

Scientific Electronic Archives

Issue ID: Sci. Elec. Arch. Vol. 15 (10)

October 2022

DOI: <http://dx.doi.org/10.36560/151020221603>

Article link: <https://sea.ufr.edu.br/SEA/article/view/1603>



Principais métodos de sequenciamento de DNA

Main methods of DNA sequencing

Corresponding author

Rowersan Cabral Silva

Universidade de Cuiabá

rowersancabral@gmail.com

Airton Lima

Universidade de Cuiabá

Liliane Cristina da Silva Souza

Universidade de Cuiabá

Resumo. O sequenciamento do ácido desoxirribonucleico (DNA) é resultado de uma série de procedimentos bioquímicos que têm como principal finalidade a determinação da sequência de bases nitrogenadas, que são elas: adenina (A), guanina (G), citosina (C) e timina (T). O sequenciamento do DNA é de extrema importância pois está empregada na identificação, diagnóstico e tratamento de doenças e principalmente na produção biotecnológica para se desenvolver novos produtos e serviços. Atualmente, os sequenciadores de DNA estão divididos em 1ª geração (Maxam-Gilbert e Método de Sanger) que consiste em marcar radioativamente alguns desoxinucleotídeos livres e adicioná-los em uma fita molde de DNA e posteriormente ler em impressão da PCR. Na 2ª geração (454-Roche, Illumina Genome e SOLID) é uma técnica que tem como princípio a detecção de um pirofosfato liberado durante a incorporação do desoxinucleotídeo à cadeia molde de DNA. Já na 3ª geração (Ion Torrent) funciona através da liberação de íons de hidrogênio, sempre que um nucleotídeo se atribui a uma fita de DNA molde, gerando uma assinatura química. Na 4ª geração (Nanopore) a técnica visa identificar o DNA que passa através de um nanoporo, onde cada nucleotídeo que atravessa é registrado um sinal elétrico diferente gerando uma assinatura molecular. Com isso é possível identificar que ao longo do tempo e com o desenvolvimento da tecnologia esse processo foi sendo modificado e aperfeiçoado buscando melhor qualidade e menor tempo de sequenciamento.

Palavras-chaves: Sequenciamento, DNA, Biotecnologia, Genoma.

Abstract. Deoxyribonucleic acid (DNA) sequencing is the result of a series of biochemical procedures whose main purpose is to determine the sequence of nitrogenous bases, which are: adenine (A), guanine (G), cytosine (C) and thymine (T). DNA sequencing is extremely important because it is used in the identification, diagnosis and treatment of diseases and mainly in biotechnological production to develop new products and services. Currently, DNA sequencers are divided into 1st generation (Maxam-Gilbert and Sanger Method) which consists of radioactively labeling some free deoxynucleotides and adding them to a DNA template strand and later reading in PCR printing. In the 2nd generation (454-Roche, Illumina Genome and SOLID) it is a technique whose principle is the detection of a pyrophosphate released during the incorporation of the deoxynucleotide to the DNA template strand. In the 3rd generation (Ion Torrent) it works by releasing hydrogen ions whenever a nucleotide is attached to a DNA template strand, generating a chemical signature. In the 4th generation (Nanopore) the technique aims to identify the DNA that passes through a nanopore, where each nucleotide that passes through is registered a different electrical signal generating a molecular signature. With this it is possible to identify that over time and with the development of technology this process was being modified and improved seeking better quality and shorter sequencing time.

Keywords: Sequencing, DNA, Biotechnology, Genome

Introdução

A partir da descoberta da estrutura do DNA em 1953 por Francis Crick e James Watson em

Cavendish na Inglaterra, foi possível compreender a complexidade e diversidade do genoma dos seres vivos.

Os primeiros métodos de sequenciamento direto do DNA só foram criados na década de 1970 e os conhecimentos existentes sobre a organização do genoma, eram baseados principalmente em estudos de genética reversa, na qual a sequência de aminoácidos do produto de interesse era retrotraduzida em uma sequência de nucleotídeos com base nos códons apropriados, porém considerando a característica degenerada do código genético este processo pode ser complicado e os resultados não correspondem sempre com a realidade.

Os dois primeiros métodos de sequenciamento de DNA foram os de Maxam-Gilbert, conhecido como método de clivagem química e o método de terminação de cadeia de Sanger, que foi utilizado em trabalhos de sequenciamento até meados dos anos 2000 (KREMER e PINTO, 2016).

