fungos filamentosos, microalgas, cultura de células de mamíferos, insetos e plantas. Cada sistema possui vantagens e desvantagens, sendo o melhor escolhido de acordo com a característica da proteína que se

deseja expressar. Quanto à análise e identificação, diversas técnicas são aplicadas para verificar a presença do gene e das proteínas nas células. A purificação das proteínas expressas pelos sistemas ocorre especialmente por cromatografia, podendo ser do tipo cromatografia de troca iônica, de afinidade, de gel filtração, de fase reversa, de imunoafinidade e cromatografia líquida de alta resolução. Sequências específicas de aminoácidos, conhecidas como 'tags', geralmente são adicionadas às proteínas para facilitar principalmente o processo de purificação, apesar de influenciar também em outras etapas do processo. Após a obtenção da proteína purificada segue-se com o processamento de acordo com o objetivo proposto com sua síntese.

**PALAVRAS-CHAVE:** Expressão heteróloga, proteína, DNA recombinante, vetores, purificação de proteína.

### **HETEROLOGOUS PROTEIN EXPRESSION**

**ABSTRACT:** Proteins are important macromolecules for the functioning of organisms because they play countless central roles, besides being widely used as reagents in research, being applied in several areas such as pharmaceutical industry, agriculture and disease diagnosis. The heterologous expression of proteins, possible through recombinant DNA technology, stands out by allowing organisms that do not normally express the protein to synthesize it in scales necessary for its proposed use. For this, expression vectors with adequate characteristics and that can correctly direct the transcription and translation of the gene of interest in the selected host cell are needed. Different systems of heterologous expression of proteins are known, especially bacteria *Escherichia coli*, yeast, filamentous fungi, microalgae, mammalian cell culture, insects and

plants. Each system has advantages and disadvantages, and the best one is chosen according to the characteristic of the protein to be expressed. As for the analysis and identification, several techniques are applied to verify the presence of the gene and proteins in the cells. The purification of the proteins expressed by the systems occurs especially by chromatography, and may be of the ion exchange, affinity, gel filtration, reverse phase, immunoaffinity, and high performance liquid chromatography types. Specific amino acid sequences, known as 'tags', are usually added to proteins to facilitate mainly the purification process, although they also influence other steps in the process. Once the purified protein is obtained, it is processed according to the proposed objective of its synthesis.

**KEYWORDS:** Heterologous expression, protein, recombinant DNA, vectors, protein purification.

## 1 | INTRODUÇÃO

Proteínas são macromoléculas formadas por um ou mais polipeptídeos, os quais são constituídos por uma cadeia de aminoácidos unidos por ligações peptídicas. A sequência exata dos aminoácidos é determinada pelo gene que codifica para esse polipeptídeo específico. Quando sintetizada, uma cadeia polipeptídica se dobra, assumindo uma conformação específica (WALSH, 2007). Assim, cada tipo de proteína possui uma estrutura tridimensional única e essa estrutura confere também uma função única. Portanto, é evidente que a sequência de aminoácidos desempenha um papel fundamental tanto na determinação da estrutura tridimensional da proteína quanto na sua função (NELSON; COX, 2011).

A conformação adotada pela proteína é amplamente estabilizada por múltiplas interações não covalentes fracas. Um processo de desnaturação pode ocorrer quando há alguma influência, como calor e produtos químicos por exemplo, que perturbe essas interações fracas resultando na interrupção da conformação nativa do polipeptídeo, o que geralmente leva a perda de atividade funcional devido a perda da estrutura. A conformação da proteína pode ser determinada por técnicas de difração de raios X e ressonância magnética nuclear (RMN), por exemplo (WALSH, 2007).

Para que as informações genéticas presentes no DNA da célula sejam expressas na forma de uma proteína, enzima ou anticorpo, um intermediário entre a proteína e o DNA é requerido: RNA mensageiro (RNAm), conforme discutido anteriormente.

Relembrando: o fluxo geral de informações, do DNA às proteínas, foi resumido por Francis Crick como o "Dogma Central" da biologia molecular em que o DNA de fita dupla é transcrito para o RNAm (em eucariotos, com processamento da transcrição), que por sua vez é traduzido pelo ribossomo na cadeia de aminoácidos que formam uma proteína. A síntese de proteínas possui três estágios principais: iniciação, alongamento e terminação (GROVES, 2006).

As proteínas possuem papéis importantes em inúmeros processos celulares, sendo sintetizadas como parte do metabolismo de todas as formas de vida. Tanto as proteínas nativas quanto as recombinantes são muito utilizadas como reagentes na pesquisa científica, sejam em processos industriais mediados por enzimas, diagnóstico de doenças, na indústria farmacêutica e em setores da agricultura (EGELKROUT; RAJAN; HOWARD,

2012).

Uma proteína heteróloga é definida como aquela produzida em organismos diferentes dos de sua origem, podendo ser expressas em sistemas procariotos ou eucariotos (CANÇADO, 2002, RAI; PADH, 2001). Assim, é denominada produção heteróloga de proteínas a expressão de proteínas recombinantes em células diferentes de sua ocorrência natural (WALSH, 2007).

A expressão heteróloga, de uso da tecnologia do DNA recombinante, permite a obtenção de produtos biofarmacêuticos como hormônios, fatores de coagulação, citocinas, anticorpos, vacinas, enzimas, imunossupressores, entre outros, em uma elevada quantidade quando comparados aos seus produtores nativos (RAMANA; XAVIER; SHARMA, 2017). Existem diferentes sistemas de expressão de proteínas heterólogas. Dentre os procariotos, estão os que utilizam células de bactérias e são os mais comuns nas indústrias biotecnológicas na produção de proteínas recombinantes; e, dentre os eucariotos, encontram-se os de leveduras, cultura de células de mamíferos e insetos, plantas e animais transgênicos (CANÇADO, 2002; WALSH, 2007), que serão discutidos posteriormente.

## 2 I TECNOLOGIA DO DNA RECOMBINANTE

Por meio de uma tecnologia de manipulação do DNA que surgiu na década de 1970, foi possível o isolamento e caracterização de genes criando uma forma de recombinar segmentos de DNA de diversas fontes em novas moléculas compostas. Essas técnicas de manipulação do DNA são conhecidas como tecnologia de DNA recombinante ou engenharia genética (GROVES, 2006). Normalmente envolve o isolamento, a manipulação e a reintrodução de trechos de DNA nas células, conferindo à célula receptora a capacidade de produzir uma proteína específica. O pedaço de DNA criado artificialmente *in vitro* que contém DNA obtido de duas ou mais fontes é chamado de DNA recombinante (DNAr) (WALSH, 2007).

