**UNIVERSIDADE FEDERAL DE MINAS GERAIS**
**Departamento de Física**
**Programa Interunidades de Pós-Graduação em Bioinformática**

Pâmella Miranda de Moura

# THERMODYNAMIC ASSESSMENT OF DNA MISMATCHES IN PCR AND LAMP PRIMERS FOR SARS-COV-2 DETECTION AND THEIR EFFECTIVENESS TO VARIANTS

Belo Horizonte
2024

Pâmella Miranda de Moura

# THERMODYNAMIC ASSESSMENT OF DNA MISMATCHES IN PCR AND LAMP PRIMERS FOR SARS-COV-2 DETECTION AND THEIR EFFECTIVENESS TO VARIANTS

Doctoral thesis presented to the Programa Interunidades de Pós-Graduação em Bioinformática at the Universidade Federal de Minas Gerais as requirement for a doctoral degree in Bioinformatics.

Supervisor: Prof. Dr. Gerald Weber
Co-supervisor: Prof. Dr. Kira Asthakova

Belo Horizonte
2024

043 Moura, Pâmella Miranda de.
Thermodynamic assessment of DNA mismatches in PCR and LAMP primers for SARS-CoV-2 detection and their effectiveness to variants [manuscrito] / Pâmella Miranda de Moura. – 2024.
154 f. : il. ; 29,5 cm.

Orientador: Gerald Weber. Coorientador: Kira Astakhova.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Betacoronavirus. 3. Primers do DNA. 4. Reação em Cadeia da Polimerase. 5. Técnicas de Amplificação de Ácido Nucleico. I. Weber, Gerald. II. Astakhova, Kira. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004

Ficha catalográfica elaborada pela bibliotecária Fabiane C M Reis – CRB 6 - 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

**ATA DE DEFESA DE TESE**

***PÂMELLA MIRANDA DE MOURA***

Às nove horas do dia **01 de março de 2024**, reuniu-se, através de videoconferência, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"Thermodynamic assessment of DNA mismatches in PCR and LAMP primers for SARS-CoV-2 detection and their effectiveness to variants"**, requisito para obtenção do grau de Doutora em **Bioinformática.** Abrindo a sessão, o Presidente da Comissão, **Dr. Gerald Weber**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

| **Professor(a)/Pesquisador** | **Instituição** | **Indicação** |
|---|---|---|
| Dr. Gerald Weber - Orientador | Universidade Federal de Minas Gerais | **Aprovada** |
| Dr. Daniel Carlos Ferreira Lanza | Universidade Federal do Rio Grande do Norte | **Aprovada** |
| Dr. Lucas Bleicher | Universidade Federal de Minas Gerais | **Aprovada** |
| Dr. Rodrigo Giglioti | Instituto de Zootecnia do Governo do Estado de São Paulo | **Aprovada** |
| Dr. Ubirajara Agero Batista | Universidade Federal de Minas Gerais | **Aprovada** |

Pelas indicações, a candidata foi considerada: **Aprovada**

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

**Belo Horizonte, 01 de março de 2024.**

Documento assinado eletronicamente por **Gerald Weber**, **Professor do Magistério Superior**, em 01/03/2024, às 14:07, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ubirajara Agero Batista**, **Servidor(a)**, em 01/03/2024, às 16:10, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Lucas Bleicher**, **Professor do Magistério Superior**, em 04/03/2024, às 16:45, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Daniel Carlos Ferreira Lanza**, **Usuário Externo**, em 08/03/2024, às 13:50, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rodrigo Giglioti**, **Usuário Externo**, em 12/03/2024, às 17:53, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3069977** e o código CRC **578BEB79**.

**Referência:** Processo nº 23072.212070/2024-46 SEI nº 3069977

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

**FOLHA DE APROVAÇÃO**

**Pâmella Miranda de Moura**

**"Thermodynamic assessment of DNA mismatches in PCR and LAMP primers for SARS-CoV-2 detection and their effectiveness to variants"**

Tese aprovada pela banca examinadora constituída pelos Professores:

Prof. Gerald Weber - Orientador

Universidade Federal de Minas Gerais

Prof. Daniel Carlos Ferreira Lanza

Universidade Federal do Rio Grande do Norte

Prof. Lucas Bleicher

Universidade Federal de Minas Gerais

Prof. Rodrigo Giglioti

Instituto de Zootecnia do Governo do Estado de São Paulo

Prof. Ubirajara Agero Batista

Universidade Federal de Minas Gerais

Belo Horizonte, 01 de março de 2024.

Documento assinado eletronicamente por **Gerald Weber**, **Professor do Magistério Superior**, em 01/03/2024, às 14:08, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ubirajara Agero Batista**, **Servidor(a)**, em 01/03/2024, às 16:10, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Lucas Bleicher**, **Professor do Magistério Superior**, em 04/03/2024, às 16:45, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Daniel Carlos Ferreira Lanza**, **Usuário Externo**, em 08/03/2024, às 13:49, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rodrigo Giglioti**, **Usuário Externo**, em 12/03/2024, às 17:54, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3070145** e o código CRC **94C9B771**.

**Referência:** Processo nº 23072.212070/2024-46 SEI nº 3070145

# Acknowledgements

I would like to thank a few people who were very important to the conclusion of this work.

First, I would like to thank Gerald for all support, guidance, friendship, patience and for believing in me. When you believed in me many years ago I do not think you imagined all the adventures we would have. Well, we are here now and it was worth it!

To Kira, thank you for your support, guidance and our talks. I learned a lot in different ways.

To the Grupo de Biofísica Computacional, thank you for everything. Each one who passed by me left something special. Especially Vivianne Basílio and Gabriel Schiess for giving me the opportunity to have an experience from a different point of view.

Gladystone and mom, thank you for all your support, love, patience and for being by my side.

Thanks to my sister Karina and my Madrinha Solange, and to my friends for love and support, especially Alessandra Lima, Wyll Nogueira, Aline Soares, Bruna Torres, Thaís Milagres, Maria Luiza Sena, Izabela Ferreira, Anna Alencar, Gabriel Bonifácio, Edgar Saldanha, Samuel Vieira, Richard Maxwell and Shifu Danillo Cocenzo.

Rubens Monte-Neto and Pedro Alves, thank you for our collaboration that was fantastic and rewarding for me.

Thanks to all of my professors and colleagues along my journey. I have learned a lot from you.

Thanks to the Programa Interunidades de Pós-Graduação em Bioinformática, especially to Sheila and Tiago, and Departamento de Física for being another home to me.

I thank CAPES for the financial support granted by the CAPES/Ação Emergencial scholarship throughout my doctoral period, which was of great supportive, and also for the CAPES-PrInt fellowship to attend a sandwich doctorate period at the Technical University of Denmark (DTU).

Finally, I thank myself for not giving up on my dreams and fighting until the end and beyond.

“The single biggest threat to man’s continued dominance on this planet is the virus.”

Joshua Lederberg

I dedicate this manuscript to everyone who has a dream.

“Always keep fighting.”

# Resumo

Em 2020, a OMS (Organização Mundial da Saúde) declarou o surto global de uma nova doença causada pelo vírus SARS-CoV-2, a qual se tornou conhecida como COVID-19 (*coronavirus disease 2019*). A aplicação em ampla escala de testes de diagnóstico é uma das formas de controle em surtos similares a esse, evitando a disseminação da doença. Dentre os testes moleculares empregados estão os métodos baseados em PCR (reação em cadeia da polimerase) e LAMP (amplificação isotérmica mediada por loop), principalmente os que envolvem a etapa de transcrição reversa (RT), na qual o RNA viral é transcrito em DNA. Um dos elementos usados por essas técnicas são os *primers*, que são sequências curtas de DNA, em geral, dentro de 18–30 pares de bases. A função desses *primers* é reconhecer a região genômica do alvo para a qual foram desenhados e hibridizarem a ela, conduzindo à detecção do vírus. A temperatura de hibridização, ou *melting*, é um parâmetro crucial para o desenho de *primers*. Tanto em PCR quanto em LAMP, esse parâmetro contribui para que ocorra uma hibridização consistente entre *primer* e alvo, levando ao sucesso da técnica. Porém incompatibilidades, conhecidas como *mismatches*, podem surgir desestabilizando o duplexo *primer*-alvo, o que pode impedir a amplificação do alvo e, consequentemente, a não detecção do vírus. Contudo, a presença de *mismatches* pode ter o efeito inverso contribuindo para a estabilidade do *primer* com o alvo. Nesse trabalho, aplicamos um modelo de física estatística para avaliarmos o impacto de *mismatches* de DNA em *primers* de PCR e LAMP desenhados para a detecção de SARS-CoV-2. Nós coletamos 19 e 18 conjuntos de *primers* publicados para PCR e LAMP, respectivamente, os alinhamos com genomas de SARS-CoV-2 e calculamos a cobertura para hibridizações completas (*perfect match*) e parciais (*mismatch*). Além disso, avaliamos reações cruzadas com genomas de outros seis coronavírus (229E, OC43, HKU1, NL63, MERS-CoV e SARS-CoV-1). Devido ao surgimento de novas variantes de SARS-CoV-2, incluímos a avaliação dos 42 conjuntos de *primers*, de forma completa e parcial, para sete variantes (Alpha, Beta, Gamma, Delta, Lambda, Mu e Omicron) e quatro subvariantes (BA.2, BA.3, BA.4 e BA.5). Nossos resultados mostraram que a maioria dos *primers* avaliados têm alta cobertura para alinhamentos completos e um número considerável obteve um aumento na cobertura considerando a presença de até três *mismatches* consecutivos, incluindo a avaliação para as variantes e subvariantes. Alguns *primers*, entretanto, não cobrem nem completa nem parcialmente os genomas coletados. Em relação à reação cruzada, um número pontual de *primers* cobrem alguns genomas de outros coronavírus e, em alguns casos, somente considerando a presença de *mismatches*.

Palavras-chave: SARS-CoV-2, *mismatches* de DNA, *primers* de PCR, *primers* de LAMP, modelos mesoscópicos.

# Abstract

In 2020, WHO (World Health Organization) declared the global outbreak of a novel disease caused by SARS-CoV-2 virus, which became known as COVID-19 (coronavirus disease 2019). The large-scale application of diagnostic tests is a way of control in outbreaks similar to this, avoiding the dissemination of the disease. Among the molecular tests used are the methods based in PCR (polymerase chain reaction) and LAMP (loop-mediated isothermal amplification), especially those involving the reverse transcription (RT) step, in which the viral RNA is transcribed into DNA. One of the elements used by these techniques are the primers, which are short sequences of DNA, usually within 18–30 base pairs. The function of these primers is to recognize the genomic region of the target to which they were designed and hybridise to it, carrying to the detection of the virus. The hybridisation temperature, or melting, is a crucial parameter to the design of the primers. Both in PCR and LAMP, this parameter contributes to a consistent hybridisation between primer and target, leading to the success of the technique. However, incompatibilities, known as mismatches, may arise destabilising the primer-target duplex, which may prevent the amplification of the target and, consequently, the non-detection of the virus. Notwithstanding, the presence of mismatches can have the reverse effect contributing to the primer-target stability. In this work, we applied a statistical physics model to evaluate the impact of DNA mismatches in PCR and LAMP primers designed to the detection of SARS-CoV-2. We collected 19 and 18 sets of published primers to PCR and LAMP, respectively, aligned them against genomes of SARS-CoV-2 and calculated the coverage to complete (perfect match) and partial (mismatch) hybridisations. In addition, we evaluated cross-reactivity to genomes of other six coronaviruses (229E, OC43, HKU1, NL63, MERS-CoV and SARS-CoV-1). Due to the emergence of new variants of SARS-CoV-2, we included the assessment of the 42 sets of primers, in a complete and partial way, to seven variants (Alpha, Beta, Gamma, Delta, Lambda, Mu and Omicron) and four subvariants (BA.2, BA.3, BA.4 and BA.5). Our results showed that most of the evaluated primers have a high coverage for complete alignments and a considerable number of them achieved an increase in the coverage when considering the presence of up to three consecutive mismatches, including the evaluation to the variants and subvariants. A few primers, however, neither completely or partially cover the collected genomes. Regarding cross-reactivity, a punctual number of primers covers some genomes of other coronaviruses and, in some cases, only when considering the presence of mismatches.



Keywords: SARS-CoV-2, DNA mismatches, PCR primers, LAMP primers, mesoscopic models.

# Used Abreviations

**ACE2** angiotensin-converting enzyme 2

**ASO** antisense oligonucleotides

**AuNR** gold nanorod

**BIP** backward inner primer

**BLOSUM** BLOcks of Amino Acid SUbstitution Matrix

**Bst** Bacillus stearothermophilus

**CDC** U.S. Centers for Disease Control and Prevention

**cDNA** complementary DNA

**COVID-19** coronavirus disease 2019

**CRISPR** clustered regularly interspaced short palindromic repeats

**Ct** cycle threshold

**DNA** deoxyribonucleic acid

**dNTP** deoxyribonucleotide triphosphate

**DTU** Technical University of Denmark

**EGFR** epidermal growth factor receptor

**FIP** forward inner primer

**FRET** fluorescence resonance energy transfer

**GISAID** Global Initiative on Sharing Avian Influenza Data

**HIV** human immunodeficiency virus

**LAMP** loop-mediated isothermal amplification

**LNA** locked nucleic acid

**MD** molecular dynamics

**MERS-CoV** Middle East Respiratory Syndrome Coronavirus

**MM** mismatch

**NCBI** National Center for Biotechnology Information

**NIID** National Institute of Infectious Diseases

**NMR** nuclear magnetic resonance

**NN** Nearest-Neighbour

**PAM** Point Accepted Mutation

**PB** Peyrard-Bishop

**PCR** polymerase chain reaction

**POC** point-of-care

**RBD** receptor-binding domain

**RdRp** RNA-dependent RNA polymerase

**RNA** ribonucleic acid

**RT** reverse transcription

**RT-LAMP** reverse transcription loop-mediated isothermal amplification

**RT-PCR** reverse transcription polymerase chain reaction

**SARS-CoV-1** Severe Acute Respiratory Syndrome Coronavirus 1

**SARS-CoV-2** Severe Acute Respiratory Syndrome Coronavirus 2

**SBE** single base extension

**SNP** single nucleotide polymorphism

**STR** short tandem repeat

**UTR** untranslated region

**UV** ultraviolet

**VOC** variant of concern

**VOI** variant of interest

**VUM** variant under monitoring

**WHO** World Health Organization

# Contents

# 1 Introduction

In the beginning of 2020, the World Health Organization (WHO) declared the pandemic of coronavirus disease 2019 (COVID-19), which was caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The drastic situation demanded several prevention and protection measures, and sparked several emergency research initiatives, including in Brazil such as the CAPES Emergency Program for Covid[1]. This thesis is the direct result of the CAPES initiative and the fellowship granted for this project is gratefully acknowledged.

To be able to contribute with a meaningful project for the CAPES initiative on such a short notice, we searched for potential projects involving our groups expertise in nucleic acids biophysics, especially hybridisation models. Polymerase chain reaction (PCR) was being the essential tool for detection of SARS-CoV-2 coronavirus, and being a molecular diagnostic technique, which depends on nuclear acids hybridisation, it quickly became clear that we should be concentrating our efforts on reverse transcription polymerase chain reaction (RT-PCR) and reverse transcription loop-mediated isothermal amplification (RT-LAMP). Each one is a variation of the basic PCR and loop-mediated isothermal amplification (LAMP) techniques, respectively, which use oligonucleotides (short DNA sequences), called primers, to identify the viral genome. Although both methods have had a good performance, there are a few factors that may interfere with the results. One such factor is the presence of one or more mismatches (MMs) between primer and target (or template), which can potentially affect the detection results. It just so happened that our group had concluded a major study on DNA MMs around the same time which provided us with a completely new tool to evaluate these effects in PCR. In view of the emergency, the importance of the PCR methods and the expertise of our group in DNA MMs hybridisation properties that the main goal of this thesis was to assess if the presence of DNA mismatches in both RT-PCR and RT-LAMP may influence the detection of SARS-CoV-2 virus.

This thesis is organized as follows: in chapter 2, we show a brief description about nucleic acids (DNA and RNA), mismatched base pairs and thermal processes, essential in both mentioned molecular methods. In chapter 3, we give an overview of the SARS-CoV-2 coronavirus and the main aspects of this virus pertaining to our discussion. The RT-PCR and RT-LAMP techniques are described in chapter 4. After presenting the objectives in chapter 5, we describe our methods to applied to the assessment of the primers in chapter 6 and the results and discussion are presented in chapter 7. In chapter 8, we also present some other projects that were developed throughout the doctoral period. In section 8.1, mismatch analyses of different types of primer sets and targets and in section 8.2, the studies of LNA-modified primers and developing thermodynamic parameters. The overall conclusion and perspectives are presented in chapters 9 and 10, respectively.

The main methods for RT-PCR developed in this thesis were published in *Molecular and*

[1]Programa Estratégico Emergencial de Prevenção e Combate a Surtos, Endemias, Epidemias e Pandemias, CAPES Edital 09/2020.

*Cellular Probes* [1], and has received 15 citations according to Google Scholar. During the first year of this project, we established a collaboration with Prof. Rubens L. do Monte-Neto and Prof. Pedro A. Alves (Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte), who were working with RT-LAMP. They gave us access to their primer sequences which we analysed with the methods developed in this thesis and which resulted in a publication in *Frontiers in Microbiology* [2] in 2021, having received 31 citations according to Google Scholar. An extended analysis for the WHO RT-PCR primers for SARS-CoV-2 variants were published as a bookchapter in *Advances in Bioinformatics and Computational Biology* [3] in 2021, with 2 citations according to Google Scholar. The application of our methods to RT-LAMP primers was published in *The Open COVID Journal* [4] in 2024.

In 2022/23 I had the opportunity to spend 6 months at the Technical University of Denmark (DTU), with my co-supervisor Prof. Kira Astakhova. This was made possible with a CAPES-PrInt fellowship. During this internship, I pursued hybridisation studies in LNA which are reported in section 8.2 and resulted in a co-authorship published this year [5].

# 2 Nucleic Acids

Here we describe the structural and thermodynamic properties of DNA and RNA molecules that are necessary to understand the hybridisation models and results. Also we present the mismatch effects on nucleic acid structure and primer-target hybridisation. Last, we discuss the thermal processes of nucleic acids that are used in some molecular techniques.

## 2.1 DNA and RNA Molecules

Nucleic acids are essential molecules of living organisms. There are two types: deoxyribonucleic acid (DNA), discovered in 1869 by Johann Friedrich Miescher, and ribonucleic acid (RNA), classified in the 1920s [6]. Both are formed by polinucleotide chains, also called strands, which are the combination of a phosphate group, a pentose and a nitrogenous base, see Figure 1. The nucleotides are connected by phosphodiester bonds, which occurs between the 5′-phosphate group of the pentose of one nucleotide and the 3′-hydroxyl group of the pentose of the next nucleotide [7–10]. They may contain one, two or three phosphate group, respectively, mono-, di- or triphosphate [7, 11]. The nucleic acid strand is oriented (its polarity) in the direction 5′ $\longrightarrow$ 3′ terminal, that is, from 5′ end (free phosphate group) to 3′ end (free hydroxyl group), see Figure 2. A characteristic difference between DNA and RNA is that the first has a deoxyribose and the other a ribose sugar. The common nitrogenous bases are adenine (A), cytosine (C), guanine (G), thymine (T) and uracil (U) [7, 9, 10, 12]. The latter is found in RNA in place of thymine, although occasionally thymine may be found in RNA molecules [7]. Also uracil may be found in DNA molecule, e.g. as a result of spontaneous deamination of cytosine [13]. Those bases are divided in two groups: purines (adenine and guanine), and pyrimidines (cytosine, thymine and uracil). Purines are bases with two carbon-nitrogen aromatic rings and pyrimidines are those with a single ring, see Figure 3.

Usually DNA and RNA are characterised as a double strand (ds), formed by two nucleotide chains, and a single strand (ss), respectively [6], see Figure 4. However, both can be single or double strand, sometimes even triple strand [14, 15], e.g. viral genomes, which are classified in different groups, such as ssRNA, dsRNA, ssDNA and dsDNA [16, 17]. The duplexes

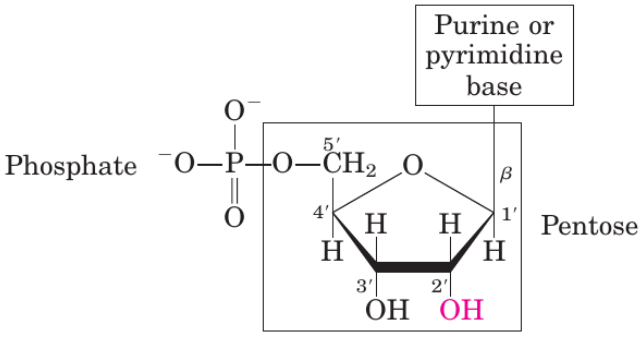



**Figure 1**
Nucleotide structure. A phosphate group, a pentose, here a deoxyribose, and a nitrogenous base, which can be a purine or pyrimidine. Figure adapted from Ref. [7].

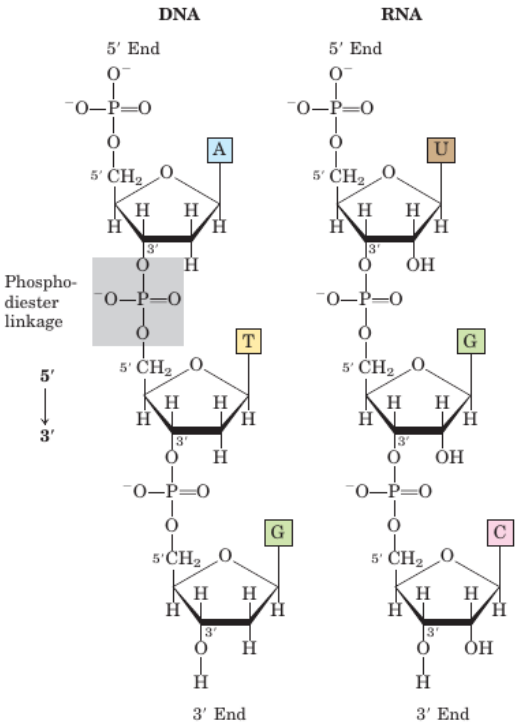



**Figure 2**
Single-stranded DNA and RNA. A nucleotide chain formed by a phosphate, a pentose and a nitrogenous base. Each nucleotide is connected to the next by a phosphodiester linkage (highlighted block). The polarity of the single strands are 5′ end (or terminal) to 3′ end. Figure taken from Ref. [7].

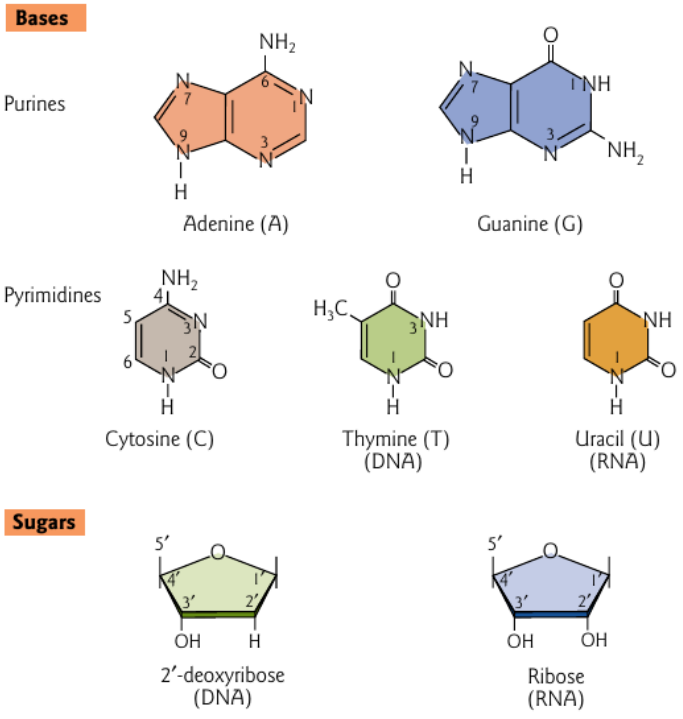
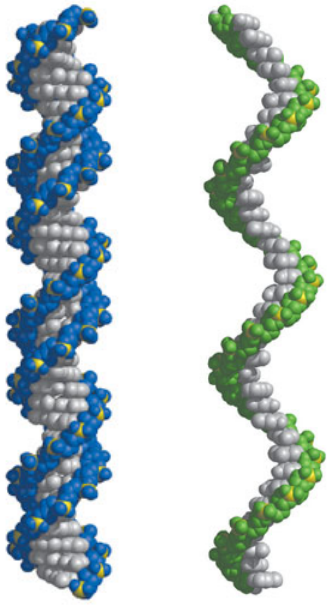



**Figure 3**
The five nitrogenous bases and DNA and RNA sugars. Above (Bases), the five bases: adenine (A) and guanine (G) are purine bases, and cytosine (C), thymine (T) and uracil (U) are pyrimidine bases. Below (Sugars), the deoxyribose sugar for DNA and the ribose sugar for RNA. Figure adapted from Ref. [11].



**(a)** DNA **(b)** RNA

**Figure 4**
(a) Double-stranded DNA (dsDNA) and (b) single-stranded RNA (ssRNA) molecules. Both in a helix conformation. Figure adapted from Ref. [7].

can be of three types: DNA/DNA, as human genome; RNA/RNA, as some viral genomes, or DNA/RNA (hybrid), as the product of transcription process. For the three types, the double strand is antiparallel, one strand is in the $5' \longrightarrow 3'$ direction and the other in $3' \longrightarrow 5'$, and complementary, the nitrogenous bases in one strand bind to their counterparts on the other, see Figure 5. The bases interact pair to pair with cytosine bonding to guanine and adenine bonding to thymine (or uracil), which are called canonical base pairs or Watson-Crick (WC) base pairs, that is, purine-pyrimidine pairs [7, 11]. The CG base pair (bp) has three hydrogen bonds and AT and AU pairs have two hydrogen bonds [7, 8, 10, 12], see Figure 6. Due to the base pair interaction along the direction of hydrogen bonds, according to Chargaff's rules, the number of residues of guanine in a double-stranded DNA is equal to the number of cytosine and the same is valid to the number of residues of adenine and thymine [7–9, 12].

Hydrogen bonds are very weak bonds formed by the sharing of a hydrogen atom between two electronegative atoms, e.g. oxygen and nitrogen. Although they are very weak, a large number of them offers a structural stability to a molecule [11]. In addition to the hydrogen bonds interaction, the bases interact with each other on the same strand like a staircase, one stacked on top of another. This interaction is called "stacking interaction" or "base stacking", see Figure 7. The planar configuration of bases leads to $\pi$–$\pi$ electrostatic interactions between their aromatic rings and they contribute to ultraviolet (UV) absorption around 260 nm [7, 18, 19]. Both hydrogen bonds and stacking interactions contribute to the thermodynamic stability of DNA duplex with the base stacking contributing most [7, 9, 19]. Since the stacking interaction eliminates any gaps between bases, the interior of molecule is hydrophobic, which provides a driving force for DNA molecule to form a double helix [11]. Furthermore, RNA molecules are stabilised by stacking interactions, as well as DNA molecules [7].

DNA and RNA are deeply related by the central dogma of molecular biology, see Figure 8.

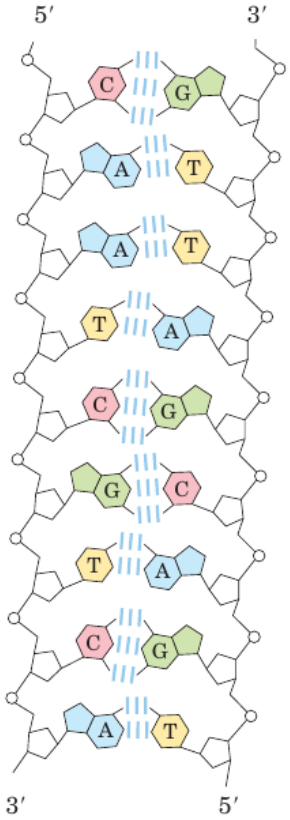



**Figure 5**
A double-stranded DNA showing the polarity of both strands ($5' \longrightarrow 3'$ and $3' \longrightarrow 5'$), and two and three hydrogen bonds of AT and CG base pairs, respectively. Figure taken from Ref. [7].

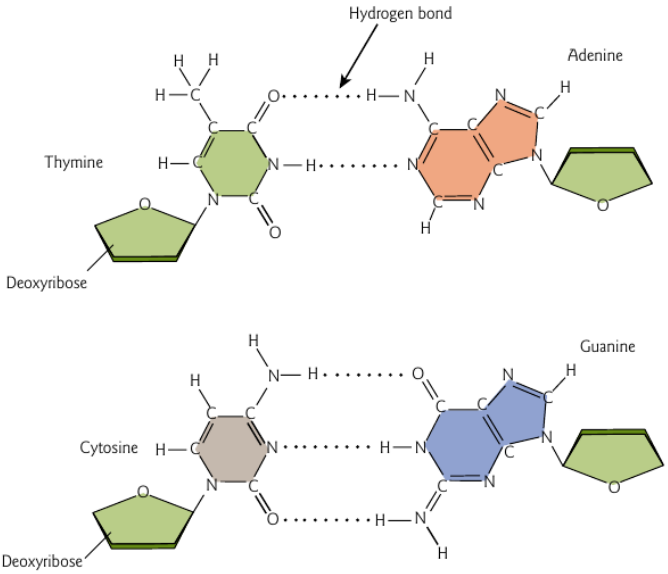



**Figure 6**
Hydrogen bonds of Watson-Crick base pairs. Above, the two hydrogen bonds for AT pair and below the three hydrogen bonds for CG pair. The hydrogen bonds of AU pair are not shown here, but it is similar to the interaction of AT pair. Figure taken from Ref. [11].

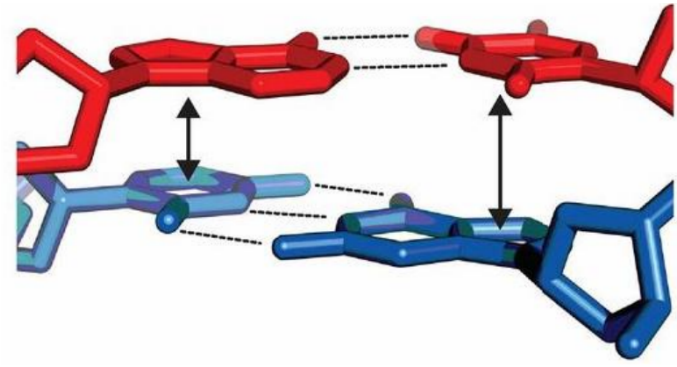


**Figure 7**
The stacking interaction between nitrogenous bases. Each one is stacked on the other like a staircase, while their heterocyclic rings are interacting with each other. Figure taken from Ref. [20].

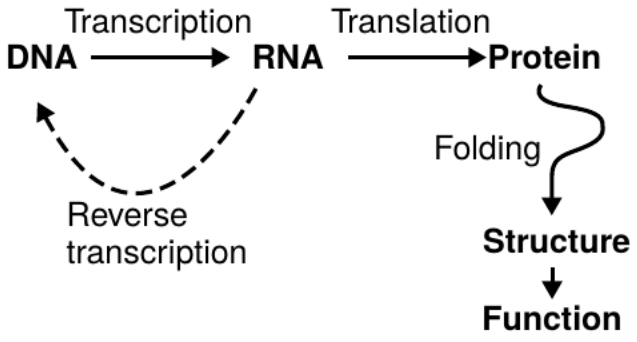



**Figure 8**
The central dogma of molecular biology. DNA is transcribed into RNA, which is translated into protein. As well as RNA is reverse transcribed into DNA. Figure taken from Ref. [21].

In transcription step, DNA is used as a template to generate RNA molecule — the complement to DNA template. One of the elements that mediates this process is the DNA polymerase. Next, in translation step, the RNA template carries the information for designing the protein that folds into a proper structure. There is a parallel step — the reverse transcription — which is exactly what its name means. The RNA molecule is now the template to generate a DNA molecule and, among other components, is mediated by reverse transcriptase [21–23]. This step revolutioned molecular biology elucidating biology processes, and promoting new study tools, e.g. transcriptome analysis, and molecular techniques like reverse transcription polymerase chain reaction (RT-PCR) and reverse transcription loop-mediated isothermal amplification (RT-LAMP), based in the reverse transcription (RT), see chapter 4.

## 2.2 Mismatched Base Pairs

Mutations and mismatches can interfere in the stability of nucleic acid molecules [24], which can take place in both DNA and RNA. There is a subtle difference between mutations and mismatches. Mutations occur when a base pair in a DNA, or RNA, molecule has been changed to another base pair, e.g. CG $\longrightarrow$ AT. Or in a simple way, the change of one base in just one strand results in the complementary strand with the correct counterpart changed, as well as occur with viral genome [16, 25]. In the case of mismatch (MM), an incorrect pairing occurs when one base binds to a non-counterpart base, e.g. AA, CT, GU and GT [11, 25], which is also called noncanonical or non-Watson-Crick base pair, see Figure 9. It can take place during DNA replication [7], genetic recombination [26] or primer-template hybridisation in PCR reactions [27], which may lead to false-negative results [28].

There are eight possible mismatched base pairs: AA, AC, AG, CC, CT, GG, GT and TT, which may influence the stability and structural properties of DNA helix. Mismatches change hydrogen bonds conformation in comparison to canonical base pairs, as well as the stacking interactions. However, some of them show a similar overall shape to a canonical pair and a relatively stable configuration, e.g. GT mismatched pair [11, 29], see Figures 10 and 11. DNA Watson-Crick base pairs are in an *anti-anti* conformation, while MM base pairs can also be *anti-syn* or *syn-anti*. Some molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) experiments have shown that mismatches do not impact on the global conformation of

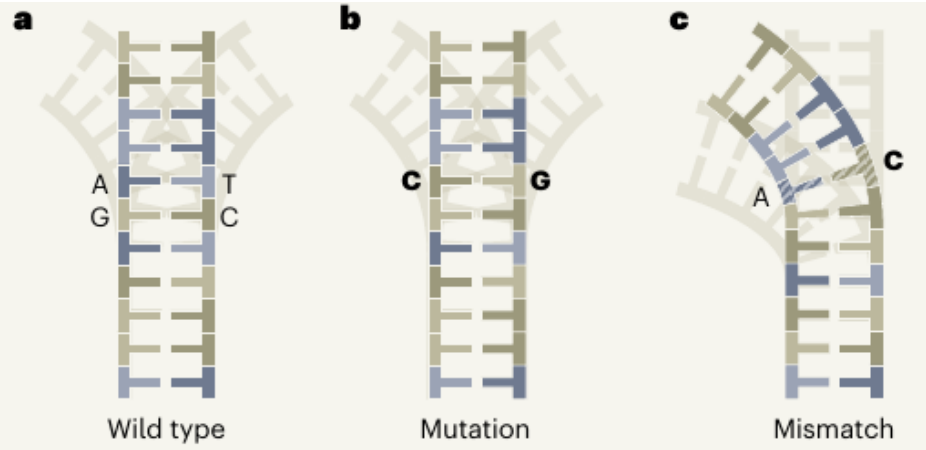



**Figure 9**
Mutation × Mismatch. (a) A double-stranded wild type, which is the "original" sequence; (b) a mutation in the sequence (AT ⟶ CG), when a base pair changes to another Watson-Crick base pair; (c) a mismatch pair (AT ⟶ AC), a base binds to an incorrect base (non-Watson-Crick base pair), which may destabilise the duplex. Figure adapted from Ref. [25].

the DNA duplex, such as for AA and TT pairs [30, 31]. For instance, a consecutive GG and CC pairs showed a local distortion impacting only their AT neighbours [32]. Yet in a local conformation the influence of mismatches varies from weakly bound (CC pair) to strongly bound (GG pair) cases [30]. In addition, guanine (GT, GG, AG) and cytosine (AC, CC) mismatches are the most and least stable base pairs, respectively [33–35].