O Projeto Genoma Humano propiciou o desenvolvimento de soluções tecnológicas mais avançadas tanto para a geração dos dados quanto para a análise; assim, estes avanços ajudaram a responder aos novos questionamentos que surgiam, porém, as principais barreiras como a produção limitada e os altos custos de sequenciamento permaneciam (KREMER e PINTO, 2016).

O lançamento da primeira plataforma de sequenciamento de alto rendimento (high throughput), em meados da década de 2000, propiciou uma redução de 50.000 vezes no custo do sequenciamento. Com isso, essa Nova Geração de Sequenciadores (NGS) de DNA continuou a evoluir, aumentando a capacidade por um fator de 100-1.000. Cabe salientar que, devido ao grande número de sequências geradas, a tecnologia de processamento dos dados também teve que evoluir,

incluindo a capacidade computacional associada e softwares (KREMER e PINTO, 2016).

Os métodos mais modernos de sequenciamento, não se limitam em analisar pequenos fragmentos de DNA, eles passaram a avaliar milhões deles de uma única vez, assim esse avanço tecnológico permitiu que fosse realizado um sequenciamento mais eficiente e rápido de genomas inteiros através de uma única reação (KREMER e PINTO, 2016).

Assim, este trabalho visa exemplificar o funcionamento das quatro gerações de sequenciadores, evidenciando passo a passo de cada método.

Materiais e métodos

A abordagem desse estudo é referente a uma revisão de literatura, utilizando artigos científicos, dissertações e teses, disponibilizadas online e selecionadas através dos principais bancos de periódicos: Scielo, PubMed, Science-Direct e Google Acadêmico. Para tanto, foram utilizados variados trabalhos publicados preferencialmente nos últimos dez anos, bem como aqueles publicados fora desse período que contemplaram a discussão da temática.

Ao final do levantamento bibliográfico, como critério de inclusão foram utilizados os trabalhos que apresentaram relevância a temática abordada e os achados tecnológicos e biotecnológicos fundamentais para tal desenvolvimento.

O método de busca, análise e seleção dos dados ocorreram nos idiomas espanhol, inglês e português, por meio da combinação das palavras-chave, conforme demonstrado na Tabela 1.

Tabela 1: Distribuição do número de trabalhos encontrados por banco de periódicos online.

Termos	Scielo	Pubmed	Sci-D	Google Acad.
Sequenciamento	663	456	161	15.600
DNA	8.750	1.639,57	1.520,34	1.810.00
Biotecnologia	1.968	20.597	13,923	57.100
Genoma	1.073	921	3.689	25.500

Legenda: SciELO = Scientific Electronic Library Online; Pubmed= National Library of Medicine; Sci-D = Science-Direct; Google Acad. = Google Acadêmico.

Resultados e discussão

Métodos de 1ª geração

Método de Maxam–Gilbert e Sanger

O primeiro método de sequenciamento de DNA foi o método Maxam–Gilbert que se baseia na clivagem química da molécula de DNA proposto por Allan Maxam e Walter Gilbert no início dos anos 70. Frederick Sanger em 1977, com base nos estudos do sequenciador Maxam–Gilbert aperfeiçoou o método e popularizou o seu uso.

O processo de sequenciamento pelo método de Sanger funciona da seguinte maneira: é realizado a partir de uma cadeia “simples de DNA” a ser sequenciado, que é obtida por desnaturação da molécula nativa através do calor. Esta fita “simples de DNA” servirá de molde para gerar a outra

metade complementar da dupla hélice. Essa técnica também utiliza marcação radioativa que marca fragmentos de DNA e a síntese desses fragmentos a partir de um DNA molde é possível graças a PCR. Por fim, é necessário ainda a presença de didesoxinucleotídeos (ddATP, ddCTP, ddGTP e ddTTP), que atuam como terminadores da síntese do DNA que poderá ser lido posteriormente através de eletroforese (SANTOS e OLIVEIRA, 2013).

O princípio da técnica consiste em marcar radioativamente alguns dos desoxinucleotídeos livres em solução ou o primeiro desoxinucleotídeo do primer com P^{32} ou S^{35} . Após incorporação na cadeia de DNA nascente, estes átomos marcados emitem radiação que é utilizada para impressão de uma chapa radiográfica, permitindo, desta forma,

DNA molde (MOROZOVA e MARRA, 2008).

