

Características genómicas dos principais microrganismos produtores de DNA polimerase: Aplicação para PCR

Genomic characteristics of major DNA polymerase producing microorganisms: PCR application

Darling de Andrade Lourenço

Universidade Federal de Pelotas, Brasil

darlinglourenco@gmail.com

Altino Branco Choupina

Instituto Politécnico de Bragança, Portugal

albracho@ipb.pt

Resumo

O conjunto de enzimas denominadas DNA polimerases são responsáveis pela síntese e reparação de ácido desoxirribonucleico (DNA). Essa característica é muito útil para ensaios em biologia molecular nos quais há a necessidade de amplificar fragmentos de DNA, como na Reação em Cadeia da Polimerase (PCR). Na PCR, um fragmento de DNA delimitado por *primers* específicos é amplificado pela ação de DNA polimerase num processo com vários ciclos. Nos primeiros anos, a realização da técnica necessitava de grandes volumes dessa enzima, uma vez que a mesma não resistia as altas temperaturas. Após a descoberta das DNA polimerases em microrganismos termófilos e hipertermófilos, ou seja, resistentes à alta temperatura, a técnica de PCR foi refinada até chegar ao que é conhecido atualmente. Nesse sentido, o objetivo deste trabalho foi realizar pesquisas nas bases de dados de literatura científica PubMed e Science Direct e identificar os principais microrganismos produtores das DNA polimerases comercializadas atualmente. Foram escolhidos os seguintes microrganismos: *Thermus aquaticus*, *Pyrococcus furiosus*, *Thermococcus litoralis* e *Thermotoga maritima*. A partir da utilização do banco de dados de genoma do NCBI, as características de genómica estrutural e funcional de cada um são apresentadas, bem como a genómica comparativa entre os quatro microrganismos. As sequências de aminoácidos das DNA polimerases destes microrganismos foram alinhadas através da ferramenta ClustalW para caracterização filogenética. Apesar das diferenças evolutivas, as DNA polimerases apresentam a mesma atividade base em todos os organismos, variando apenas em características de acordo com a família à qual pertencem. No entanto, a DNA polimerase de *Thermus aquaticus*, Taq polimerase, é a mais utilizada devido ao seu custo/benefício ser melhor em comparação com as outras.

Palavras-chave: *Genómica estrutural; Genómica funcional; Genómica comparativa; Termófilos.*

Abstract

The group of enzymes called DNA polymerases are responsible for the synthesis and repairment of deoxyribonucleic acid (DNA). This feature is useful for molecular biology assays in which there is a need to amplify DNA fragments, such as in the Polymerase Chain Reaction (PCR). In the PCR, a DNA fragment delimited by specific primers is amplified by the action of DNA polymerase in a process with several cycles. In the first years, to accomplish the amplification, a great amount of this enzyme was demanded since the ones used did not resist to the high temperatures. After the discovery of DNA polymerases in thermophilic and hyperthermophilic microorganisms, i. e., resistant to high temperature, the PCR technique was refined to how is known today. In line with this, this work aimed to identify, through research on PubMed and Science Direct databases, the major microorganisms that produce the

DNA polymerases currently commercialized. The following microorganisms were chosen: *Thermus aquaticus*, *Pyrococcus furiosus*, *Thermococcus litoralis* and *Thermotoga maritima*. Using the genome database from NCBI, the genomics structural and functional characteristics of each one is presented, besides, the comparative genomics between the four microorganisms also is presented. The amino acid sequences of the DNA polymerases from these microorganisms were aligned with the ClustalW tool for phylogenetic characterization. Despite the evolutionary differences, DNA polymerases have the same essential activity in all organisms, diversifying only in characteristics according to the family in which these enzymes belong. Nevertheless, DNA polymerase from *Thermus aquaticus*, Taq polymerase, is still the most widely used due to its cost/benefit ratio that is better when compared to the others.