Mismatches are commonly associated to primer/probe-target hybridisation, mainly for polymerase chain reaction (PCR), which will be discussed in section 4.1. Ideally, there should be no mismatches between primer and target, however, this is unavoidable [37, 38], e.g. as result from viral genome mutations [39, 40]. Primer-target mismatches can impact the PCR assay in a few ways as disrupt PCR chemistry, impair detection of the target and produce false negative results [28, 41]. As well mismatches in loop-mediated isothermal amplification (LAMP) primers may reduce the amplification efficiency [42], decrease the sensitivity of detection and lead to false negative results [43]. Internal and terminal mismatches influence primer/probe hybridisation in different ways [44], see Figure 12. Mismatches located far from 3′ end are not so relevant and will have a moderate effect without influencing PCR performance [27]. Those near 3′ terminal are critical for PCR good performance interfering in its accuracy [45–47] and may lead to non amplification of the target [44, 48], as well as a single mismatch at 3′ terminal may results in the same impact [49]. According to Bru et al. [50], MMs in the direction of 3′ terminal are more detrimental to the PCR reaction and a single internal mismatch has a variable effect, probably due to the dependency on mismatch position and sequential environment of the primer [50]. Notwithstanding, Mitsuhashi [46] suggests that mismatches either near or at 3′ terminal may avoid false priming unlike 5′ terminal and internal mismatches. Kwok et al. [51] show that mismatches in the three last termini positions have no effect and a thymine at 3′ ter-

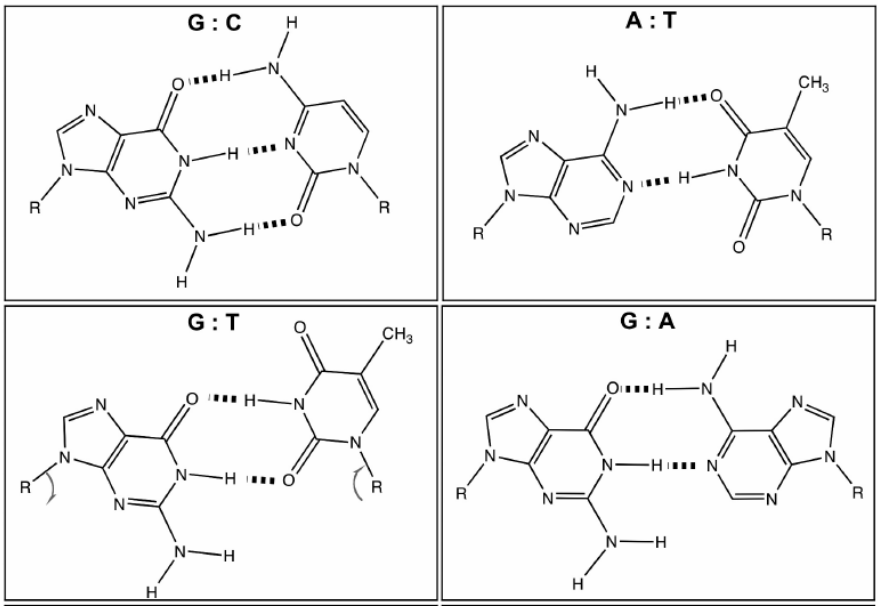



**Figure 10**
Chemical structures of GC and AT canonical and GT and GA mismatched base pairs. Figure adapted from Ref. [36].

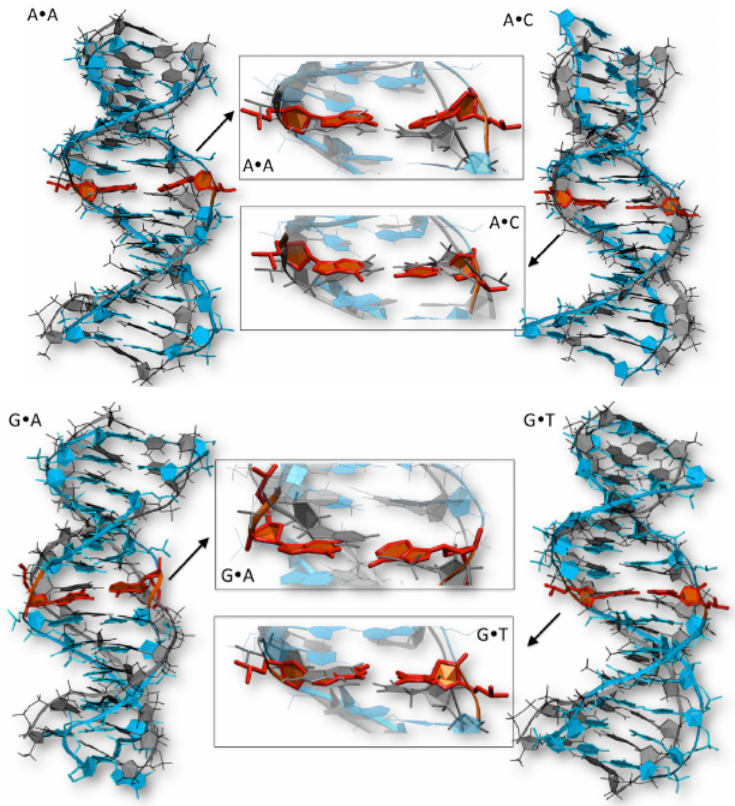



**Figure 11**
Overlap of canonical (grey) and mismatched (blue) duplexes. A close-up is shown in the compared region: canonical AT pair (grey) and AA, AC, GA and GT mismatched pairs (red). Figure from Ref. [30].

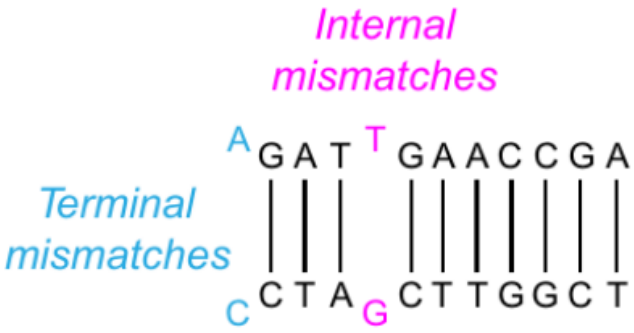



**Figure 12**
Example of terminal and internal mismatch position. Figure adapted from Ref. [44].

minal hybridising to any base may improve the amplification efficiency and suggest that some mismatched pairs may amplify efficiently as matched pairs. Furthermore, mismatch stability at 3′ terminal depends on the mismatched pair and the stability of neighbouring base pairs [52].

A large number of mismatches can destabilise primer-target duplex or even block the hybridisation altogether, and the presence of just three mismatches in the primer would turn the PCR non functional [53]. Four internal mismatches do not impact in PCR assay [51] in contrast to other evidences [50, 54], and two mismatches show little impact on assay sensitivity [55]. Since PCR assays use at least a primer pair (forward and reverse primers), the MM impact may be different in both. A single mismatch at 3′ terminal of a forward primer has little impact in opposite to two or three mismatches in the last three positions of 3′ terminal. Up to five consecutive mismatches at 5′ terminal show no influence in the forward primer, however two to three mismatches far from 3′ terminal, along the primer, affect the assay [56]. Mismatches may also have an additive effect, that is, mismatches in both forward and reverse primers may add to their effects [54, 56, 57]. Moreover, it is suggested that a single mismatch in the primers may not influence the PCR performance, whereas in the probe it may lead to a false-negative [48]. In contrast, a few mismatches in the probe may have little impact [58]. In an assay to detect HIV virus, the primers showed tolerance of up to four mismatches and the probe up to five [59]. Mismatches located near 3′ terminal in forward primers also show no obvious effect on sensitivity assay. When located at the third base near 5′ terminal of the probe, mismatch presence have a greater influence on sensitivity than those in the middle [55]. Also mismatches at both 3′ and 5′ terminals in the displacement LAMP primers do not interfere, whereas in both forward inner primer (FIP) and backward inner primer (BIP) may hamper the amplification [60].

It is suggested that some types of mismatches are more stable than others and, in some cases, than canonical AT pair [37, 61]. Some mismatches can contribute to the stability of RNA duplexes [62], e.g. GU mismatched pair [11], and the sequential environment and their position affect their behaviour in both nucleic acid types [63–65]. The stability of primer-target hybridisation may be greater in presence of few MMs in comparison to AT-rich primers with no mismatches [46]. A few Cas12 enzymes in CRISPR technique have shown mismatch tolerance [60]. LAMP assay may detect resistant mutants after an antibiotic administration incorporating a mismatch at 3′ terminal [66]. Unexpectedly findings showed that mismatches are able to enhance the protein-DNA binding [24]. Furthermore, it is known that a few mismatches in

PCR primers can contribute to the design of antisense oligonucleotides (ASO) [67], and allele-specific [37] and single nucleotide polymorphism (SNP) [68] identification, as well as LAMP based-methods for the latter [69]. The melting temperature analysis could be used to detect specific variants due to the difference in the temperature [70]. Recently, PCR based-methods have used mismatches at or near 3′ terminal to identify Delta variant [71] and Omicron subvariants (BA.1, BA.2, BA.4, and BA.5) [72].

It is known that mismatch presence may hamper the activity of DNA polymerases. However, they are able to bind to a mismatched pair with about equal affinity to a matched pair [52]. The ability of DNA polymerases to elongate the 3′ terminal mismatch depends on the concentration of next nucleotide and affinity of primer-target [52, 73, 74]. In a study comparing the efficiency of two polymerases [74], one was able to easily extend CT, TT and GT mismatched pairs, the latter as efficiently as for canonical AT pair, whereas the other polymerase easily extended AG and GG pairs. However, both polymerases were less efficient to elongate GG pair, in comparison to canonical CG pair, and CC mismatched pair. A third DNA polymerase [52] showed low efficiency to extend AC mismatched pair and high for the CT pair.

Finally, the effect of mismatches is not yet standardized as a few factors influence their behaviour, such as pH values, salt concentration, mismatch type, number and position, melting temperature, reaction condition, oligonucleotide length and sequential environment, especially flanking base pairs, that is, MMs are context dependent [27, 37, 48, 50–52, 56, 61, 75–81]. Mismatch tolerance observed for some primers [51, 82] may be due to their longer length [57]. Two AG mismatched pairs at 5′ terminal failed the assay, whereas CT pair successful detected the target [83], as well as previous evidences of AG pair is more detrimental than CT pair [37, 55]. Also lower annealing temperature may optimise the mismatch effect [54, 84, 85] and the amplification conditions may support mismatch tolerance [55].

## 2.3 Thermal Processes of Nucleic Acids

When DNA molecule has its strands connected (double-stranded), it is in a “natural or native state”, that is, the duplex is in a stable state. This state is found in an appropriate temperature for the double strand stability. However, if the temperature changes, usually increases, the DNA duplex destabilises. An increase in the temperature affects the hydrogen bonds between base pairs leading to a separation of the strands, even to a completely dissociation [8, 9, 86, 87], see Figure 13. This process is called “denaturation”, or “melting”, of double-stranded nucleic acid [11, 22].

The main parameter of this process is the hybridisation or melting temperature $T_m$ in which 50% of the strands in the solution are associated and the other 50% are dissociated — completely or partially denatured [7, 77, 88, 89]. This temperature characterises the denaturation process and, in a proper environment, may describe the stability of DNA duplex as a physical property [88, 90]. Melting temperature data can improve computational models to stability and

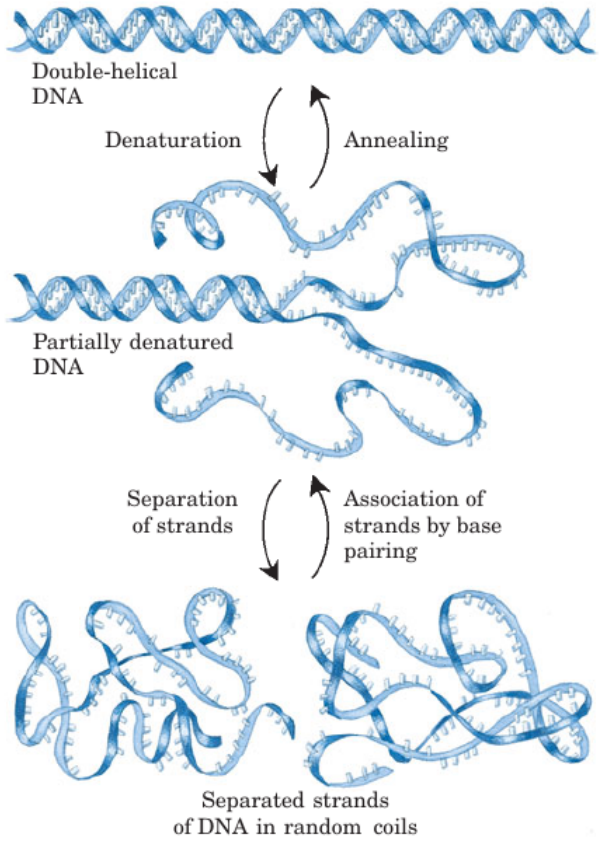



**Figure 13**
The denaturation and renaturation (annealing) processes. In the denaturation, increasing the temperature, the double strands begin the dissociation until a transition temperature is reached, after which the duplex is totaly dissociated. Decreasing the temperature, the single strands begin to anneal and, after a certain time, they will be associated again — the renaturation. Figure taken from Ref. [7].

hybridisation energies of DNA and RNA [90], and contribute to the experimental design of PCR primers [91], silencing RNA [88] and single nucleotide polymorphism (SNP) [89]. The DNA denaturation and, consequently, the melting temperature may be affected by several factors, such as pH values, denaturants (e.g. urea and formamide), salt concentration, GC content and sequence length [11, 23, 92]. The GC content is directly related to the hybridisation temperature. The higher the GC content in the duplex, the higher is the melting temperature of DNA, see Figure 14. This is due to the extra hydrogen bond in comparison to AT base pair [7, 11, 23]. In addition, the stacking interactions between GC pair and neighbouring base pairs are more favorable energetically than between AT pair and its neighbouring base pairs [11].

The dissociation of DNA double strand changes the absorption of UV light. Around 260 nm, double-stranded state absorbes less UV light (about 40%) than in a single-stranded state [11, 22]. While DNA molecule denatures — temperature increases —, the UV absorption increases, which is called “hyperchromicity”. It is due to nitrogenous bases are more exposed to environment, which allows to absorve the maximum of light. Hence about 37% more UV light is absorbed than in duplex state [19]. The temperature change is gradual and as the temperature increases DNA becomes dissociated, that is, DNA becomes single-stranded. The melting temperature is the middle point of the denaturation curve in which there are DNA molecules in both states (double and single) and even partially associated, see Figure 15. The value of $T_m$ depends on conditions of the environment of the nucleic acid molecules as salt concentration and pH value [7].

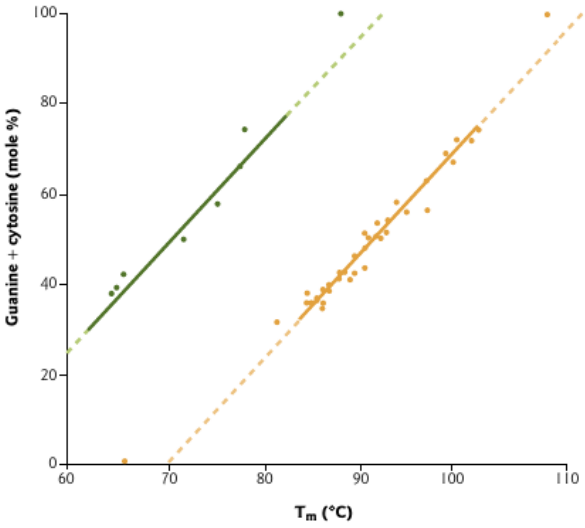



**Figure 14**
The relation between melting temperature $T_m$ and GC content in different salt concentrations. The green line (left) is a DNA solution at low salt concentration and the orange line (right) is at high salt concentration. Figure taken from Ref. [11].

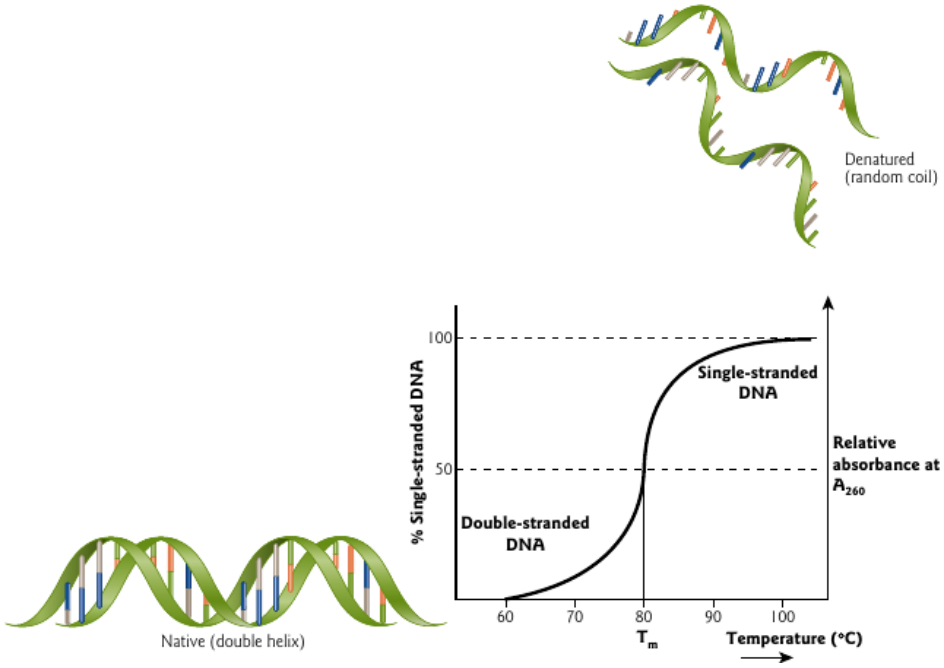



**Figure 15**
Denaturation curve. Temperature (°C) values to DNA molecule in absorbance at 260 nm ($A_{260}$). In the beginning, DNA is a double strand. Increasing temperature, the duplex denatures and the molecule is a single strand in the end. The middle point is the melting temperature ($T_m$) of the duplex, half molecules double-stranded and half single-stranded. Figure taken from Ref. [11].

The denaturation process is a reversible process. The DNA strands dissociated can be associated again by gradually decreasing the temperature. It contributes to a proper environment to bases in different strands interact and anneal over again [11, 91]. This reverse process is called “renaturation” or “annealing”, see Figure 13. In this case, the stacked bases are getting closer due to the duplex renaturation, decreasing the absorption of UV light, which is called “hypochromic effect” [7]. It is applied, for example, in molecular techniques which require the hybridisation (or annealing) of the primers and/or probes to target sequences [11], such as in PCR- and LAMP-based methods.

# 3 SARS-CoV-2 Coronavirus

The first human coronavirus was confirmed in 1965 by June Almeida from St Thomas's Hospital Medical School, see Figure 16. In a collaboration with David Tyrrell from Common Cold Research Unit, they published a work about uncharacterized human respiratory viruses [93]. Tyrrell and his team collected samples of flu-like virus, but they were unable to cultivate it in the laboratory. Then Tyrrell contacted June Almeida to analyse these samples by electron microscopy. She realised that the samples contained the same viruses which she identified in 1964 watching at bronchitis in chickens and hepatitis liver inflammation in mice. However, her work about this remark had been rejected by the reviewers at the time, who said her pictures probably were poor quality pictures of influenza virus particles [94, 95]. Due to the crown-like structure, the virus was called as coronavirus (from latin, *corona* means crown).

Since the identification of the first coronavirus, the number of coronaviruses has grown, including human and non human viruses. These viruses are classified in the *Nidovirales* order, *Coronaviridae* family, *Coronavirinae* subfamily, which is subdivided in four genera (*Alpha-*, *Beta-*, *Gamma-* and *Deltacoronavirus*). *Alpha-* (229E and NL63) and *Betacoronavirus* (OC43, HKU1, SARS-CoV-1, MERS-CoV and SARS-CoV-2) are the coronaviruses that are able to infect human and the last three are the most pathogenics [97–101]. Human coronavirus (hCoV) causes broad range of pulmonary conditions ranging from mild to severe conditions, such as a common cold to pneumonia [99, 102]. Until the beginning of 2000's, only four human coronaviruses were known — hCoV-229E, hCoV-NL63, hCoV-OC43, hCoV-HKU1 — , which are seasonal and have not caused concern [99, 103]. However, in 2003, a novel coronavirus emerged spreading all around the world and causing a high mortality rate. It was the first SARS coronavirus, known as SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), or SARS-CoV-1. In 2012, another coronavirus emerged in Saudi Arabia spreading in the Middle East. It was the time to the discovery of the sixth coronavirus — the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) [98–100]. In March 2020, World Health Organization (WHO) declared the pandemic of the disease COVID-19 (coronavirus disease 2019), arising from a new coronavirus. The seventh human coronavirus known until now is genetically similar to SARS-CoV-1 and was called SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2), see Figure 17.

SARS-CoV-2 coronavirus is an enveloped virus with positive-sense, single strand RNA (+ssRNA), that is, the genetic code can be translated directly into protein. It has a diameter ranging 60–140 nm and its genome is non-segmented and approximately 30000 bases (30 kb) in size, which encodes about 27 proteins, which are structural and non-structural. In general, the SARS-CoV-2 genome is divided as 5′-UTR, open reading frame (ORF), spike (S), envelope (E), membrane (M), nucleocapsid (N) and 3′-UTR-poly A [98, 101, 103–105], see Figure 18. Typically, it causes respiratory, digestive and nervous system diseases [106].

Coronaviruses have a high mutation rate [98, 105], usually $10^{-4}$ nucleotide substitutions per

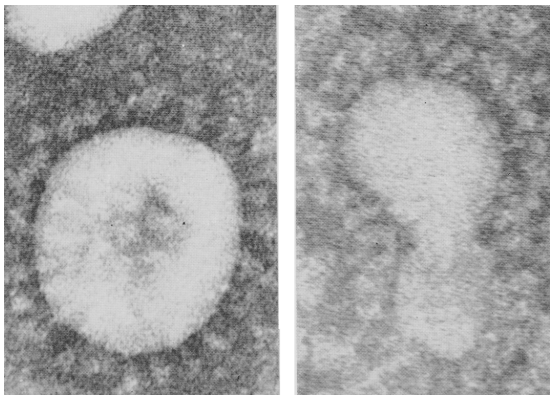


**Figure 16**
The first human coronaviruses. In 1965, June Almeida identified this virus in flu-like virus samples by electron microscopy. Its name is due to its crown appearance. Figure adapted from Ref. [93].

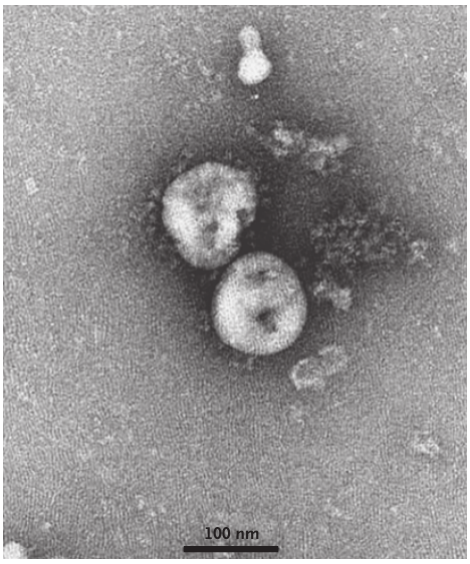



**Figure 17**
SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2) coronaviruses showed by transmission electron microscopy (TEM). Figure adapted from Ref. [96].

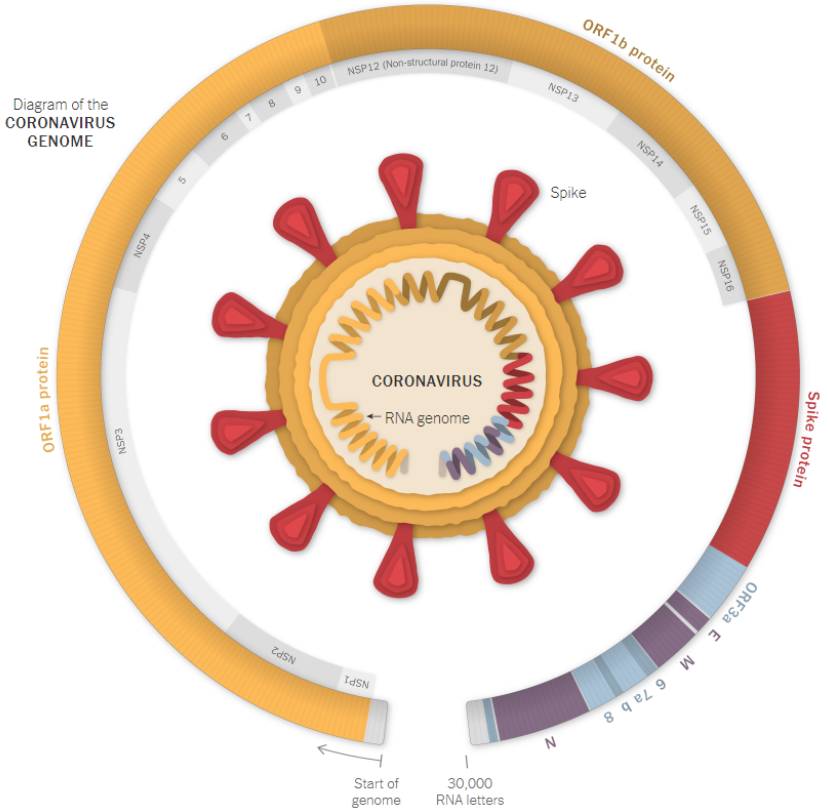



**Figure 18**
SARS-CoV-2 structure and genome. In the center, a conceptual design of coronavirus. Around of it, its divided genome. The colors in the conceptual design correspond to the part of virus encoded by its color counterpart in the genome. Figure taken from Ref. [107].

site per year [108]. It is likely that the novel coronavirus emerged from genetic mutations due to similarities between SARS-CoV-1 and SARS-CoV-2 genomes [109], which shares approximately 80% of similarity with the first one [110, 111]. SARS-CoV-2 genome is susceptible to frequent mutations [112, 113], around $10^{-3}$ mutations per site per year [114–116], and is highly transmissible [117, 118]. They may occur naturally or through interaction with the host immune system surveillance [119]. Sometimes mutations have no effect on the virus, on the other hand, they may lead to changes in virulence, infectivity and transmissibility [120].

Since the beginning of the pandemic, several SARS-CoV-2 variants and subvariants have emerged. Different genes have undergone mutations, such as ORF1A, RdRp, N and M genes, which may affect the infectivity and antigenicity. The most mutated gene, with approximately 4000 mutations, and the target of therapies and vaccines is the S gene that encodes the S (spike) protein [121]. Mutations in the S protein, especially in the receptor-binding domain (RBD), have shown a great impact in the transmissibility and infectivity [122, 123], see Figure 19. The S protein is subdivided in the S1 and S2 subunits. The S1 subunit has the RBD encode region and is the bridge to the virus entrance in the host cells through the angiotensin-converting enzyme 2 (ACE2). The fusion between the viral and cellular membrane is mediated by the S2 subunit [124]. Note in Figure 19 that the greatest variability between the variant genomes is in the S1 subunit, especially in the RBD region for Omicron subvariants.

Commonly mutations in SARS-CoV-2 strains are substitutions and indels (insertion/deletion). Errors during the viral genome replication and repair may induce single base substitutions. Deletion or insertion are a loss or an addition of one or more bases, respectively, and is the largest occurrence in SARS-CoV-2 mutation. Omicron variant is characterized due to a great amount of deletions [126, 127], which may have led this variant to become extremely transmissible. For instance, Alpha variant had undergone some deletions and the N501Y mutation increases the receptor binding affinity which may have caused its rapid spread [128]. Lambda variant shows several deletions which may resulted in an increased infectivity and immune evasion from antibodies [129].

Mutations in the viral genome also affect the dianostic tests. The emergence of variants have promoted mismatched hybridisations between primers and targets, which may hamper the amplification and lead to a no-detection of the virus, see section 2.2. Also mismatch (MM) presence may produce false negative results, reduce the amplification efficiency and sensitivity of detection [28, 41–43]. However, mismatches may be applied to identify a specific variant as the case of diagnostic assays developed to detect Delta variant [71] and Omicron subvariants [72]. The following examples of primer-target hybridisations show a canonical hybridisation for wild-type SARS-CoV-2 and the occurrence of mismatches in different variants. All $3' \longrightarrow 5'$ sequences are the same primer hybridised to five different genomes. Note the large number of mismatches (red) in Beta variant hybridisation.

5′ – CCATAACATGACCATGAGGT – 3′ $\longrightarrow$ SARS-CoV-2 wild type
3′ – GGTATTGTACTGGTACTCCA – 5′

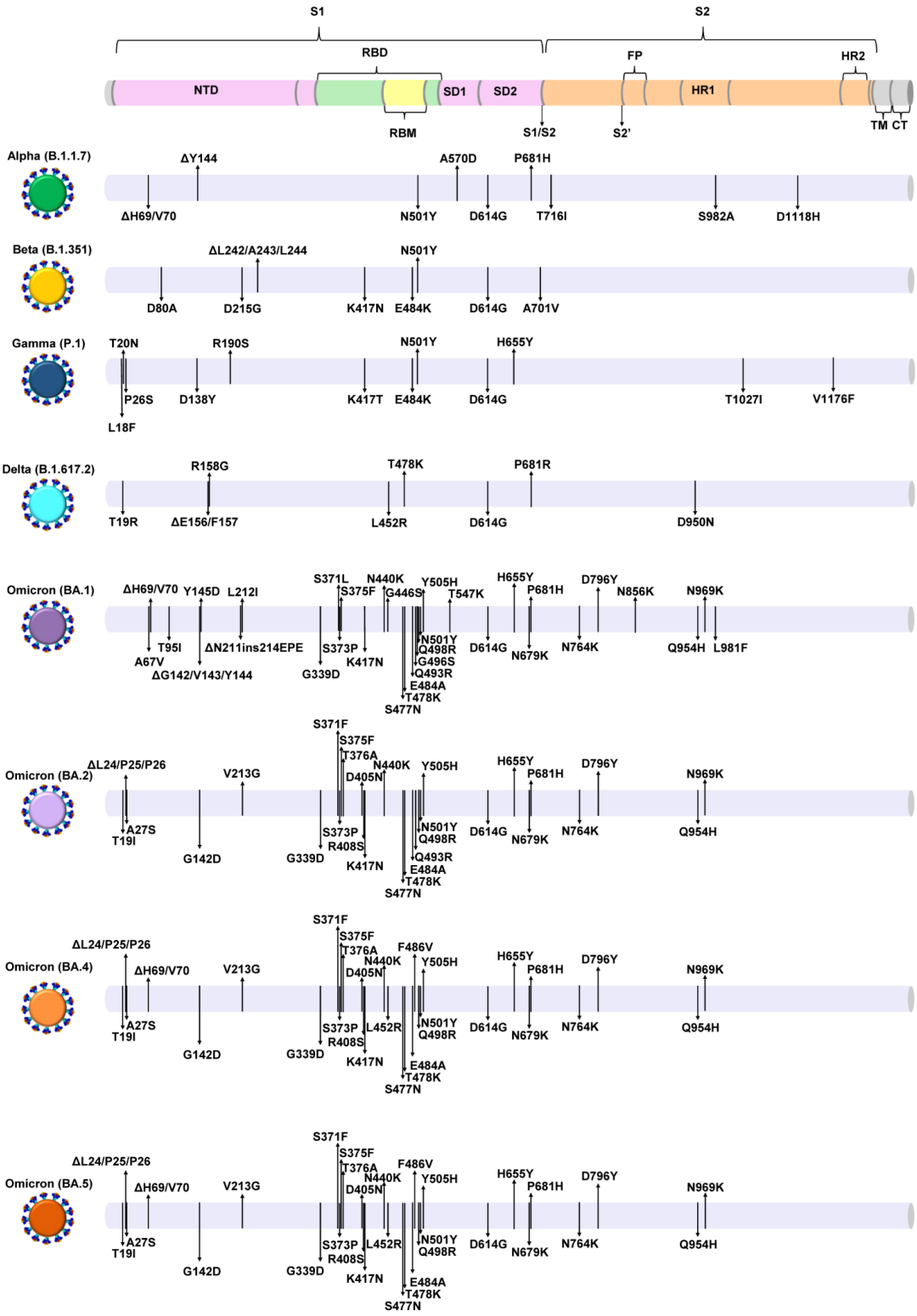



**Figure 19**

SARS-CoV-2 variants and subvariants. S protein mutations along the Alpha, Beta, Gamma and Delta variants and Omicron subvariants. At the top, the S protein domains and in front of each variant, the mutations indicated by black arrows. Figure taken from Ref. [125].

5′ – CAACAACATGACCATGAGGT – 3′ ⟶ Alpha variant
3′ – GGTATTGTACTGGTACTCCA – 5′

5′ – ATATACCTGCACCAATGGGT – 3′ ⟶ Beta variant
3′ – GGTATTGTACTGGTACTCCA – 5′

5′ – CCATAACGTGACCATGAGGT – 3′ ⟶ Gamma variant
3′ – GGTATTGTACTGGTACTCCA – 5′

5′ – GCTCAACATGACCATGAGGT – 3′ ⟶ Delta variant
3′ – GGTATTGTACTGGTACTCCA – 5′

SARS-CoV-2 variants and subvariants have been classified as variants of concern (VOCs), variants of interest (VOIs) and variants under monitoring (VUMs). Their classification has been constantly updated and the previously VOCs were de-escalated, including Omicron parent variant (B.1.1.529), since 15 March 2023 [130, 131]. In this work, we used a considerable set of SARS-CoV-2 genomes collected in the first 9 months of the pandemic, and seven variants and four subvariants sets, which were VOCs at the time, see section 7.1.3. The variants are Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Lambda (C.37), Mu (B.1.621) and Omicron (B.1.1.529), and the subvariants BA.2, BA.3, BA.4 and BA.5.

# 4 Molecular Techniques

Here we show a short review about two techniques — PCR and LAMP — to the amplification of nucleic acids and, consequently, the detection of SARS-CoV-2 virus. The work described in chapter 7 used data from both methods. Both have some variations with a few peculiarities. We will briefly describe the standard process within the reverse transcription polymerase chain reaction (RT-PCR) and reverse transcription loop-mediated isothermal amplification (RT-LAMP) versions.

## 4.1 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) is an important molecule technique that allows a small number of molecules in a sample be amplified to a considerable amount of molecules [44]. It contributes to the detection in samples with a low viral load and point mutations in genes [23]. There are several PCR-based methods such as RT-PCR [59], multiplex PCR [132], digital PCR [133], real-time PCR [134] and quantitative PCR (qPCR) [44]. The PCR method was first developed to detect the DNA nucleic acid. Later, it was introduced a short step to detect RNA nucleic acid. Prior to the standard process of PCR, the step of reverse transcription (RT) is carried out, in which the RNA template is transcribed into complementary DNA (cDNA). The main version of the technique that uses this step is the RT-PCR, which is widely applied, e.g. in assays and diagnosis of viral genomes. It has been the gold standard technique to the detection of SARS-CoV-2 virus [135, 136] and has been applied to oncogenes [137], allele-specific [138] and SNP identification [37], hepatitis [139] and forensic analysis [11].