Podemos resumir esta técnica como a detecção de fótons que produz uma quantidade

proporcional e igual ao número de nucleotídeos que iram ser incorporados a cadeia de DNA que está sendo sintetizada.

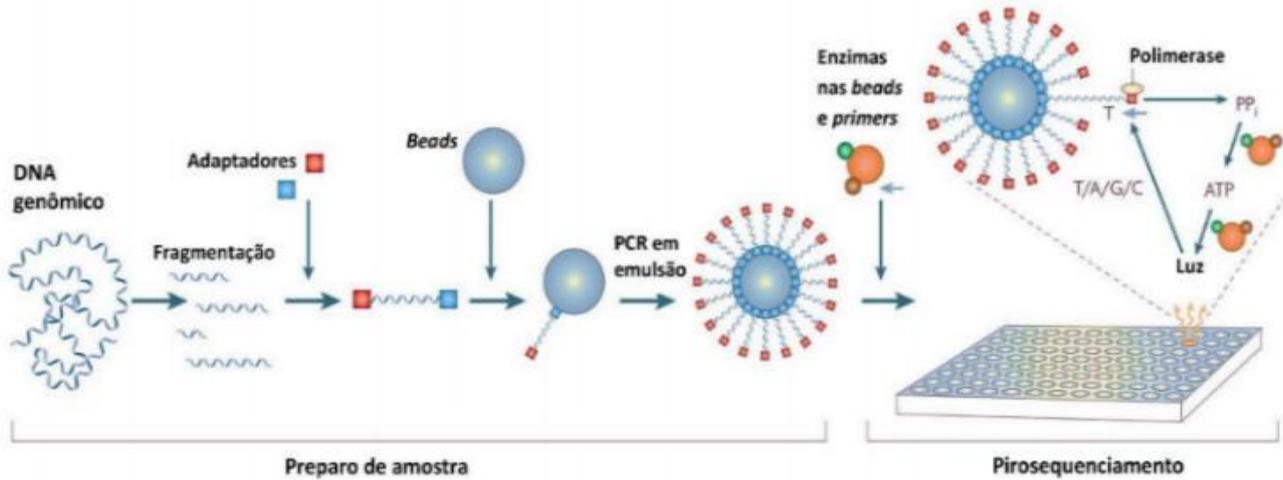


Figura 2: Método de Pirosequenciamento. Macmillan Publishers Ltd, Nature Reviews Microbiology (Medini et al., 2008), copyright 2008.

Métodos de 3° geração

Método Ion Torrent

A Plataforma Ion Torrent PGM™ foi lançada em 2011 pela Life Technologies e Thermo Scientific, ela foi apresentada como um tipo de sequenciador de bancada (*benchtop*), que é direcionado para sequenciamento de pequenos

genomas ou grupos específicos de genes. Esse método foi o primeiro a realizar um sequenciamento totalmente independente de moléculas pré-modificadas como os nucleotídeos marcados com fluorescência, tornando o equipamento e reagentes mais baratos e principalmente acessíveis.