No caso do desenvolvimento de uma proteína heteróloga, realiza-se inicialmente a identificação e o isolamento do gene responsável por codificar a proteína alvo, a geração de um pedaço apropriado de DNAr contendo a sequência de codificação da proteína e a introdução deste DNAr em uma célula hospedeira apropriada, de modo que a proteína alvo seja produzida em grandes quantidades por essa célula manipulada (WALSH, 2007).

### 2.1 Clonagem

A inserção e multiplicação de moléculas idênticas de DNA em um organismo hospedeiro refere-se à etapa de clonagem na técnica de DNAr (GROVES, 2006). Para clonar um fragmento de DNA de interesse é necessário vinculá-lo a uma molécula de DNA vetorial, que pode se replicar dentro de uma célula hospedeira. Então, após uma única molécula de DNAr, composta por um vetor e um fragmento de DNA inserido, ser introduzida em uma célula hospedeira, o DNA inserido é replicado junto com o vetor, gerando um grande número de moléculas de DNA idênticas (ALBERTS et al., 2007).

As enzimas de restrição (ERs), endonucleases produzidas por bactérias que normalmente reconhecem sequências específicas de 4 a 8 pb nos chamados locais de

restrição, facilitam a produção das moléculas de DNAr. Tais enzimas cortam o DNA de uma fonte específica em um conjunto reproduzível de fragmentos de restrição para serem introduzidos no vetor. Como exemplos podemos citar: BamHI, EcoRI, HindIII, KpnI e PstI. A inserção dos fragmentos de DNA gerados em vetores recebe o auxílio de DNA ligases (ALBERTS et al., 2007).

Graças às informações sobre diversas sequências de genes disponíveis em bancos de dados, é possível obter grandes quantidade do gene de interesse usando a técnica de Reação em Cadeia da Polimerase (PCR). Primeiramente, o DNA genômico da fonte de interesse é extraído e depois realizada a PCR, gerando cópias do gene de interesse. Locais de reconhecimento para ERs podem ser incorporados nos oligonucleotídeos para permitir a clonagem do gene amplificado (WALSH, 2007).

O corte do DNA pode ser também realizado sem o uso de ERs, no caso de cisalhamento mecânico da solução de DNA. Pode-se ainda partir de um extrato de RNA, que após algumas etapas geram o DNA complementar (cDNA) apto a ser clonado da mesma maneira que o DNA genômico (GROVES, 2006).

## 2.2 Vetores

O fragmento de DNA que contém o gene de interesse será transferido para uma célula hospedeira através dos chamados vetores. Uma característica essencial do vetor de clonagem utilizado é que ele deve ser capaz de se autorreplicar na célula em que será introduzido, além de possuir marcador de seleção e locais de restrições únicos para inserção das sequências de interesse (FERRIER, 2017). Esses vetores, que podem ser plasmídeos, bacteriófagos  $\lambda$ , cosmídeos, cromossomos artificiais bacterianos (BACs) ou cromossomos artificiais de leveduras (YACs) por exemplo, devem possuir pelo menos um local de clonagem e uma sequência que pode ser clivada por uma ER para permitir sua ligação a um fragmento de DNA clivado de maneira semelhante (GROVES, 2006; WALSH, 2007). Alguns vetores de clonagem são conhecidos como: pBR322, pUC8, pEMBL8 e  $\lambda$ gt10 (BROWN, 2013).

Os plasmídeos, vetores muito utilizados, são moléculas circulares de DNA de fita dupla que são separadas do DNA cromossômico de uma célula. Ocorrem naturalmente em bactérias ou em células eucarióticas inferiores, seu DNA é duplicado antes de cada divisão celular e suas cópias compõem cada célula filha. Os plasmídeos frequentemente carregam um ou mais genes para resistência a antibióticos e os mais comumente usados no DNAr são os que se replicam em *Escherichia coli* (ALBERTS et al., 2007; GROVES, 2006).

O vetor escolhido deve ser aberto com a mesma ER utilizada para cortar o fragmento de DNA contendo a sequência gênica, seguido por uma incubação para promover a adesão das extremidades do fragmento de DNA com o vetor. Após, o DNAr deverá ser transferido para as células hospedeiras, em um processo conhecido como transformação (WALSH, 2007), que pode acontecer por diversas técnicas, entre as quais se destacam: a eletroporação como método de transfecção, o qual consiste em sujeitar as células a um campo elétrico forte, curto e pulsado, abrindo poros na célula e permitindo que a molécula exógena seja inserida (LI; LIN, 2011); microinjeção (SAMPATH KUMAR; PUTTARAJU, 2012); bombardeamento de partículas ou biobalística (CARNEIRO; CARNEIRO; PAIVA, 2004); e, também através de vetores, como bactérias e vírus (CARNEIRO; CARNEIRO;

PAIVA, 2004; FELBERBAUM, 2015; STUDART-GUIMARÃES; LACORTE; BRASILEIRO, 2003).

## 2.3 Células hospedeiras

Diversas células podem ser usadas na técnica do DNA recombinante. Entre elas encontram-se comumente células bacterianas, leveduras e mamíferas. De acordo com o tipo de produto e a quantidade desejada baseia-se a escolha da célula e a cepa específica (GROVES, 2006).

A replicação do plasmídeo dentro da célula hospedeira começa na origem de replicação (ORI), uma sequência específica de DNA de 50 a 100 pares de bases, e continua em torno do plasmídeo circular. Dessa forma, ocorre a replicação da sequência de DNA inserida no plasmídeo, seja ela qual for (ALBERTS et al., 2007).

Para identificar se as células hospedeiras, como exemplo *E. coli*, absorveram corretamente o plasmídeo, as mesmas são incubadas em condições apropriadas para que se possa identificar qual colônia abriga o fragmento de DNA que contém o gene de interesse. Dentre as estratégias para identificação do DNA em células de *E. coli* usando o vetor pUC18, destaca-se uma em que as células são espalhadas em placas de ágar contendo o antibiótico ampicilina e um produto químico chamado X-Gal. Devido as células de *E. coli* não possuírem o gene de resistência à ampicilina (*ampR*), as células que não foram transformadas não crescerão nesse meio. Já as células que contém o plasmídeo com o gene *ampR*, mesmo sem o fragmento de DNA inserido a ele, crescerão no meio. Porém, devido a essa célula produzir a enzima  $\beta$ -galactosidase haverá a quebra do X-Gal liberando um produto de cor azul, colorindo essas colônias. Finalmente, as células contendo o plasmídeo no qual um fragmento de DNA foi inserido corretamente no gene *lacZ* não produzirão  $\beta$ -galactosidase e, assim, colônias derivadas dessas células crescerão no meio e apresentarão uma cor branca normal, as quais serão identificadas e selecionadas (WALSH, 2007).