Each PCR-based method has a particular characteristic in its process and achievement of results. However, all of them varies from the standard process, which will be described here inside the RT-PCR method. In the RT-PCR, the first stage is the reverse transcription of the RNA template into cDNA, which is the complementary strand of RNA as a DNA strand [136]. It is similar to the process of reverse transcription mentioned in section 2.1, see Figure 8. After RNA is converted into DNA, the standard process of PCR begins. The sample is heated, usually at a temperature of 94 °C for 30 seconds to 5 minutes [23]. It is the "denaturation stage" in which double strands in the sample are dissociated into single strands [44], as described in section 2.3. The next stage is the "annealing stage", the primer pair — two short DNA sequences (usually 18–24 bp in length) — hybridises to the target [140]. These primers are designed to hybridise to a specific region flanking the target region of the DNA template [141]. The temperature is lowered to the annealing temperature $T_a$, commonly 5 °C below the melting temperature $T_m$. Usually the annealing temperature is in a range 40–65 °C. It depends on the design of the primers, which will be discussed ahead. Finally, the products are amplified, which are called amplicons, by the action of a thermostable DNA polymerase enzyme (usually Taq polymerase) and a probe at an approximate temperature of 72 °C [23]. This is the "extension stage", the last

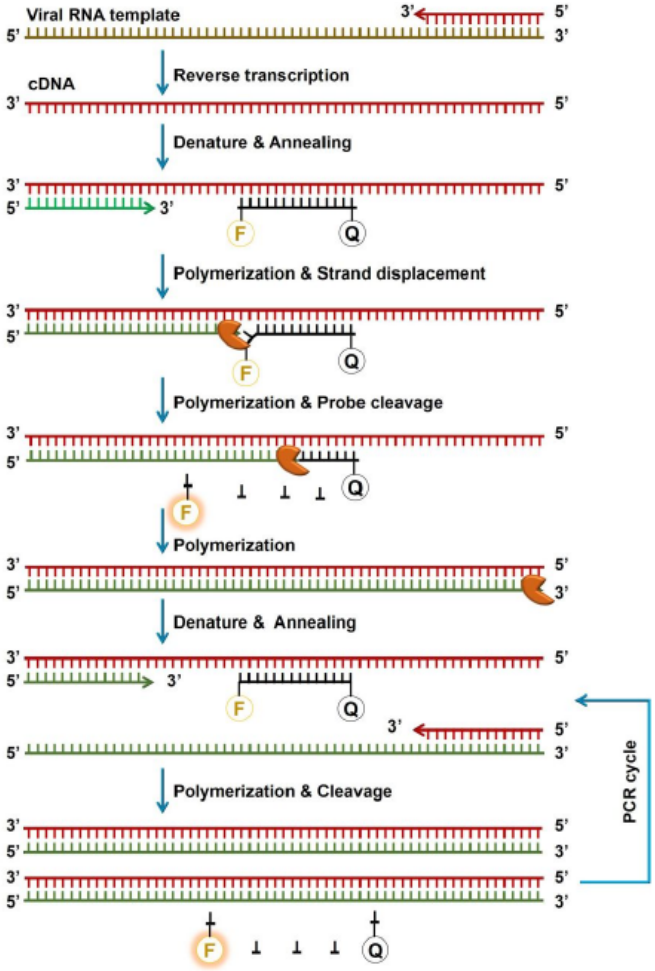



**Figure 20**
RT-PCR process. The viral RNA template is reverse transcribed into cDNA. Then the duplex is heated and dissociated. Primer and probe hybridise to the cDNA and it is amplified by DNA polymerase as its complementary strand. This newly DNA duplex is denatured, the single strands resulted are annealed by primers and probe, DNA polymerase amplifies them and a new cycle begins. In the figure, "Denature and Annealling" corresponds to denaturation and annealing stages and "Polymerization & Strand displacement/Probe cleavage" to extension stage. Figure taken from Ref. [106].

one. The probe is a short DNA sequence with fluorescent dyes in each terminal and designed to a specific region of the target, as well as the primers. The difference is its melting temperature (about 70 °C) is higher than the primers and its length is longer. From this step, the cycle is repeated (denaturation, annealing and extension stage) for a specific number of times, typically 30–50 cycles [23, 44]. Figure 20 summarizes the RT-PCR process. The temperatures described to each stage are not standard. They vary due to the conditions of the assay, e.g. GC content and primer length [23, 142–144]. The detection of the amplicons can be made by several methods, usually agarose gel electrophoresis, intercalating chemical dyes and silver ions, DNA-binding dyes and melt curve analysis [23, 136, 143].

PCR reaction requires a few components, such as buffer, primers, probe, DNA polymerase and deoxyribonucleotide triphosphates (dNTPs) [23]. The primers have a particular importance

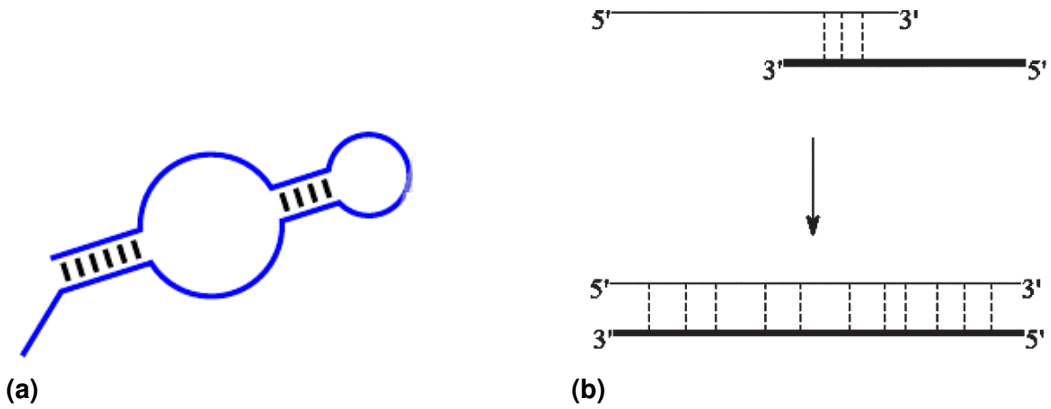



**Figure 21**
Secondary structure (a) and primer-dimer (b). Two primers hybridise in a complementary region and are amplified resulting in a non-specific product of PCR reaction. Figures adapted from Ref. [44] and [23], respectively.

due to their role in the amplification of the target. After they hybridise to the target, the DNA polymerase begins the extension of the DNA copy from the 3′ end of each primer. For this hybridisation to proceed correctly, the primers must be well designed to promote a high specificity and efficiency. This is what characterizes the PCR method as a gold standard for molecular diagnosis and other applications. Specificity is related to the frequency with a mispriming event takes place, as primers are designed for specific regions, whereas efficiency is how close the primers are able to amplify the target in order to double the product in each cycle [145]. Usually it uses two primers (a primer pair) — forward and reverse or sense and antisense —, which will hybridise in a specific region of the DNA molecule. If the conditions for primers are different, errors can occur, e.g. only one of them to hybridise [78]. Poor primer design can lead to undesirable effects, such as secondary structures [37] and primer-dimers [40, 146], which are two primers hybridised to each other completely or partially, leading to a non-specific amplification, see Figure 21. There are several factors that influence the efficiency of primers, leading to a loss in the target-specificity and, consequently, in the amplification of the target. A few factors are association and dissociation kinetics of primer-target duplex [37, 51, 147], the presence of one or more mismatches between primer and target [147], see section 2.2, DNA polymerase functionality due to mismatch presence in the 3′ terminal [47, 51] and the concentration of primers [148]. When the latter is larger than optimal, primer-dimers can be formed and interfere with the target-specificity of PCR [148].

The primer design depends on a proper length, base composition, GC content, possible primer-dimers, secondary structure and melting temperature $T_m$ [23, 78, 149]. The latter is related to the annealing temperature $T_a$, as we mentioned earlier, which is the ideal temperature to hybridise to the target. If $T_a$ is too low, non-specific DNA fragments can be amplified or secondary structure formed [44, 146], whereas if $T_a$ is too high, the amount of product is reduced due to the poor annealing of primers [146]. The melting temperature of the primers must be

designed in a such way to lead a good performance and, consequently, the PCR amplification. For this goal, the $T_m$ of primers should be ideally the same or quite similar [44, 78, 141, 145], which would avoid false hybridisation [78], since both primers should hybridise to the target in the same PCR stage, that is, at same temperature. Depending on the author, a difference of temperatures within 1–5 °C is considered acceptable [23, 140, 142]. Here we have treated only one primer pair, however, more than one can be used to detect other targets in the same reaction, e.g. multiplex PCR [23, 141]. In this case, more than one specific region can be detected and each primer pair should be follow the same primer design guidelines. Moreover, the performance of PCR may be affected by several factors, such as non-specific amplification [40], sample collection and contamination [39], substance interference [39], reactivity with other non-target [40, 150], manual errors [39], less precision in low concentration samples [151] and proper collection of specimens [152].

## 4.2 Loop-mediated Isothermal Amplification (LAMP)

In 2000, Notomi et al. [42] reported a new nucleic acid amplification technique called loop-mediated isothermal amplification (LAMP). This method uses a DNA polymerase and four primers — two outer, or displacement, (F3 and B3) and two inner (FIP and BIP) primers —, these latter are called forward inner primer and backward inner primer, respectively. The inner primers are formed by two different sequences — F1c and F2 for FIP, and B1c and B2 for BIP —, which correspond to sense and antisense sequences of the DNA target. The outer and inner primers act at the beginning of the reaction, but just the inner primers act in later cycles. The LAMP reaction occurs in an isothermal condition, that is, at a constant temperature, which makes the use of thermal cyclers unnecessary. It runs in a temperature at 60–65 °C [42]. Two years later of LAMP report, Nagamine et al. [153] optimised the technique including two loop primers — LF and LB, respectively, Loop F (forward) and Loop B (backward) — to accelerate LAMP reaction. The presence of loop primers decreases the time to run the reaction and increases the sensitivity [153]. Thereafter, LAMP technique can use from four to six primers and is able to recognize between six and eight different regions in the DNA target [154].

As well as PCR method, there is a LAMP-based technique that applies the reverse transcription, which is the reverse transcription loop-mediated isothermal amplification (RT-LAMP). Here the briefly description of LAMP method will be within this version, which is widely applied to the detection of viral RNA, e.g. human coronaviruses [99, 136, 155–158]. The previous step of RT-LAMP is the transcription of RNA template into cDNA. From this point, the standard LAMP process is carried out. The first step is the convertion of the RNA/DNA double strand into RNA and DNA single strands. Now, the DNA molecule is the template used to the amplification. The F3 and FIP primers hybridise to the template and are amplified by Bst DNA polymerase, similar to PCR extension stage. F3 primer dissociates the newly strand (amplicon) from the template in a zip-like behaviour. Then this newly strand is hybridised to B3 and BIP

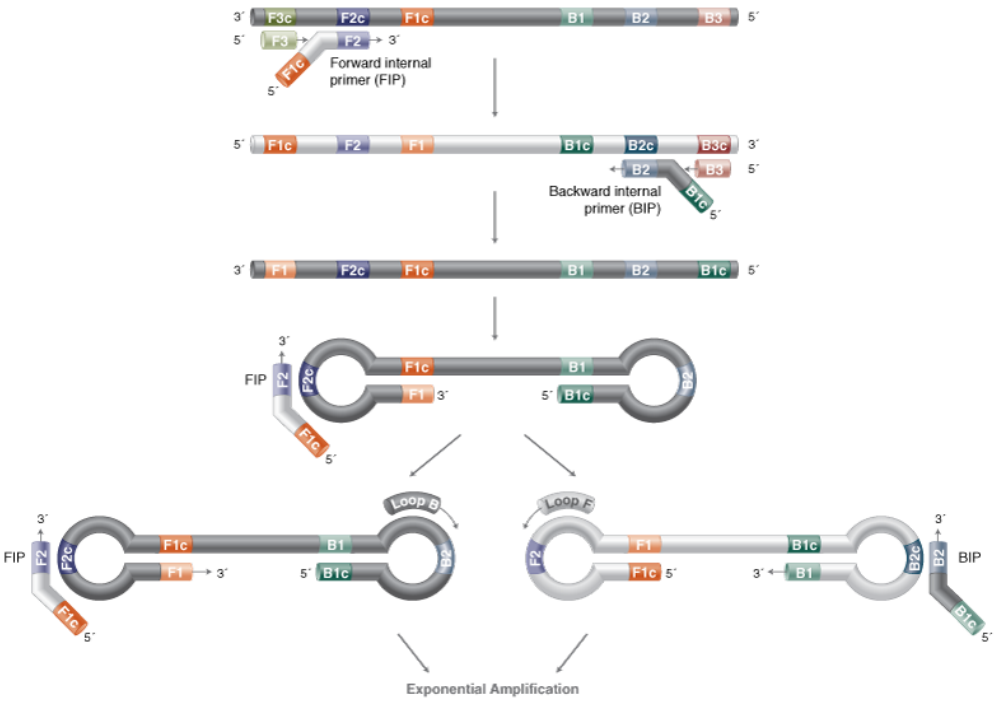



**Figure 22**
LAMP method. The markers in the template (F3c, F2c etc) show the target regions. The c's in regions and primers mean "complementary" (for instance, B3 primer hybridises to B3c region). F3 and FIP primers hybridise to the template and a newly strand is generated. B3 and BIP primers hybridise to it (the white one) and the amplification goes on. The newly formed strand folds into a loop-shaped form and Loop B (LB), Loop F (LF), FIP and BIP primers hybridise to the products proceeding to exponential amplification. Figure taken from Ref. [160].

primers and the process is similar to the previous one. DNA polymerase amplifies the BIP primer and B3 primer dissociates this newly strand. For both new strands, in each one, the two ends are hybridised by the inner regions forming loop-shaped strands. Thereafter, the amplification runs in an exponential behaviour [136, 154], see Figure 22. The entire process is carried out at a constant temperature (isothermal) [42, 153, 154] and the amplicons can be detected by precipitation of magnesium pyrophosphate, gel electrophoresis, real-time fluorescence, among other methods [99, 136, 159].

The LAMP method is very robust, fast, simple, high sensitive and specific, cost effective [154, 155, 161] and its performance is similar to conventional PCR-based methods[162, 163]. It can work at various pH values and be stable at room temperature, reagents are relatively inexpensive and it is carried out in less than an hour [164–166]. Furthermore, not using a thermal cycler contributes to a portable equipment, unlike to PCR [136, 161, 167], which is useful to LAMP method as a point-of-care (POC) [168–170]. Another difference between LAMP and PCR is the first is less sensitive to inhibitor substances than the second [154, 161, 171]. As LAMP is highly specific, it usually does not produce non-specific products [42, 153]. In particular, when loop primers are used, non-specific products are harder to be generated due to the

DNA final products include sequences that are recognized by these specific primers [153]. This feature of LAMP cooperates to a distinct recognition of multiple targets simultaneously, which is called one-pot multiplexing, and applied in LAMP-based techniques [154, 172, 173]. Also LAMP primer design depends on a few factors, such as concentration, location of primer pair, distance between target regions, temperature at 60-65°C and not form primer-dimers [174].

Until 2008, in Japan, only LAMP reagents were approved for in vitro diagnosis of SARS-CoV-1 [175]. Since its development, LAMP technique has been applied as molecular diagnostic method for several diseases, such as malaria [176], zika [177], salmonella [178], dengue [179], ebola [180] and HIV [181]. In addition, the human coronaviruses SARS-CoV-1 [157, 175], MERS-CoV [182], hCoV-NL63 [156] and SARS-CoV-2 [162, 166, 183], the responsible virus of the pandemic of 2020. For the latter, LAMP has also been applied in a combination to CRISPR-Cas12a [184–189] and nanopore sequencing technology [190].

# 5 Objectives

## 5.1 General Objective

Assessment of the impact of DNA mismatched base pairs in polymerase chain reaction (PCR) and loop-mediated isothermal amplification (LAMP) primers hybridised to genomes of SARS-CoV-2 coronavirus and primer effectiveness to variants.

## 5.2 Specific Objectives

1. Predicting the melting temperatures for matched and mismatched hybridisations between PCR/LAMP primers and SARS-CoV-2 and its variants;
2. Evaluating the influence of DNA mismatches through a coverage calculation of melting temperatures of the alignments;
3. Assessing the effectiveness of PCR/LAMP primers to SARS-CoV-2 variants;
4. Checking a possible cross-reactivity of primers to SARS-CoV-1 and other coronaviruses genomes.

# 6 Methods

In this chapter, we describe the main theoretical models used in this work. We show the Smith-Waterman algorithm used to align sequences; the two-state model, which is a theoretical analysis of denaturation and renaturation processes; the mesoscopic model applied to evaluate the thermodynamic properties of DNA duplexes; the calculation of melting temperature, and the parameters for mismatched base pairs.

## 6.1 Sequence Alignment

Sequence alignment is a method for searching similarities and differences between nucleotide or amino acid sequences. This comparison is made between two or more sequences to deduce structural and/or functional analogies and evolutionary relationships, e.g. homology (both have a common ancestor, and may show the same function or not). The aim of a sequence alignment is to reach a maximum number of associations between the bases in each sequence, that is, an optimal alignment [19].

The alignment is quantified by assigning scores to each type of base association: match, mismatch and gap. Matches (blue) are the perfect alignment when both bases are compatible; mismatches (red) are distinct bases; and gaps (green) are bases not aligned, that is, deletions or insertions:

ACACGAA-GG
ACACTAACGG

The scores used for the three types of association vary according to each alignment algorithm. A conventional scoring set is $+1$ for a match, $-1$ for a mismatch and $-2$ for a gap. See the score for the previous example alignment

A C A C G A A - G G
A C A C T A A C G G
+1 +1 +1 +1 -1 +1 +1 -2 +1 +1 = +5

An important point regarding sequence alignment is understanding the meaning of the score obtained. A high score may show an optimal alignment, but it does not determine a biological significance. To achieve a good agreement between score and biological significance, substitution matrices have been developed to take into account the substitution rates of nucleotides or amino acids throughout evolution. Two common substitution matrices are Point Accepted Mutation (PAM), developed by Margaret Dayhoff, and BLOcks of Amino Acid SUbstitution Matrix (BLOSUM), both applied to protein sequences [19].

It is possible to align sequences in a global or local way. Global alignments take an end-to-end alignment to find the sequence similarity. On the other hand, local alignments are sequence

regions of similarity. An advantage of the local alignment is the localisation of functional sites, amino acid residues that perform functional roles in proteins, in short regions of sequences. This may produce alignment results with great biological significance.

AC- -CGAATGG
ACTCCTAA-GA

Global alignment

ACCTGAGTAC- -CGAATGGAGGTACCTC
ACTCCTAA-GA

Local alignment

There are different sequence alignment algorithms. Two famous are the Needleman-Wunsch [191] and Smith-Waterman [192] algorithms. The first is a pioneer algorithm for global alignments of nucleotide and amino acid sequences. The second is a variation from the first which takes local alignments. In this work, we applied the Smith-Waterman algorithm to align primers/probes to genomes, since we can consider the short length of primers and probes as a short target region in the genome sequence, that is, a local alignment.

### 6.1.1 Smith-Waterman Algorithm

The Smith-Waterman algorithm [192] is used to align two sequences of nucleic acids or proteins. It performances local alignments to achieve the optimal alignment and is based on dynamic programming, which finds the solution of subproblems to reach the general solution of the problem. It compares two sequences

$$A = a_1, a_2 \dots a_n \tag{1}$$

and

$$B = b_1, b_2 \dots b_m \tag{2}$$

from which $s(a, b)$ is a similarity between sequence elements $a$ and $b$. A gap of lenght $k$ is given a weight $W_k$ and a scoring matrix $H$ is set up

$$H_{k0} = H_{0l} = 0 \tag{3}$$

where $0 \leq k \leq n$ and $0 \leq l \leq m$. The scoring matrix is filled

$$H_{ij} = \begin{cases} \max\{H_{i-1,j-1} + s(a_i, b_j)\}, \\ \max_{k \geq 1}\{H_{i-k,j} - W_k\}, \\ \max_{l \geq 1}\{H_{i,j-l} - W_l\}, \\ 0 \end{cases} \tag{4}$$

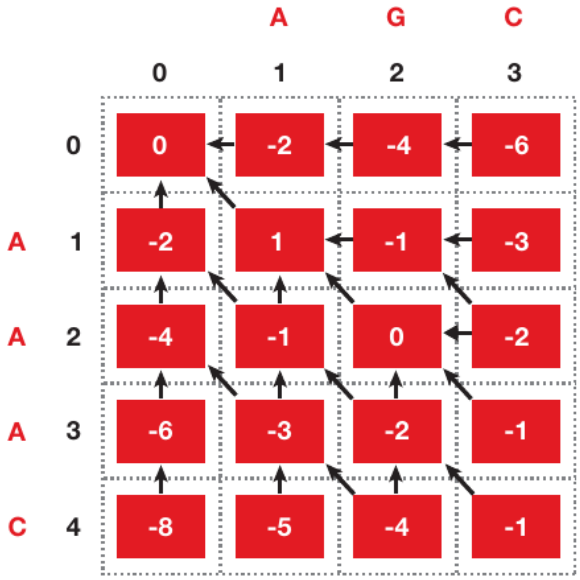



**Figure 23**
Smith-Waterman alignment traceback. From $-1$ (CC) in the inferior right space to 0 in the superior left space. The result alignment is AAA-C/--AGC. Figure adapted from Ref. [19].

where $1 \leq i \leq n$ and $1 \leq j \leq m$. $H_{ij}$ is interpreted as the maximum similarity between two subsequences ending in $a_i$ and $b_j$.

A traceback procedure is carried out that identifies the aligned subsequence and set up the best alignment. By locating the first maximum element of $H$, the subsequence with maximum similarity is found. Next, the elements with maximum value is sequentially determined until the element equal to zero, see Figure 23. Consider the sequences $g = AAAC$ e $p = AGC$ with $m = 4$ and $n = 3$. The score for a match is $+1$, a mismatch $-1$ and a gap $-2$. Given $g[i]$ and $p[j]$, and $H_{[i,j]}$ the optimal alignment between them, the score for the subsequences ending in $a_i$ and $b_j$ is given by equation 4. The maximum score determined is the score of $H_{[i,j]}$. Figure 23 shows a $(m+1) \times (n+1)$ matrix for the alignment, where the element $(i, j)$ is the alignment between subsequences $g[i]$ and $p[j]$. The result of this example alignment is

A A A - C
- - A G C

where blue is a match and green is a gap.

In this work, the Smith-Waterman algorithm was adapted to consider a match as a canonical base pair (AT and CG) and a mismatch a non-canonical base pair. We take a score 2 for matching base pairs, a score -1 for mismatched base pairs and no gaps are taken into account.

## 6.2 Two-State Model

The main idea of two-state model is DNA double strands in solution are either completely hybridised (double helical state) or completely dissociated (denatured state) [193]. It is due to the temperature of transition between full hybridisation and dissociation that occurs over a very short temperature interval. Considering a DNA duplex formed by a strand $X$ and a strand $Y$,

the process of denaturation in solution can be interpreted as

$$X \cdot Y \iff X + Y \tag{5}$$

Both in hybridisation and dissociation processes, there are strands connected and dissociated in the solution. It is called total strand concentration or "species concentration"

$$C_t = [X] + [Y] = [X] + [Y] + 2[X{\cdot}Y] \quad , \tag{6}$$

where $[X]$ and $[Y]$ are the concentration of strands dissociated and $[X \cdot Y]$ is the concentration of strands associated. The equilibrium constant $K_{eq}$ characterises the species concentration

$$K_{eq} = \frac{[X][Y]}{[X{\cdot}Y]} \tag{7}$$

The hybridisation temperature related to the concentrations of the states of DNA molecule

$$[X{\cdot}Y] = [X]_m = [Y]_m \tag{8}$$

If all DNA sequences of a dataset are non-self-complementary (both strands of the duplex are distinct when they are read from $5' \longrightarrow 3'$), the equilibrium constant is $K_{eq} = \frac{C_t}{4}$ from equation 7. In the equation of Gibbs free energy $\Delta G$

$$\Delta G = -RT_m \ln K_{eq} = \Delta H - T_m \Delta S \quad , \tag{9}$$

where $R$ is the universal constant of the gases, $\Delta H$ and $\Delta S$ are enthalpy and entropy variations, respectively, and $T_m$ is the melting temperature. Rearranging the equation 9, the straight relation to $T_m$ is

$$T_m = \frac{\Delta H}{\Delta S - R \ln C_t/4} \tag{10}$$

Usually $\Delta H < 0$ and $\Delta S < 0$ for canonical sequences (only AT and CG bonds) and the equation 10 is applied to non-self-complementary sequences.

For self-complementary sequences (both strands are identical from $5' \longrightarrow 3'$), $K_{eq} = C_t$ and the relation of concentrations of dissociated and associated strands

$$[X{\cdot}X]_m = [X_m] \tag{11}$$

From this, the equation 10 related to melting temperature changes to

$$T_m = \frac{\Delta H}{\Delta S - R \ln C_t} \tag{12}$$

Up to now, our discussion was in relation to the denaturation process. For the reverse process

(renaturation or hybridisation), the equilibrium constant $K_{eq}$ is

$$K_{\text{desn}} = \frac{1}{K_{\text{hibr}}} \quad , \tag{13}$$

where $K_{\text{desn}}$ is the equilibrium constant to the denaturation and $K_{\text{hibr}}$ is the equilibrium constant to the hybridisation. From equation 13, the equation 10 is now

$$T_m = \frac{\Delta H}{\Delta S + R \ln C_t/4} \tag{14}$$

Usually $\Delta H > 0$ and $\Delta S > 0$ for canonical pairs. For non-self-complementary, equation 14 is valid, while for self-complementary

$$T_m = \frac{\Delta H}{\Delta S + R \ln C_t} \tag{15}$$

The equation 13 can be rearranged similar to the equation of a straight line $(x = a + by)$

$$T_m{}^{-1} = \frac{\Delta S}{\Delta H} - \frac{R}{\Delta H} \ln C_t/4 \quad , \tag{16}$$

where $T_m{}^{-1}$ is $x$ and $\ln C_t/4$ is $y$ from the equation of a straight line. The melting temperature can be measured from different concentrations and $\Delta H$ and $\Delta S$ can be obtained fitting data in equation 16. From experiments, we can design the relation between melting temperature $T_m{}^{-1}$ and concentration $C_t$ as in Figure 24.

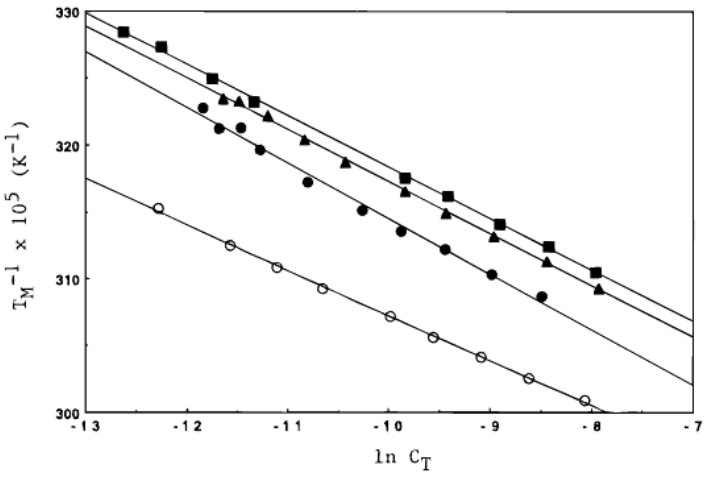



**Figure 24**
Melting temperature $T_m{}^{-1}$ can be measured from different concentrations $C_t$. In the graph, each straight line is a distinct sequence. Figure taken from Ref. [194].

## 6.3 Peyrard-Bishop (PB) Mesoscopic Model

Mesoscopic models work at a size scale between microscopic and macroscopic — the mesoscale, see Figure 25. In 1989, Peyrard and Bishop [195] proposed a statistical model to describe the physical properties of DNA molecule, particularly, the thermal denaturation.

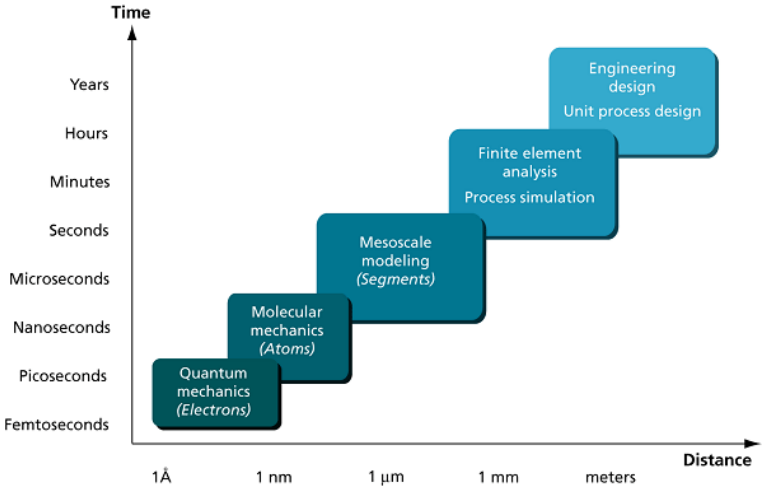



**Figure 25**
The mesoscale is a size scale between micro- and macroscale. Figure taken from Ref. [197].

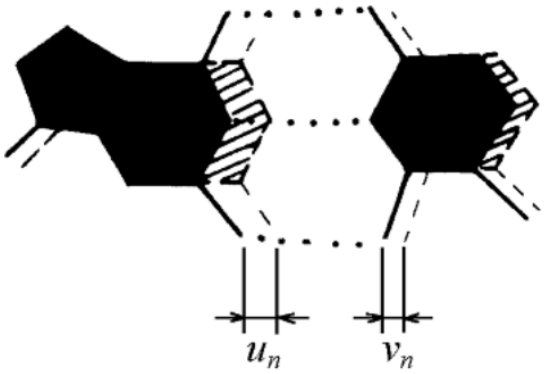



**Figure 26**
Degrees of freedom, $u_n$ and $v_n$, of a displacement between two nitrogenous bases. The dark pair is the new position of the base pair, that is, it is its displacement from an equilibrium point. Figure taken from Ref. [198].

This model is simple and increases the computational efficiency. It considers the DNA duplex as a two-dimensional planar structure formed by homogeneous bases (similar base pairs), which are concerned as particles. Then each base, as a particle, has a mass $m$ [195, 196]. The Peyrard-Bishop (PB) model considers independently the hydrogen bonds in base pairs and stacking interaction between neighbouring bases [195, 196]. In addition, it can predict melting temperatures from given sequences.

The dissociation of the DNA strands are determined by the method of transfer integral [195, 196]. There are two degrees of freedom ($u_n$ and $v_n$) to the base pair, which are the base displacements along the hydrogen bond direction from the equilibrium point [195], see Figure 26. In Figure 27, the two degrees of freedom are shown in a representation of the model and their equations are

$$u_n = \frac{x_n + y_n}{\sqrt{2}}, \quad v_n = \frac{x_n - y_n}{\sqrt{2}} \tag{17}$$

The PB model approximates the hydrogen bonds potential of the base pair to the Morse potential, which represents two or three bonds [195]. The Morse potential $V(y)$ [199] is shown in Figure 28 and

$$V(y) = D\left(e^{-y/\lambda_a} - 1\right)^2 \quad , \tag{18}$$

where $D$ is the dissociation energy of the base pair and $\lambda_a$ is the parameter which defines the reach of the potential [200]. The stacking interaction is approached to an harmonic coupling,

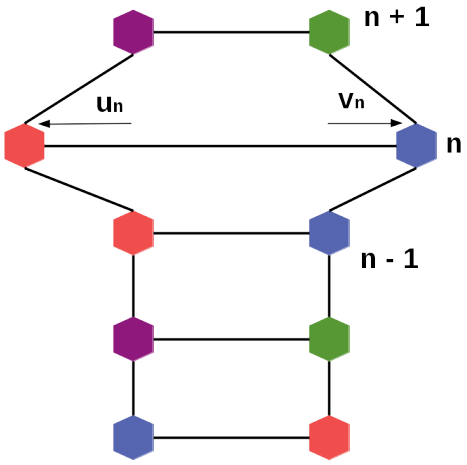



**Figure 27**
Representation of the degrees of freedom, $u_n$ e $v_n$, in the PB model. From top to bottom, in the second base pair, the base displacements along the hydrogen bonds direction from the equilibrium point.

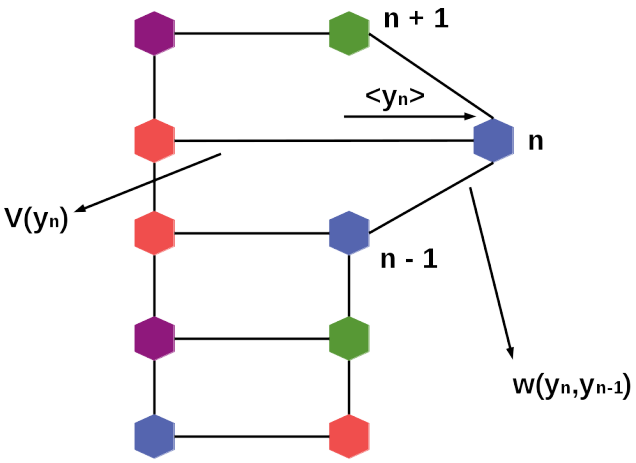



**Figure 28**
Representation of PB model. In relation to a base $n$: Morse potential, as the hydrogen bonds, is $V(y_n)$, base stacking along the strands is $w(y_n, y_{n_1})$ and the base displacement in the direction of hydrogen bonds is $\langle y_n \rangle$.

see Figure 28. The elastic constant $k$ is the same to all base interactions along the strand [195]

$$w(y_n, y_{n-1}) = \frac{k}{2}(y_n - y_{n-1})^2 \tag{19}$$

A Hamiltonian combines the potential to the hydrogen bonds (Morse potential) and the stacking interactions (harmonic potential)

$$H = \sum_n \frac{1}{2}m\left[\dot{u}_n^2 + \dot{v}_n^2\right] + \frac{1}{2}k_u\left(u_n - u_{n-1}\right)^2 + \frac{1}{2}k_v\left(v_n - v_{n-1}\right)^2 + D\left[e^{-\frac{a}{\sqrt{2}}(u_n - v_n)} - 1\right]^2 \tag{20}$$

The first term is the kinetic energy of DNA molecule, the second and third terms describe the stacking interactions as coupled harmonic oscillators and the fourth term describes the hydrogen bonds as the Morse potential. The motion of the double-stranded DNA can be described by a coordinate transformation by $x_n$ and $y_n$

$$x_n = \frac{u_n + v_n}{\sqrt{2}}, \quad y_n = \frac{u_n - v_n}{\sqrt{2}} \quad , \tag{21}$$

where $y_n$ is the displacement along the hydrogen bond direction [195]. Then, the new Hamiltonian is obtained from these variables

$$H = \sum_n \left[\frac{1}{2m}p_n^2 + \frac{1}{2m}q_n^2 + w(x_n, x_{n-1}) + w(y_n, y_{n-1}) + V(y_n)\right] \quad , \tag{22}$$

where $p_n = m\dot{x}_n$ and $q_n = m\dot{y}_n$ are the momenta to $x_n$ and $y_n$. From equation 22 comes the classical partition function $Z$ calculated at a fixed temperature for a DNA molecule with $N$ base

pairs

$$Z = \int \prod_n e^{-\beta H} dp_n dq_n dx_n dy_n = Z_p Z_q Z_x Z_y \quad , \tag{23}$$

where $\beta = 1/(k_B T)$, in which $k_B$ is the Boltzmann constant. From the integrals of momenta $p_n$ and $q_n$

$$Z_p = Z_q = (2\pi m k_B T)^{N/2} \tag{24}$$

The follow equation characterizes the contribution of the potential energy in an "harmonic strand" $Z_x$

$$Z_x = \left(\frac{2\pi}{\beta k}\right)^{N/2} \tag{25}$$

The stacking interaction considers only the interaction between neighbouring bases $Z_y$ [195, 200]

$$Z_y = \int dy_1 \int dy_2 \ldots \int dy_N \prod_n e^{-\beta w(y_n, y_{n-1})} e^{-\beta V(y_n)} \tag{26}$$

Finally, the PB model is able to calculate the average opening $\langle y_m \rangle$, or average displacement between the bases, of DNA strands [195, 201]. The value of average opening indicates that DNA molecule is denatured or not. The average displacement between the hydrogen bonds can be calculated from Morse and harmonic potentials

$$\langle y_m \rangle = \frac{1}{Z} \int \prod_n y_m e^{-\beta H} dy_n = \frac{1}{Z_y} \int dy_1 \int dy_2 \ldots \int dy_N y_m \prod_n e^{-\beta w(y_n, y_{n-1})} e^{-\beta V(y_n)} \tag{27}$$

As we mentioned earlier, Peyrard and Bishop considered the DNA molecule as homogeneous (similar base pairs) in their model [195]. This means that Morse potential and stacking parameters are the same for all base pairs. However, a review of the model led to new versions in which the DNA molecule is considered as heterogeneous [202] — the base pairs are distinct (AT and CG pairs). This is the configuration that we used in this work. We considered all sequences analysed as heterogeneous DNA.