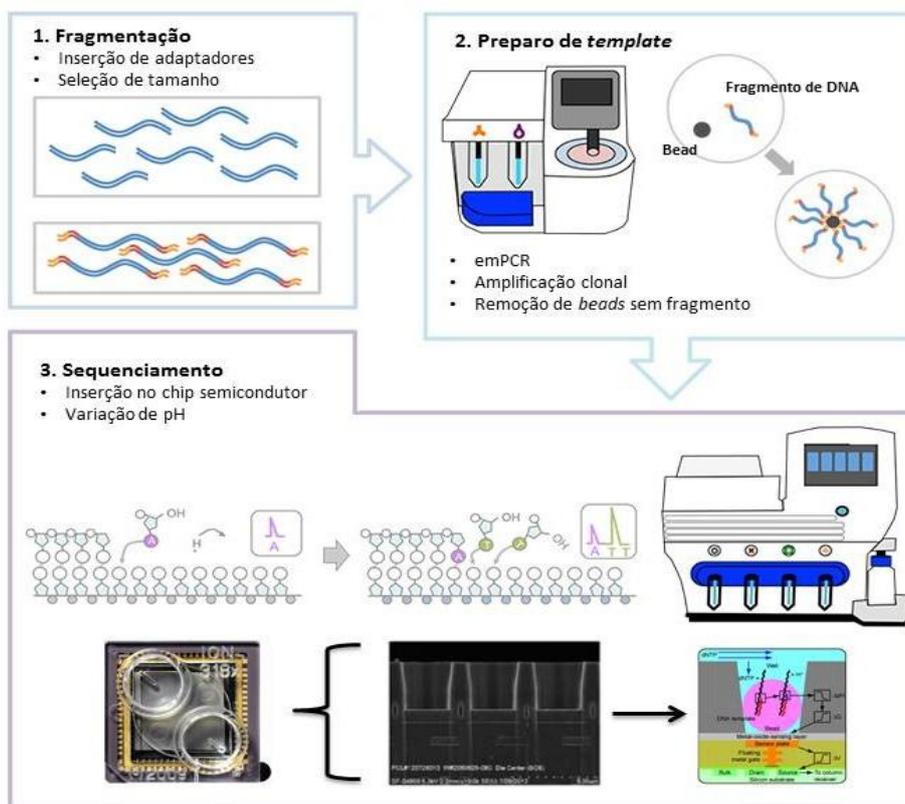


Figura 03: Método Ion Torrent PGM™. 8. BRAGG LM, STONE G, BUTLER MK, HUGENHOLTZ P, TYSON GW, Shining a light on dark sequencing: Characterising errors in Ion Torrent PGM data. Plos Computer Biology. 9(4). 2013.

A Ion Torrent desenvolveu um método que utiliza uma abordagem diferente para o sequenciamento de DNA com base em nucleotídeos dispensando os nucleotídeos fluorescente-etiquetados e explorando o íon de hidrogênio que é liberado sempre que um nucleotídeo se atribui a uma fita de DNA, foram projetados sensores em microchip, para detectar esta assinatura química (HUI, 2012).

O fluxo de bancada completo inclui três etapas: 1) Construção da biblioteca, 2) Preparo de *template* e 3) Sequenciamento em chip de semicondutor. A construção de biblioteca inicia-se com a fragmentação do DNA de interesse, por via enzimática ou mecânica. Em seguida, os fragmentos recebem em suas extremidades dois adaptadores de DNA com sequência conhecida (denominados A e P1), os quais são essenciais para garantir o seu sequenciamento. Caso haja intenção de se processar mais de uma amostra por corrida de sequenciamento, são utilizados juntos a esses adaptadores sequências denominadas *barcodes*, as quais são específicas para cada amostra, permitindo que no momento da análise cada amostra seja devidamente separada. Geralmente nesse passo, também se realiza uma seleção de tamanho de fragmentos para que a maioria dos fragmentos a serem sequenciados esteja entre 200 e 400pb (GODONI, 2018).

Preparados os fragmentos, estes são encaminhados para o preparo de *template*, cuja principal etapa é a PCR em emulsão, a qual consiste na amplificação clonal dos fragmentos dentro de micelas geradas em uma solução emulsionada. Dentro de cada microrreator há um grânulo magnético (*bead*) com sequências de DNA na superfície, no qual um fragmento se acoplará através de seu adaptador P1. A amplificação clonal dos fragmentos ocorre através do pareamento dos sítios de ligação dos primers presentes nos adaptadores e, conseqüentemente, com a extensão de uma nova fita de DNA. Terminado o tempo da PCR, uma fase de enriquecimento separará os *beads* acoplados à fragmentos dos não acoplados, e aqueles selecionados serão inseridos no chip semicondutor e encaminhados para o equipamento, onde acontecerá a reação de sequenciamento (OLIVEIRA, 2015).

Na última etapa, a amostra já devidamente preparada é inserida em um chip semicondutor, o qual é constituído por poços de diâmetro micrométrico nos quais as *beads* com DNA serão depositados para sequenciamento. Cada poço recebe uma *bead*, e embaixo de cada um existe um aparato semicondutor capaz de detectar variações elétricas. Assim, o equipamento passa a fornecer às *beads* os quatro nucleotídeos (A, C, T ou G) soltos, porém sempre um de cada vez, nunca misturados. Caso haja a incorporação de algum nucleotídeo em algum fragmento de DNA, a ligação fosfodiéster gerada liberará íons de hidrogênio no meio, alterando o pH e fazendo com que essa variação

seja detectada pelos semicondutores. Assim, sempre que houver o fornecimento de um nucleotídeo seguido de variação de pH, significa que aquela base foi incorporada, permitindo a geração de sequências ao final do processo (OLIVEIRA, 2015).