Segundo Groves (2006), para escolher a melhor célula hospedeira deve-se observar alguns aspectos, entre eles: se o crescimento da célula é lento ou rápido; o custo; o nível de expressão de produto alcançado pela célula; a facilidade da purificação do produto levando em conta a secreção dele no meio; se o produto, no caso uma proteína, será dobrada corretamente e modificada para atingir sua atividade; e, se os vetores são adequados para ela.

## 3 | EXPRESSÃO HETERÓLOGA DE PROTEÍNAS

Devido às técnicas de DNA recombinante, que transformam células em fábricas para sintetizar proteínas, a produção comercial de insulina, hormônio do crescimento, fator estimulador de colônias de granulócitos, vacinas e outras proteínas humanas com usos terapêuticos tornou-se possível (ALBERTS et al., 2007; RAMANA; XAVIER; SHARMA, 2017). A expressão heteróloga de proteínas provenientes da manipulação de genes específicos representa a última etapa da tecnologia do DNA recombinante (CARUSO, 2007).

Inicialmente, obtém-se um clone do gene que codifica a proteína de interesse, após, os vetores que expressarão as proteínas são projetados com promotores para a transcrição do DNA e inseridos nas células hospedeiras. Para facilitar a purificação da proteína expressa, entre outros benefícios, normalmente implanta-se uma sequência nucleotídica curta (podendo ser seis resíduos de histidina no carbono terminal, por exemplo) ao final do DNA complementar, conhecida também como 'tag' (ALBERTS et al., 2007), que será discutida posteriormente.

Os vetores citados anteriormente foram projetados para isolamento e propagação de fragmentos de DNA com genes que codificam uma proteína específica, porém, para se alcançar altos níveis de expressão da proteína requerida, o vetor em questão deve suportar alto nível de transcrição e tradução. Os vetores de expressão devem possuir além de alguns elementos como origem de replicação, marcador selecionável com um gene de resistência a antibiótico e locais de restrição para clonagem, também região com promotor da transcrição, região de início da tradução, bem como terminadores da transcrição e da tradução, sequências reguladoras, sequência Shine-Dalgarno (RBS - *Ribosome Binding Site*) no caso de procaríotos, proteínas de fusão para facilitar dobramento e purificação (opcional), ou seja, deve conter todas as sequências de DNA específicas que irão direcionar a transcrição e tradução do gene de interesse. Esses vetores estão disponíveis comercialmente e cada um é adaptado para funcionar melhor em um tipo específico de célula hospedeira (KAUR; KUMAR; KAUR, 2018; SØRENSEN; MORTENSEN, 2005; WALSH, 2007).

A clonagem de genes para expressão de proteínas geralmente começa com um cDNA, pois a maioria dos genes eucarióticos contém íntrons. Esse cDNA é usado como modelo na PCR onde o gene de interesse será amplificado, clonado, sequenciado e subclonado em um ou mais vetores de expressão. Alguns motivos levam a preferência da clonagem inicial em um vetor de não expressão antes do uso do vetor de expressão, como por exemplo o tamanho maior do vetor de expressão e seu baixo número de cópias complicando um pouco mais a clonagem e sequenciamento, a ocorrência de erros de sequência no gene ou polimorfismos de nucleotídeo único e também erros nos iniciadores de PCR. Assim, para melhorar a expressão e posterior purificação das proteínas esse processo pode ser realizado, confirmando antes a sequência do gene em questão e corrigindo algum erro que possa ter ocorrido (HARTLEY, 2006).

Muitos sistemas de expressão comerciais desenvolvidos para inúmeras aplicações e compatibilidades estão disponíveis, como exemplos pode-se citar o sistema de expressão pET baseado no promotor do fago T7 e também sistemas que utilizam os promotores *trc*, *lac*, *tac* e promotor PL do fago  $\lambda$  (RESENDE, 2015; SØRENSEN; MORTENSEN, 2005). Os promotores são elementos de um vetor que estão relacionados diretamente ao rendimento proteico, devido seu efeito na força e duração da transcrição. A RNA polimerase se liga a sequência promotora, localizada ao lado do gene alvo, iniciando a síntese de RNAm. Um promotor efetivo para expressão heteróloga de proteínas deve ser forte, apresentar baixa atividade de transcrição basal e sua indução deve ser simples e de baixo custo (FRANCIS; PAGE, 2010; RESENDE, 2015).

Espera-se de uma expressão heteróloga que a proteína de interesse seja estável, solúvel, expressa em grande quantidade, não tóxica para a célula hospedeira e facilmente

purificada. Não existe um sistema ideal para todas as proteínas possíveis, cada um possui características favoráveis e limitações (WALSH, 2007).

### 3.1 Sistemas de expressão heteróloga de proteínas

O uso de bactérias como *E. coli* como sistema de expressão heteróloga de proteínas é vantajoso devido sua biologia molecular ser bem caracterizada, possuir rápido crescimento, baixos custos de produção (GOMES et al., 2016; WALSH, 2007), escalonamento direto para maiores produções, além de fácil manipulação. Como desvantagens têm-se a falta de processamentos pós-traducionais, como a glicosilação; acúmulo intracelular de corpos de inclusão, que são agregados constituídos de proteínas desnaturadas ou incorretamente enoveladas; o processamento inadequado de transcritos de genes eucarióticos que contêm sequências não codificantes ou íntrons; a presença de lipopolissacarídeo em sua superfície; capacidade limitada de formar pontes dissulfídicas; viés de códon; acúmulo de endotoxinas e a falta de um sistema para secretar as proteínas ao meio extracelular, sendo que essa última desvantagem faz com que seja necessária uma purificação mais extensa para separar a proteína de interesse das demais proteínas homólogas adicionais produzidas pelas células de *E. coli* (BROWN, 2013; GOMES et al., 2016; GROVES, 2006; WALSH, 2007).

Alternativas para contornar algumas desvantagens desse sistema têm sido implantadas, como o uso de proteínas de fusão, uso de peptídeo sinal para direcionamento da proteína, condições especiais de crescimento, uso de aditivos, co-expressão com chaperonas e escolha da linhagem mais adequada (KAUR; KUMAR; KAUR, 2018). Deve-se selecionar a cepa de *E. coli* mais interessante e viável para clonagem e expressão de determinada proteína (SØRENSEN; MORTENSEN, 2005), pois existem diversas disponíveis comercialmente, como exemplo BL21 (Novagen), BL21 (DE3) (Novagen), BLR (DE3) (Addgene), Origami B (Novagen), T7 Express (Novagen), entre outras (KAUR; KUMAR; KAUR, 2018). A insulina humana recombinante foi o primeiro produto biofarmacêutico feito por engenharia genética, sendo produzida em *E. coli* com aprovação de marketing em 1982 (WALSH, 2007).