## 6.4 Calculation of Melting Temperature

The PB model is able to predict melting, or hybridisation, temperature $T_m$. It takes the thermal equivalence of DNA sequences applying a index $\tau$, which clusters these sequences by similar melting temperatures [203]. The indexes of experimental melting temperatures are fitted, then new temperatures can be calculated by a linear regression of the indexes. It is also the way that the model refines the parametrization predicting melting temperatures from experimental data.

The classical partition function $Z$ is calculated only once and expanded in non-diagonal terms of the integral transfer matrix. The index $\tau_i$ to each sequence $i$ is the largest expansion

term and the expansion of partition function is

$$Z = Tr(C^{(1,2)}C^{(2,3)}...C^{(N,1)}) \quad , \tag{28}$$

where $C^{(n,n+1)}$ is the interaction matrix between neighbouring base pairs and $C^{(N,1)}$ is the boundary condition (the first and last base pair bound) [203]. The boundary condition considers two situations: (1) the DNA sequence as a ring in which the first and last base pairs are connected, and (2) this last bonding does not exist.

Considering the hydrogen bonds to AT and CG base pairs as weak ($w$) and strong ($s$), respectively. There are four types of neighbours and the corresponding matrices are $C^{(w,w)}$, $C^{(w,s)}$, $C^{(s,w)}$ and $C^{(s,s)}$ [203]. The matrix $C^{(s,s)}$ uses, as the orthonormal base, the matrices of CG pair to convert to a diagonal matrix $\Lambda$, which keeps the eigenvalues $\lambda_i$ of homogeneous sequence of CG. The matrix $\Delta^{(a,b)}$ to the interaction difference between neighbouring bases of type (a,b) and type (s,s)

$$C^{(a,b)} = \Lambda + \Delta^{(a,b)} \tag{29}$$

From equation 29, rearranging the partition function

$$Z = Tr[(\Lambda + \Delta^{(1,2)})(\Lambda + \Delta^{(2,3)})...(\Lambda + \Delta^{(N,1)})] \tag{30}$$

When trace properties are applied, the partition function is

$$Z = \sum_{\omega=0}^{N} Z_\omega(\Lambda) = \sum_{\omega=0}^{N} Tr[M(\Lambda^\omega)] \quad , \tag{31}$$

where $M(\Lambda^\omega)$ are all terms which have $\omega$ multiplications of matrix $\Lambda$. As it was considered the CG homogeneous sequence, we have $\Delta^{(s,s)} = 0$ and only the $\Lambda^N$ terms.

Figure 29a shows four sequences of 10 base pairs, from 40% to 60% of GC content, and a normal distribution to each sequence, calculated at 370 K, which is a "temperature of calculation". Figure 29b shows the maximum point $\omega_{max}$ of each curve, which is dependent on the sequence content, but it is not strongly dependent on temperature [203]. As an example, Figure 29c shows sequences of different lengths that were fitted by linear regression to each set of the same length. We can observe the relation between the parameter $(\omega_{max})^{1/2}$ and the experimental melting temperature. From this, $(\omega_{max})^{1/2}$ can be used as a dimensionless value to the thermal equivalence, which is called melting index $\tau$ [204]. In addition, the regression is dependent on $N^{1/2}$.

As shown in Figure 29b, the melting indexes $\tau_i$ are correlated to experimental melting temperatures. Then the linear regression is applied for each set of sequences of the same length

$$T'_i = a_0(N) + a_1(N)\tau_i \quad , \tag{32}$$

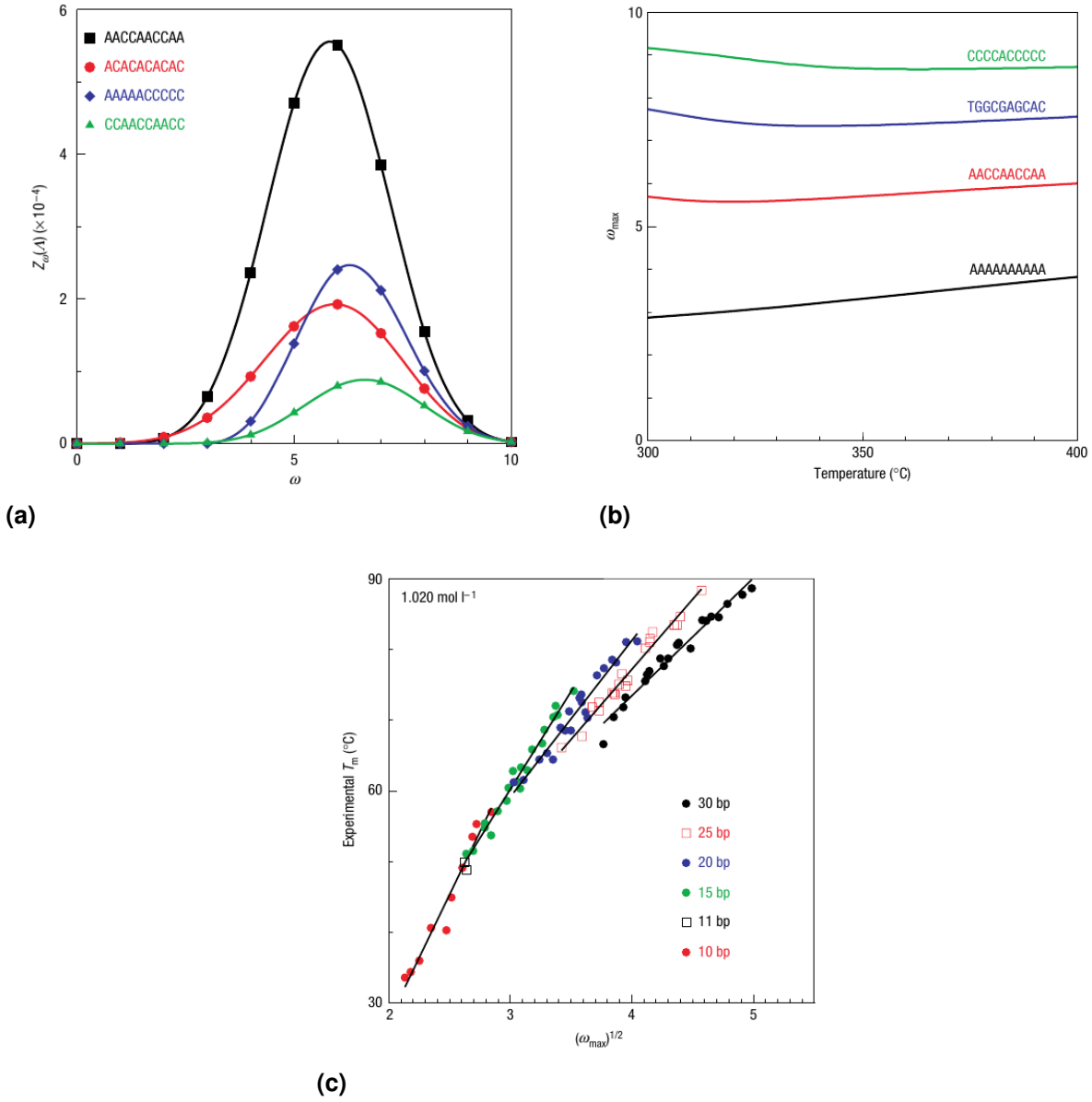



**Figure 29**

(a) Normal distribution to four different sequences formed by 10 base pairs. The GC content of each sequence is between 40% and 60% and was calculated at the temperature of 370 K. (b) Relation between parameter $\omega_{max}$ and temperature. Four different sequences formed by 10 base pairs. (c) Parameter $(\omega_{max})^{1/2}$ versus experimental temperature $T_m$. The sequences are 10–30 base pairs in size. The lines are the linear regression to each group of sequences in the same length, except to two points of data from sequences of 11 base pairs. Figures taken from Ref. [203].

where $T'_i$ is the temperature predicted to sequence $i$, $N$ is the sequence length and $a_0(N)$ and $a_1(N)$ are the linear regression coefficients to each set $N$. If there is not a reference in the dataset for a sequence of different length, due to the dependency with $N^{1/2}$, a new linear regression is applied

$$a_k(N) = b_{0,k} + b_{1,k}N^{1/2} \tag{33}$$

## 6.5 Parameters to Mismatched Base Pairs

In this work, we used parameters for up to three consecutive mismatched base pairs obtained from Ref. [61]. From equation 32, the authors modeled an experimental melting temperatures set of 4096 sequences of which 4032 containing single, double and triple consecutive mismatches. The buffer conditions of sequences melting were of 50 mM sodium chloride, 10 mM sodium phosphate, pH 7.4 and total strand concentration 1.0 $\mu$M. They analysed the sequences in nearest-neighbour and next-nearest neighbour contexts and compared with 64 sequences containing only canonical base pairs under the same conditions. The parameters obtained were shared in two groups: (1) context independent, which considers all mismatches as independent from the flanking pairs, and (2) context dependent, which considers the flanking pairs of mismatches, see Figure 30.

Canonical – Mismatch – Canonical ⟶ single mismatch

5′ – ATCTGAGAGTCCACCAAACG – 3′ ⟶ $T_m = 52.0°C$
3′ – TAGACTCCCAGGTGGTTTGC – 5′

Mismatch – Mismatch – Canonical ⟶ double mismatch

5′ – AGTTGAGGGTCCACCAAACG – 3′ ⟶ $T_m = 55.5°C$
3′ – TAGACTCCCAGGTGGTTTGC – 5′

Mismatch – Mismatch – Mismatch ⟶ triple mismatch

5′ – CCAGCTCATGACCATGAGGT – 3′ ⟶ $T_m = 43.5°C$
3′ – GGTATTGTACTGGTACTCCA – 5′

**Figure 30**
A central mismatch in three different context: canonical, mismatch/-canonical and mismatch contexts, in other words, single, double and triple mismatches. Note that first two hybridisations show the same 3′ ⟶ 5′ primer.

Oliveira et al. [61] found most of the Morse potentials to mismatched base pairs are small, which is expected considering the role of mismatches in the destabilisation of DNA duplex. However, almost all the largest Morse potentials were found in the case the mismatch (MM) pair is flanked on both sides by other mismatches, except the GT base pair, see Figure 31. Morse potentials for AA mismatches in CAC/GAG and AAC/TAG trimers suggest a single hydrogen bond to mismatched pair, which is comparable to previous works [31, 79]. For TT pair flanked by CG pair, nuclear magnetic resonance (NMR) experiments have shown none or a single hydrogen bond [79, 205], which is consistent with the results found for this configuration. AG pair in AAC/TGG trimer were reported to be in both *anti-syn* and *syn-anti* conformations with two hydrogen bonds by X-ray diffraction [206]. Morse potential found to AG pair is con-

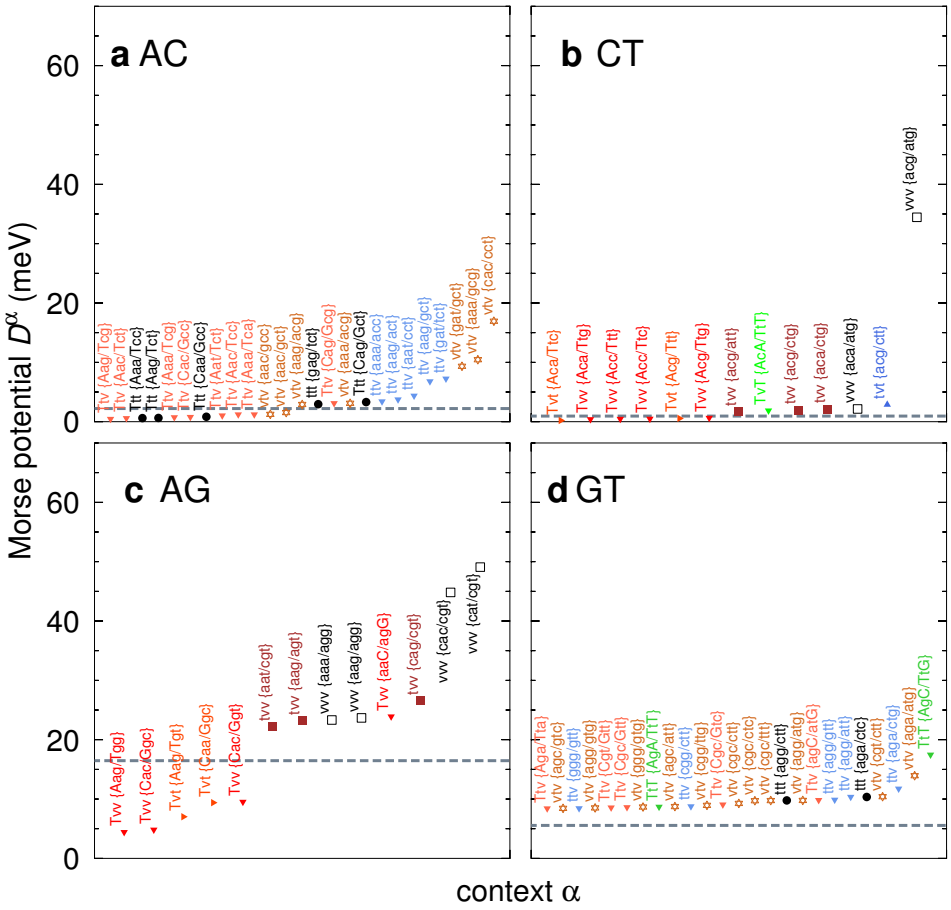



**Figure 31**
Morse potentials ($D^{\alpha}$) for (a) AC, (b) CT, (c) AG and (d) GT mismatched base pairs in a context dependent ($\alpha$). Shown are the Morse potentials that deviate by more than 30% (a–c) and 50% (d) from the seed context independent potentials (dashed grey lines). T represents AT and CG base pairs; v a transversion (purine $\longrightarrow$ pyrimidine or pyrimidine $\longrightarrow$ purine) and t a transition (purine $\longrightarrow$ purine or pyrimidine $\longrightarrow$ pyrimidine). Figure taken from Ref. [61].

sistent to a double hydrogen bond as well. Tandem GA-AG mismatches have been reported to have a strong stacking interaction [207–209], which was shown by X-ray diffraction experiments [207]. A large stacking potential was found for GA-AG mismatches in comparison to canonical base pairs, 17.3 eV nm$^{-2}$ and 2–4 eV nm$^{-2}$, respectively.

Melting temperatures obtained to single, double and triple contiguous mismatches showed different situations. Oliveira et al. [61] found that GG mismatched pair in G<u>G</u>C/C<u>G</u>G trimer has a melting temperature $T_m = 65.9$ °C lower than the canonical trimer GGC/CCG with $T_m = 68.9$ °C. It is the case of most mismatched pairs, however, they found double mismatches with a melting temperature above a single one. For instance, the double mismatched A<u>C</u>T/<u>G</u>G<u>G</u> with melting temperature $T_m = 54.4$ °C, while the single mismatched A<u>C</u>T/T<u>T</u>A has $T_m = 51.5$ °C. A considerable number of mismatched pairs obtained low melting temperatures, such as the trimers <u>ACC</u>/<u>GAA</u> and A<u>GC</u>/T<u>GC</u> with melting temperature $T_m = 44.1$ °C. Mismatched pairs with a melting temperature similar to canonical trimers were also found. In

particular, the triple mismatched trimers GAA/AGG and GTA/AGG with $T_m = 60.7$ °C, the same temperature as the canonical ATT/TAA. In addition, their results show that 15% of single mismatches have higher melting temperature than canonical AT base pair, which suggest that MM may contribute to some applications, such as single nucleotide polymorphism (SNP) identification and primer/probe design [37].

# 7 Evaluation of PCR/LAMP Primers for SARS-CoV-2

In this chapter, we describe the work developed to evaluate the presence of DNA mismatches in polymerase chain reaction (PCR) and loop-mediated isothermal amplification (LAMP) primers, which have been applied to the detection of SARS-CoV-2 virus. We show the data collected; the workflow carried out to analyse the primers and probes; the results that we found and the discussion about them.

As the tables of Appendix are too large to be viewed here, by clicking on their coloured names you will be redirected to a web page of our group where the referenced material is stored.

## 7.1 Dataset

### 7.1.1 PCR Primers/Probes

We collected 19 primer/probe sets from Refs. 40, 132, 148, 210–225, representing a total of 340 oligonucleotides (297 primers and 43 probes). A few of these sets included primers from earlier published sets. For instance, primers from U.S. Centers for Disease Control and Prevention (CDC) [210] have been used in several publications. Therefore, we considered only primers were not repeated in earlier works. The sequences of all PCR primers and probes are shown in Table 10 in Appendix A.1.

### 7.1.2 LAMP Primers

We collected 18 LAMP primer sets designed to wild-type SARS-CoV-2 genomes that underwent clinical validation from Refs. 2, 118, 168, 169, 173, 183, 185, 226–236. Primers from earlier published sets were not considered. FIP and BIP primers were divided in F1c and F2, and B1c and B2 primers, respectively, except those from three sets [2, 227, 233], which indicated the division of primers, see section 4.2. We found all possible combinations of primers and selected those according to the temperatures of the same primer pair type from the three sets just mentioned. All the sets include 436 primers in total and their sequences are shown in Table 11 in Appendix A.2.

### 7.1.3 Viral Genomes

From the National Center for Biotechnology Information (NCBI) page [237] dedicated to SARS-CoV-2 coronavirus, we collected 21665 wild-type genomes on 8 October 2020. Considering some possible cross-reactivity, we collected 10 genomes of SARS-CoV-1 and 313 genomes of five other coronaviruses (229E, NL63, OC43, HKU1 and MERS-CoV) from NCBI main page [238]. Together with Vivianne Basílio Barbosa, an undergraduate student I co-supervised, we collected genomes of four variants: 7247 genomes of Alpha variant, 7497 of

Beta variant, 2308 of Gamma variant and 7943 of Delta variant. Alpha, Beta and Gamma genomes were collected on 7 April 2021 and Delta genomes on 5 June 2021 from GISAID page dedicated to the tracking of variants [239]. In addition, 7029 of Omicron variant on 16 December 2021; 6610 of Mu variant, 9340 of Lambda variant, 7393 and 348 of Omicron subvariants BA.2 and BA.3 on 11 February 2022; 629 and 1231 of Omicron subvariants BA.4 and BA.5 on 19 September 2022, at GISAID [239]. The accession codes of all groups of genomes are shown in Table 12–Table 25 in Appendix B.

## 7.2 Workflow

Here we describe in details the workflow developed to assess PCR and LAMP primers to the detection of SARS-CoV-2 virus. A software package to carry it out is available in Ref. 240 and the guidelines are in the documentation (`analyse-user-manual.pdf`), which is located in `analyse_mismatch_primers.tar.gz` in the same web page. Also step-by-step instructions on how to download, install, run and perform the analysis are described in the supplementary material in Ref. 1.

### 7.2.1 Primer/Genome Alignment

All genome sets mentioned in the last section 7.1 were aligned against each oligonucleotide of all 37 sets applying the Smith-Waterman algorithm [192], see section 6.1.1. There are two strand configurations in the alignment of primers/genomes. The first one is the genome sequence in the direction $5' \longrightarrow 3'$ as in the FASTA file collected from the database, and primer/probe in the direction $3' \longrightarrow 5'$ as a complementary strand

$$5' - \text{(genome sequence)} - 3'$$
$$3' - \text{(primer/probe sequence)} - 5'$$

The other strand configuration is the alignment of the complementary genome sequence in the direction $5' \longrightarrow 3'$ and primer/probe in the direction $3' \longrightarrow 5'$ as a complementary strand

$$5' - \text{(complementary genome sequence)} - 3'$$
$$3' - \text{(primer/probe sequence)} - 5'$$

We called "strictly matched" the completely alignments between genome and primer/probe, while "partially matched" was for up to three contiguous mismatches in the alignments. The limit of this maximum number is due to the available parameter only covering up to three consecutive mismatches [61]. Here four or more mismatches in the alignment is considered as not aligned, since it is known that four or more mismatches in a single primer may prevent the amplification [54, 241]. Also deletions in the viral genome, e.g. Omicron variant [126, 127], may lead to no alignment of the primers. An example of partial alignment is nCoV_IP2-12759Rv primer (bottom strand) to MW065008.1 genome

$$5' - \text{ACAACACGACAAAGGGAG} - 3'$$
$$3' - \text{TGTTGTGTTGTTTCCCTC} - 5'$$

where the mismatched base pair is underlined in red.

### 7.2.2 Hybridisation Temperature Prediction

Given all alignments, we predicted the hybridisation temperature $T_m$ to each alignment found in the conditions of strictly matched and single, double and triple consecutive mismatches. We applied the method described in section 6.4. We used parameters for DNA Watson-Crick base pairs and DNA mismatched base pairs at $Na^+$ 60 mM concentration.

For reference, the melting temperatures of primers and probes as a perfectly matched duplex were also predicted with the same method. We generated the complementary strand of each primer and then calculated the $T_m$ as a double perfect strand. Their melting temperatures were called reference temperature $T_{\text{ref.}}$, which will perform an essential role in the next section 7.2.3. All melting temperatures of PCR primers/probes and LAMP primers are shown in Table 10 and Table 11, respectively, in Appendices A.1 and A.2.

### 7.2.3 Coverage Calculation

We analysed how perfect matched and mismatched hybridisations cover or not the genomes collected. We selected the alignments in the conditions mentioned earlier and calculated the coverage for strict and partial alignments. First, for strictly alignment

$$C_{\text{strict}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{MM}}}{N_G} \tag{34}$$

where $N_G$ is the total number of genomes, considered only those above 25000 bp in size, $N_{\text{n.a.}}$ is the number of genomes not aligned to any primer/probe and $N_{\text{MM}}$ is the number of genomes aligned with single, double and triple consecutive mismatches. The limit for genome size is arbitrary to consider only complete genomes. Considering the melting temperature $T_{\text{MM}}$ for mismatched alignments, the difference to the reference temperature $T_{\text{ref.}}$

$$\Delta T_{\text{MM}} = T_{\text{ref.}} - T_{\text{MM}} \tag{35}$$

In general, $T_{\text{MM}}$ is lower than $T_{\text{ref.}}$ [61]. Given a melting temperature range $\Delta T_{\text{lim.}}$, we calculated the coverage for partial alignments

$$C_{\text{part.}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{low}}(\Delta T_{\text{lim.}})}{N_G} \tag{36}$$

where $N_{\text{low}}$ is the number of mismatched alignments according to

$$\Delta T_{\text{MM}} \leq \Delta T_{\text{lim.}} \tag{37}$$

where we used $\Delta T_{\text{lim.}} = 5$ °C, that is, we considered only mismatched alignments with $T_{\text{MM}}$ no more than 5 °C below the reference temperature $T_{\text{ref.}}$. All calculated coverages for both PCR and LAMP sets are shown in Appendices C and D, respectively.

The workflow described in this section 7.2 was performed by Vivianne Barbosa to the first four variants (Alpha, Beta, Gamma and Delta) regarding PCR primer/probe sets. The results found are shown in Table 26 and for SARS-CoV-1 and non-SARS genomes in Table 27, both in Appendix C. All genome sets, including seven variants and four subvariants, were evaluated in the same way to the LAMP sets and the results are shown in Table 28–Table 31 in Appendix D.

## 7.3 Results and Discussion

### 7.3.1 PCR Evaluation

We collected 19 PCR primer/probe sets and aligned them to genomes of SARS-CoV-2, Alpha, Beta, Gamma and Delta variants, SARS-CoV-1 and five other coronaviruses (229E, NL63, OC43, HKU1 and MERS-CoV). Given the alignments, the hybridisation temperatures of double strand of primers and probes and all alignments found for perfect matched and up to three consecutive mismatched alignments were calculated. In Table 1, we show the range of reference melting temperature of primers/probes and both strict and partial coverages for SARS-CoV-2 and its variants.

**Table 1**
Summary of results for 19 PCR primer/probe sets. Shown are the number of primers/probes $N_{pp}$ and the range of reference temperatures $T_{\text{ref.}}$ for each set, and the range of strict and partial coverages for SARS-CoV-2 and Alpha, Beta, Gamma and Delta variants, respectively.

| | | | SARS-CoV-2 | | Alpha Variant | | Beta Variant | | Gamma Variant | | Delta Variant | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Name of Set | $N_{pp}$ | $T_{\text{ref.}}$ (°C) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
| CDC [210] | 6 | 61.1–75.5 | 98.4–99.3 | 99.1–99.4 | 98.4–99.8 | 98.7–99.9 | 92.7–98.5 | 98.0–99.3 | 98.7–99.8 | 98.8–99.9 | 98.6–100 | 99.7–100 |
| WHO [211] | 21 | 51.3–70.3 | 97.9–99.4 | 97.9–99.5 | 90.1–99.9 | 90.1–99.9 | 98.1–99.5 | 98.1–99.5 | 98.9–100 | 98.9–100 | 98.0–100 | 98.0–100 |
| Luminex [212] | 6 | 59.4–81.0 | 64.9–99.5 | 65.1–99.5 | 0.152–99.5 | 0.152–99.5 | 0.120–99.5 | 0.120–99.5 | 0.563–99.7 | 0.563–99.7 | 0.0–100 | 0.0–100 |
| Corman et al. [213] | 21 | 61.7–81.7 | 0.0–99.4 | 0.0–99.4 | 0.0–99.9 | 0.0–99.9 | 0.0–99.4 | 0.0–99.6 | 0.0–99.9 | 0.0–100 | 0.0–100 | 0.0–100 |
| Davda et al. [214] | 16 | 56.5–70.3 | 95.3–99.4 | 95.3–99.4 | 97.6–99.8 | 97.6–99.8 | 95.4–99.3 | 95.4–99.3 | 96.3–99.8 | 97.0–99.8 | 98.1–99.9 | 98.1–99.9 |
| Grant et al. [215] | 2 | 62.4–79.9 | 97.5–99.3 | 97.5–99.5 | 0.138–99.7 | 0.138–99.7 | 96.0–98.2 | 96.0–98.4 | 99.3–99.6 | 99.3–99.6 | 99.4–99.9 | 99.4–100 |
| Hirotsu et al. [216] | 3 | 60.6–68.8 | 0.0–99.4 | 0.0–99.4 | 0.0–99.7 | 0.0–99.8 | 0.0–98.3 | 0.0–98.3 | 0.0–99.8 | 0.0–99.8 | 0.0126–99.8 | 0.0126–99.8 |
| Jalali et al. [217] | 6 | 61.3–64.2 | 98.7–99.4 | 98.7–99.4 | 99.5–99.8 | 99.5–99.8 | 98.6–99.8 | 98.6–99.8 | 99.7–100 | 99.7–100 | 99.8–100 | 99.8–100 |
| Lanza et al. [218] | 27 | 60.5–75.0 | 98.5–99.4 | 98.8–99.5 | 0.787–99.8 | 98.9–99.9 | 97.7–99.7 | 97.7–99.7 | 80.5–100 | 80.5–100 | 73.2–100 | 97.7–100 |
| Li et al. [132] | 2 | 67.5–70.4 | 98.2–99.2 | 98.3–99.2 | 99.7–99.8 | 99.8 | 97.9–99.3 | 97.9–99.3 | 99.0–99.2 | 99.1–99.2 | 99.7–99.9 | 99.7–100 |
| Lu et al. [219] | 3 | 64.0–74.7 | 98.6–99.4 | 99.3–99.4 | 99.6–99.7 | 99.7–99.8 | 99.2–99.4 | 99.3–99.4 | 99.8–99.9 | 99.8–99.9 | 99.9–100 | 99.9–100 |
| Munnink et al. [220] | 171 | 65.4–74.9 | 45.9–99.5 | 46.0–99.7 | 1.28–4.10 | 1.32–22.5 | 0.0934–99.7 | 0.0934–99.7 | 0.997–100 | 40.6–100 | 0.0378–100 | 0.0755-100 |
| Nalla et al. [221] | 4 | 51.7–68.0 | 0.0–99.5 | 0.0–99.5 | 0.0–99.8 | 0.0–99.8 | 0.0–99.1 | 0.0–99.1 | 0.0–99.5 | 0.0–99.5 | 0.0–100 | 0.0–100 |
| Niu et al. [40] | 6 | 59.2–84.2 | 0.0–99.4 | 0.0–99.4 | 0.0–99.9 | 0.0–99.9 | 0.0–99.4 | 0.0–99.4 | 0.0–99.8 | 0.0–99.9 | 0.0–100 | 0.0–100 |
| Park et al. [148] | 20 | 59.3–65.4 | 94.9–99.5 | 94.9–99.5 | 80.4–99.7 | 96.2–99.7 | 40.3–99.8 | 40.3–99.8 | 94.2–100 | 94.2–100 | 98.2–100 | 98.2–100 |
| Rahman et al. [222] | 6 | 64.2–76.1 | 97.6–99.4 | 97.6–99.4 | 96.7–99.8 | 96.7–99.8 | 98.2–91.7 | 98.5–91.7 | 93.9–99.8 | 93.9–99.8 | 98.8–99.9 | 98.8–100 |
| Toptan et al. [223] | 6 | 62.2–65.4 | 98.8–99.5 | 98.8–99.5 | 99.2–99.9 | 99.2–99.9 | 98.1–99.7 | 98.1–99.7 | 99.4–100 | 99.4–100 | 98.8–100 | 98.8–100 |
| Vogels et al. [224] | 11 | 58.6–65.5 | 0.0–99.3 | 0.0–99.3 | 0.0–99.9 | 0.0–99.9 | 0.0–99.4 | 0.0–99.4 | 0.0–99.9 | 0.0–99.9 | 0.0–100 | 0.0–100 |
| Yip et al. [225] | 2 | 61.4–63.4 | 99.1–99.2 | 99.1–99.3 | 99.3–99.5 | 99.3–99.5 | 97.9–98.7 | 97.9–98.7 | 98.6–99.7 | 98.6–99.7 | 99.7–99.9 | 99.7–99.9 |

Several authors recommend a PCR primer design in which the range between hybridisation temperatures of the primer pair should be no more than 5 °C [23, 142], whereas others suggest a difference less than 1 °C [140]. Despite this, we observed that primer pairs from a few sets have a $T_m$ range a little larger than the recommendation. The primer pair of Institute Pasteur [211], nCoV_IP4-14059Fw and nCoV_IP4-14146Rv, has the same $T_m = 54.8$ °C, and nsp2f and nsp2r pair from Yip et al. [225] set has 2 °C in difference. However, SARS-CoV-2-84_LEFT

and SARS-CoV-2-84_RIGHT pair from Munnink et al. [220] set has a difference in the melting temperature of 7.7 °C. We show in detail the temperature ranges pair to pair in Table 10 in Appendix A.1.

The strict coverage $C_{\text{strict}}$ for SARS-CoV-2 is generally greater than 90%, which is expected since the primers and probes were designed specifically to this virus. However, a few oligonucleotides do not show strict coverage to SARS-CoV-2 at all, even when mismatches are taken into account. The exception are a few primers and probes from Corman et al. [213] set. For instance, RdRp_SARSr-FG primer shows a strict coverage $C_{\text{strict}} = 0\%$, while considering the presence of mismatches the partial coverage is $C_{\text{part.}} = 99.3\%$. It is consistent with the point of Pillonel et al. [242] that several probes of Corman et al. [213] do not match completely to SARS-CoV-2 genomes available at the time. As Corman and Drosten [243] explained, the mismatch presence within their set was due to the incomplete genomes available at the time when their primers/probes were designed. Moreover, RdRp primers/probes from Corman et al. [213] set were designed to detect both SARS-CoV-2 and SARS-CoV-1, only RdRp_SARSr-P2 probe is specific to SARS-CoV-2. Also experiments have shown that mismatches in RdRp primers/probes did not affect their ability to bind to the target [58, 241, 244]. When mismatches are taken into account for this subset, there are also an increase in the partial coverages for variants, SARS-CoV-1 and non-SARS genomes, see Table 26 and Table 27 in Appendix C. In particular, a few RdRp primers/probes have a partial coverage of 100% to the Delta variant.

In a few cases, the presence of mismatches leads to a substantially decrease of melting temperature, which reflects in an absence of coverage. An example is the four primers from Vogels et al. [224] set that do not completely align to any wild-type SARS-CoV-2 genomes. Both strict and partial coverages are absent, since the mismatched alignments show a hybridisation temperature too low in comparison to reference temperature $T_{\text{ref.}}$. Otherwise, several cases show that the mismatch presence increases the partial coverage. For instance, 2019-nCoV_N1-P probe from CDC [210] set has 223 further partial alignments, which increases the strict coverage of 98.4% to partial coverage of 99.4%. Similar cases are shown for SARS-CoV-2_6_LEFT primer from Munnink et al. [220] set and NIID_WH-1_F501 primer from WHO [211] set. Furthermore, the Hirotsu et al. [216] set includes NIID-N2 primers from National Institute of Infectious Diseases (NIID) from Japan, which were evaluated by Shirato et al. [245] concerning the mismatch presence in the detection of variants of concern at the time (Alpha, Beta, Gamma and Delta variants). The authors showed that single mismatches do not prevent the amplification and the primers are able to detect the variants, which is in agreement with our results of high partial coverages for SARS-CoV-2 and its variants for this set, except to NIID-N2-R primer that showed null coverages. Regarding all sets evaluated here, a considerable number of primers achieved strict and partial coverages of 100% for both Gamma and Delta variants, especially to the latter.

The primer-target mismatches may affect the hybridisation temperatures decreasing or increasing it. In Figure 32a, we can see a displacement profile for a single AC mismatch in 2019-nCoV_N2-F primer from CDC [210] set. It shows a disorder in the surrounding AT

base pairs and its melting temperature is $T_{\text{MM}} = 48.5$ °C, while the reference temperature is $T_{\text{ref.}} = 61.1$ °C. Considering the fact that AC pair is a weak mismatch and AT pair is weaker than CG pair due to its two hydrogen bonds, the high perturbation in this sequence region makes sense. However, in Figure 32b, SARS-CoV-2_89_RIGHT primer from Munnink et al. [220] set has two consecutive mismatches around the 5′ terminal. Although these mismatches induce a small end fraying, its hybridisation temperature $T_{\text{MM}} = 68.7$ °C is similar to its reference temperature $T_{\text{ref.}} = 68.3$ °C. Figure 32c shows also a small end fraying from an AG mismatched pair near 5′ terminal. Even so the temperature decreases from $T_{\text{ref.}} = 67.1$ °C to $T_{\text{MM}} = 59.8$ °C. Note the end fraying driven by AG pair is smaller in comparison to AT pair at last three positions of 3′ terminal, which may hamper the DNA polymerase activity. Also the AG pair is one of the strongest mismatched pair due to its double hydrogen bonds [61, 206]. In contrast, GT pair at second position of 3′ terminal shows a small end fraying and a fairly low difference in temperatures, $T_{\text{ref.}} = 65.6$ °C and $T_{\text{MM}} = 62.0$ °C, see Figure 32d. Figures 32e and 32f show two internal mismatches that induce an increase in both temperatures: $T_{\text{ref.}} = 79.2$ °C to $T_{\text{MM}} = 81.7$ °C and $T_{\text{ref.}} = 72.3$ °C to $T_{\text{MM}} = 73.4$ °C, respectively. GG and GT mismatched pairs show a slight shift what is consistent with evidences that these pairs are some of the strongest mismatches [26, 30, 34, 61, 246]. Also TT pair shows a destabilising behaviour in both cases, probably due to its weak stability [61, 246], especially because when it is flanked in both sides by canonical CG pairs it may have none or single hydrogen bond [79, 205]. Yet the mismatch behaviour is not straightforward since it is context dependent [26, 246, 247]. Additional average displacement profiles are shown in Figure 46 in Appendix E.1.