Keywords: *Structural genomic; Functional genomic; Comparative genomic; Thermophilic.*

INTRODUÇÃO

O grupo de enzimas denominado DNA polimerases é expresso constitutivamente nas células e é responsável pela síntese de fitas complementares de ácido desoxirribonucleico (DNA), de acordo com a fita molde. A função desse processo é fundamental para manter a integridade do DNA durante o processo de replicação, além de também realizar a reparação a danos. Com a descoberta da DNA polimerase I de *Escherichia coli*, em 1958, obteve-se a primeira evidência de atividade enzimática capaz de sintetizar DNA (Lehman *et al.*, 1958). As DNA polimerases são classificadas em cinco famílias que embora relacionadas evolutivamente, são divergentes de acordo com suas sequências de aminoácidos. Apesar das diferenças nas sequências, a função das enzimas é a mesma nas diferentes famílias (Braithwaite & Ito, 1993; Garcia-Diaz & Bebenek, 2007; Ito & Braithwaite, 1991). Desde a década de 1980, as DNA polimerases são muito utilizadas para a manipulação de DNA *in vitro*, principalmente na técnica de amplificação PCR (*Polymerase Chain Reaction* - Reação em Cadeia da Polimerase) (Ishino & Ishino, 2014).

O PCR é uma técnica de amplificação de segmentos específicos de DNA (amplicons) a partir de um pequeno volume de material (aproximadamente 1 µL) de partida (DNA molde ou sequência alvo). Foi descrito e utilizado experimentalmente por Kary Mullis em 1985, que na época, utilizava DNA polimerase de *E. coli*, que não é estável nas temperaturas elevadas para o procedimento. O refinamento da técnica deu-se quando as DNA polimerases de termófilos foram inseridas no procedimento, uma vez que as enzimas provindas dessas bactérias são termoestáveis. A primeira delas foi a DNA polimerase da bactéria *Thermus aquaticus*, denominada Taq DNA polimerase (Lorenz, 2012).

As DNA polimerases de organismos termofílicos são estáveis em altas temperaturas e, por isso, tornaram-se populares como uma fonte de DNA polimerase e outras enzimas para aplicação laboratorial. Esses organismos são classificados como termófilos extremos, quando crescem em

temperaturas acima de 75°C e, em termófilos moderados, quando crescem na faixa de temperatura de 55-75°C. Também há os hipertermófilos, que crescem em temperaturas acima de 80°C. Muitos dos representantes dessa classificação pertencem ao domínio Archaea (Ishino & Ishino, 2014).

Inicialmente, a Taq polimerase era purificada de *T. aquaticus*, entretanto, o gene desta DNA polimerase foi clonado em *E. coli* e a versão recombinante foi lançada no mercado sob o nome de AmpliTaq DNA polimerase. A manipulação da enzima resultou numa DNA polimerase recombinante de melhor qualidade, quando comparada com a Taq polimerase de *T. aquaticus* (Ishino & Ishino, 2014; Lawyer *et al.*, 1989). A DNA polimerase de *Thermotoga maritima* (ULTIMA DNA polimerase) foi a primeira enzima provinda de bactérias hipertermófilas. Apresenta atividade exonuclease 3'-5', no entanto, não apresenta diferença de fidelidade quando comparada com a Taq DNA polimerases (Diaz & Sabino, 1998). Outras DNA polimerases utilizadas na PCR são originárias de Archaeas hipertermófilas, como a Pfu polimerase de *Pyrococcus furiosus*, que é mais estável do que Taq polimerase (Ishino & Ishino, 2014). Outro exemplo é Vent DNA polimerase de *Thermococcus litoralis*, que é extremamente termoestável, podendo operar na faixa de 95-100°C (Kong, Kucera, & Jack, 1993).

Considerando a temática abordada neste trabalho, o objetivo do mesmo foi demonstrar e discutir as características genômicas, estruturais e funcionais dos quatro principais microrganismos produtores de DNA polimerase para aplicações no PCR: *Thermus aquaticus*, *Thermotoga maritima*, *Pyrococcus furiosus* e *Thermococcus litoralis*. Além disso, também foi estabelecida a genômica comparativa desses microrganismos.