O tempo total de sequenciamento pode chegar a 5 horas, entretanto, ao finalizar o processo podem ser lidas mais de um bilhão de pares de bases (>1Gb), fazendo dessa plataforma uma técnica extremamente eficaz no que diz respeito ao consumo de tempo na geração de dados brutos (RIBEIRO, 2020).

Métodos de 4ª geração

Método Oxford Nanopore

A Oxford Nanopore Technologies apresentou o seu próprio método inteligente para o sequenciamento onde é incorporado um poro de proteína em uma membrana de polímero sintético dentro do microchip. A cadeia de DNA a ser sequenciada passa através do poro e, quando cada nucleotídeo atravessa, é registrado como um sinal elétrico diferente detectado pela máquina Oxford (SHANG, 2012).

O princípio básico deste método de sequenciamento da Oxford Nanopore não requer nenhum tipo de marcação fluorescente, o que reduz os custos e aumenta sua velocidade. Em desenvolvimento por mais de 20 anos, esta metodologia, inicialmente baseava-se principalmente em um poro de α -hemolisina modificado, isolado de estafilococos, pelo qual uma molécula de DNA fita simples (ssDNA) passava sob uma diferença de potencial e a corrente iônica passando pelo poro era registrada. Cada base nucleotídica que passava pelo poro era registrada em sequência, de acordo com o decréscimo na amplitude da corrente iônica no poro (COSTA, 2020).

Inicialmente, a velocidade com que as moléculas de DNA passavam pelo poro era um fator limitante para o sequenciamento, pois, esta velocidade era muito alta (na ordem de microssegundos por base) para permitir que a alteração na corrente iônica fosse detectada pelo poro. Assim, em 2009, uma molécula adaptadora (ciclodextrina) foi covalentemente associada à região de barril da proteína, próximo ao centro do poro, de forma a aperfeiçoar a discriminação das bases nucleotídicas em taxas de aquisição de dados mais elevadas. Esta diferenciação de bases é mantida sob condições operantes, mesmo quando uma exonuclease próxima ao poro é utilizada. A exonuclease I, utilizada em associação ao complexo do nanoporo, permite uma redução da velocidade em que o DNA é translocado pelo poro, melhorando a qualidade da captura de sinal e o nanoporo por sua vez é capaz de detectar de forma contínua os nucleosídeos monofosfato (dNMPs), com auxílio da molécula de ciclodextrina que se liga a mononucleotídeos, fazendo com que eles

permaneçam no poro por aproximadamente 10 ms. (SBG, 2017).

Este nanoporo é capaz de discriminar os quatro dNMPs de forma acurada, com

aproximadamente 99% de confiança, em condições ótimas, sendo capaz de reconhecer inclusive dCMPs metilados. (SBG, 2017).

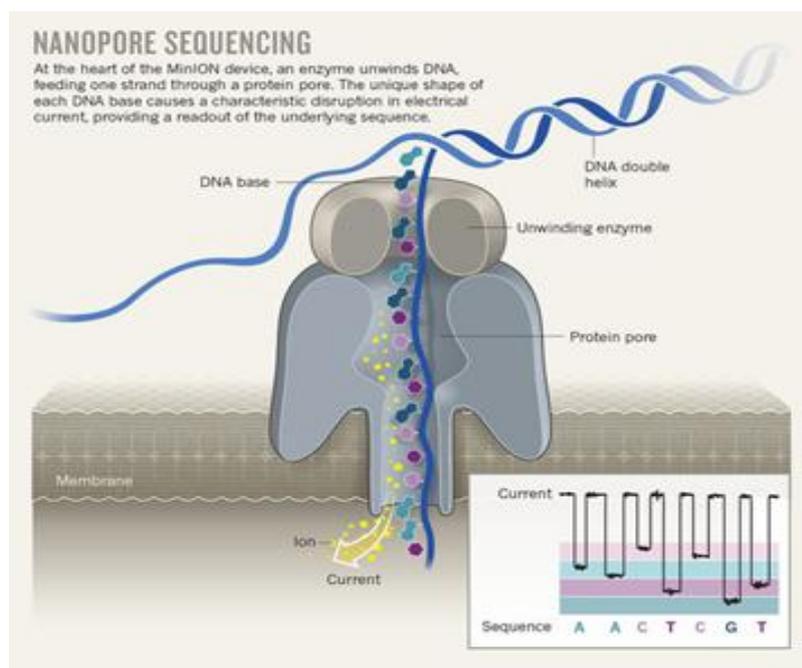


Figura 4: sequenciamento pelo método Oxford Nanopore Technologies. Eisenstein, M. Um ás no buraco para sequenciamento de DNA. *Nature* 550, p. 285–288 (2017).