We also considered cross-reactivity of PCR primers/probes with SARS-CoV-1 and non-SARS genomes. We aligned all sets against genomes of other coronaviruses separating SARS-CoV-1 from others, as described in section 7.1. In case a primer or probe binds to other target not that SARS-CoV-2, it is not a specific oligonucleotide [142], which may affect the detection of the target virus. It could affect the results of PCR leading to false conclusions [40]. In this work, when mismatches are taken into account, most of primer/probe sets show a considerable coverage. From all 19 primer/probe sets, only 5 sets do not show any cross-reactivity at all, see Table 2.

The outcomes to the evaluation of PCR primer/probe sets for SARS-CoV-2, SARS-CoV-1 and non-SARS, and those concerning to the variants of concern at the time were published in *Molecular and Cellular Probes* and *Advances in Bioinformatics and Computational Biology* in February and November 2021, respectively, see Appendix K.

### 7.3.2 LAMP Evaluation

We collected 18 LAMP primer sets that underwent clinical validation and aligned them to the genomes described earlier. The hybridisation temperature of the primers as a double strand and the alignments found were predicted. Next, we calculated the strict and partial coverages following the considerations mentioned in section 7.2.3. The range of temperatures

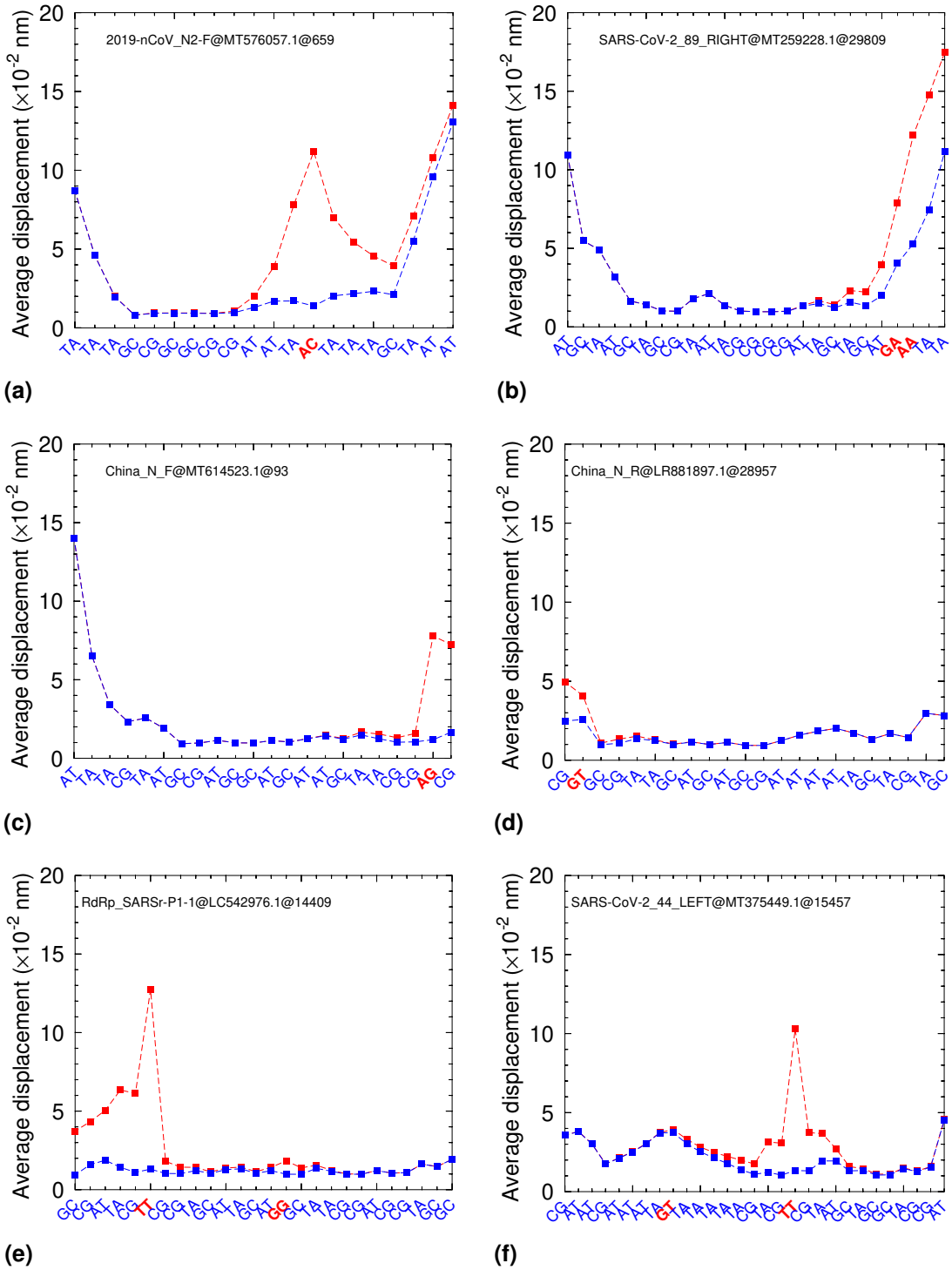



**Figure 32**

Displacement profiles for PCR primers. The blue line shows the displacement for matched alignment and the red one for mismatched alignment. The genome/primer hybridisation direction is $5' \longrightarrow 3'/3' \longrightarrow 5'$. (a) 2019-nCoV_N2-F primer [210] aligned to MT576057.1 genome at position 659 and an AC mismatched pair; (b) SARS-CoV-2_89_RIGHT primer [220] aligned to MT259228.1 genome at position 2909 and consecutive GA and AA mismatched pairs; (c) China_N_F primer [212] aligned to MT614523.1 genome at position 934 and an AG mismatched pair; (d) China_N_R primer [212] aligned to LR881897.1 genome at position 28957 and a GT mismatched pair; (e) RdRp_SARSr-P1-1 primer [213] aligned to LC542976.1 genome at position 14409 and TT and GG mismatched pairs; (f) SARS-CoV-2_44_LEFT primer [220] aligned to MT375449.1 genome at position 15457 and GT and TT mismatched pairs.

**Table 2**
Summary of results for 19 PCR primer/probe sets. Shown are the number of primers/probes $N_{pp}$ and the range of reference temperature $T_{\text{ref.}}$ for each set and the range of strict and partial coverages for SARS-CoV-1 and non-SARS genomes.

| | | | SARS-CoV-1 | | non-SARS | |
|---|---|---|---|---|---|---|
| Name of Set | $N_{pp}$ | $T_{\text{ref.}}$ (°C) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
| CDC [210] | 6 | 61.1–75.5 | 0.0–80.0 | 0.0–80.0 | 0.0 | 0.0 |
| WHO [211] | 21 | 51.3–70.3 | 0.0 | 0.0 | 0.0 | 0.0 |
| Luminex [212] | 6 | 59.4–81.0 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Corman et al. [213] | 21 | 61.7–81.7 | 0.0–100 | 0.0–100 | 0.0 | 0.0–19.2 |
| Davda et al. [214] | 16 | 56.5–70.3 | 0.0–80.0 | 0.0–90 | 0.0 | 0.0 |
| Grant et al. [215] | 2 | 62.4–79.9 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Hirotsu et al. [216] | 3 | 60.6–68.8 | 0.0 | 0.0 | 0.0 | 0.0 |
| Jalali et al. [217] | 6 | 61.3–64.2 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Lanza et al. [218] | 27 | 60.5–75.0 | 0.0 | 0.0–90.0 | 0.0 | 0.0 |
| Li et al. [132] | 2 | 67.5–70.4 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Lu et al. [219] | 3 | 64.0–74.7 | 0.0 | 0.0 | 0.0 | 0.0 |
| Munnink et al. [220] | 171 | 65.4–74.9 | 0.0–100 | 0.0–100 | 0.0 | 0.0–0.639 |
| Nalla et al. [221] | 4 | 51.7–68.0 | 0.0–90.0 | 0.0–100 | 0.0–68.7 | 0.0–68.7 |
| Niu et al. [40] | 6 | 59.2–84.2 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Park et al. [148] | 20 | 59.3–65.4 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Rahman et al. [222] | 6 | 64.2–76.1 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Toptan et al. [223] | 6 | 62.2–65.4 | 0.0 | 0.0 | 0.0 | 0.0 |
| Vogels et al. [224] | 11 | 58.6–65.5 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Yip et al. [225] | 2 | 61.4–63.4 | 0.0 | 0.0 | 0.0 | 0.0 |

and coverages for each primer set are shown in Tables 3 and 4, respectively. All coverages for SARS-CoV-2, its variants, SARS-CoV-1 and non-SARS genomes are shown in Table 28–Table 31 in Appendix D.

According to Yang et al. [248] and a common protocol for LAMP primer design [249], the hybridisation temperature of primers should be in the range 60–65 °C, since the LAMP process is carried out usually at 65 °C [183]. Here there is a variation in the primer melting temperature, or reference temperature $T_{\text{ref.}}$. Most are around 55–67 °C, but a few show a high temperature around 70 °C. The primers N15-B1c from Huang et al. [183] set and N2-F1c from Ji et al. [228] set show $T_{\text{ref.}} = 76.6$ °C and 75.6 °C, respectively. In particular, the two lowest hybridisation temperatures found are $T_{\text{ref.}} = 48.4$ °C for NEB_E1-LF primer from Lalli et al. [169] set and $T_{\text{ref.}} = 49.2$ °C for S447-B2 primer from Diego et al. [226] set. The reference temperature $T_{\text{ref.}}$ for each primer are shown in Table 11 in Appendix A.2.

LAMP primers showed high strict and partial coverages for SARS-CoV-2 and its variants. A high number of primers obtained coverages greater than 90% considering only complete alignments. For instance, nCoV-N-F1c primer from Jiang et al. [118] set with 99.8% and RdRp_LF primer from Alves et al. [2] set with 99.5% for wild-type SARS-CoV-2. For variants, we found a considerable number of 100% strict coverages, mainly for subvariants. S3-F2 primer from Mohon et al. [232] set completely covered all subvariant sets. However, a few primers showed no coverages for matched alignments, including for SARS-CoV-2 genomes. Only Alves et al. [2], Garcia-Venzor et al. [185], Jang et al. [173], Lau et al. [229], Luo et al. [230], Mautner

**Table 3**
Summary of all 18 LAMP primer sets. Shown are the number of primers $N_{pp}$ and the range of reference temperature $T_{\text{ref.}}$ for each set.

| Name of Set | $N_{pp}$ | $T_{\text{ref.}}$ (°C) |
|---|---|---|
| Alekseenko et al. [168] | 20 | 54.8–72.3 |
| Alves et al. [2] | 32 | 56.6–70.6 |
| Diego et al. [226] | 40 | 49.2–72.5 |
| Ganguli et al. [227] | 81 | 55.2–73.4 |
| Garcia-Venzor et al. [185] | 8 | 53.8–66.7 |
| Huang et al. [183] | 32 | 55.2–76.6 |
| Jang et al. [173] | 15 | 57.8–69.3 |
| Ji et al. [228] | 32 | 55.2–75.6 |
| Jiang et al. [118] | 8 | 59.9–69.8 |
| Lalli et al. [169] | 32 | 48.4–69.2 |
| Lau et al. [229] | 10 | 52.7–69.7 |
| Luo et al. [230] | 48 | 59.9–73.3 |
| Mautner et al. [231] | 8 | 57.2–66.4 |
| Mohon et al. [232] | 16 | 57.0–67.0 |
| Rodriguez-Manzano et al. [233] | 8 | 60.5–68.8 |
| Yan et al. [234] | 15 | 56.1–71.0 |
| Yang et al. [235] | 15 | 57.6–69.5 |
| Yoshikawa et al. [236] | 16 | 58.2–69.9 |

**Table 4**
Summary of the results for all 18 LAMP primer sets. Shown is the range of strict and partial coverages for SARS-CoV-2 and Alpha, Beta, Gamma, Delta, Lambda, Mu, Omicron, BA.2, BA.3, BA.4 and BA.5 variants.

| Name of Set | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SARS-CoV-2 | | Alpha Variant | | Beta Variant | | Gamma Variant | | Delta Variant | | Lambda Variant | |
| Alekseenko et al. [168] | 0.0–99.5 | 96.0–99.5 | 0.0–99.8 | 98.9–99.8 | 0.0–99.2 | 97.9–99.2 | 0.0–99.8 | 99.4–99.8 | 0.0–100 | 98.7–100 | 0.0–100 | 99.1–100 |
| Alves et al. [2] | 97.9–99.5 | 97.9–99.5 | 96.1–99.9 | 96.1–99.9 | 92.6–99.4 | 92.6–99.5 | 0.737–99.9 | 1.08–99.9 | 97.4–100 | 98.0–100 | 0.664–99.9 | 0.664–100 |
| Diego et al. [226] | 0.0–99.6 | 83.8–99.6 | 0.0–99.9 | 0.911–100 | 0.0–99.8 | 0.293–99.8 | 0.0–100 | 0.780–100 | 0.0–100 | 0.0126–100 | 0.0–99.9 | 95.0–99.9 |
| Ganguli et al. [227] | 0.0–99.5 | 0.0–99.5 | 0.0–99.8 | 0.0–99.8 | 0.0–99.3 | 0.0–99.4 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.9 | 0.0–99.9 |
| Garcia-Venzor et al. [185] | 98.8–99.3 | 98.8–99.3 | 96.8–99.7 | 96.8–99.7 | 97.3–98.6 | 97.3–98.6 | 99.0–99.9 | 99.1–99.9 | 98.8–100 | 98.8–100 | 1.04–99.5 | 1.04–99.5 |
| Huang et al. [183] | 0.0–99.5 | 97.3–99.5 | 0.0–99.9 | 97.0–99.9 | 0.0–99.3 | 92.6–99.4 | 0.0–99.7 | 0.563–99.9 | 0.0–100 | 38.0–100 | 0.0107–99.8 | 0.0107–99.9 |
| Jang et al. [173] | 0.0–99.4 | 98.6–99.4 | 0.0–99.8 | 99.1–99.9 | 0.0–99.8 | 0.293–99.8 | 0.0–100 | 97.8–100 | 0.0–100 | 98.8–100 | 0.0–99.9 | 98.4–99.9 |
| Ji et al. [228] | 0.0–99.5 | 0.0–99.5 | 0.0–99.9 | 0.0–99.9 | 0.0–99.8 | 0.0–99.8 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.9 | 0.0–99.9 |
| Jiang et al. [118] | 97.9–99.6 | 98.4–99.6 | 99.4–99.9 | 99.4–100 | 97.1–99.4 | 97.1–99.5 | 0.780–99.8 | 98.7–99.8 | 99.5–99.9 | 99.7–100 | 94.8–99.9 | 95.0–99.9 |
| Lalli et al. [169] | 0.0–99.5 | 97.7–99.5 | 0.0–99.8 | 96.1–99.9 | 0.0–99.7 | 97.2–99.7 | 0.0–100 | 98.2–100 | 0.0–100 | 98.0–100 | 0.0–100 | 1.03–100 |
| Lau et al. [229] | 98.5–99.4 | 98.5–99.4 | 0.221–99.8 | 0.235–99.8 | 0.814–99.3 | 0.814–99.3 | 96.7–99.2 | 96.7–99.2 | 0.768–100 | 0.768–100 | 0.0535–99.5 | 0.664–99.5 |
| Luo et al. [230] | 98.5–99.4 | 98.5–99.4 | 0.221–99.8 | 0.235–99.8 | 0.814–99.3 | 0.814–99.3 | 96.7–99.2 | 96.7–99.2 | 0.768–100 | 0.768–100 | 0.107–99.9 | 0.107–99.9 |
| Mautner et al. [231] | 93.7–99.3 | 93.7–99.3 | 0.0552–99.8 | 0.0552–99.8 | 97.4–99.1 | 97.4–99.1 | 2.21–93.1 | 2.21–93.1 | 28.8–99.9 | 28.8–99.9 | 92.9–99.0 | 93.9–99.0 |
| Mohon et al. [232] | 96.0–99.5 | 96.0–99.5 | 0.110–99.8 | 0.110–99.9 | 1.99–99.1 | 1.99–99.1 | 93.4–99.5 | 93.4–99.5 | 2.19–100 | 2.19–100 | 2.74–100 | 2.87–100 |
| Rodriguez-Manzano et al. [233] | 0.0–99.4 | 0.0–99.5 | 0.0–99.8 | 0.0–99.9 | 0.0–98.2 | 0.0–98.3 | 0.0–99.7 | 0.0–99.7 | 0.0–99.9 | 0.0–100 | 0.0–99.8 | 0.0–100 |
| Yan et al. [234] | 98.6–99.5 | 98.6–99.5 | 0.138–99.9 | 0.138–99.9 | 94.0–99.6 | 94.0–99.6 | 99.0–100 | 99.0–100 | 97.9–100 | 97.9–100 | 0.642–99.9 | 0.642–100 |
| Yang et al. [235] | 0.0–99.3 | 66.6–99.3 | 0.0–99.7 | 0.0690–99.7 | 0.0–98.8 | 96.4–98.8 | 0.0–99.8 | 0.563–100 | 0.0–99.9 | 0.0–100 | 0.0–99.7 | 0.107–99.8 |
| Yoshikawa et al. [236] | 98.7–99.4 | 98.8–99.5 | 96.1–99.8 | 96.1–99.9 | 96.9–98.3 | 97.2–98.3 | 98.1–99.7 | 98.2–99.7 | 99.4–99.9 | 99.5–99.9 | 0.0–99.8 | 0.0–99.8 |
| | Mu Variant | | Omicron Variant | | BA.2 Subvariant | | BA.3 Subvariant | | BA.4 Subvariant | | BA.5 Subvariant | |
| Alekseenko et al. [168] | 0.0–100 | 90.2–100 | 0.0–99.6 | 90.3–99.6 | 0.0–100 | 99.7–100 | 0.0–100 | 98.6–100 | 0.0–100 | 99.2–100 | 0.0–100 | 95.0–100 |
| Alves et al. [2] | 0.756–100 | 0.953–100 | 0.199–99.8 | 0.199–99.9 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.159–100 | 0.159–100 | 0.162–100 | 0.162–100 |
| Diego et al. [226] | 0.0–100 | 88.8–100 | 0.0–99.9 | 0.341–99.9 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 |
| Ganguli et al. [227] | 0.0–99.9 | 0.0–99.9 | 0.0–99.7 | 0.0–99.7 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.8 | 0.0–99.9 |
| Garcia-Venzor et al. [185] | 98.9–99.8 | 98.9–99.9 | 86.5–99.6 | 86.5–99.6 | 99.2–100 | 99.2–100 | 18.8–100 | 18.7–100 | 97.6–99.8 | 97.6–99.8 | 98.1–100 | 98.1–100 |
| Huang et al. [183] | 0.0–99.8 | 0.877–99.9 | 0.0–99.5 | 0.341–99.5 | 0.0–100 | 0.0–100 | 0.0–100 | 0.862–100 | 0.0–100 | 0.318–100 | 0.0–99.9 | 0.569–100 |
| Jang et al. [173] | 0.0–99.9 | 99.4–99.9 | 0.0–99.9 | 0.185–99.9 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.9 | 0.0812–100 |
| Ji et al. [228] | 0.0–99.9 | 0.0–99.9 | 0.0–99.8 | 0.0–99.8 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 |
| Jiang et al. [118] | 88.8–99.7 | 88.8–99.8 | 0.555–94.9 | 0.555–94.9 | 1.31–100 | 1.31–100 | 0.862–96.0 | 0.862–96.0 | 0.318–99.7 | 0.318–99.7 | 0.975–99.4 | 0.975–99.4 |
| Lalli et al. [169] | 0.0–99.9 | 85.7–99.9 | 0.0142–99.7 | 0.199–99.7 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.159–100 | 0.0–100 | 0.162–100 |
| Lau et al. [229] | 0.197–99.7 | 88.8–99.7 | 0.142–94.9 | 0.555–94.9 | 0.0–99.9 | 0.0–99.9 | 0.0–96.0 | 0.862–96.0 | 0.318–99.8 | 0.318–99.8 | 0.569–99.4 | 0.569–99.5 |
| Luo et al. [230] | 88.7–99.8 | 88.8–99.9 | 0.0854–99.3 | 0.0854–99.3 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.9 | 0.0–100 |
| Mautner et al. [231] | 0.0908–99.4 | 0.0908–99.4 | 81.7–91.9 | 81.7–91.9 | 99.4–99.8 | 99.4–99.8 | 82.8–95.4 | 87.1–95.4 | 97.1–99.8 | 97.1–99.8 | 15.5–99.4 | 15.5–99.6 |
| Mohon et al. [232] | 90.0–100 | 90.0–100 | 0.512–99.6 | 0.512–99.6 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.318–100 | 0.318–100 | 0.325–100 | 0.325–100 |
| Rodriguez-Manzano et al. [233] | 0.0–99.6 | 0.0–99.8 | 0.0–99.2 | 0.0–99.2 | 0.0–100 | 0.0–100 | 0.0–99.7 | 0.0–99.7 | 0.0–100 | 0.0–100 | 0.0–99.6 | 0.0–99.7 |
| Yan et al. [234] | 98.6–100 | 98.6–100 | 0.726–99.9 | 0.726–99.9 | 0.135–100 | 0.135–100 | 0.0–100 | 0.0–100 | 0.636–100 | 0.795–100 | 0.569–100 | 0.569–100 |
| Yang et al. [235] | 0.0–99.8 | 97.9–99.8 | 0.0–99.7 | 0.0854–99.7 | 0.0–100 | 0.0947–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–99.8 | 0.0–99.9 |
| Yoshikawa et al. [236] | 0.0–99.8 | 0.0–99.8 | 0.0–99.9 | 0.0–99.9 | 0.0–100 | 0.0–100 | 0.0–99.1 | 0.0–99.1 | 0.0–100 | 0.0–100 | 0.0–100 | 0.0–100 |

et al. [231], Mohon et al. [232] and Yan et al. [234] sets showed strict coverage from all primers for SARS-CoV-2. They achieved coverages greater than 90%, except three primers from Luo et al. [230] set (around 65%). Considering the presence of up to three contiguous mismatches the coverage increased to around 90% or more for some cases. For example, E-B2 and As1_F1c primers from Jang et al. [173] and Alekseenko et al. [168] sets, respectively, showed a coverage from 0% to 99.0% or more for all genome sets. Another example is the Yang et al. [235] set which showed ORF1e-B1c primer with a coverage from 0% for complete matched to 100% for mismatched alignments in the evaluation of Gamma and Delta variants and BA.2 subvariant. In contrast, in some cases, the mismatch presence either did not change or little increased the coverage. Only one primer from Garcia-Venzor et al. [185] set slightly increased the coverage for SARS-CoV-2 genomes. M-B1c and S555-B2 primers from Diego et al. [226] set covered no subvariant sets. CU-N2-F1c primer from Yang et al. [235] set did not cover subvariant sets, while its CU-N2-F2 pair covered 100%, except for BA.5 subvariant (99.4%). Furthermore, a few primers showed no coverage at all for both matched and mismatched aligments. For instance, ORF1ab-1-LB and F2-N primers from Ji et al. [228] and Rodriguez-Manzano et al. [233] sets, respectively, and three primers from Ganguli et al. [227] set did not cover any genome set. In addition, only one set did not show both strict and partial coverages of 100% and Lau et al. [229] and Mohon et al. [232] sets showed 100% for strict coverage.

We previously evaluated the primers from Alves et al. [2] set using a smaller set of SARS-CoV-2 genomes (3364), see Ref. 2. These primers were developed by the Fiocruz group led by Rubens Monte-Neto, which performed the clinical validation. In their experiments, RT-LAMP assay showed a good performance of primers in the detection of SARS-CoV-2 virus, which showed herein coverages greater than 90% for both strict and partial alignments. In addition, they carried out the assay for the Gamma variant, also known as P.1 variant, and achieved success detecting it. Our results for their primers against the Gamma variant also showed high coverages, see Table 28 in Appendix D. E_Set1_LF primer increased its coverage from 0% to 97.8% for Gamma variant and E_Set1_B2 primer showed no coverage at all for subvariant sets. Also a few primers covered 100% of Delta and Mu variants and the entire subvariant set. Furthermore, Almeida et al. [250] showed E1 and N2 primer subsets did not prevent the target amplification of Omicron variant, despite the presence of a single mismatch. In our assessment, those primer subsets covered Omicron variant set in more than 90%. Only N_Set2_F1c, N_Set2_F2 and E_Set1_B2 primers did not achieve both strict and partial coverages for this variant.

We also investigated the cross-reactivity to SARS-CoV-1 and non-SARS genomes. A substantial number of primers showed a high coverage for SARS-CoV-1, while only one primer from Alekseenko et al. [168] set showed a non-zero coverage for both strict and partial alignments for the five other coronaviruses. In particular to SARS-CoV-1 alignments, a couple of primers covered the complete alignments around 90-100% and only two sets showed no coverages at all [231, 234]. The summary of the cross-reactivity is showed in Table 5.

**Table 5**
Summary of the results for all 18 LAMP primer sets. Shown are the number of primers $N_{pp}$ and the range of reference temperature $T_{\text{ref.}}$ for each set and the range of strict and partial coverages for SARS-CoV-1 and non-SARS genomes, respectively.

| | | | SARS-CoV-1 | | non-SARS | |
|---|---|---|---|---|---|---|
| Name of Set | $N_{pp}$ | $T_{\text{ref.}}$ (°C) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
| Alekseenko et al. [168] | 20 | 54.8–72.3 | 0.0–10 | 0.0–100 | 0.0–12.8 | 0.0–94.2 |
| Alves et al. [2] | 32 | 56.6–70.6 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Diego et al. [226] | 40 | 49.2–72.5 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Ganguli et al. [227] | 81 | 55.2–73.4 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Garcia-Venzor et al. [185] | 8 | 53.8–66.7 | 0.0–90 | 0.0–90 | 0.0 | 0.0 |
| Huang et al. [183] | 32 | 55.2–76.6 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Jang et al. [173] | 15 | 57.8–69.3 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Ji et al. [228] | 32 | 55.2–75.6 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Jiang et al. [118] | 8 | 59.9–69.8 | 0.0–90.0 | 0.0–90.0 | 0.0 | 0.0 |
| Lalli et al. [169] | 32 | 48.4–69.2 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Lau et al. [229] | 10 | 52.7–69.7 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Luo et al. [230] | 48 | 59.9–73.3 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Mautner et al. [231] | 8 | 57.2–66.4 | 0.0 | 0.0 | 0.0 | 0.0 |
| Mohon et al. [232] | 16 | 57.0–67.0 | 0.0–90 | 0.0–90 | 0.0 | 0.0 |
| Rodriguez-Manzano et al. [233] | 8 | 60.5–68.8 | 0.0–90 | 0.0–90 | 0.0 | 0.0 |
| Yan et al. [234] | 15 | 56.1–71.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Yang et al. [235] | 15 | 57.6–69.5 | 0.0–90 | 0.0–100 | 0.0 | 0.0 |
| Yoshikawa et al. [236] | 16 | 58.2–69.9 | 0.0–90 | 0.0–100 | 0.0 | 0.0 |

The average opening along the direction of hydrogen bonds can illustrate qualitatively the behaviour of the base pairs. In Figure 33a, the AC mismatched pair in N_Set2_F2 primer from Alves et al. [2] set shows a small destabilisation, probably due to surrounding GC base pairs since AC is not a strong mismatched pair. However, its presence decreases the reference temperature $T_{\text{ref.}} = 62.8$ °C to melting temperature of mismatched pair $T_{\text{MM}} = 48.1$ °C, which removes a feasible coverage to this alignment, once the $T_{\text{MM}}$ is lower than the ideal range, i.e. 60–65 °C, according to LAMP primer design guidelines. On the other hand, a GA mismatch in S17-B1c primer from Huang et al. [183] set increases the temperature from $T_{\text{ref.}} = 68.7$ °C to $T_{\text{MM}} = 70.2$ °C. Figure 33b shows a slight shift for this mismatched pair which may not hamper the amplification. In Figure 33c, a CC mismatched pair shows a large shift as well as decreases the temperature from $T_{\text{ref.}} = 68.5$ °C to $T_{\text{MM}} = 58.9$ °C. Perhaps due to the CC pair to be the weakest mismatched pair [30, 34, 61, 246] and show a severe impact to primer-target duplex stability [37, 51, 82]. In Figure 33d, double consecutive mismatched pairs at 3′ terminal show a large end fraying in comparison to the canonical pairs, although the temperature increases from $T_{\text{ref.}} = 72.3$ °C to $T_{\text{MM}} = 74.5$ °C. Note that the vertical scale of the figures is adapted according to the value of sequence average opening. Additional displacement profiles are shown in Figure 47 in Appendix E.2.

It is expected that over time mismatches should increasingly occur within the primer regions due to the continuous mutation of the SARS-CoV-2 genomes, see section 3. The decreasing coverage as variants appear are shown in Figure 34. In comparison to the wild type (ws) coverage, all variants decrease their coverage. Considering partial coverages with $\Delta T_{\text{lim}} = 0$°C,

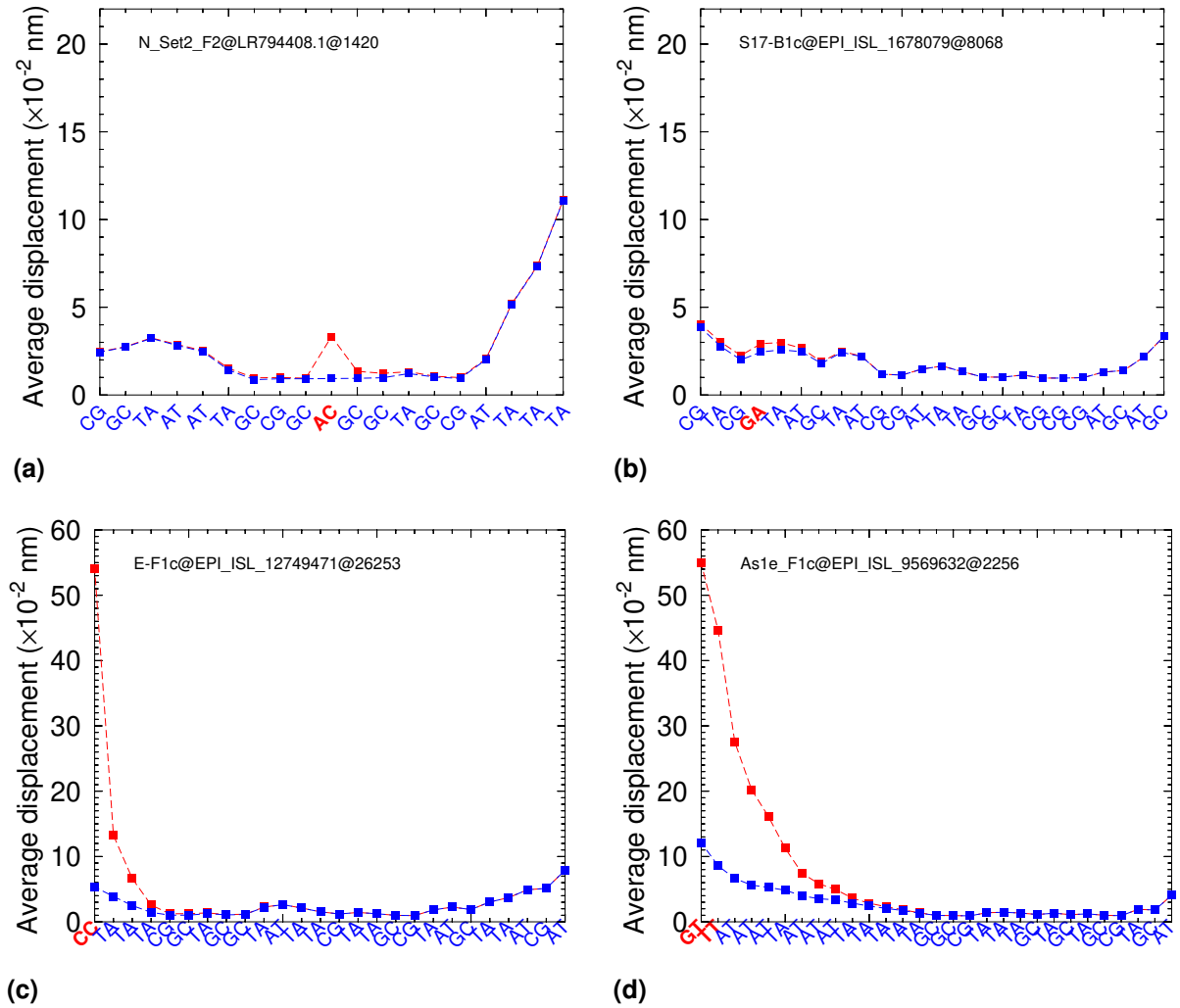



**Figure 33**

Displacement profiles for LAMP primers. The blue line shows the displacement for matched alignment and the red one for mismatched alignment. The genome/primer hybridisation direction is $5' \longrightarrow 3'/3' \longrightarrow 5'$. (a) N_Set2_F2 primer [2] aligned to LR794408.1 genome at position 1420 and an AC mismatched pair; (b) S17-B1c primer [183] aligned to EPI_ISL_1678079 genome at position 8068 and a GA mismatched pair near to $3'$ terminal; (c) E-Fic primer [226] aligned to EPI_ISL_12749471 genome at position 26253 and a CC mismatched pair at $3'$ terminal; (d) As1e_F1c primer [168] aligned to EPI_ISL_9569632 genome at position 2256 and GT and TT consecutive mismatched pairs at $3'$ terminal.

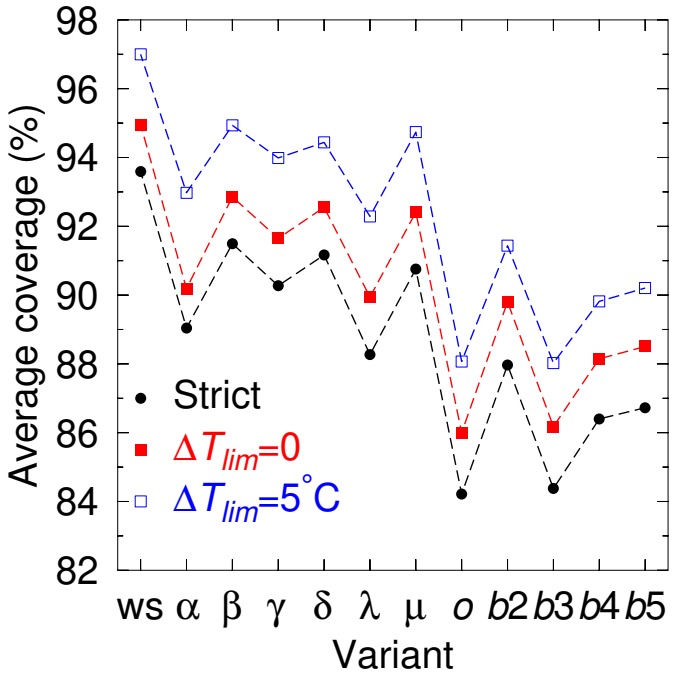



**Figure 34**
Coverage averaged over all primers for wild-type SARS-CoV-2 (ws), Alpha ($\alpha$), Beta ($\beta$), Gamma ($\gamma$), Delta ($\delta$), Lambda ($\lambda$), Mu ($\mu$), Omicron ($o$) variants and BA.2 (b2), BA.3 (b3), BA.4 (b4) and BA.5 (b5) subvariants. Black bullet are for $C_{\text{strict}}$, red (filled) boxes are for $\Delta T_{\text{lim}} = 0°\text{C}$ and blue (unfilled) boxes are for ($5°\text{C}$). The dashed line connecting the data point is only intended as a guide to the eye.

that is, primers with $T_{\text{MM}} = T_{\text{ref.}}$, the curve is uniformly shifted upwards. For $\Delta T_{\text{lim}} = 5°\text{C}$, partial coverage of Beta, Gamma, Delta and Mu variants becomes slightly higher than the ws strict coverage. However, the rate of decrease is not uniform and some variants have higher coverage than their presumed predecessor variants, such as Beta, Delta, Mu, BA.2, BA.4 and BA.5 variants. For Omicron variant, which has a larger number of mutations [251], we observe a sharp drop in the coverage. BA.3 subvariant shows a low coverage similar to Omicron variant, while BA.2, BA.4 and BA.5 subvariants have a higher coverage. The reason for this oscillation for Omicron variant and its subvariants is not clear.