Um sistema de expressão eucariótico interessante e antigo é o de leveduras, representado na maioria das vezes por *Saccharomyces cerevisiae*, mas também por *Hansenulla polymorpha* e *Picha pastoris* (BALAMURUGAN et al., 2006). Os sistemas eucarióticos permitem a modificação e o processamento de proteínas ao contrário dos bacterianos, o que constitui vantagem (CANÇADO, 2002; GOMES et al., 2016). Como vantagens também caracterizam-se o crescimento rápido em meios relativamente de baixo custo, não há produção de endotoxinas, parede externa resistente a danos físicos, possibilidade de secretar produtos de proteína para o meio mais rapidamente do que a *E. coli*, organismos listados como seguros pela grande aplicação industrial já utilizada, capacidade de processar transcritos de RNA para remover os íntrons e promover glicosilação de proteínas, esta última pode representar também uma desvantagem devido ao padrão de glicosilação variar um pouco das proteínas nativas podendo ocorrer também hiperglicosilação (GROVES, 2006; WALSH, 2007). Outra desvantagem observada é em relação aos níveis de expressão de proteínas heterólogas que situam-se abaixo dos alcançados por *E. coli* (WALSH, 2007) e também a possibilidade de viés do códon (GOMES

et al., 2016). Os principais biofarmacêuticos obtidos através de leveduras são a albumina sérica humana, insulina e seus análogos, e vacinas contra papilomavírus e hepatite (NIELSEN, 2013).

Fungos filamentosos também possuem capacidade de produzir grandes quantidades de proteína e são muito utilizados na indústria de alimentos principalmente como produtores de enzimas, como por exemplo *Aspergillus niger*, *Aspergillus oryzae* e *Trichoderma reesei* (GOMES et al., 2016; LANDOWSKI et al., 2016). Possuem vantagens em relação ao crescimento rápido e robusto, secreção das proteínas nos meios extracelulares facilitando a purificação e também a capacidade de realizar modificações pós-traducionais como os demais sistemas eucarióticos (HAVLIK et al., 2017; WALSH, 2007). Porém, como desvantagem tem-se a maior complexidade desse sistema e a relativa falta de compreensão de sua fisiologia em comparação às bactérias (SU et al., 2012).

As microalgas, principalmente da espécie *Chlamydomonas reinhardtii*, representam outro sistema para a produção de proteínas recombinantes, devido ao baixo custo e tecnologia simples (DYO; PURTON, 2018). *C. reinhardtii* é um organismo modelo cujos três genomas foram sequenciados, sendo possível então direcionar a modificação genética (TAUNT; STOFFELS; PURTON, 2018). A introdução do gene de interesse no genoma de pequenos cloroplastos é desejada, pois garante uma expressão estável e de alto nível (DYO; PURTON, 2018). Outras vantagens apresentadas são a formação de ligações dissulfeto e dobragem adequada. Muitas proteínas recombinantes, incluindo anticorpos monoclonais, fatores de crescimento e vacinas são produzidas em *C. reinhardtii* (TAUNT; STOFFELS; PURTON, 2018).

Outras células eucarióticas usadas incluem linhagens celulares derivadas de tumor, células de ovário de hamster chinês e células de rim de hamster bebê. O crescimento dessas células é muito mais lento do que leveduras ou bactérias além de serem mais sensíveis às mudanças no pH, temperatura, nível de oxigênio, metabólitos e às forças de cisalhamento. O cultivo para produção em larga escala é mais complicado e apresentam níveis variados de expressão das proteínas heterólogas (GOMES et al., 2016; GROVES, 2006), além de geralmente exigirem a suplementação de fatores de crescimento, aminoácidos, agentes redutores ou vitaminas (ZHU, 2012). Porém, dentre as vantagens encontram-se a capacidade de secretar o produto proteico no meio, de processar adequadamente os transcritos de RNA para expressão gênica e modificar adequadamente a proteína expressa por clivagem, glicosilação e redobragem, por exemplo (GROVES, 2006). A cultura de células de mamíferos geralmente é usada apenas na fabricação de proteínas terapêuticas que mostram modificações pós-traducionais extensas e essenciais, incluindo gonadotrofinas, algumas citocinas, anticorpos monoclonais, fatores de coagulação, enzimas, entre outros (DUMONT et al., 2016; WALSH, 2007). Os promotores do citomegalovírus são amplamente utilizados na construção de cassetes de expressão para células de mamíferos (KHAN, 2014).

Sobre os sistemas de cultura de células de insetos envolve-se a infecção dessas células cultivadas com o vetor de expressão baculovírus geneticamente modificado, que carrega integrado ao seu genoma o gene que codifica a proteína desejada controlado por um potente promotor viral e pelo seu terminador correspondente (FELBERBAUM, 2015; STOGER et al., 2005). Alguns insetos utilizados são *Spodoptera frugiperda*, *Drosophila*

*melanogaster* e *Autographa californica* (GECHELE et al., 2015). São células que crescem mais rápido comparadas a de mamíferos, não há risco de contaminação por príons e DNA oncogênico, garantem alto rendimento (CONTRERAS-GÓMEZ et al., 2014; GECHELE et al., 2015) e servem de ferramenta para produção de glicoproteína recombinante. Porém, como desvantagens encontram-se as modificações pós-traducionais que podem ser incompletas, podem diferir dos padrões associados às glicoproteínas humanas nativas e culturas com condições mais exigentes. O antígeno de superfície da hepatite B e IFN- $\gamma$  são exemplos de proteínas terapêuticas produzidas com sucesso em escala laboratorial em linhas celulares de insetos (WALSH, 2007).

A produção de proteínas heterólogas em animais transgênicos também ganhou um certo destaque, pois como cada célula do animal transgênico resultante abrigará uma cópia do DNA transferido, essa nova informação genética introduzida pode ser transmitida de uma geração para a seguinte, produzindo a proteína de interesse continuamente. Para facilitar a liberação da proteína, uma alternativa foi direcionar a produção de proteínas para a glândula mamária, na qual cabras e ovelhas provaram ser interessantes devido a características como altas capacidades de produção de leite e facilidade de manuseio e criação (WALSH, 2007). A primeira proteína recombinante aprovada em 2006 foi a antitrombina, secretada no leite de cabras (MAKSIMENKO et al., 2013). Além do leite, anticorpos e outras proteínas já foram produzidos no sangue de porcos e coelhos transgênicos. Porém, várias razões tornam esse sistema inviável industrialmente, entre elas a possibilidade de coletar pequenos volumes de sangue, a complexidade do soro que contém uma variedade de proteínas nativas tornando a purificação mais difícil, a falta de estabilidade das proteínas e os efeitos colaterais fisiológicos negativos que podem ser causados pela proteína no animal produtor (WALSH, 2007).