DESENVOLVIMENTO

Para o desenvolvimento deste trabalho, foram realizadas pesquisas na literatura científica da área utilizando as plataformas PubMed do *National Center for Biotechnology Information* (NCBI) (www.ncbi.nlm.nih.gov/pubmed) e Science Direct da Elsevier (www.sciencedirect.com). Além disso, foram realizadas consultas nas ferramentas de bancos de dados disponíveis NCBI e *European Molecular Biology Laboratory* (EMBL) nos dias 04, 05 e 06 de Novembro de 2018. Foram ainda utilizadas ferramentas de alinhamento, como o BLASTn e tBLASTn. e a ferramenta de alinhamento múltiplo ClustalW para a construção de árvores filogenéticas para realizar genômica comparativa.

RESULTADOS E DISCUSSÃO

Thermus aquaticus

O microrganismo *Thermus aquaticus* (Figura 1) pertence ao domínio Bacteria e ao gênero *Thermus*. Foi isolada pela primeira vez em 1969 no Lower Geyser Basin do Parque Nacional de Yellowstone pelo pesquisador Thomas Brock (Brock & Freeze, 1969). É um bacilo Gram-negativo que apresenta diversas morfologias, anaeróbico facultativo e termófilo, com temperatura ótima de crescimento entre 65°C e 70°C, mas também é capaz de crescer entre 50°C e 80°C (Brumm *et al.*, 2015).

No banco de dados do NCBI existem dois projetos de sequenciamento de *T. aquaticus*, um para a cepa YT1 e outro para a cepa Y51MC23. Para este trabalho, a cepa Y51MC3 foi escolhida, visto que o artigo do sequenciamento da mesma encontra-se publicado nas databases. O sequenciamento desta cepa deu-se através da parceria entre as empresas C5-6 Technologies LLC e Lucigen Corporation, ambas dos Estados Unidos da América. O trabalho foi publicado em 2015 e descreve o sequenciamento do genoma e dos quatro plasmídeos de *T. aquaticus* (Brumm *et al.*, 2015).

O sequenciamento fez-se pela estratégia de *Whole Genome Shotgun* (WGS), utilizando um sequenciador de segunda geração MiSeq (Illumina, San Diego, CA) com uma cobertura de 60 x e a estratégia de montagem *de novo*. Esse sequenciamento foi capaz de validar o genoma rascunho de *T. aquaticus* Y51MC23 que havia sido depositado em 2008 pela *Joint Genome Institute* (JGI) (Brumm *et al.*, 2015).

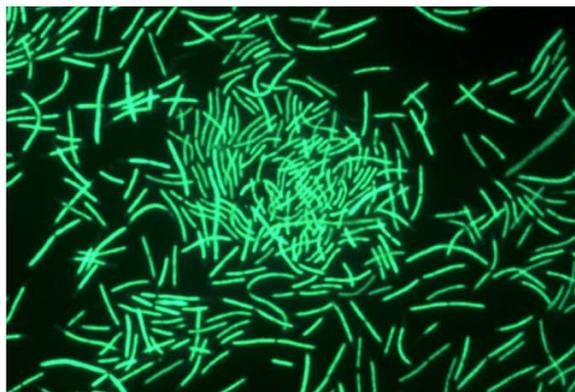


Figura 1 - *Thermus aquaticus*. Fonte: Brumm *et al.*, 2015.

Genômica Estrutural

Os dados da genômica estrutural de *T. aquaticus*, *P. furiosus*, *T. litoralis* e *T. maritima* podem ser observados na Tabela 1. Os dados encontrados referem-se ao número total de pares de bases (pb) do DNA, genes, pseudogenes, proteínas, conteúdo guanina-citosina (GC), quantidade de

genes para RNA transportador (tRNA) e RNA ribossômico (rRNA), além da quantidade de outros RNAs.

Tabela 1 – Dados estatísticos dos genomas dos microrganismos avaliados. Fonte: NCBI em 06 novembro de 2018.