In some cases, the presence of mismatches does not destabilise the duplex, that is, when we consider $\Delta T_{\text{lim}} = 0°\text{C}$. In Table 6 are shown a few examples of primers that show zero strict coverage, but achieve a greater than 90% coverage if stabilising mismatches are considered. A curious case is N1_F1c primer from Ji et al. [228] set which shows only partial coverages for Omicron variant and two of its subvariants (BA.2 and BA.3), since all primers were designed to wild-type SARS-CoV-2. Although it suggests to be an opposite trend to the overall decline of Omicron variant, note that higher coverage for Omicron is quite exceptional and only occurs for very few primers. On the other hand, it clearly highlights that the assessment of mismatch impact is far from trivial.

The continuous mutation of SARS-CoV-2 has impacted the diagnostic assays. Nucleotide substitutions in different gene regions, especially S gene, have produced false negative re-

**Table 6**
Examples of primer coverages with stabilising mismatches, $T_{\text{lim.}} = 0°\text{C}$, which have $C_{\text{part.}}(0°\text{C}) > 90\%$ while having $C_{\text{strict}} = 0$. Only those primers which have stabilising mismatches for given variant shown.

| | $C_{\text{part.}}(0°\text{C})$ (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Primer | ws | Alpha | Beta | Gamma | Delta | Lambda | Mu | Omicron | BA.2 | BA.3 | BA.4 | BA.5 |
| As1e_F1c [168] | 99.2 | 99.4 | 98.7 | 99.7 | 99.2 | 99.0 | 99.7 | 98.5 | 100 | 98.6 | 99.5 | 99.7 |
| iLACO-F1c [168] | 99.2 | — | 99.1 | 99.5 | 99.3 | 99.8 | 99.8 | 99.5 | 99.8 | 99.7 | 100 | 99.9 |
| N15-B1c [183] | 97.7 | 98.8 | 96.5 | 98.4 | 99.9 | 99.4 | 99.6 | 99.2 | 99.9 | 99.7 | 98.9 | 98.9 |
| N1-B1c [228] | 98.4 | 99.5 | 97.1 | 99.6 | 99.6 | 99.7 | 99.3 | 93.8 | 99.8 | 94.3 | 99.0 | 98.7 |
| N1-F1c [228] | — | — | — | — | — | — | — | 92.7 | 97.5 | 94.0 | — | — |
| N2-F1c [228] | 99.1 | 99.1 | 96.6 | 99.1 | 99.9 | 98.7 | 99.7 | 99.3 | 99.9 | 99.7 | 99.0 | 99.0 |
| NEB_orf1a-A-F1c [169] | 98.9 | 99.5 | 99.2 | 99.5 | 99.8 | 99.6 | 99.1 | 99.4 | 99.7 | 99.1 | 99.2 | 96.9 |
| F1c [229] | — | — | — | — | — | 98.9 | 99.2 | 93.9 | 99.5 | 93.1 | — | — |

sults [252–255]. Although most primers have achieved high coverages, some primers fail to achieve significant coverage for at least one variant and may represent a potential dropout. A summary of the number of primers found that could potential represent dropouts are shown in Table 7. Here we are considering a very stringent threshold of 5% at $\Delta T_{\text{lim}} = 5°\text{C}$, that is, even considering a maximal 5°C melting temperature below the reference temperature, those primers covered less than 5% of each genome set. Only Alekseenko et al. [168] set had no potential drop-out primers at all. To avoid the dropout is suggested the use of at least two gene targets [252, 255], which could be evaluated by applying the workflow presented here. The complete list of potential drop-out primers for each genome set is shown in Table 32–Table 43 in Appendix F.

**Table 7**
Sets that have at least one potential drop-out primer for any of the variants. Only the reference number is shown for each set. Drop-out primers are considered as those with a partial coverage $C_{\text{part.}}$ below 5% ($\Delta T_{\text{lim.}} = 5°\text{C}$), $\text{N}_{\text{drop}}$, for at least one variant. $N_{\text{primers}}$ is the number of separate primers for each set.

| Ref. | $N_{\text{primers}}$ | $N_{\text{drop}}$ | ws | Alpha | Beta | Gamma | Delta | Lambda | Mu | Omicron | BA.2 | BA.3 | BA.4 | BA.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [2] | 32 | 6 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 3 | 3 | 3 | 4 | 3 |
| [226] | 40 | 7 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 3 | 3 | 3 | 3 | 3 |
| [227] | 81 | 28 | 4 | 11 | 5 | 5 | 5 | 6 | 12 | 9 | 5 | 10 | 7 | 6 |
| [185] | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [183] | 32 | 6 | 0 | 0 | 0 | 1 | 0 | 3 | 1 | 3 | 2 | 3 | 2 | 2 |
| [173] | 15 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 2 |
| [228] | 32 | 5 | 1 | 3 | 1 | 3 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 2 |
| [118] | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| [169] | 32 | 5 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 2 | 2 | 4 | 2 |
| [229] | 10 | 6 | 0 | 2 | 1 | 0 | 2 | 3 | 0 | 3 | 3 | 3 | 3 | 3 |
| [230] | 48 | 12 | 0 | 3 | 0 | 5 | 3 | 6 | 0 | 7 | 7 | 7 | 8 | 7 |
| [231] | 8 | 5 | 0 | 2 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| [232] | 16 | 5 | 0 | 1 | 2 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| [233] | 8 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| [234] | 15 | 3 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| [235] | 15 | 2 | 0 | 1 | 0 | 1 | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 |
| [236] | 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Mismatched pairs in 5′ or 3′ terminals of FIP and BIP primers may hamper the amplification. Here FIP and BIP primers were divided into F1c/F2 and B1c/B2 primers, respectively, and assessed individually. A few alignments with either 5′ or 3′ terminal mismatches show a melting

temperature within the threshold and contribute to the increase of partial coverage. Clearly, in some cases, mismatches in both terminals decreased the temperature. Since for the LAMP technique the F1c and B1c primers depend on their respective F2 and B2 complements, the dropout may in practice be higher. FIP and BIP primers with terminal mismatches which have an increase in their coverage are shown in Table 44–Table 55 in Appendix G for each genome set. In Figure 35, we observe four primers from Ji et al. [228] set that show single, double and triple contiguous mismatches at 3′ terminal. Although it is obvious that these mismatches destabilise the terminal, in this case, only the double mismatched pairs show a decrease in the temperature. The TT and CT double pairs show $T_{\text{ref.}} = 69.0$ °C to $T_{\text{MM}} = 66.6$ °C, see Figure 35b, and the GT and TT double pairs, $T_{\text{ref.}} = 71.0$ °C to $T_{\text{MM}} = 70.1$ °C, see Figure 35c. The triple mismatches (GT, CT and TT pairs) increase the temperature from $T_{\text{ref.}} = 75.6$ °C to $T_{\text{MM}} = 79.9$ °C. Albeit GT mismatched pair has been reported as a strong pair [26, 34, 61, 246], it is influenced by its TT and CT mismatched neighbouring pairs, see Figure 35a. In comparison, the shift caused by the single GT mismatched pair in the same position (3′ terminal) is smaller, and perhaps perturbed by neighbouring AT pairs, see Figure 35d. Also for this primer hybridisation the temperature increases from $T_{\text{ref.}} = 73.9$ °C to $T_{\text{MM}} = 80.6$ °C.

The outcomes of LAMP primers assessment to variants and subvariants were published in *The Open COVID Journal*, see Appendix K.

## 7.4 Applications

We have seen that mismatch presence in molecular techniques may contribute to different applications, such as allele-specific and single nucleotide polymorphism (SNP) identification [37, 68]. Also mismatches could be inserted in primers and probes designed to detect oncogenes. Since it is known that some types of mismatches are more stable than others, even more than canonical base pairs, it could extend their application.

Along with mismatch features known so far, our developed workflow could be applied to a continuously assessment of primers/probes not only for SARS-CoV-2 coronavirus as well as other pathogens. Oligonucleotides designed with mismatches could be analysed considering single or up to three consecutive mismatches, different positions and length size. For instance, mismatches have been applied to identify SARS-CoV-2 variants [71, 72]. Moreover, the application of this workflow may reduce the oligonucleotide synthesis costs, reducing the need to produce several different primers/probes to obtain the most viable one.

## 7.5 Conclusion

Here we showed an assessment workflow to investigate the mismatch impact in PCR and LAMP primers designed to detect the SARS-CoV-2 coronavirus. In addition, we extended this analysis to SARS-CoV-2 variants to evaluate the primers effectiveness. Our findings showed a considerable number of primers achieves a low coverage to the wild-type SARS-CoV-2 genomes.

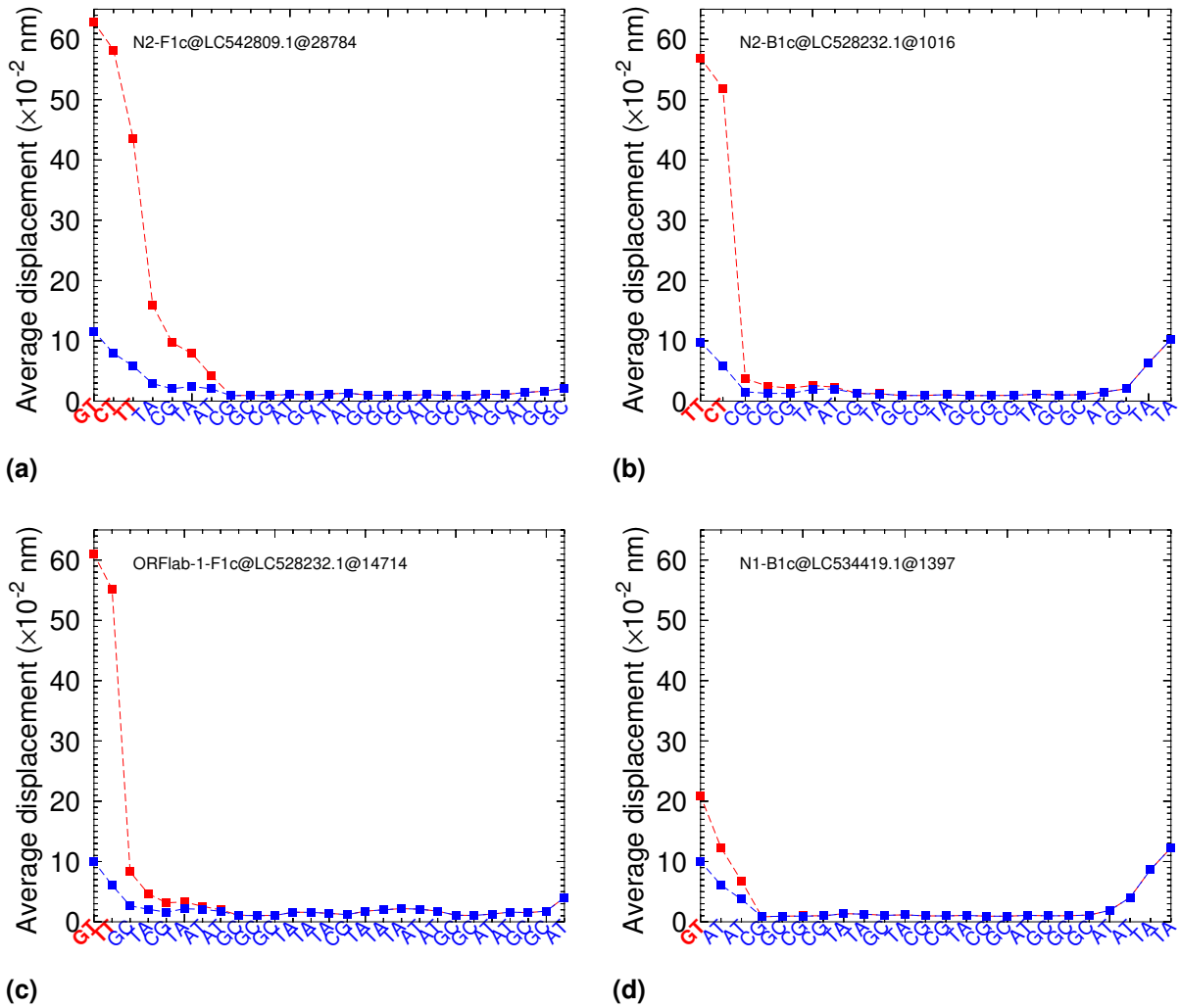



**Figure 35**

Displacement profiles for LAMP primers. The blue line shows the displacement for matched alignment and the red one for mismatched alignment. The genome/primer hybridisation direction is $5' \longrightarrow 3'/3' \longrightarrow 5'$. (a) N2-F1c primer aligned to LC542809.1 genome at position 28784 and GT, CT and TT triple consecutive mismatched pair; (b) N2-B1c primer aligned to LC528232.1 genome at position 1016 and TT and CT double consecutive mismatched pair; (c) ORFlab-1-F1c primer aligned to LC528232.1 genome at position 14714 and GT and TT double consecutive mismatched pair; (d) N1-B1c primer aligned to LC534419.1 genome at position 1397 and a GT mismatched pair. The mismatched pairs are at $3'$ terminal and all primers are from Ref. 228.

In contrast, the coverage increases substantially, even 100%, when mismatches are taken into account. In particular, a few primers show their higher partial coverages for Omicron subvariants. Yet the mismatch presence is not totally beneficial destabilising both temperature and locally structure of primer-target duplex. Furthermore, our outcomes are in agreement with the reported evidences concerned to context dependence and mismatch type, position and number.

# 8 Side Projects

Here we show the side projects based on mismatch effect inquiring, section 8.1, and locked nucleic acid (LNA) applications, section 8.2.

## 8.1 Mismatch Analyses

We analysed three LAMP primer sets to different targets (SARS-CoV-2, SARS-CoV-1 and Bos taurus beta-casein (CSN2) gene); gold-nanorod probes to detect SARS-CoV-2 and we carried out an assessment of mismatch position.

### 8.1.1 Gold Nanorod (AuNR) DNA Probes

Gold nanorods (AuNRs) are very small-sized, easily synthesized nanoparticles with unique optical properties, potentially acting in drug delivery and gene therapy [256]. They have been applied in studies of metal ions, amino and nucleic acids and protein [257–261]. In addition, diagnostic assays based in gold nanorods have been developed such as antiepidermal growth factor receptor (anti-EGFR) antibodies [256], p53 gene and single nucleotide polymorphism (SNP) [262] and SARS-CoV-2 biosensors [263].

Lívia Siman Gomes, from Department of Physics (UFMG), works with AuNR DNA probes. She shared, in a collaboration with us, a probe-target set designed to detect SARS-CoV-2 coronavirus. These DNA probes are attached to AuNR particles for hybridising to the viral RNA. We assessed seven probes and nine targets. Applying our workflow from section 7.2, we investigated the mismatch presence in these probes against the nine targets and SARS-CoV-2, SARS-CoV-1 and non-SARS genomes. We found a high coverage for SARS-CoV-2 and for two probes mismatch presence increased the coverage. Only one probe showed coverage for SARS-CoV-1 and other achieved 100% when mismatches were taken into account. For non-SARS genomes, no coverage at all was found. Also we analysed the probes on ViennaRNA [264] with the contribution of Vivianne Basílio Barbosa, a scientific initiation student in our group at the time. ViennaRNA is a package of stand-alone programs and libraries applied to the prediction and comparison of RNA secondary structures. We applied our parameter sets and the default set of ViennaRNA. A few results showed a discrepancy in the secondary structure. The data are not shown here for confidentiality reasons.

### 8.1.2 CRISPR LAMP Primers

In summary, clustered regularly interspaced short palindromic repeats (CRISPR) is a technique in which Cas enzymes bind to a specific target substrate, then they are hyperactivated to cleave all neighbouring nucleic acids. The Cas nuclease is designed to recognize target sequences, e.g. cancer mutations. Furthermore, CRISPR assays are typically combined with an isothermal amplification step to increase sensitivity [60].

We collected a LAMP primer set developed to amplify S gene SARS-CoV-2 target in combination to a CRISPR-Cas detection from Ref. 60. The FIP and BIP primers were divided as described in section 7.1.2. The difference in primers derived from the same code is a specific mismatch incorporated to improve the assay. In total, we worked on 42 primers and those with "MM" in their name are the mismatched ones. Their sequences and hybridisation temperatures are shown in Table 56 in Appendix H.1.

We applied our primer/probe assessment workflow, section 7.2, to these primers regarding to wild-type SARS-CoV-2 and Alpha, Beta, Gamma and Delta variants. We found null coverages for a few primers to SARS-CoV-2, mainly the FIP and BIP primers. Perhaps it is due to the selection we carried out, although it was made carefully. Despite it, the mismatched F3 primers also achieved null coverages. It is unexpected since the experimental results [60] show the mismatched F3 and B3 primers had a good performance in spite the mismatch presence. Similar results were obtained to the variants. The coverages for all genomes are shown in Table 57 in Appendix H.2.

Last, we compared the predicted hybridisation temperatures to the cycle threshold (Ct) values. Since a few factors of the diagnostic assay may adapted to achieve Ct values of interest [265], we are interested in investigating if there is a relation between Ct values and hybridisation temperatures. We took the Ct values as the black horizontal bars (the mean) of Fig.3 from Ref. 60. We plotted these mean Ct values and the predicted temperatures to all matched and mismatched primers, see Figure 36. However, it is an initial attempt to develop this idea and more data is required to a further inquiring.

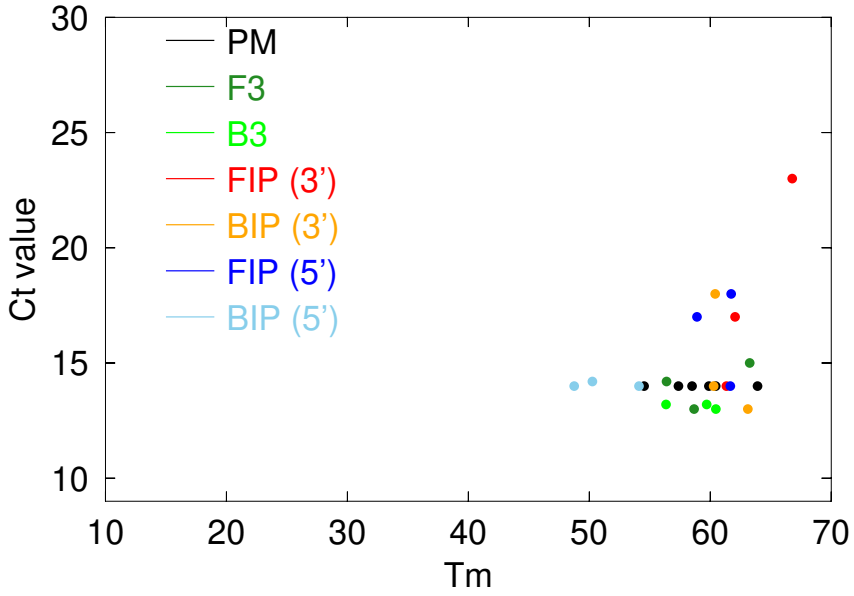



**Figure 36**
Hybridisation Temperatures ($T_m$) $\times$ Ct Values. We predicted hybridisation temperatures for all matched and mismatched primers and collected the Ct values from Fig.3 from Ref. 60. FIP and BIP labels represent all F1c/F2 and B1c/B2 primers, respectively, and the mismatch at $3'$ or $5'$ terminal.

### 8.1.3 SARS-CoV-1 LAMP Primers

We applied the workflow developed in previous section 7.2 to a LAMP primer set from Ref. 157 designed to identify SARS-CoV-1 coronavirus. The 8 primers were evaluated for three genome groups: SARS-CoV-1, SARS-CoV-2 and non-SARS, that is, the other five coronaviruses (229E, NL63, OC43, HKU1 and MERS-CoV). The primer sequences and their predicted melting temperature are shown in Table 8.

**Table 8**
Summary of SARS-CoV-1 LAMP primers [157] and their predicted hybridisation temperature $T_{\text{ref.}}$.

| Primer | Sequence $5' \rightarrow 3'$ | $T_{\text{ref.}}$ (°C) |
|---|---|---|
| B1c | GCTGTGGGTACTAACCTACCT | 62.2 |
| B2 | GTCAACATAACCAGTCGG | 55.6 |
| B3 | CTCTGGTGAATTCTGTGTT | 55.6 |
| LoopB | CCAGCTAGGATTTTCTACAGG | 62.2 |
| F1c | GTTGCATGACAGCCCTCTAC | 63.7 |
| F2 | AGAAGCTATTCGTCACGT | 56.1 |
| F3 | AATATGTTTATCACCCGCG | 57.3 |
| LoopF | CAAAGCCAATCCACGCAC | 61.0 |

We followed each step: primer/genome alignment, hybridisation temperature prediction and coverage calculation. We found the primers are very specific for SARS-CoV-1, although no primer achieved 100% coverage. Two primers achieved 80% and the other six 90%. A few mismatches were identified and did not interfere on the coverage. Also there is no cross-reactivity at all. All coverages are shown in Table 9. It shows the excellence of primer set and the potential of the workflow developed herein to assess different primer sets and targets.

**Table 9**
Both strict and partial coverages for SARS-CoV-1 LAMP primers from Ref. 157.

| | SARS-CoV-1 | | SARS-CoV-2 | | non-SARS | |
|---|---|---|---|---|---|---|
| Primer | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
| B1c | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| B2 | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| B3 | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| LoopB | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| F1c | 80 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| F2 | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| F3 | 80 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |
| LoopF | 90 | 90 | 0.0 | 0.0 | 0.0 | 0.0 |

### 8.1.4 LAMP Primers for Mutant Identification

We analysed two LAMP primer sets shared by Rodrigo Giglioti from Instituto de Zootecnia (IZ). These sets was designed to identify the wild-type and mutant targets of Bos taurus beta-casein (CSN2) gene. Both sets have six primers and are identical with only one difference in

the 3′ terminal base of F2 primers. This differentiates each set: one for wild type and the other for mutant.

All primers from both sets have hybridisation temperatures $T_m$ within the LAMP range temperature (60–65°C), except the F2 primer for wild-type sequence (66.7°C). Even though, its temperature are around the temperature range and could be performed successfully. Considering the 3′ terminal base difference, we observe a drop in the temperature, 60.0°C and 60.4°C, when the F2 primers hybridise to the other target, which may differentiate them.

Also we expanded the F2 primers incorporating a second mismatch. We generated all possible mismatch for all positions, except for the 3′ terminal position, which we considered the same mismatch to all primers. For instance, for a 18-length primer we generated 52 different primers, including the matched primer. We found double mismatched primers, for wild-type and mutant targets, decrease the temperature contributing to differentiate both types. The data are not shown here for private reasons.

### 8.1.5 Mismatch Position

We collected a dataset of 111 primers of DNA from Ref. 266 to reproduce its results and use the data to study the mismatch effects. Wu et al. [266] used two sets of primers to the target 16S rRNA gene, a conserved molecular marker used for phylogenetic analysis in microbial communities [266]. The first set (primer MX), which we called "set1", had one perfect matched primer and 54 mismatched primers. These primers were 18 bp in length and had all possible types and positions of mismatches. The same conditions was valid for the other set (primer 1492R), called here as "set2". The difference is that set2 had one perfect match and 57 mismatched primers with 19 bp in length. All primers and their predicted melting temperature are shown in Table 58 in Appendix I.1.

Wu et al. [266] used the single base extension (SBE), or minisequencing, experiment to analyse the mismatch effect in all possible positions in the primers. They considered the position 0 in the 3′ terminal and the last position in the 5′ terminal. The authors calculated in silico the melting temperatures of primers using the IDT Oligo Analyzer [267] for secondary structure of primer and DINAMelt web server [268] for primer/template hybridisation.

First, we collected the primers in two sets (set1 and set2) and predicted the melting temperature $T_m$ of each primer-target hybridisation. This prediction was done from the method described in section 6.4. Then we calculated the difference $\Delta T_m$ between the $T_m$ of perfect match and the mismatched match. We made, using our calculated temperatures, the same graphs of position versus $\Delta T_m$ from Wu et al. [266], that is, their Fig. 5a and Fig. 4a, respectively, set1 and set2. In Figures 37 and 38, it is possible to visualise the comparison of results with our melting temperatures above and in silico melting temperatures of Wu et al. [266] below. We see that the graphs are not similar. Their interval is within –5 and 15 in axis y for both sets while we found an interval within 0 and 20 for set1 and 0 and 30 for set2. It is due to the difference in conditions, e.g. salt concentration, of the parameters for melting temperature prediction. While

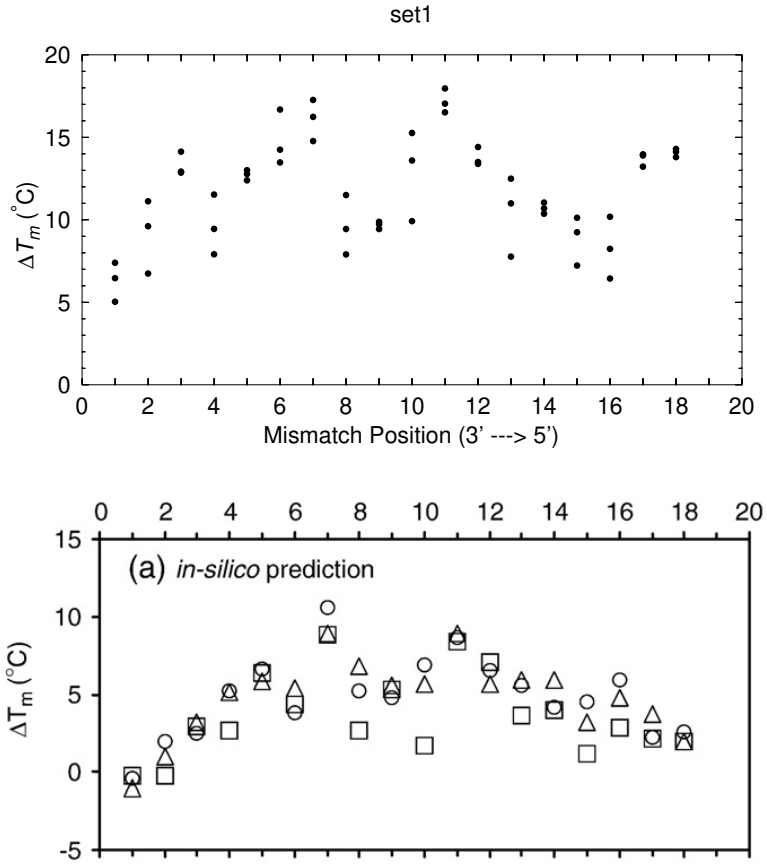



**Figure 37**
Correlation between position of the mismatches and the difference of melting temperature for set1. The position axis is oriented from 3′ to 5′ end. Above, the data predicted by the Peyrard-Bishop (PB) model. Below, the original graph (Fig. 5a) from Wu et al. [266].

they calculated in silico with salt concentration of $Na^+$ 1M, we used parameters for DNA in $Na^+$ 60mM. However, when we calculated the melting temperatures using the same tools and parameters used by authors and compared with a few $T_m$'s mentioned in their work, we found another small discrepancy. Wu et al. [266] mentioned that the melting temperature for perfect match of set2 was below 36.9°C. When we used the DINAMelt [268], as they indicated, we found 75.0°C. The authors also mentioned melting temperatures for a few mismatches of set2 below 40.0°C, but we found all melting temperatures for set2 above 60.0°C. Although they do not mention melting temperatures for set1, we found 75.2°C for perfect match and above 60.0°C for mismatched primers. We verified with the tool they used to calculate secondary structure, IDT Oligo Analyzer [267], maybe it was an exchange in the name of tools in the paper, but also the results did not check. The authors did not provide the in silico melting temperatures that they found for each sequence. In spite of the difference in the in silico $T_m$'s, the range for $\Delta T_m$ for both sets are identical to the work of Wu et al. [266]. We calculated these data and compared the range: for set1, -1.0–10.6°C and for set2, -0.2–9.4°C.

The next step was to reproduce the Fig. 6 from Wu et al. [266] to verify the quality of the data. In this figure, the authors used the $\Delta T_m$ obtained by in silico calculations and the ex-

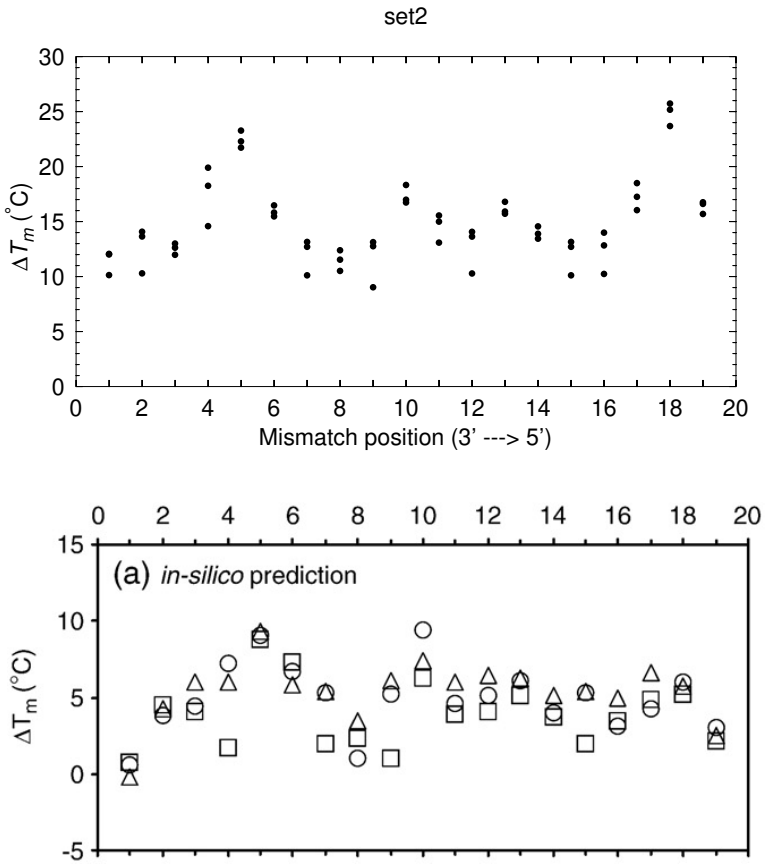



**Figure 38**
Correlation between position of the mismatches and the difference of melting temperature for set2. The position axis is oriented from 3′ to 5′ end. Above, the data predicted by the PB model. Below, the original graph (Fig. 4a) from Wu et al. [266].

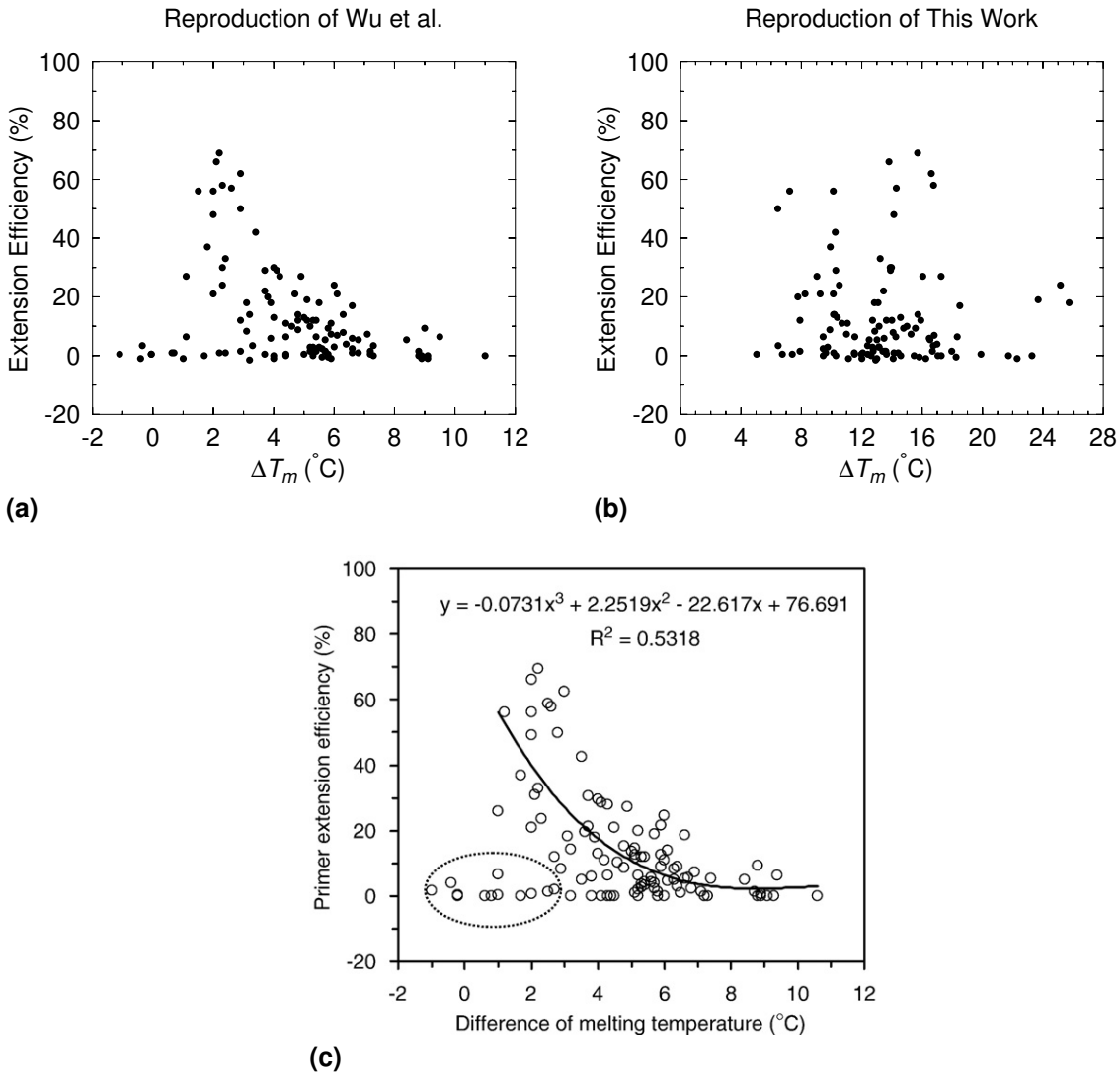



**Figure 39**
Difference of melting temperatures in relation to the extension efficiencies. (a) The version made by us after the extraction of data from the paper; (b) the same idea using the melting temperatures calculated by the PB model; (c) the original graph showed in Ref. 266.

periment efficiencies shown in Fig. 5b (set1) and 4b (set2) from their reference [266]. They correlated the difference in melting temperatures and extension efficiencies for all 111 primers (both set1 and set2). We collected their $\Delta T_m$ and the extension efficiencies from the same figures using an online tool [269] and generated a similar graph. We also designed the same graph — $\Delta T_m$ versus extension efficiencies — using the extension efficiencies that we collected and the $\Delta T_m$ that we calculated. In Figures 39a and 39c, the reproduction made by us and the original figure, respectively. In this case, the figures are very similar. A few discrepancies is probably due to the error in the extraction of the data using the online tool mentioned earlier [269]. However, when we made the same graph using the $\Delta T_m$'s which we calculated the result is very distinct, see Figure 39b.