Também foi demonstrada a produção de  $\beta$ -lactamase na clara de ovos de galinhas transgênicas (HARVEY et al., 2002). Algumas vantagens como dobragem adequada das proteínas, glicosilação correta e modificações pós-traducionais adequadas são destacadas (GOMES et al., 2016), porém, o custo mais alto, tempo de produção elevado e baixo rendimento representam problemas, além ainda do risco de contaminação por vírus (CANÇADO, 2002; GOMES et al., 2016; LARRICK; THOMAS, 2001).

O uso de plantas como sistema de expressão de proteínas heterólogas, em comparação aos demais eucariotos, apresenta-se com maior segurança, menor tempo, baixo custo e é superior em termos de armazenamento e distribuição. Constituem um sistema ideal para expressar proteínas heterólogas que requerem modificações pós-traducionais um pouco mais complicadas. Tais proteínas expressas podem ser localizadas em diferentes órgãos da planta e existe a possibilidade de manipular o tempo de expressão para estágios de crescimento específicos (ŁOJEWSKA et al., 2016). Oferecem uma fonte de material menos custosa para seu cultivo em comparação a outros microrganismos e devido a facilidade pode-se cultivar um número maior para aumentar o rendimento total das proteínas (ESPÍRITO SANTO, 2013).

Em relação ao uso de plantas transgênicas, uma gama de proteínas terapêuticas podem ser expressas pelo tecido da planta, pois a planta pode receber a transferência dos genes de interesse através de vetores principalmente pelo patógeno vegetal *Agrobacterium*, no qual uma porção do plasmídeo é translocada e integrada ao genoma da célula vegetal,

ou também através de vetores virais de plantas (BUYEL; TWYMAN; FISCHER, 2017; WALSH, 2007). No primeiro método, plantas transgênicas com uma expressão estável são obtidas, enquanto no segundo caso a expressão transitória do gene é obtido (YAO et al., 2015). Algumas desvantagens do sistema de plantas transgênicas são padrões de glicosilação diferentes da proteína humana nativa (WALSH, 2007), presença de metabólitos que podem contaminar o produto bruto, natureza geográfica/sazonal do crescimento da planta (BALAMURUGAN et al., 2006), a variação dos níveis de expressão de acordo com o destino e a necessidade de melhorar o desenvolvimento dos ensaios (GOMES et al., 2016)invasion and metastasis. Tissue and serum levels of MMP2 and MMP9 correlate with disease prognosis. Real-time PCR is emerging as an alternative or supplementary technique to immunohistochemistry (IHC).

### 3.2 Análise e seleção dos recombinantes

Quanto ao processo de seleção, o mesmo pode envolver a resistência a antibióticos ou a habilidade das células hospedeiras de crescerem na ausência de certos suplementos nutricionais, capacidade gerada pelos vetores inseridos, como previamente discutido a respeito da inserção do fragmento de interesse no gene lacZ, o que determinará quais colônias foram transformadas pela cor apresentada (WALSH, 2007). Pode-se realizar também uma PCR seguida de sequenciamento para confirmar a presença do gene de interesse.

Além disso, técnicas imunológicas podem ser usadas para detectar a presença da proteína que foi produzida pelos clones (GROVES, 2006), como a verificação da expressão da proteína pela técnica de *Western blotting* acompanhada por eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (SDS-PAGE) (LIU; YANG, 2012), e ainda, dependendo do produto em questão, existe a possibilidade de encontrar o clone adequadamente transformado que expressou uma enzima, por exemplo, procurando pela atividade dela na célula (GROVES, 2006).

### 3.3 Purificação de proteínas

Antes da purificação é necessário que a proteína seja retirada da célula hospedeira e o processo vai depender se o sistema de expressão libera o produto no meio intracelular ou extracelular. Nos dois casos, as células podem ser coletadas por centrifugação ou microfiltração (WALSH, 2007). Ainda, no caso de produto intracelular, realiza-se a ruptura celular para obter a proteína de interesse. Essa ruptura é frequentemente alcançada por métodos mecânicos como homogeneização ou agitação vigorosa com abrasivos, mas também pode-se recorrer ao uso de produtos químicos como detergentes, solventes como tolueno, agentes caotrópicos como ureia, através de enzimas como lisozima, exposição a condições alcalinas e sonicação (AHMAD et al., 2014). Após a ruptura das células, ocorre a concentração do produto bruto, podendo ocorrer principalmente por ultrafiltração ou também através de sais que induzem a precipitação (WALSH, 2007).

No caso do produto intracelular produzido em procariotos, a secreção no espaço extra citoplasmático pode ser alcançada fundindo a proteína recombinante a um peptídeo sinal adequado. Alguns exemplos de peptídeos são: OmpA (proteína A da membrana externa), PhoE (proteína E dos poros da membrana externa), Lpp (lipoproteína mureína) e OmpT

(protease VII) (CHOI; LEE, 2004). O peptídeo sinal é clivado e a proteína recombinante se dobra no periplasma com a ajuda de chaperonas e formam-se ligações dissulfeto gerando a conformação correta (KAUR; KUMAR; KAUR, 2018). Geralmente a proteína é expressa em maiores quantidades no citoplasma bacteriano, porém na forma de corpos de inclusão. Então, é necessário um processamento adequado desses agregados para se obter a forma solúvel da proteína, passando pelo isolamento, purificação, solubilização, renaturação e dobramento, através de reagentes específicos e técnicas diversas (CLARK, 2001).

Para purificar proteínas recombinantes são utilizados os mesmos meios de purificação de proteínas tradicionais, porém, a purificação pode ser até mais direta devido a capacidade de alcançar altos níveis de expressão da proteína desejada. Técnicas de cromatografia que as separam considerando suas propriedades específicas e diferenças são amplamente utilizadas, como exemplo cita-se a cromatografia de troca iônica, de afinidade, de imunoafinidade e cromatografia líquida de alta resolução (HPLC), sendo que essa última poderá ser utilizada desde que a alta pressão aplicada não afete o produto proteico (WALSH, 2007).