Características	<i>T. aquaticus</i>	<i>P. furiosus</i>	<i>T. litoralis</i>	<i>T. maritima</i>
DNA (pb)	2.33864	190.983	222.517	186.961
Genes	2.529	2.133	2.453	1.949
Pseudogenes	61	65	96	22
Proteínas	2394	2.001	2.306	1.875
Conteúdo CG	68%	40.8%	43.1%	46.2
rRNA	6	4	4	3
tRNA	59	46	45	46
Outros RNAs	4	17	2	3

Genômica Funcional

O gene responsável DNA polimerase I encontra-se dentre os 2.529 genes que compõem o cromossoma de *T. Aquaticus*. Esse gene é denominado *polA* e a sequência começa na posição 618862 do cromossoma estendendo-se até à posição 621360, produzindo uma proteína com 832 aminoácidos. A função dessa proteína, como já citado, é sintetizar DNA a partir de desoxirribonucleotídeos trifosfatos na presença de um DNA molde e um grupo 3'OH livre, atuando nos processos biológicos de replicação e reparação de DNA.

Pyrococcus furiosus

Os microrganismos do género *Pyrococcus* spp. são hipertermófilos e obrigatoriamente anaeróbicos. Eles crescem em temperatura ótima de 100°C (Leigh, Albers, Atomi, & Allers, 2011). A *P. furiosus* (Figura 2) é o representante melhor caracterizado e foi o primeiro a ser isolado por Karl Stetter e colaboradores (Fiala & Stetter, 1986). Esse microrganismo pertence ao domínio Archaea e a família Thermococcaceae.

Está disponível comercialmente a Pfu DNA polimerase, provinda de *Pyrococcus furiosus* sendo atualmente, expressa de forma recombinante em *E. coli*. Na database do NCBI há dois projetos de genoma rascunho de *P. furiosus*, a escolhida para ser descrita neste trabalho foi a cepa COM1, que possui trabalho publicado. O sequenciamento foi feito pelo método WGS, utilizando um

sequenciador de 2ª geração da Illumina. O método de montagem adotado foi o *de novo* (Bridger *et al.*, 2012).



Figura 2 – *Pyrococcus furiosus*. Fonte: Kengen, 2017.

Genômica Estrutural

Os dados do genoma de *P. furiosus* encontram-se da Tabela 1. Os dados encontrados referem-se ao número total de pares de bases (pb) do DNA, número de genes e pseudogenes, número de proteínas codificadas, conteúdo GC, número de genes para rRNA e tRNA, além de genes para outros RNAs.

Genômica Funcional

A DNA polimerase de *P. furiosus* é transcrita a partir do gene *pol*, localizado entre os pares de base 1097163 e 1097753 do cromossoma, dando origem a uma proteína de 196 aminoácidos.

Thermococcus litoralis

Thermococcus litoralis é um microrganismo que pertence ao domínio Archaea e à família Thermococcaceae. Foi isolado pela primeira vez de uma fossa termal submarina na praia de Lucrino, na Itália (Belkin & Jannasch, 1985). Sua DNA polimerase comercial é denominada VENT e também é expressa na forma recombinante em *E. coli*. No NCBI há a indicação de dois projetos de sequenciamento de *Thermococcus litoralis*, o escolhido foi o da cepa DSM 5473, uma vez que o trabalho se encontra publicado.

Genômica Estrutural

As características do genoma de *T. litoralis* encontram-se na Tabela 1. Os dados encontrados referem-se ao número total de pares de bases (pb) do DNA, número de genes e pseudogenes, número de proteínas codificadas, conteúdo GC, número de genes para rRNA e tRNA, além de genes para outros RNAs.

Genômica Funcional

A DNA polimerase de *T. litoralis* é transcrita a partir do gene *pol*, localizado entre os pares de base 298517 e 299266 do seu cromossoma, dando origem a uma proteína de 249 aminoácidos.

Thermotoga marítima

O microorganismo *T. marítima* pertence ao domínio Bacteria, ao filo Thermatogae, a Ordem Thermotogales e a família Thermotogaceae. Sua DNA polimerase comercial para PCR é denominada ULTMA e também é obtida através de DNA recombinante e expressa em *E. coli*. Foi descoberta em 1986 em uma área geotérmica marinha perto de Volcano, na Itália (Diaz & Sabino, 1998).