We also correlated the difference of melting temperatures from Wu et al. [266] and which

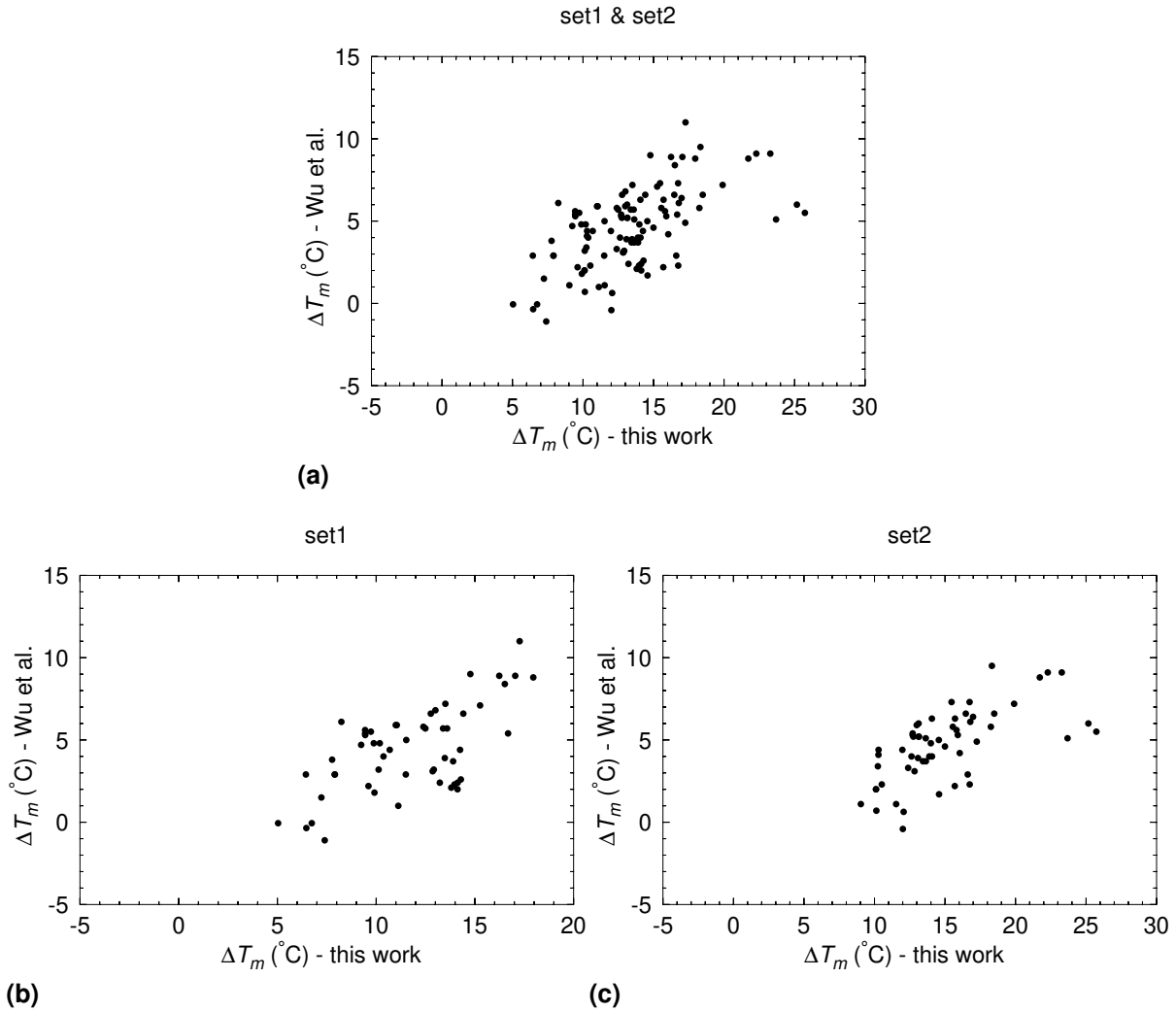



**Figure 40**
Difference of melting temperatures $\Delta\, T_m$ calculated in this work in relation to the melting temperatures extracted from Wu et al. [266]. (a) $\Delta\, T_m$ for both sets together; (b) only set1; (c) only set2.

we calculated. Taken the total differences $\Delta T_m$'s for both sets, we found a Pearson correlation of $r = 0.6$, see Figure 40a. Figures 40b and 40c show the correlation between differences of melting temperatures for both sets, separately. Although the Pearson correlation is considerable, both figures show a few outliers which is a little suspicious.

Using the Peyrard-Bishop (PB) model, we calculated the average opening $\langle y_m \rangle$, which is the displacement of the bases in the direction of the hydrogen bonds, for all primers. With these data, we designed graphs to analyse the behaviour of mismatches in comparison to the position and the correlation between extension efficiencies and the average opening of both terminals. We collected the terminal average opening of each sequence for each set in four different temperature of calculation (180, 200, 220 and 280 K), as described in section 6.3. It was carried out to $3'$ and $5'$ terminals separately. In Figures 41a and 41b, we observe the correlation in 180 K for both terminals of set1. The concentration of data is around 1 for terminal $3'$ and 0 for terminal $5'$. However, there are a few outliers in both graphs and when

we increase the temperature of calculation to 280 K, the data are considerable disperse, see Figures 41c and 41d. The figures of the other temperatures of calculation for both set1 and set2 are shown in Figures 48 and 49 in Appendix I.2. It shows us that there is no correlation between average opening and extension efficiencies what led us, among other results, to conclude that data from Wu et al. [266] is not an appropriate dataset to work in relation to our purpose for this project.

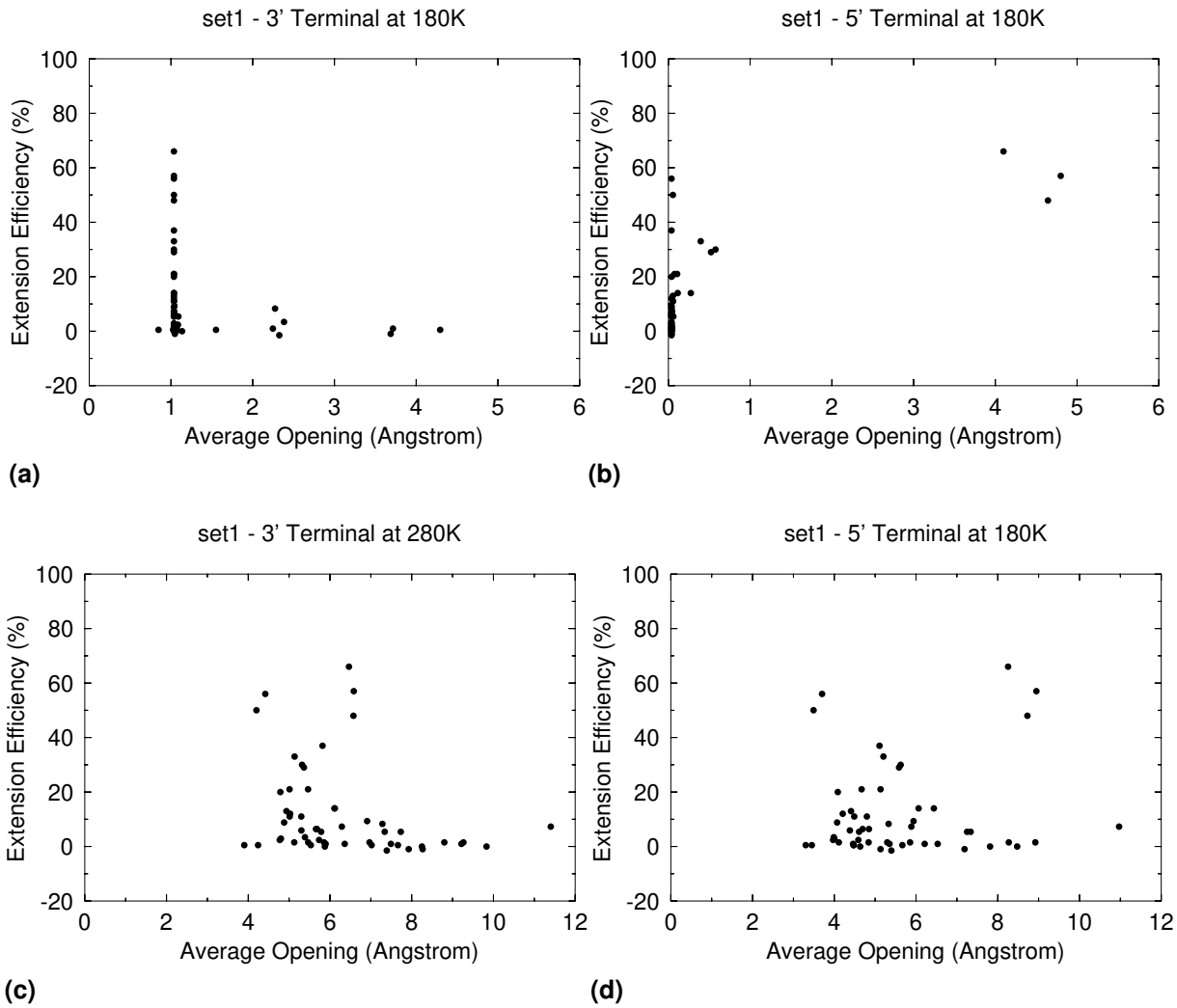



**Figure 41**
Correlation between average opening and extension efficiency for set1. In a temperature of calculation at 180 K to both 3′ (a) and 5′ (b) terminals, and in a temperature of calculation at 280 K in both 3′ (c) and 5′ (d) terminals.

Although we conclude it is not an appropriate dataset to use, we compared the average opening within each mismatch position set. We input the perfect matched primer and the three mismatched primers according to the position on the sequence. Figure 42 shows a few subsets from set1. Figures 42a and 42b show the mismatch influence within the first three positions at 3′ terminal. In the first, CT and TT mismatched pairs show a larger end fraying. In the other hand, GT mismatched pair is very stable even more than the perfect match. Here we can ob-

serve the behaviour of the strongest mismatch and two mismatched pairs which are considered weaker mismatched pairs. Even though, the mismatch presence decrease the melting temperature of perfect match $T_{\text{pm}} = 61.6°\text{C}$ to around $T_{\text{mm}} = 52.0°\text{C}$. In Figure 42b, the weakest CC mismatched pair destabilises in the direction of 3′ terminal. The other direction is more stable perhaps due to the triple neighbouring GC pairs. We observe the same behaviour of CC mismatched pair in Figure 42c agreeing to the fact of this pair is the weakest mismatch. Figure 42d shows the higher destabilisation of all three mismatched primers at 5′ terminal. For all figures, the melting temperature of mismatched primers within each subset achieved a very close temperature and lower than the perfect match temperature. Figure 43 shows the mismatched primer subsets for the set2. In Figure 43a, TT and CT mismatched pairs are more unstable than in Figure 43b. Perhaps the proximity to the double AT pairs at 3′ terminal increases this destabilisation. Also in both figures, we observe the higher stability of GT mismatched pair in comparison to the other corresponding pairs in the subset. Although the mismatches decrease the perfect match temperature, the temperature of GT pair in both figures is higher than the others in the subset, see Table 58 in Appendix I.1. Figure 43d show the strongest mismatched pairs (GG and GT pairs) represented in a small destabilisation. Note the GT pair is the most stable and show the highest temperature within this subset. Last, Figure 43d show a great destabilisation although the GC canonical pairs franking in both sides.

## 8.2 Locked Nucleic Acid (LNA) Analyses

We briefly describe LNA modification; optimisation of new thermodynamic parameters for LNA with mismatches, which is still in progress; LNA probes for identification of repeat units on human DNA; participation in a validation of LNA nearest-neighbour parameter set, and application of this set on a LNA-modified LAMP primer set.

### 8.2.1 A Chemical Modification

Locked nucleic acid (LNA) is a nucleic acid chemical modification characterised by a methylene bridge that connects the 2′-oxygen of ribose with the 4′-carbon [270], see Figure 44. This bridge "locks" the ribose in a C3-endo/N conformation reducing its flexibility and mimicking a RNA helix [271]. It also makes the nucleotide more stable, increasing the organization of the phosphate backbone and inducing stable duplexes with DNA and RNA. These both duplexes may have an increase in their melting temperature by +4.0 – +9.3°C per LNA modification in comparison to the unmodified reference duplexes [272, 273].

LNA nucleotides exhibit resistance to nuclease degradation, chemical stability, convenient synthesis using standard reagents and minimal nonspecific interactions with nucleic acid binding proteins [275]. LNA effects are local and depend on probe length, sequence context and mismatch identity [270, 273]. Notwithstanding, the addition of LNA bases has shown to improve mismatch discrimination [273].

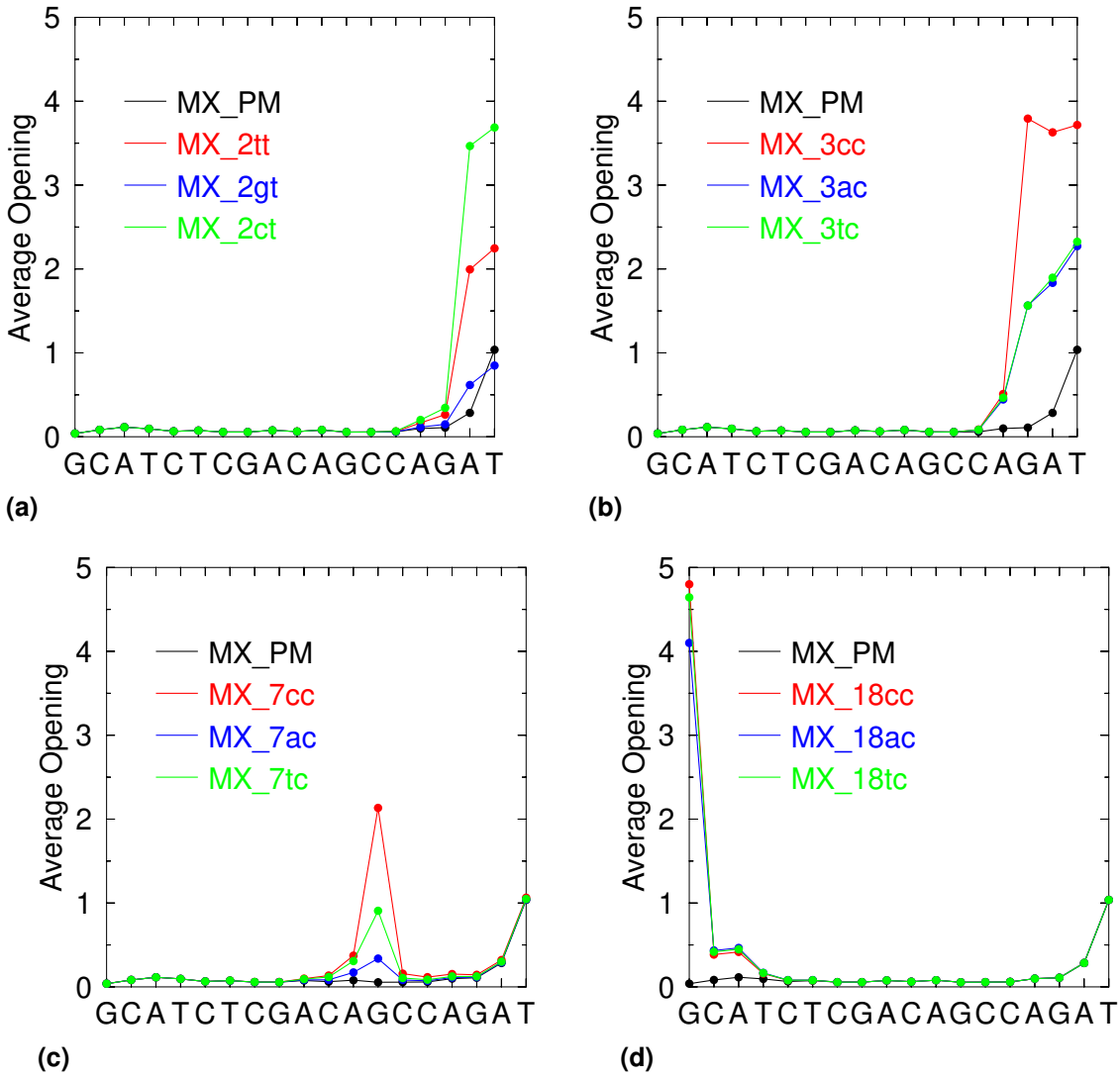



**Figure 42**

Average opening for subsets from set1 containing the perfect matched and mismatched primers according to the position. The sequences are oriented from $5' \longrightarrow 3'$ and the position 0 (the first one) is at $3'$ terminal and the last at $5'$ terminal.

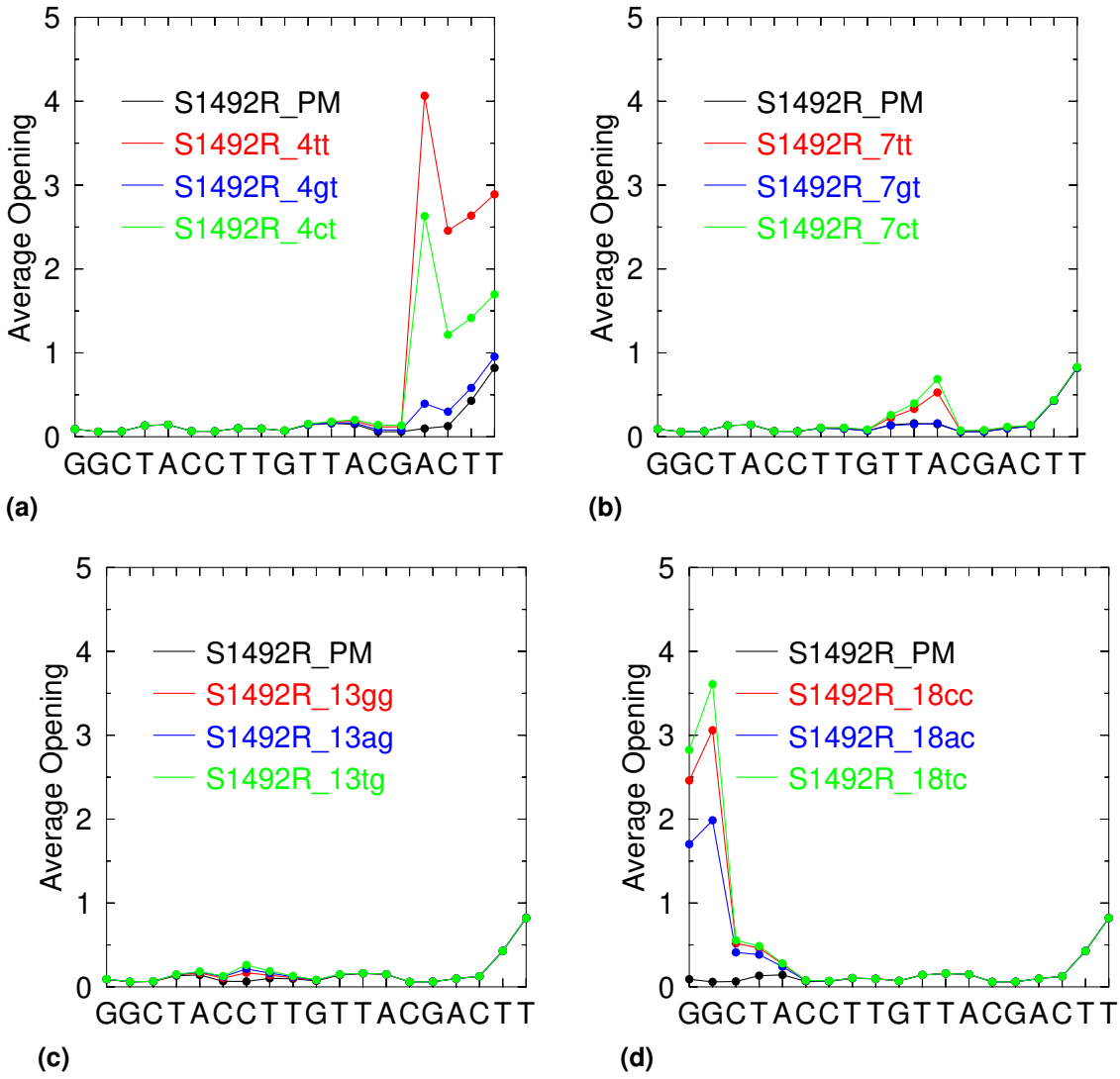



**Figure 43**

Average opening for subsets from set2 containing the perfect matched and mismatched primers according to the position. The sequences are oriented from $5' \longrightarrow 3'$ and the position 0 (the first one) is at $3'$ terminal and the last at $5'$ terminal.

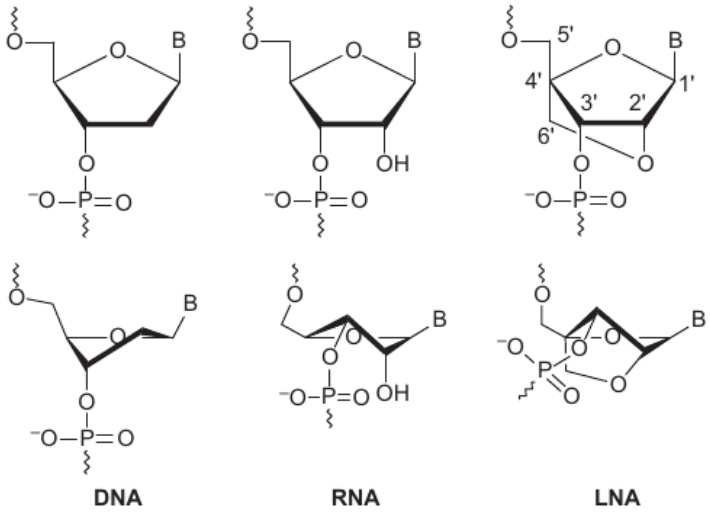



**Figure 44**
Comparison of DNA, RNA and LNA chemical structures. Note the C3-endo conformation of the LNA nucleotide similar to the RNA nucleotide. Figure adapted from Ref. 274.

LNA modifications have been used in different applications, such as specific PCR detection [276], high binding diagnostic probes [277], LAMP primers [278], antisense oligonucleotides (ASO) [279], aptamers [280] and molecular beacons [281].

### 8.2.2 LNA Mismatch

Our group has already developed thermodynamic parameters for LNA bases [282]. This parameter set describes the hydrogen bonds and the base stacking. It is valid for canonical LNA base pairs, e.g. +AT, +CG, +GC, A+T. Focusing on the mismatch (MM) effect and LNA ability in mismatch discrimination, we have initiated a project to optimise LNA mismatch (LNA-MM) parameters. We collected 279 sequences including canonical DNA and LNA duplexes, and mismatched LNA from Ref. 283. The mismatched LNA sequences show all possible mismatched LNA pairs. For the optimisation, we used the available DNA and LNA parameter sets at high salt concentration (1021 mM $Na^+$). We have carried out the optimisation for the initial evaluation of the data. This project is still on progress.

### 8.2.3 STR LNA Probes

Short tandem repeats (STRs), also known as microsatellites, are short nucleotide sequence motifs from 2 to 6 base pairs in size [284]. These repeat units vary in numbers in the human genome which leads to an individual identification, see Figure 45. This feature has contributed to forensic applications [285] and disease studies [286]. Also the STRs high mutation rates [287] have been related to some pathologies, such as cancer [288] and neurodegenerative diseases [289].

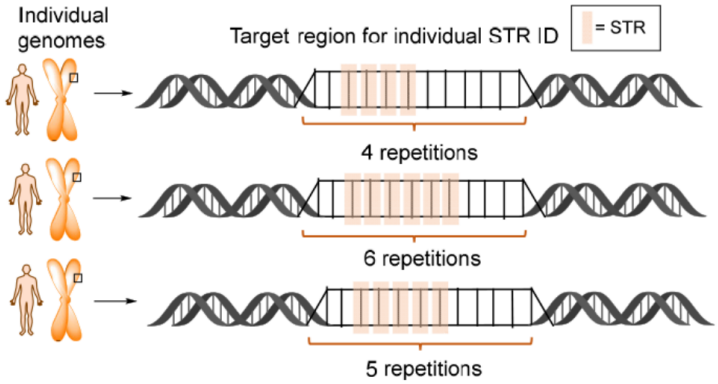



**Figure 45**
STR regions in human genome. The STR ID is the individual identification by the number of repeat units. Figure adapted from Ref. 290.

During my fellowship time at the Technical University of Denmark (DTU), I worked on STR probes. I analysed a LNA-modified probe set developed to identify STR in DNA sequences. Also these probes were modified with perylene dyes attached to three different nitrogen bases (A, C and U).

I applied our DNA and LNA thermodynamic parameter sets to feature the probes. I optimised the probes to obtain a new DNA/LNA parameters exclusively for them. Since the perylene is attached to the base, I consider it as a modified base. Therefore, I used the newly DNA/LNA-derived data to optimise the perylene dye in the duplexes. Also I inquired the mismatch presence in the STR probes.

The probe set only had 8 probes, which is very poor for deriving a completely new parameter set since there are not enough number of matched and mismatched perylene-modified pairs. Even though, I obtained a restrict parameter set which could be used as an initial seed for optimisation of perylene modified bases and LNA uracil base. The data are not shown here for private reasons.

### 8.2.4 LNA-NN Parameters

The Nearest-Neighbour (NN) model subdivides the enthalpy $\Delta H$ and the entropy $\Delta S$ for each combination of neighbouring base pairs, promoting a method to predict melting temperatures for unknown sequences. It was proposed from the contribution of neighbouring base interaction to the DNA and RNA molecule stability, since the base sequence strongly affects it. Both the bases inside the strand and at the ends are greatly influenced [291, 292]. The sum of the base interactions with its neighbours is thermodynamically considered as the DNA duplex structure [293] and indicates the free energy of a double strand [294].

I participated in the validation of a new LNA-NN parameter set, see Ref. 5. We optimised 306 DNA sequences containing single LNA modifications. We validated the new parameters for two groups of sequences in a high and low salt concentration. We obtained a better predictive

accuracy than previous parameter sets and found a general increase of stability compared to unmodified DNA duplexes. The except is **+A** base which destabilises in most contexts. Since this is not one of my direct works, the data are not shown here.

### 8.2.5 LNA-modified LAMP Primers

We analysed a LNA-modified primer set from Ref. 295 with 24 primers designed to detect SARS-CoV-2 coronavirus. We applied the LNA-NN parameter set described in previous section 8.2.4 to predict melting temperatures $T_{\text{lna}}$ of the primers.

We also predicted the melting temperature $T_{\text{ref.}}$ for the corresponding reference duplexes, which we called as control primer set. Comparing the $T_m$'s for reference and LNA duplexes, we observed LNA bases increase the temperature, that is, stabilise the primer-target hybridisation. Especially two or more LNA adenosine achieved the largest increase. For instance, E2L_B3 primer showed a $T_{\text{ref.}} = 68.6°\text{C}$ for reference duplex and a $T_{\text{lna}} = 83.9°\text{C}$ taken into account the three **+A** bases. E2L_LF primer also with three **+A** bases showed $T_{\text{ref.}} = 72.1°\text{C}$ and $T_{\text{lna}} = 82.7°\text{C}$. In contrast to double consecutive **+T** bases in the E3L_B1c primer which achieved a small difference between predicted temperatures: $T_{\text{ref.}} = 74.1°\text{C}$ and $T_{\text{lna}} = 75.8°\text{C}$. All primers and their melting temperatures are shown in Table 59 in Appendix J.

The LNA-NN parameters performed well, since it is known LNA bases tend to stabilise the duplex greater than canonical ones. This stabilisation could contribute to an improvement in the viral detection and other types of targets. Also this LNA-NN parameter set shows a potential as a new tool to the design of both PCR and LAMP primers as well as other oligonucleotide design. Even so, we now intend to evaluate other LNA-modified primer sets, as well as different techniques, e.g. LNA LAMP [296] and PCR [297] primers. Furthermore, an experimental assessment would be of great value to the application of this parameter set.

# 9 Conclusions

Mismatches have been reported to impact in the stability of nucleic acid molecules. Studies have shown that mismatch presence in primer-target duplex may disrupt molecular assays, impair target detection and generate false negative results [28, 41, 43]. Even just a single mismatch may hamper the target amplification [44, 48, 49]. However, mismatches may have no effect or even avoid false priming [46, 51], as well as may amplify efficiently as matched base pairs [51]. Also some types of mismatches are more stable than others and than canonical AT base pair in some cases [37, 61]. Notwithstanding, the mismatch impact depends on a few factor, such as mismatch type, number and position, melting temperature, sequential environment and reaction condition [37, 48, 50–52, 61, 76, 79, 81].

Studies regarding the interactions of single, double and triple consecutive mismatches are rarely performed and a more quantitative assessment is complicated since a lack in parameters. Yet a recent parameter set for up to three contiguous mismatched base pairs were developed by our group [61]. An experimental melting temperatures set containing single, double and triple contiguous mismatches were modeled and compared to a canonical set under the same conditions. The parameters were shared in context independent and dependent, respectively, all mismatches are independent from the flanking pairs and the flanking pairs of mismatches are considered. This work showed that some types of mismatches may stabilise the mismatched duplex, perhaps higher than canonical AT base pair.

The emergence of SARS-CoV-2 coronavirus and its constant mutations prompted the question of how these mutations would interfere in the primer-target hybridisation. Since mismatches have been successful applied to diagnostic assays, e.g. allele-specific [37] and single nucleotide polymorphism (SNP) [68] identification, may be they play a valuable role in COVID-19 diagnostic. Therefore, the application of mismatch parameters to primers and probes could assist their specificity and continuous effectiveness.

This work shows how mismatches in primer-target hybridisation may affect PCR and LAMP primer coverages to detect SARS-CoV-2 coronavirus. Melting temperature $T_m$ is an important parameter to assess the efficiency of primers to hybridise to the specific region of the genome. From the calculation of $T_m$ for each alignment within the proper conditions, we were able to calculate the coverage of all primer sets to wild-type SARS-CoV-2, its variants, SARS-CoV-1 and five other coronaviruses (non-SARS) genomes. For both PCR and LAMP sets, we obtained similar outcomes for all genome sets. We found that mismatched alignments increase the coverage for SARS-CoV-2 in a considerable number of cases, which may contribute to enhance the detection of the virus. However, a few primers showed no coverage at all. The cross-reactivity was verified for SARS-CoV-1 and non-SARS coronaviruses. For the first, in both PCR and LAMP assessment, a substantial number of primers showed high coverages even when mismatches are considered. Yet some primers achieved no coverages. In the case of non-SARS genomes, only few primers showed some coverage.

We found that a significant number of primers may contribute to detect both wild-type and mutated SARS-CoV-2 genomes. Both PCR and LAMP primers achieved low and high coverages, even null coverages for a few variants. RdRp primers/probes from Corman et al. [213] set achieved null coverages for wild-type SARS-CoV-2, but high coverages when mismatches are taken into account. Experiments have shown that mismatches in this primer subset did not affect their ability to bind to the target [58, 241, 244]. In particular, a few RdRp primers achieved 100% partial coverage to the Delta variant. S3-F2 primer from Mohon et al. [232] set achieved 100% strict coverages for all subvariant sets, in contrast to two primers from Diego et al. [226] set which covered no subvariant sets. Alves et al. [2] set achieved both strict and partial coverages greater than 90% for variants, as well as experiments showed that E1 and N2 primer subsets did not prevent the target amplification of Omicron variant considering the mismatch presence [250]. In some cases, although the mismatched temperature is similar or greater than the reference temperature, the mismatched base pair shows a considerable destabilisation in the primer-target duplex. Highlight for a consecutive double mismatch from Munnink et al. [220] set that has a similar temperature to reference temperature while a small end fraying. Also a GA pair achieved an increase in the temperature showing a slight shift which may not hamper the amplification. In summary, some primers showed null strict coverages to greater than 90% partial coverages.

Finally, when we consider a null threshold $\Delta T_{\text{lim.}} = 0\ ^{\circ}\text{C}$ between mismatched temperature $T_{\text{MM}}$ and reference temperature $T_{\text{ref.}}$, the mismatch presence does not destabilise the duplex. A few primers show zero strict coverage, while achieve a greater than 90% coverage if stabilising mismatches are considered. Highlight for N1_F1c primer from Ji et al. [228] set that shows only partial coverages for Omicron variant and BA.2 and BA.3 subvariants. Note that all primers assessed here were designed to wild-type SARS-CoV-2.

Our findings are also in agreement with evidences which have been reported that mismatches depend on number, position and type, sequential environment, especially the flanking base pairs, and their impact is locally higher than globally. Nevertheless, this work shows that the impact of DNA mismatches in primer-target hybridisation is not straightforward and depends on a fully evaluation with detailed calculation with up-to-date model parameters. A detailed and in-depth investigation of the behaviour of entire collection of mismatches would be of great interest and application.

In addition, we applied our workflow to a SARS-CoV-1 LAMP primer set which showed good results. We could verify the specificity of a few primers and the mismatch effect on hybridisation temperatures. The analysis of LNA-modified LAMP primers also achieved considerable results showing the potential application of nucleic acid modification parameters to be incorporated in our workflow and an improvement in primer design in the future.

# 10 Perspectives

The workflow developed in this work to evaluate PCR and LAMP primers designed to SARS-CoV-2 allows us to analyse other primer sets to different types of targets, such as human immunodeficiency virus (HIV), oncogenes and other coronaviruses. We showed an example in section 8.1.3 applying our workflow to LAMP primers designed to detect SARS-CoV-1 coronavirus. It motivated us to extend this type of study to other primer sets and pathogens and, with collaboration, analyse experimental data for matched and mismatched hybridisations in both techniques. Moreover, molecular beacons are a special type of probes that has been used in PCR and LAMP methods, usually to quantificate the final product of the reaction and other techniques, such as fluorescence resonance energy transfer (FRET). Studying molecular beacons behaviour using our tools could improve the primer design and maintenance of the techniques similar to the work showed here. We are also interested in studying different structural modifications of DNA and RNA to amplify our parameter collection, which may be useful for several types of primer design. An example is the application of LNA modifications in PCR and LAMP primers to SARS-CoV-2 detection [298–301]. We have worked in an initial analysis of a published LNA-modified LAMP primer set designed to detect SARS-CoV-2 [295], see section 8.2.5. Also LNA mismatches are studied to amplify the assessment of modified primers and probes.

# References


[1] Pâmella Miranda and Gerald Weber. Thermodynamic evaluation of the impact of DNA mismatches in PCR-type SARS-CoV-2 primers and probes. *Mol. Cell. Probes*, 2021. doi:10.1016/j.mcp.2021.101707. in press.

[2] Pedro A. Alves, Ellen G. de Oliveira, Ana Paula M. Franco-Luiz, Letícia T. Almeida, Amanda B. Gonçalves, Iara A. Borges, Flávia de S. Rocha, Raissa P. Rocha, Matheus F. Bezerra, Pâmella Miranda, Flávio D. Capanema, Henrique R. Martins, Gerald Weber, Santuza M. R. Teixeira, Gabriel Luz Wallau, and Rubens L. do Monte-Neto. Optimization and clinical validation of colorimetric reverse transcription loop-mediated isothermal amplification, a fast, highly sensitive and specific COVID-19 molecular diagnostic tool that is robust to detect SARS-CoV-2 variants of concern. *Frontiers in Microbiology*, 12:3172, 2021. ISSN 1664-302X. doi:10.3389/fmicb.2021.713713. URL https://www.frontiersin.org/article/10.3389/fmicb.2021.713713.

[3] Pâmella Miranda, Vivianne Basílio Barbosa, and Gerald Weber. Mesoscopic evaluation of DNA mismatches in PCR primer-target hybridisation to detect SARS-CoV-2 variants of concern. In *Brazilian Symposium on Bioinformatics*, pages 145–150. Springer, 2021. doi:10.1007/978-3-030-91814-9_15.

[4] Pâmella Miranda, Pedro A Alves, Rubens L do Monte-Neto, and Gerald Weber. In silico thermodynamic evaluation of the effectiveness of RT-LAMP primers SARS-CoV-2 variants detection. *The Open COVID Journal*, 4(1), 2024. doi:10.2174/0126669587279780240130063422.

[5] Izabela Ferreira, Pâmella Miranda, Kira Astakhova, and Gerald Weber. Improved melting temperature predictions for single LNA modifications in synthetic DNA oligonucleotides. *Chemical Physics*, 579:112204, 2024. ISSN 0301-0104. doi:https://doi.org/10.1016/j.chemphys.2024.112204. URL https://www.sciencedirect.com/science/article/pii/S0301010424000338.