O procedimento de purificação consiste na separação da proteína alvo, mantendo sua estrutura química e atividade biológica (LABROU, 2014), e representam entre 45 a 92% dos custos totais da fabricação de proteínas recombinantes (SARASWAT et al., 2013). Como princípio básico da separação cromatográfica tem-se uma fase móvel fluida que transporta espécies em direção a uma fase estacionária sólida, baseando-se nas diferenças de afinidade. Ao longo da coluna, alguns componentes da amostra transportada pelo eluente apresentarão interações mais fortes com a fase estacionária do que outros, gerando perfis de concentração e diferentes velocidades de eluição. Assim, espécies mais retidas eluirão por último, permitindo a coleta do produto de interesse com maior pureza (FARIA; RODRIGUES, 2015).

Cada metodologia de purificação utiliza uma característica particular das proteínas, por exemplo, quando se trata da especificidade do ligante, a cromatografia de afinidade é usada, tendo como base a interação bioespecífica entre uma proteína e um ligante apropriado, sendo o método mais popular; quando as proteínas possuem diferenças na carga a um determinado pH a cromatografia de troca iônica é preferida; quando as proteínas variam em tamanho é utilizada a cromatografia por exclusão de tamanho ou gel filtração; e, quando se trata de hidrofobicidade, são utilizadas a cromatografia de fase reversa e a cromatografia de interação hidrofóbica (OWCZAREK; GERSZBERG; HNATUSZKO-KONKA, 2019). Dependendo da proteína em questão, várias etapas de cromatografia podem ser requeridas (WALSH, 2007).

Algumas considerações importantes sobre o processo de purificação de proteínas são: deve ser simples e com o mínimo de etapas; uso de reagentes mais baratos; deve haver monitoramento constante; o produto final deve ser altamente purificado; gastar pouco tempo; permitir alta reprodutibilidade e sempre utilizar aparelhagem confiável (KHAN, 2014).

A incorporação de 'tags' ou proteínas de fusão podem facilitar o processo de purificação. Utiliza-se da engenharia genética para incorporação de tags específicas de peptídeos ou proteínas à proteína de interesse, que ligam-se às proteínas alvo por

um sítio de reconhecimento de uma protease específica (KOSOBOKOVA; SKRYPNIK; KOSORUKOV, 2016). Foram projetadas com sucesso tags que permitem a purificação rápida e direta da proteína híbrida por técnicas como cromatografia de afinidade, troca iônica, interação hidrofóbica ou imunoafinidade. Dependendo do marcador adicionado pode ser gerada uma carga positiva na proteína o que facilita sua purificação por troca catiônica, por exemplo (WALSH, 2007).

Após a purificação da proteína híbrida, é necessária a remoção do marcador, que pode ser realizada por meios químicos ou enzimáticos (WALSH, 2007), adicionando outra etapa e custo ao processo de produção de proteína recombinante (LI, 2011), porém, dependendo da sequência adicionada pode não afetar a estrutura final da proteína (OWCZAREK; GERSZBERG; HNATUSZKO-KONKA, 2019). Essas proteínas de fusão que se ligam às proteínas recombinantes facilitam não só a purificação, como também podem proteger as proteínas da proteólise, aumentar a solubilidade, facilitar o reconhecimento e melhorar a expressão. Podem ser tanto sequências curtas de aminoácidos como resíduo de histidina (poli-His), resíduo de arginina (poli-Arg) e epítipo FLAG, quanto proteínas como tioredoxina (TRX), proteína ligante da maltose (MBP), glutationa S-transferase (GST) e proteína verde fluorescente (GFP) (BUCHER; EVDOKIMOV; WAUGH, 2002).

Para acompanhar o progresso da purificação de proteínas ensaios analíticos são essenciais, entre eles estão métodos físicos e químicos, para determinar a qualidade da preparação final e a caracterização da proteína (VEDADI et al., 2010). Um método de quantificação do produto final é a medição da absorbância da proteína utilizando a ligação com corantes, e, para avaliar a pureza utiliza-se a técnica de SDS-PAGE, podendo ser acompanhada de *Western blotting*, além também de outras técnicas como eletroforese capilar, HPLC e espectrometria de massas (WALSH, 2007).

Adicionalmente, podem ser realizadas diferentes análises que não somente detectam impurezas proteicas, mas verificam se a substância está em total conformidade com o que foi proposto inicialmente, como mapeamento de peptídeos, análise de aminoácidos, sequenciamento N-terminal e análises espectrofotométricas. O processamento seguinte vai depender do objetivo com a proteína obtida, o que pode envolver a adição de excipientes, filtração estéril, liofilização ou preparo de solução, por exemplo (WALSH, 2007).

## 4 | CONSIDERAÇÕES FINAIS

A expressão heteróloga de proteínas, que surgiu através da tecnologia do DNA recombinante, permitiu um avanço na produção em larga escala de muitos produtos importantes como hormônios, vacinas, entre outros. No presente capítulo foram descritos de forma geral o histórico, aplicação e todo processo que envolve desde a manipulação do gene de interesse, escolha de vetores, características dos sistemas de expressão heteróloga de proteínas, métodos de análise e identificação até a completa purificação das proteínas.

## REFERÊNCIAS

- AHMAD, M. *et al.* **Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production.** Applied Microbiology and Biotechnology, v. 98, p. 5301–5317, 2014.
- ALBERTS, B. *et al.* **Molecular Biology of the Cell.** 5. ed. New York: Garland Science, 2007.
- BALAMURUGAN, V.; SEN, A.; SARAVANAN, P.; SINGH, R. K. **Biotechnology in the Production of Vaccine or Antigen for animal health.** Journal of Animal and Veterinary Advances, v. 5, n. 6, p. 487–495, 2006.
- BROWN, T. A. **Gene Cloning and DNA Analysis: An Introduction.** 6. ed. Hong Kong: John Wiley & Sons, 2013.
- BUCHER, M. H.; EVDOKIMOV, A. G.; WAUGH, D. S. **Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrinbinding protein.** Acta Crystallographica Section D Biological Crystallography, v. 58, p. 392–397, 2002.
- BUYEL, J. F.; TWYMAN, R. M.; FISCHER, R. **Very-large-scale production of antibodies in plants: The biologization of manufacturing.** Biotechnology Advances, v. 35, n. 4, p. 458–465, 2017.
- CANÇADO, L. J. **Utilização de Sementes de Tabaco Transgênico como Biorreatores para Produção de um Fragmento scFv de um Anticorpo Monoclonal.** [s.l.] Universidade Estadual de Campinas UNICAMP, 2002.
- CARNEIRO, A. A.; CARNEIRO, N. P.; PAIVA, E. **Transformação genética de milho utilizando o bombardeamento de partículas.** Sete Lagoas: Embrapa Milho e Sorgo, 2004.
- CARUSO, C. S. **Clonagem, expressão e caracterização de proteínas recombinantes de Xylella fastidiosa.** [s.l.] Universidade de São Paulo, 2007.
- CHOI, J. H.; LEE, S. Y. **Secretory and extracellular production of recombinant proteins using *Escherichia coli*.** Applied Microbiology and Biotechnology, v. 64, n. 5, p. 625–635, 2004.
- CLARK, E. D. B. **Protein refolding for industrial processes.** Current Opinion in Biotechnology, v. 12, n. 2, p. 202–207, 2001.
- CONTRERAS-GÓMEZ, A. *et al.* **Protein production using the baculovirus-insect cell expression system.** Biotechnology Progress, v. 30, n. 1, p. 1–18, 2014.
- DUMONT, J. *et al.* **Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives.** Critical Reviews in Biotechnology, v. 36, n. 6, p. 1110–1122, 2016.
- DYO, Y. M.; PURTON, S. **The algal chloroplast as a synthetic biology platform for production of therapeutic proteins.** Microbiology, v. 164, n. 2, p. 113–121, 2018.
- EGELKROUT, E.; RAJAN, V.; HOWARD, J. A. **Overproduction of recombinant proteins in plants.** Plant Science, v. 184, p. 83–101, 2012.
- ESPÍRITO SANTO, P. F. R. **Pesquisa de estirpes de *Saccharomyces cerevisiae* para expressão de proteína recombinante.** [s.l.] Universidade de Coimbra, 2013.