Genômica Estrutural

As características estruturais do genoma de *T. marítima* encontram-se na Tabela 1. Os dados encontrados referem-se ao número total de pares de bases (pb) do DNA, número de genes e pseudogenes, número de proteínas codificadas, conteúdo GC, número de genes para rRNA e tRNA, além de genes para outros RNAs.

Genômica Funcional

A DNA polimerase de *T. marítima* é codificada pelo gene *pol*, entre os genes 644001 e 644351 do cromossoma, dando origem a uma proteína de 116 aminoácidos.

Genômica Comparativa

Para realizar a genômica comparativa entre as DNA polimerases de todos os microrganismos tratados neste trabalho, foi construída uma árvore filogenética (Figura 3) utilizando o programa de alinhamento múltiplos ClustalW (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).



Figura 3 – Árvore filogenética montada a partir das sequências de DNA polimerases de *T. aquaticus*, *P. furiosus*, *T. litoralis* e *T. maritima*.

Como pode ser observado na Figura 3, a *Thermococcus litoralis* e a *Thermotoga maritima* são os microrganismos mais próximos evolutivamente, uma vez que compartilham o mesmo ancestral comum mais recente. O resultado obtido mostra que apesar de *T. litoralis* pertencer ao domínio Archaea e de *T. maritima* pertencer ao domínio Bacteria, elas apresentam o ancestral comum mais recente. Esse dado é reforçado por Doolittle (2000), que discute que as bactérias termófilas e Archaeas termófilas estão próximas evolutivamente e apresentam ancestrais em comum mais recentes. Por outro lado, *P. furiosus* e *T. aquaticus*, pertencentes aos domínios Bacteria e Archaea, respectivamente, não apresentam relações evolutivas entre si nem com *T. litoralis* e *T. maritima*.

Tabela 2 – Tamanho da cadeia polipeptídica da DNA polimerase nos diferentes microrganismos.

Microrganismo	Cadeia polipeptídica DNA polimerase
<i>Thermus aquaticus</i>	832 a.a.
<i>Pyrococcus furiosus</i>	196 a.a.
<i>Thermococcus litoralis</i>	249 a.a.
<i>Thermotoga maritima</i>	116 a.a.

As enzimas DNA polimerase dos organismos abordados neste trabalho, embora possuam relação evolutivamente, são divergentes no seu tamanho (Tabela 2). A partir da construção da árvore filogenética utilizando as sequências de aminoácidos de cada enzima, é possível analisar quais os organismos que possuem maior parentesco. *Thermococcus litoralis* e *Thermotoga maritima* são mais próximas evolutivamente, enquanto *Pyrococcus furiosus* e *Thermus aquaticus* são filogeneticamente mais distantes. Nesse sentido, é possível perceber que apesar de exercerem praticamente a mesma função, essas enzimas não são próximas filogeneticamente. Entretanto, vale a pena destacar, que os dados do genoma não são fixos, estão sempre sendo alterados de acordo com o que as novas tecnologias oferecem. A Taq DNA polimerase ainda é a enzima mais utilizada comercialmente, devido às suas boas características, entretanto, as outras enzimas disponíveis no mercado também são muito utilizadas em experiências com requisitos específicos, por apresentarem tempos de meia vida em reação consideráveis, como se pode observar na Tabela 3.

Tabela 3 – Dados de tempos de meia vida das DNA polimerases abordadas no trabalho.
Fonte: BioLabs (2018).

Enzimas	Meia vida (h)	Temperatura (°C)
Taq	0,9	95
Pfu	19	95
Tma	2	94
Vent	6,7	95

CONCLUSÃO

Este trabalho apresenta os microrganismos responsáveis pela produção das principais DNA polimerases mais utilizadas atualmente nas pesquisas laboratoriais de investigação e análise biomédica e clínica. Cada uma possui qualidades e aplicações bem definidas, no entanto, a TaqDNA polimerase proveniente de *Thermus aquaticus* ainda é enzima mais utilizada, visto que apresenta ótimas características e uma boa relação preço qualidade por ser a primeira a ter sido utilizada. Conhecer as características genômicas de cada microrganismo é fundamental para que a tecnologia possa melhorar essas enzimas, oferecendo assim, melhores produtos biotecnológicos para as pesquisas.