Conclusão

O conhecimento sobre o sequenciamento de DNA pode ser útil em praticamente qualquer área da biologia, medicina e biotecnologia. Com ela, é possível descrever estudos evolutivos e filogenéticos, buscar doenças genéticas, realizar clonagem gênica e melhorar a reprodução dos seres vivos.

A evolução das técnicas para se sequenciar os genomas cresceu juntamente com o avanço da tecnologia, facilitou e aperfeiçoou todo o processo que antes era caro e demorado. Do método de Maxam–Gilbert ao Nanopore é possível observar quão aperfeiçoado o método de sequenciamento se tornou, e o que antes era demorado, hoje em poucos minutos milhares de pares de bases são sequenciados.

Referências

KREMER, F.S.; PINTO, L.S. Plataformas de Sequenciamento de Nova Geração & Pré-processamento de dados. Laboratório de Bioinformática e Proteômica, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas. 2016. Disponível em: <http://labbioinfo.ufpel.edu.br/aulas_2016/Capitulo%203.pdf>

SANTOS. W.F.; OLIVEIRA. M.S et al. Sequenciamento de DNA: Métodos e Aplicações. XIII Safety, Health and Environment World Congress. 2013. Disponível em: <<https://copec.eu/congresses/shewc2013/proc/works/33.pdf>>

MOREIRA, L.M. Ciências Genômicas: Fundamentos e Aplicações. Sociedade Brasileira de Genética, Ribeirão Preto, 2015. ISBN 978-85-89265-22-5

CARVALHO, M.C.C.G.; Silva, D.C.G. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. *Revista de Ciência Rural*. Santa Maria. 2010.

Disponível em: <<https://repositorio.unesp.br/bitstream/handle/11449/29442/S0103-84782010000300040.pdf?sequence=1&isAllowed=y>>

MOROZOVA, O.; MARRA, M. A. Applications of next-generation sequencing technologies in functional genomics. *Genomics*. 2008. DOI: 10.1016/j.ygeno.2008.07.001

HUI, P. Next Generation Sequencing: Chemistry, Technology and Applications. Springer-Verlag. Berlin Heidelberg. 2012. DOI: 10.1007/128_2012_329.

GODONE, R.L.N. Identificação de marcadores moleculares para diagnóstico, predição e prognóstico de câncer de mama. Universidade Federal do Pernambuco. 2018. Disponível em: <<https://repositorio.ufpe.br/handle/123456789/32876>>.

OLIVEIRA, T.G.M. Aplicação de sequenciamento de nova geração no diagnóstico molecular de cardiomiopatia hipertrófica. Faculdade de Medicina da Universidade de São Paulo. São Paulo. 2015. DOI: 10.11606/D.5.2015.tde-29102015-115502.

RIBEIRO. I.H. Identificação das variants dos genes RHD e RHCE em pacientes com doença falciforme usando estratégia de sequenciamento de nova geração. Faculdade de Medicina da Universidade

de São Paulo. São Paulo. 2020. DOI: <https://doi.org/10.11606/D.5.2020.tde-23082021-103036>.

SHANG, F.; GUIHEN, E.; GLENNON, J.D. Recent advances in miniaturization. The role of microchip electrophoresis in clinical analysis. Electrophoresis. 2012. DOI: <https://doi.org/10.1002/elps.201300359>.

COSTA. C.N. O uso de ferramenta de data analytics para o desenvolvimento de metodologia visando o mapeamento genético de doenças raras – Estudo de caso no rastreamento do gene causador da cistinose. Universidade Federal Fluminense – UFF. Rio de Janeiro. 2020. Disponível em: <https://app.uff.br/riuff/handle/1/16381>.

TURCHETTO-ZOLET. A.C.; TURCHETTO. C.; ZANELLA, C.M. Passaia, G. Marcadores Moleculares na Era genômica: Metodologias e Aplicações. Sociedade Brasileira de Genética. Ribeirão Preto, 2017. Disponível em: <<https://www.lume.ufrgs.br/bitstream/handle/10183/206113/001056154.pdf?sequence=1>>