FARIA, R. P. V.; RODRIGUES, A. E. **Instrumental aspects of Simulated Moving Bed chromatography.** Journal of Chromatography A, v. 1421, p. 82–102, 2015.

FELBERBAUM, R. S. **The baculovirus expression vector system: A commercial manufacturing platform for viral vaccines and gene therapy vectors.** Biotechnology Journal, v. 10, n. 5, p. 702–714, 2015.

FERRIER, D. R. **Bioquímica Ilustrada.** 7. ed. Porto Alegre: Artmed, 2017.

FRANCIS, D. M.; PAGE, R. **Strategies to optimize protein expression in *E. coli*.** Current Protocols in Protein Science, n. SUPPL. 61, 2010.

GECHELE, E. *et al.* **A comparative analysis of recombinant protein expression in different biofactories: Bacteria, insect cells and plant systems.** Journal of Visualized Experiments, n. 97, 2015.

GOMES, A. R. *et al.* **An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins.** Advances in Animal and Veterinary Sciences, v. 4, n. 7, p. 346–356, 2016.

GROVES, M. J. **Pharmaceutical Biotechnology.** 2. ed. New York: Taylor & Francis Group, LLC, 2006.

HARTLEY, J. L. **Cloning technologies for protein expression and purification.** Current Opinion in Biotechnology, v. 17, n. 4, p. 359–366, 2006.

HARVEY, A. J. *et al.* **Expression of exogenous protein in the egg white of transgenic chickens.** Nature Biotechnology, v. 20, n. 4, p. 396–399, 2002.

HAVLIK, D. *et al.* **Establishment of *Neurospora crassa* as a host for heterologous protein production using a human antibody fragment as a model product.** Microbial Cell Factories, v. 16, n. 1, p. 128, 2017.

KAUR, J.; KUMAR, A.; KAUR, J. **Strategies for optimization of heterologous protein expression in *E. coli* : Roadblocks and reinforcements.** International Journal of Biological Macromolecules, v. 106, p. 803–822, 2018.

KHAN, K. H. **Gene Expression Systems and Recombinant Protein Purification.** Research Journal of Pharmaceutical, Biological and Chemical Sciences, v. 5, n. 6, p. 450–463, 2014.

KOSOBOKOVA, E. N.; SKRYPNIK, K. A.; KOSORUKOV, V. S. **Overview of fusion tags for recombinant proteins.** Biochemistry (Moscow), v. 81, n. 3, p. 187–200, 2016.

LABROU, N. E. **Protein purification: An overview.** Methods in Molecular Biology, v. 1129, p. 3–10, 2014.

LANDOWSKI, C. P. *et al.* **Enabling low cost biopharmaceuticals: high level interferon alpha-2b production in *Trichoderma reesei*.** Microbial Cell Factories, v. 15, n. 104, 2016.

LARRICK, J. W.; THOMAS, D. W. **Producing proteins in transgenic plants and animals.** Current Opinion in Biotechnology, v. 12, n. 4, p. 411–418, 2001.

LI, J.; LIN, H. **Numerical simulation of molecular uptake via electroporation.** Bioelectrochemistry, v. 82, n. 1, p. 10–21, 2011.

LI, Y. **Self-cleaving fusion tags for recombinant protein production.** *Biotechnology Letters*, v. 33, n. 5, p. 869–881, 2011.

LIU, Z. Q.; YANG, P. C. **Construction of pET-32  $\alpha$  (+) vector for protein expression and purification.** *North American Journal of Medical Sciences*, v. 4, n. 12, p. 651–655, 2012.

ŁOJEWSKA, E. *et al.* **Extraction and purification methods in downstream processing of plant-based recombinant proteins.** *Protein Expression and Purification*, v. 120, p. 110–117, 2016.

MAKSIMENKO, O. G. *et al.* **Use of Transgenic Animals in Biotechnology: Prospects and Problems.** *Acta Naturae*, v. 5, n. 1, p. 33–46, 2013.

NELSON, D. L.; COX, M. M. **Princípios de Bioquímica de Lehninger.** 5. ed. Porto Alegre: Artmed, 2011.

NIELSEN, J. **Production of biopharmaceutical proteins by yeast: Advances through metabolic engineering.** *Bioengineered*, v. 4, n. 4, p. 207–211, 2013.

OWCZAREK, B.; GERSZBERG, A.; HNATUSZKO-KONKA, K. **A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals.** *BioMed Research International*, v. 2019, 2019.

RAMANA, K. V.; XAVIER, J. R.; SHARMA, R. K. **Recent Trends in Pharmaceutical Biotechnology.** *Pharmaceutical Biotechnology: Current Research*, v. 1, n. 1, p. 1–10, 2017.

RAI, M.; PADH, H. **Expression systems for production of heterologous proteins.** *Current Science*, v. 80, n. 9, p. 1121–1128, 2001.

RESENDE, R. R. **Biotecnologia aplicada à saúde - vol. 2: fundamentos e aplicações.** São Paulo: Blucher, 2015.

SAMPATH KUMAR, S.; PUTTARAJU, H. P. **Improved microinjection technique for mosquito vectors.** *Indian Journal of Medical Research*, v. 136, n. 6, p. 971–978, 2012.