De acordo com os resultados encontrados, é possível concluir que apesar das diferenças evolutivas, as DNA polimerases apresentam a mesma atividade base em todos os organismos, variando apenas em características de acordo com a família a qual pertencem. Permitindo assim que essas enzimas possam ser utilizadas na amplificação de fragmentos genômicos tanto de procariontes como de eucariontes. As perspectivas futuras apontam que novas cepas desses microrganismos possam prover enzimas com maior estabilidade e maior eficiência para técnicas moleculares, como PCR. Além disso, a ferramenta de DNA recombinante também permite que essas enzimas possam ser melhoradas e expressas em sistemas de mais fácil cultivo e menor custo.

Referências

- Belkin, S., & Jannasch, H. W. (1985). A new extremely thermophilic, sulfur-reducing heterotrophic, marine bacterium. *Archives of Microbiology*, *141*(3), 181–186. <https://doi.org/10.1007/BF00408055>
- BioLabs, N. E. (2018). DNA Polymerase Selection Chart.
- Braithwaite, D. K., & Ito, J. (1993). Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Research*, *21*(4), 787–802. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8451181>
- Bridger, S. L., Lancaster, W. A., Poole, F. L., Schut, G. J., & Adams, M. W. W. (2012). Genome sequencing of a genetically tractable *Pyrococcus furiosus* strain reveals a highly dynamic genome. *Journal of Bacteriology*, *194*(15), 4097–4106. <https://doi.org/10.1128/JB.00439-12>
- Brock, T. D., & Freeze, H. (1969). *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *Journal of Bacteriology*, *98*(1), 289–297. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/5781580>
- Brumm, P. J., Monsma, S., Keough, B., Jasinovica, S., Ferguson, E., Schoenfeld, T., ... Mead, D. A. (2015). Complete Genome Sequence of *Thermus aquaticus* Y51MC23. *PLOS ONE*, *10*(10), e0138674.

- <https://doi.org/10.1371/journal.pone.0138674>
- Diaz, R. S., & Sabino, E. C. (1998). Accuracy of replication in the polymerase chain reaction. Comparison between *Thermotoga maritima* DNA polymerase and *Thermus aquaticus* DNA polymerase. *Brazilian Journal of Medical and Biological Research*, 31(10), 1239–1242. <https://doi.org/10.1590/S0100-879X1998001000001>
- Doolittle, W. F. (2000). Uprooting the Tree of Life. *Scientific American*, 282(6), 90–95.
- Fiala, G., & Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Archives of Microbiology*, 145(1), 56–61. <https://doi.org/10.1007/BF00413027>
- Garcia-Diaz, M., & Bebenek, K. (2007). Multiple Functions of DNA Polymerases. *Critical Reviews in Plant Sciences*, 26(2), 105–122. <https://doi.org/10.1080/07352680701252817>
- Ishino, S., & Ishino, Y. (2014). DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Frontiers in Microbiology*, 5, 465. <https://doi.org/10.3389/fmicb.2014.00465>
- Ito, J., & Braithwaite, D. K. (1991). Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Research*, 19(15), 4045–4057. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1870963>
- Kong, H., Kucera, R. B., & Jack, W. E. (1993). Characterization of a DNA polymerase from the hyperthermophile archaea *Thermococcus litoralis*. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. *The Journal of Biological Chemistry*, 268(3), 1965–1975. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8420970>
- Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R., & Gelfand, D. H. (1989). Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *The Journal of Biological Chemistry*, 264(11), 6427–6437. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2649500>
- LEHMAN, I. R., BESSMAN, M. J., SIMMS, E. S., & KORNBERG, A. (1958). Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *The Journal of Biological Chemistry*, 233(1), 163–170. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/13563462>
- Leigh, J. A., Albers, S.-V., Atomi, H., & Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. *FEMS Microbiology Reviews*, 35(4), 577–608. <https://doi.org/10.1111/j.1574-6976.2011.00265.x>
- Lorenz, T. C. (2012). Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *Journal of Visualized Experiments*, (63), e3998. <https://doi.org/10.3791/3998>