SARASWAT, M. *et al.* **Preparative Purification of Recombinant Proteins: Current Status and Future Trends.** *BioMed Research International*, v. 2013, p. 1–18, 2013.

SØRENSEN, H. P.; MORTENSEN, K. K. **Advanced genetic strategies for recombinant protein expression in *Escherichia coli*.** *Journal of Biotechnology*, v. 115, p. 113–128, 2005.

STOGER, E. *et al.* **Sowing the seeds of success: Pharmaceutical proteins from plants.** *Current Opinion in Biotechnology*, v. 16, n. 2, p. 167–173, 2005.

STUDART-GUIMARÃES, C.; LACORTE, C.; BRASILEIRO, A. C. M. **Transformação Genética em Espécies Florestais.** *Ciência Florestal*, v. 13, n. 1, p. 167–178, 2003.

SU, X.; SCHMITZ, G.; ZHANG, M.; MACKIE, R. I.; CANN, I. K. **Heterologous gene expression in filamentous fungi.** In: GADD, G. M.; SARIASLANI, S. *Advances in Applied Microbiology*. USA: Elsevier, 2012. p 1-61.

TAUNT, H. N.; STOFFELS, L.; PURTON, S. **Green biologics: The algal chloroplast as a platform for making biopharmaceuticals.** *Bioengineered*, v. 9, n. 1, p. 48–54, 2018.

VEDADI, M. *et al.* **Biophysical characterization of recombinant proteins: A key to higher structural genomics success.** *Journal of Structural Biology*, v. 172, n. 1, p. 107–119, 2010.

WALSH, G. **Pharmaceutical Biotechnology Concepts and Applications.** England: John Wiley & Sons Ltd, 2007.

YAO, J. *et al.* **Plants as factories for human pharmaceuticals: Applications and challenges.** *International Journal of Molecular Sciences*, v. 16, n. 12, p. 28549–28565, 2015.

ZHU, J. **Mammalian cell protein expression for biopharmaceutical production.** *Biotechnology Advances*, v. 30, n. 5, p. 1158–1170, 2012.

## **SOBRE OS ORGANIZADORES**

**ANDERSON BARROS ARCHANJO** – Farmacêutico pela Universidade Federal do Espírito Santo (2014). Doutor em Biotecnologia (Área de concentração: Biotecnologia na Saúde) pela Rede Nordeste de Biotecnologia ponto focal Universidade Federal do Espírito Santo (2018) e pesquisador em Biotecnologia (Nível: Pós-doutorado) pela mesma universidade (2018-2021) com fomento pela FAPES/CAPES (Processo: 83757490). Atuou como professor voluntário, para a graduação, na Universidade Federal do Espírito Santo nas disciplinas de Bioquímica, Bioquímica III, Microbiologia Básica e Biologia Celular, e para a pós-graduação as disciplinas de Metodologia da Pesquisa Científica e Tópicos Especiais em Biotecnologia II: Avanços Metodológicos em Biologia Molecular e Biotecnologia. Tem experiência nas áreas de Patologia Experimental e Laboratorial e Biologia Celular e Molecular, atuando nos seguintes temas: câncer de cabeça e pescoço, regulação de expressão gênica, vias de hipóxia e estresse oxidativo, tabagismo e caracterização elementar em câncer oral.

**ALINE RIBEIRO BORÇOI** - Pesquisadora FAPES/CNPq-Desenvolvimento Científico Regional. Bacharel em farmácia pela Universidade Federal de Juiz de Fora (UFJF), mestre em farmacologia pela Universidade Federal de São Paulo (UNIFESP), doutorado e pós-doutorado em biotecnologia em saúde pela Universidade Federal do Espírito Santo (UFES)/RENORBIO. Atua na área de neurociência, psicofarmacologia e epigenética, com ênfase em comportamento e transtornos mentais.

**SUZANNY OLIVEIRA MENDES** - Pesquisadora de Pós-Doutorado em Biotecnologia na Universidade Federal do Espírito Santo. Possui Bacharelado em Ciências Biológicas, Mestrado e Doutorado em Biotecnologia pela mesma instituição. Foi Professora Voluntária da disciplina de Imunologia por 2 semestres. Possui 2 especializações na área da Docência. Tem experiência na área de Biologia Molecular, atuando principalmente nos seguintes temas: câncer de cabeça e pescoço, câncer intestinal, regulação de expressão gênica, vias de hipóxia, inflamação e tabagismo. Atualmente desenvolve trabalhos na área de epigenética, hábitos de vida e adoecimento.

**ADRIANA MADEIRA ÁLVARES-DA-SILVA** - Possui graduação em Ciências Biológicas pela Universidade Estadual Paulista Júlio de Mesquita Filho UNESP São José do Rio Preto (1989), Mestrado em Genética pelo Departamento de Anatomia da Universidade Federal de São Paulo (2000), Doutorado em Ciências na área de Genética e Biologia molecular aplicada à Endocrinologia Clínica pela Universidade Federal de São Paulo (2003), Pós-Doutorado em Biologia Molecular pelo projeto Genoma Clínico do Câncer de Cabeça e Pescoço no Hospital Heliópolis. Professora do Departamento de Biologia do CCENS/UFES (2010-2020). Atualmente é Professora Associada I do Departamento de Morfologia do CCS/UFES e docente do Programa de Pós-Graduação em Biotecnologia e RENORBIO da UFES. Tem experiência na área de Genética, com ênfase em Biologia Molecular e Epigenética, atua principalmente nos seguintes temas: carcinoma epidermoide de cabeça e pescoço, epigenética, práticas integrativas e complementares na redução do estresse, estilo de vida, biologia molecular. Na área de extensão é responsável pela criação da Rede do Bem Capixaba - atividade de organização da sociedade civil. Coordenadora do projeto de implantação de práticas integrativas no SUS e criação do ambulatório de redução de estresse em Alegre - ES.

# AVANÇOS METODOLÓGICOS EM BIOLOGIA MOLECULAR E BIOTECNOLOGIA

 [www.atenaeditora.com.br](http://www.atenaeditora.com.br)

 [contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br)

 @atenaeditora

 [www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br)

# AVANÇOS METODOLÓGICOS EM BIOLOGIA MOLECULAR E BIOTECNOLOGIA



[www.atenaeditora.com.br](http://www.atenaeditora.com.br)



[contato@atenaeditora.com.br](mailto:contato@atenaeditora.com.br)



[@atenaeditora](https://www.instagram.com/atenaeditora)



[www.facebook.com/atenaeditora.com.br](https://www.facebook.com/atenaeditora.com.br)