[6] Neil C Jones, Pavel A Pevzner, and Pavel Pevzner. *An introduction to bioinformatics algorithms*. MIT press, 2004.

[7] David L Nelson and Michael M Cox. *Princípios de Bioquímica de Lehninger*. Artmed Editora, 2014.

[8] Tom Strachan and Andrew Read. *Genética Molecular Humana*. Artmed Editora, 2013.

[9] Benjamin A Pierce. *Genética: um enfoque conceitual*. Guanabara Koogan, 2013.

[10] Jack A Tuszynski and Michal Kurzynski. *Introduction to molecular biophysics*. CRC press, 2003.

[11] Lizabeth A Allison. *Fundamental Molecular Biology*. Blackwell Publishing Ltd, 2007.

[12] Arthur M Lesk. *Introduction to genomics*. Oxford University Press, 2007.

[13] R. Olinski, M. Jurgowiak, and T. Zaremba. Uracil in DNA–its biological significance. *Mutation Research/Reviews in Mutation Research*, 705(3):239–245, 2010.

[14] Alice LB Pyne, Agnes Noy, Kavit HS Main, Victor Velasco-Berrelleza, Michael M Piperakis, Lesley A Mitchenall, Fiorella M Cugliandolo, Joseph G Beton, Clare EM Stevenson, Bart W Hoogenboom, et al. Base-pair resolution analysis of the effect of supercoiling on DNA flexibility and major groove recognition by triplex-forming oligonucleotides. *Nature communications*, 12(1):1–12, 2021. doi:https://doi.org/10.1038/s41467-021-21243-y.

[15] Jessica A. Brown. Unraveling the structure and biological functions of RNA triple helices. *WIREs RNA*, 11(6):e1598, 2020. doi:https://doi.org/10.1002/wrna.1598. URL https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wrna.1598.

[16] Jane Flint, Vincent Racaniello, Glenn Rall, Anna Marie Skalka, and Lynn Enquist. *Principles of Virology*. ASM PRESS, 2015.

[17] David Baltimore. Expression of animal virus genomes. *Bacteriological reviews*, 35(3): 235, 1971.

[18] Vassili Ivanov, Yan Zeng, and Giovanni Zocchi. Statistical mechanics of base stacking and pairing in DNA melting. *Phys. Rev. E*, 70(5):051907, 2004.

[19] Arnaldo Zaha, Henrique Bunselmeyer Ferreira, and Luciane MP Passaglia. *Biologia Molecular Básica-5*. Artmed Editora, 2014.

[20] Fabian Kilchherr, Christian Wachauf, Benjamin Pelz, Matthias Rief, Martin Zacharias, and Hendrik Dietz. Single-molecule dissection of stacking forces in DNA. *Science*, 353 (6304):aaf5508, 2016.

[21] Michael Moorhouse and Paul Barry. *Bioinformatics Biocomputing and Perl*. Wiley Online Library, 2004.

[22] Giuseppe Gaeta, Claude Reiss, Michel Peyrard, and Thierry Dauxois. Simple models of non-linear DNA dynamics. *Rivista del Nuovo cimento*, 17(4):1–48, 1994.

[23] Elizabeth van Pelt-Verkuil, Alex Van Belkum, and John P Hays. *Principles and technical aspects of PCR amplification*. Springer Science & Business Media, 2008.

[24] Ariel Afek, Honglue Shi, Atul Rangadurai, Harshit Sahay, Alon Senitzki, Suela Xhani, Mimi Fang, Raul Salinas, Zachery Mielko, Miles A Pufall, et al. DNA mismatches reveal conformational penalties in protein–DNA recognition. *Nature*, 587(7833):291–296, 2020. doi:https://doi.org/10.1038/s41586-020-2843-2.

[25] Kale Kundert and James S Fraser. DNA-binding proteins meet their mismatch. *Nature*, 587:199–200, 2020. doi:https://doi.org/10.1038/d41586-020-02658-x.

[26] Song-Hua Ke and Roger M. Wartell. Influence of nearest neighbor sequence on the stability of base pair mismatches in long DNA: determination by temperature-gradient gel electrophoresis. *Nucleic Acids Research*, 21(22):5137–5143, 11 1993. ISSN 0305-1048. doi:10.1093/nar/21.22.5137. URL https://doi.org/10.1093/nar/21.22.5137.

[27] Kashif Aziz Khan and Peter Cheung. Presence of mismatches between diagnostic PCR assays and coronavirus SARS-CoV-2 genome. *Royal Society Open Science*, 7(6):200636, 2020. doi:10.1098/rsos.200636. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsos.200636.

[28] Kevin S. Kuchinski, Agatha N. Jassem, and Natalie A. Prystajecky. Assessing oligonucleotide designs from early lab developed PCR diagnostic tests for SARS-CoV-2 using the PCR_strainer pipeline. *Journal of Clinical Virology*, 131:104581, 2020. ISSN 1386-6532. doi:https://doi.org/10.1016/j.jcv.2020.104581. URL https://www.sciencedirect.com/science/article/pii/S1386653220303231.

[29] Sean J Johnson and Lorena S Beese. Structures of mismatch replication errors observed in a DNA polymerase. *Cell*, 116(6):803–816, 2004. doi:10.1016/S0092-8674(04)00252-1.

[30] Giulia Rossetti, Pablo D Dans, Irene Gomez-Pinto, Ivan Ivani, Carlos Gonzalez, and Modesto Orozco. The structural impact of DNA mismatches. *Nucleic Acids Res.*, 43(8):4309–4321, 2015.

[31] Virginie Gervais, Jean AH Cognet, Marc Le Bret, Lawrence C Sowers, and G Victor Fazakerley. Solution structure of two mismatches A·A and T·T in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. *Eur. J. Biochem.*, 228(2):279–290, 1995.

[32] Anirban Ghosh, Rajiv Kumar Kar, Janarthanan Krishnamoorthy, Subhrangsu Chatterjee, and Anirban Bhunia. Double GC:GC mismatch in dsDNA enhances local dynamics retaining the DNA footprint: A high-resolution NMR study. *ChemMedChem*, 9(9):2059–2064, 2014.

[33] Nicodemus Tibanyenda, Simon H De Bruin, Cornelis AG Haasnoot, Gijs A Marel, van der, Jacques H Boom, van, and Cornelis W Hilbers. The effect of single base-pair mismatches on the duplex stability of d(T-A-T-T-A-A-T-A-T-C-A-A-G-T-T-G)· d (C-A-A-C-T-T-G-A-T-A-T-T-A-A-T-A). *Eur. J. Biochem.*, 139(1):19–27, 1984.

[34] Fareed Aboul-Ela, David Koh, Ignacio Tinoco Jr, and Francis H Martin. Base-base mismatches. thermodynamics of double helix formation for dCA3XA3G+ dCT3YT3G (X, Y= A, C, G, D). *Nucleic Acids Res.*, 13(13):4811–4824, 1985. doi:10.1093/nar/13.13.4811.

[35] Ying Li, Gerald Zon, and W David Wilson. Thermodynamics of DNA duplexes with adjacent G·A mismatches. *Biochem.*, 30(30):7566–7572, 1991.

[36] C Kunz, Y Saito, and P Schär. Dna repair in mammalian cells: Mismatched repair: variations on a theme. *Cellular and molecular life sciences*, 66:1021–1038, 2009.

[37] Ralph Stadhouders, Suzan D Pas, Jeer Anber, Jolanda Voermans, Ted HM Mes, and Martin Schutten. The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5′ nuclease assay. *J. Mol. Diagn.*, 12(1):109–117, 2010.

[38] Sofia Persson, Måns Karlsson, Henrik Borsch-Reniers, Patrik Ellström, Ronnie Eriksson, and Magnus Simonsson. Missing the match might not cost you the game: primer-template mismatches studied in different hepatitis A virus variants. *Food and Environmental Virology*, 11(3):297–308, 2019.

[39] Giuseppe Lippi, Ana-Maria Simundic, and Mario Plebani. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). *Clinical Chemistry and Laboratory Medicine (CCLM)*, 58(7):1070–1076, 2020. doi:doi:10.1515/cclm-2020-0285. URL https://doi.org/10.1515/cclm-2020-0285.

[40] Peihua Niu, Roujian Lu, Li Zhao, Huijuan Wang, Baoying Huang, Fei Ye, Wenling Wang, and Wenjie Tan. Three novel real-time rt-pcr assays for detection of covid-19 virus. *China CDC Weekly*, pages 1–5, 2020.

[41] Swati Gupta, Amit Kumar, Nivedita Gupta, Deepak R. Bharti, Neeraj Aggarwal, and Vasanthapuram Ravi. A two-step process for in silico screening to assess the performance of qRTPCR kits against variant strains of SARS-CoV-2. *BMC Genomics*, 23(1), nov 2022. doi:10.1186/s12864-022-08999-3. URL https://doi.org/10.1186%2Fs12864-022-08999-3.

[42] Tsugunori Notomi, Hiroto Okayama, Harumi Masubuchi, Toshihiro Yonekawa, Keiko Watanabe, Nobuyuki Amino, and Tetsu Hase. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.*, 28(12):e63–e63, 2000.

[43] Yi Zhou, Zhenzhou Wan, Shuting Yang, Yingxue Li, Min Li, Binghui Wang, Yihong Hu, Xueshan Xia, Xia Jin, Na Yu, and Chiyu Zhang. A mismatch-tolerant reverse transcription loop-mediated isothermal amplification method and its application on simultaneous detection of all four serotype of dengue viruses. *Frontiers in Microbiology*, 10, 2019. ISSN 1664-302X. doi:10.3389/fmicb.2019.01056. URL https://www.frontiersin.org/article/10.3389/fmicb.2019.01056.

[44] Michaela Hendling and Ivan Barišić. In-silico design of DNA oligonucleotides: Challenges and approaches. *Computational and Structural Biotechnology Journal*, 17: 1056–1065, 2019. ISSN 2001-0370. doi:https://doi.org/10.1016/j.csbj.2019.07.008. URL https://www.sciencedirect.com/science/article/pii/S2001037019300844.

[45] Mitsuhashi M, Cooper A, Ogura M, Shinagawa T, Yano K, and Hosokawa T. Oligonucleotide probe design — a new approach. *Nature*, 367:759–761, 1994. doi:https://doi.org/10.1038/367759a0.

[46] Masato Mitsuhashi. Technical report: Part 1. Basic requirements for designing optimal oligonucleotide probe sequences. *Journal of Clinical Laboratory Analysis*, 10(5):277–284, 1996. doi:10.1002/(SICI)1098-2825(1996)10:5<277::AID-JCLA8>3.0.CO;2-5.

[47] Jun-Sub Kim, Jun-Hyeong Jang, Jeong-Min Kim, Yoon-Seok Chung, Cheon-Kwon Yoo, and Myung-Guk Han. Genome-wide identification and characterization of point mutations in the SARS-CoV-2 genome. *Osong Public Health and Research Perspectives*, 11 (3):101, 2020.

[48] Ahmed Elaswad and Mohamed Fawzy. Mutations in animal SARS-CoV-2 induce mismatches with the diagnostic PCR assays. *Pathogens*, 10(3), 2021. ISSN 2076-0817. doi:10.3390/pathogens10030371. URL https://www.mdpi.com/2076-0817/10/3/371.

[49] Srinivas Ayyadevara, John J Thaden, and Robert J Shmookler Reis. Discrimination of primer 3′-nucleotide mismatch by Taq DNA polymerase during polymerase chain reaction. *Analytical biochemistry*, 284(1):11–18, 2000.

[50] D. Bru, F. Martin-Laurent, and L. Philippot. Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Applied and Environmental Microbiology*, 74(5):1660–1663,

2008. doi:10.1128/AEM.02403-07. URL https://journals.asm.org/doi/abs/10.1128/AEM.02403-07.

[51] S Kwok, DE Kellogg, N McKinney, D Spasic, L Goda, C Levenson, and JJ Sninsky. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. *Nucleic acids research*, 18(4):999–1005, 1990.

[52] Mei-Mei Huang, Norman Arnheim, and Myron F Goodman. Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Research*, 20(17):4567–4573, 1992.

[53] Derek Toms, Julang Li, and Hugh Y. Cai. Evaluation of WHO listed COVID-19 qPCR primers and probe in silico with 375 SERS-CoV-2 full genome sequences. *medRxiv*, 2020. doi:10.1101/2020.04.22.20075697. URL https://www.medrxiv.org/content/early/2020/04/28/2020.04.22.20075697.

[54] Steve Lefever, Filip Pattyn, Jan Hellemans, and Jo Vandesompele. Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clinical chemistry*, 59(10):1470–1480, 2013.

[55] Chonticha Klungthong, Piyawan Chinnawirotpisan, Kittinun Hussem, Thipwipha Phonpakobsin, Wudtichai Manasatienkij, Chuanpis Ajariyakhajorn, Kamonthip Rungrojcharoenkit, Robert V Gibbons, and Richard G Jarman. The impact of primer and probe-template mismatches on the sensitivity of pandemic influenza A/H1N1/2009 virus detection by real-time RT-PCR. *J. Clin. Virol.*, 48(2):91–95, 2010.

[56] David M Whiley and Theo P Sloots. Sequence variation in primer targets affects the accuracy of viral quantitative PCR. *Journal of Clinical Virology*, 34(2):104–107, 2005.

[57] Brett M Ledeker and K Susan. The effect of multiple primer–template mismatches on quantitative PCR accuracy and development of a multi-primer set assay for accurate quantification of pcra gene sequence variants. *Journal of microbiological methods*, 94 (3):224–231, 2013.

[58] Stephen Bustin, Sara Kirvell, Jim F. Huggett, and Tania Nolan. RT-qPCR diagnostics: The "Drosten" SARS-CoV-2 assay paradigm. *Int. J. Mol. Sci.*, 22(16), 2021. ISSN 1422-0067. doi:10.3390/ijms22168702. URL https://www.mdpi.com/1422-0067/22/16/8702.

[59] Cindy Christopherson, John Sninsky, and Shirley Kwok. The effects of internal primer-template mismatches on RT-PCR: HIV-1 model studies. *Nucleic Acids Research*, 25 (3):654–658, 02 1997. ISSN 0305-1048. doi:10.1093/nar/25.3.654. URL https://doi.org/10.1093/nar/25.3.654.

[60] Kean Hean Ooi, Mengying Mandy Liu, Jie Wen Douglas Tay, Seok Yee Teo, Pornchai Kaewsapsak, Shengyang Jin, Chun Kiat Lee, Jingwen Hou, Sebastian Maurer-Stroh, Weisi Lin, et al. An engineered CRISPR-Cas12a variant and DNA-RNA hybrid guides enable robust and rapid COVID-19 testing. *Nature Communications*, 12(1):1739, 2021. doi:10.1038/s41467-021-21996-6.

[61] Luciana M. Oliveira, Adam S. Long, Tom Brown, Keith R. Fox, and Gerald Weber. Melting temperature measurement and mesoscopic evaluation of single, double and triple DNA mismatches. *Chem. Sci.*, 11:8273–8287, 2020. doi:10.1039/d0sc01700k. URL https://pubs.rsc.org/en/content/articlelanding/2020/SC/D0SC01700K.

[62] Sarah E Morse and David E Draper. Purine–purine mismatches in RNA helices: evidence for protonated G·A pairs and next-nearest neighbor effects. *Nucleic Acids Res.*, 23(2): 302–306, 1995.

[63] Amy E. Walter, Ming Wu, and Douglas H. Turner. The stability and structure of tandem GA mismatches in RNA depend on closing base pairs. *Biochemistry*, 33(37):11349–11354, 1994. doi:10.1021/bi00203a033. URL https://doi.org/10.1021/bi00203a033. PMID: 7537087.

[64] Ming Wu, Jeffrey A McDowell, and Douglas H Turner. A periodic table of tandem mismatches in RNA. *Biochem.*, 34(10):3204–3211, 1995.

[65] Thomas Naiser, Oliver Ehler, Jona Kayser, Timo Mai, Wolfgang Michel, and Albrecht Ott. Impact of point-mutations on the hybridization affinity of surface-bound DNA/DNA and RNA/DNA oligonucleotide-duplexes: comparison of single base mismatches and base bulges. *BMC biotechnology*, 8(1):1–23, 2008. doi:https://doi.org/10.1186/1472-6750-8-48.

[66] Deguo Wang. Effect of internal primr-template mismatches on loop-mediated isothermal amplification. *Biotechnology & Biotechnological Equipment*, 30(2):314–318, 2016. doi:10.1080/13102818.2015.1125765. URL https://doi.org/10.1080/13102818.2015.1125765.

[67] Naoki Sugimoto, Mariko Nakano, and Shu-ichi Nakano. Thermodynamics- structure relationship of single mismatches in RNA/DNA duplexes. *Biochemistry*, 39(37):11270–11281, 2000.

[68] Neil J. Gibson. The use of real-time PCR methods in DNA sequence variation analysis. *Clinica Chimica Acta*, 363(1):32–47, 2006. ISSN 0009-8981. doi:https://doi.org/10.1016/j.cccn.2005.06.022. URL https://www.sciencedirect.com/science/article/pii/S0009898105004237.

[69] Marcelino Varona and Jared L Anderson. Advances in mutation detection using loop-mediated isothermal amplification. *ACS omega*, 6(5):3463–3469, 2021.

[70] Margherita Scapaticci, Andrea Bartolini, Francesca Vitone, Vincenzo Cerreta, Monica Vignoli, Elena Gnudi, Alessandra Frazzoni, Barbara Sitta, Silvia Capitani, Annamaria Lopriore, Mariapina Donadio, Stefania Chiarastella, Marina Bioli, and Rita Mancini. Detection of a characteristic melting profile of a SARS-CoV-2 Kappa variant in Italy using the SARS-CoV-2 variants ELITe MGB® kit. page 114458, 2022. doi:10.1016/j.jviromet.2022.114458.

[71] Jeremy A. Garson, Samuel Badru, Eleanor Parker, Richard S. Tedder, and Myra O. McClure. Highly sensitive and specific detection of the SARS-CoV-2 Delta variant by double-mismatch allele-specific real time reverse transcription PCR. *Journal of Clinical Virology*, 146:105049, 2022. ISSN 1386-6532. doi:https://doi.org/10.1016/j.jcv.2021.105049. URL https://www.sciencedirect.com/science/article/pii/S1386653221003164.

[72] Petros Bozidis, Eleni Petridi, and Konstantina Gartzonika. An ARMS-Multiplex PCR targeting SARS-CoV-2 Omicron sub-variants. *Pathogens*, 12(8):1017, 2023. doi:10.3390/pathogens12081017.

[73] S Creighton, MM Huang, HONG Cai, N Arnheim, and MF Goodman. Base mispair extension kinetics. Binding of avian myeloblastosis reverse transcriptase to matched and mismatched base pair termini. *Journal of Biological Chemistry*, 267(4):2633–2639, 1992.

[74] L V Mendelman, J Petruska, and M F Goodman. Base mispair extension kinetics. comparison of DNA polymerase alpha and reverse transcriptase. *Journal of Biological Chemistry*, 265(4):2338–2346, feb 1990. doi:10.1016/s0021-9258(19)39981-8. URL https://doi.org/10.1016/s0021-9258(19)39981-8.

[75] Brian Boyle, Nancy Dallaire, and John MacKay. Evaluation of the impact of single nucleotide polymorphisms and primer mismatches on quantitative PCR. *BMC biotechnology*, 9:1–15, 2009.

[76] Thomas Ehlen and Louis Dubeau. Detection of ras point mutations by polymerase chain reaction using mutation-specific, inosine-containing oligonucleotide primers. *Biochemical and biophysical research communications*, 160(2):441–447, 1989.

[77] John SantaLucia. Physical principles and visual-OMP software for optimal PCR design. *PCR primer design*, pages 3–33, 2007.

[78] Masato Mitsuhashi. Technical report: Part 2. Basic requirements for designing optimal pcr primers. *Journal of Clinical Laboratory Analysis*, 10(5):285–293, 1996.

[79] Chulhyun Lee, Hae-Kap Cheong, Jee-Hyun Cho, and Chaejoon Cheong. AA mismatched DNAs with a single base difference exhibit a large structural change and a propensity for the parallel-stranded conformation. *Journal of Analytical Science & Technology*, 1:37–48, 2010.

[80] Nick A Rejali, Endi Moric, and Carl T Wittwer. The effect of single mismatches on primer extension. *Clinical chemistry*, 64(5):801–809, 2018.

[81] Myong-Rim Ri, Jin-Sok Kang, Myong-Ryong Ri, and Song Nam U. Models for prediction of single base primer extension efficiency from position and type of single mismatch in primer-template duplex. *Heliyon*, 9(8):e18829, aug 2023. doi:10.1016/j.heliyon.2023.e18829. URL https://doi.org/10.1016/j.heliyon.2023.e18829.

[82] Beate Süß, Gabriele Flekna, Martin Wagner, and Ingeborg Hein. Studying the effect of single mismatches in primer and probe binding regions on amplification curves and quantification in real-time PCR. *Journal of Microbiological Methods*, 76(3): 316–319, 2009. ISSN 0167-7012. doi:https://doi.org/10.1016/j.mimet.2008.12.003. URL https://www.sciencedirect.com/science/article/pii/S0167701208004387.

[83] Carmen Koo, Simrandeep Kaur, Zhi-Yong Teh, Helen Xu, Amna Nasir, Yee-Ling Lai, Erum Khan, Lee-Ching Ng, and Hapuarachchige C Hapuarachchi. Genetic variability in probe binding regions explains false negative results of a molecular assay for the detection of dengue virus. *Vector-Borne and Zoonotic Diseases*, 16(7):489–495, 2016.

[84] Rita Sipos, Anna J Székely, Márton Palatinszky, Sára Révész, Károly Márialigeti, and Marcell Nikolausz. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS microbiology ecology*, 60(2):341–350, 2007.

[85] Kousuke Ishii and Manabu Fukui. Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR. *Applied and Environmental Microbiology*, 67(8):3753–3755, 2001.

[86] Charlotte W Pratt and Kathleen Cornely. *Essential Biochemistry*. Guanabara Koogan, 2004.

[87] Lauren A Levine, Matthew Junker, Myranda Stark, and Dustin Greenleaf. A DNA melting exercise for a large laboratory class. *J. Chem. Educ.*, 92:1928–1931, 2015. doi:10.1021/acs.jchemed.5b00049.

[88] Sherrie Schreiber-Gosche and Robert A Edwards. Thermodynamics of oligonucleotide duplex melting. *J. Chem. Educ.*, 86(5):644, 2009.

[89] Kathleen P Howard. Thermodynamics of DNA duplex formation: A biophysical chemistry laboratory experiment. *J. Chem. Educ.*, 77(11):1469, 2000.

[90] Richard Owczarzy. Melting temperatures of nucleic acids: discrepancies in analysis. *Biophys. Chem.*, 117(3):207–215, 2005.

[91] Jean-Louis Mergny and Laurent Lacroix. Analysis of thermal melting curves. *Oligonucleotides*, 13(6):515–537, 2003.

[92] Nicolas von Ahsen, Michael Oellerich, and Ekkehard Schutz. Limitations of genotyping based on amplicon melting temperature. *Clin. Chem.*, 47(7):1331–1332, 2001. URL http://www.clinchem.org.

[93] June D. Almeida and D. A. J. Tyrrell. The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture. *Journal of General Virology*, 1(2):175–178, 1967. ISSN 1465-2099. doi:https://doi.org/10.1099/0022-1317-1-2-175. URL https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-1-2-175.

[94] 2021. https://www.bbc.com/news/uk-scotland-52278716.

[95] 2021. https://www.nationalgeographic.com/history/article/june-almeida-discovered-coronaviruses-decades-ago-little\-recognition.

[96] Na Zhu, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, Baoying Huang, Weifeng Shi, Roujian Lu, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, 2020. doi:https://doi.org/10.1056/NEJMoa2001017.

[97] Ying Yan, Le Chang, and Lunan Wang. Laboratory testing of SARS-CoV, MERS-CoV, and SARS-CoV-2 (2019-nCoV): Current status, challenges, and countermeasures. *Reviews in Medical Virology*, 30(3):e2106, 2020. doi:https://doi.org/10.1002/rmv.2106. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rmv.2106.

[98] Sahoo C, Mishra A, and Moharana A. Novel coronavirus (SARS-CoV-2) 2019: A systematic review. *European Journal of General Medicine*, 17, 05 2020. doi:10.29333/ejgm/8258.

[99] Yasin Orooji, Hessamaddin Sohrabi, Nima Hemmat, Fatemeh Oroojalian, Behzad Baradaran, Ahad Mokhtarzadeh, Mohamad Mohaghegh, and Hassan Karimi-Maleh. An overview on SARS-CoV-2 (COVID-19) and other human coronaviruses

and their detection capability via amplification assay, chemical sensing, biosensing, immunosensing, and clinical assays. *Nano-Micro Letters*, 13(1):1–30, 2021. doi:https://doi.org/10.1007/s40820-020-00533-y.

[100] Piepoli S, Shamloo B, Bircan A, Adebali O, and Erman B. Molecular biology of SARS-CoV-2. *Turkish Journal of Immunology*, 8(2):73–88, 2020. doi:https://doi.org/10.25002/tji.2020.1293.

[101] Ankur Das, Raja Ahmed, Suraiya Akhtar, Khaleda Begum, and Sofia Banu. An overview of basic molecular biology of SARS-CoV-2 and current COVID-19 prevention strategies. *Gene Reports*, 23:101122, 2021. ISSN 2452-0144. doi:https://doi.org/10.1016/j.genrep.2021.101122. URL https://www.sciencedirect.com/science/article/pii/S2452014421001072.

[102] Hiya Islam, Ahsab Rahman, Jaasia Masud, Dipita Shweta, Yusha Araf, Md Ullah, Syed Muktadir Al Sium, and Bishajit Sarkar. A generalized overview of SARS-CoV-2: Where does the current knowledge stand? *European Journal of General Medicine*, 17, 05 2020. doi:10.29333/ejgm/8258.

[103] Grigore Mihaescu, Mariana Carmen Chifiriuc, Ciprian Iliescu, Corneliu Ovidiu Vrancianu, Lia-Mara Ditu, Luminita Gabriela Marutescu, Raluca Grigore, Șerban Bertești eanu, Marian Constantin, and Gratiela Gradisteanu Pircalabioru. SARS-CoV-2: From structure to pathology, host immune response and therapeutic management. *Microorganisms*, 8(10), 2020. ISSN 2076-2607. doi:10.3390/microorganisms8101468. URL https://www.mdpi.com/2076-2607/8/10/1468.

[104] Buddhisha Udugama, Pranav Kadhiresan, Hannah N. Kozlowski, Ayden Malekjahani, Matthew Osborne, Vanessa Y. C. Li, Hongmin Chen, Samira Mubareka, Jonathan B. Gubbay, and Warren C. W. Chan. Diagnosing COVID-19: The disease and tools for detection. *ACS Nano*, 14(4):3822–3835, 2020. doi:10.1021/acsnano.0c02624. URL https://doi.org/10.1021/acsnano.0c02624. PMID: 32223179.

[105] SK Manirul Haque, Omar Ashwaq, Abdulla Sarief, and Abdul Kalam Azad John Mohamed. A comprehensive review about SARS-CoV-2. *Future Virology*, 15(9):625–648, 2020. doi:10.2217/fvl-2020-0124. URL https://doi.org/10.2217/fvl-2020-0124. PMID: 33224265.

[106] Zilong Zhao, Yaling Wang, Liping Qiu, Ting Fu, Yu Yang, Ruizi Peng, Mengyu Guo, Lichun Mao, Chunying Chen, Yuliang Zhao, et al. New insights from chemical biology: molecular basis for transmission, diagnosis and therapy of SARS-CoV-2. *CCS Chemistry*, pages 1–41, 2020.

[107] 2021. https://www.nytimes.com/interactive/2021/health/coronavirus-variant-tracker.html.

[108] Shuo Su, Gary Wong, Weifeng Shi, Jun Liu, Alexander C.K. Lai, Jiyong Zhou, Wenjun Liu, Yuhai Bi, and George F. Gao. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends in Microbiology*, 24(6):490–502, 2016. ISSN 0966-842X. doi:https://doi.org/10.1016/j.tim.2016.03.003. URL https://www.sciencedirect.com/science/article/pii/S0966842X16000718.

[109] Manish Tiwari and Divya Mishra. Investigating the genomic landscape of novel coronavirus (2019-nCoV) to identify non-synonymous mutations for use in diagnosis and drug design. *Journal of Clinical Virology*, 128:104441, 2020. ISSN 1386-6532. doi:https://doi.org/10.1016/j.jcv.2020.104441. URL https://www.sciencedirect.com/science/article/pii/S1386653220301839.

[110] Yuxuan Zhang, Zhiwei Huang, Jiajie Zhu, Chaonan Li, Zhongbiao Fang, Keda Chen, and Yanjun Zhang. An updated review of SARS-CoV-2 detection methods in the context of a novel coronavirus pandemic. *Bioengineering & Translational Medicine*, jun 2022. doi:10.1002/btm2.10356. URL https://doi.org/10.1002%2Fbtm2.10356.

[111] João Carneiro, Catarina Gomes, Cátia Couto, and Filipe Pereira. CoV2ID: Detection and therapeutics oligo database for SARS-CoV-2. *bioRxiv*, 2020. doi:10.1101/2020.04.19.048991. URL https://www.biorxiv.org/content/early/2020/06/12/2020.04.19.048991.

[112] Rafael Sanjuán and Pilar Domingo-Calap. Mechanisms of viral mutation. *Cellular and molecular life sciences*, 73(23):4433–4448, 2016. doi:https://doi.org/10.1007/s00018-016-2299-6.

[113] Tyson E. Graber, Élisabeth Mercier, Patrick M. D'Aoust, Huy-Dung Hoang, Xin Tian, Syeda Tasneem, Kamya Bhatnagar, and Robert Delatolla. An allele-specific primer extension assay to quantify the proportion of B.1.1.7-specific SARS-CoV-2 RNA in wastewater. *medRxiv*, 2021. doi:10.1101/2021.02.22.21252041. URL https://www.medrxiv.org/content/early/2021/02/23/2021.02.22.21252041.

[114] Takahiko Koyama, Daniel Platt, and Laxmi Parida. Variant analysis of SARS-CoV-2 genomes. *Bulletin of the World Health Organization*, 98(7):495, 2020.

[115] Xingguang Li, Wei Wang, Xiaofang Zhao, Junjie Zai, Qiang Zhao, Yi Li, and Antoine Chaillon. Transmission dynamics and evolutionary history of 2019-nCoV. *Journal of Medical Virology*, 92(5):501–511, feb 2020. doi:10.1002/jmv.25701. URL https://doi.org/10.1002/jmv.25701.

[116] Sebastian Duchene, Leo Featherstone, Melina Haritopoulou-Sinanidou, Andrew Rambaut, Philippe Lemey, and Guy Baele. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evolution*, 6(2):veaa061, 08 2020. ISSN 2057-1577. doi:10.1093/ve/veaa061. URL https://doi.org/10.1093/ve/veaa061.

[117] A. Shirvani, L. Azimi, R. Mansour Ghanaie, M. Alebouyeh, F. Fallah, S. Rafiei Tabatabaei, M. Rajabnejad, S. Armin, S. A. Fahimzad, F. Shirvani, S. Maham, and A. Karimi. Utility of available methods for diagnosing SARS-CoV-2 in clinical samples. *Archives of Pediatric Infectious Diseases*, 8(3):e103677, 2020. ISSN 2322-1836. doi:10.5812/pedinfect.103677. URL https://sites.kowsarpub.com/apid/articles/103677.html.

[118] Minghua Jiang, Weihua Pan, Amir Arasthfer, Wenjie Fang, Liyan Ling, Hua Fang, Farnaz Daneshnia, Jian Yu, Wanqing Liao, Hao Pei, Xiaojing Li, and Cornelia Lass-Flörl. Development and validation of a rapid, single-step reverse transcriptase loop-mediated isothermal amplification (RT-LAMP) system potentially to be used for reliable and high-throughput screening of COVID-19. *Frontiers in Cellular and Infection Microbiology*, 10:331, 2020. ISSN 2235-2988. doi:10.3389/fcimb.2020.00331. URL https://www.frontiersin.org/article/10.3389/fcimb.2020.00331.

[119] Rui Wang, Yuta Hozumi, Changchuan Yin, and Guo-Wei Wei. Decoding SARS-CoV-2 transmission and evolution and ramifications for COVID-19 diagnosis, vaccine, and medicine. *Journal of Chemical Information and Modeling*, 60(12):5853–5865, 2020. doi:10.1021/acs.jcim.0c00501.

[120] Zijie Shen, Yan Xiao, Lu Kang, Wentai Ma, Leisheng Shi, Li Zhang, Zhuo Zhou, Jing Yang, Jiaxin Zhong, Donghong Yang, Li Guo, Guoliang Zhang, Hongru Li, Yu Xu, Mingwei Chen, Zhancheng Gao, Jianwei Wang, Lili Ren, and Mingkun Li. Genomic diversity of severe acute respiratory syndrome–coronavirus 2 in patients with coronavirus disease 2019. *Clinical Infectious Diseases*, 71(15):713–720, 03 2020. ISSN 1058-4838. doi:10.1093/cid/ciaa203. URL https://doi.org/10.1093/cid/ciaa203.

[121] Begum Cosar, Zeynep Yagmur Karagulleoglu, Sinan Unal, Ahmet Turan Ince, Dilruba Beyza Uncuoglu, Gizem Tuncer, Bugrahan Regaip Kilinc, Yunus Emre Ozkan, Hikmet Ceyda Ozkoc, Ibrahim Naki Demir, et al. SARS-CoV-2 mutations and their viral variants. *Cytokine & growth factor reviews*, 63:10–22, 2022. doi:10.1016/j.cytogfr.2021.06.001. URL http://dx.doi.org/10.1016/j.cytogfr.2021.06.001.

[122] Marta Itarte, Sílvia Bofill-Mas, Sandra Martínez-Puchol, Helena Torrell, Adrià Ceretó, Marina Carrasco, Eva Forés, Núria Canela, Rosina Girones, and Marta Rusiñol. Looking for a needle in a haystack. SARS-CoV-2 variant characterization in

sewage. *Current Opinion in Environmental Science & Health*, 24:100308, dec 2021. doi:10.1016/j.coesh.2021.100308. URL https://doi.org/10.1016/j.coesh.2021.100308.

[123] Jiahui Chen, Rui Wang, Menglun Wang, and Guo-Wei Wei. Mutations strengthened SARS-CoV-2 infectivity. *Journal of molecular biology*, 432(19):5212–5226, 2020. doi:https://doi.org/10.1016/j.jmb.2020.07.009.

[124] Anshumali Mittal, Kavyashree Manjunath, Rajesh Kumar Ranjan, Sandeep Kaushik, Sujeet Kumar, and Vikash Verma. COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2. *PLOS Pathogens*, 16(8): e1008762, August 2020. ISSN 1553-7374. doi:10.1371/journal.ppat.1008762. URL http://dx.doi.org/10.1371/journal.ppat.1008762.

[125] Andreas C. Chrysostomou, Antonia Aristokleous, Johana Hezka Rodosthenous, Christina Christodoulou, Georgia Stathi, and Leondios G. Kostrikis. Detection of circulating SARS-CoV-2 variants of concern (VOCs) using a multiallelic spectral genotyping assay. *Life*, 13(2), 2023. ISSN 2075-1729. doi:10.3390/life13020304. URL https://www.mdpi.com/2075-1729/13/2/304.

[126] Divya Sharma, Chengjin Ye, Giusppe Lippi, Jordi B. Torrelles, Luis Martinez-Sobrido, M. Michael Gromiha, and Brandon Michael Henry. In silico evaluation of the impact of the omicron variant on the sensitivity of RT-qPCR assays for SARS-CoV-2 detection using whole genome sequencing. *Research Square Preprint*, jan 2022. doi:10.21203/rs.3.rs-1220446/v1. URL https://doi.org/10.21203%2Frs.3.rs-1220446%2Fv1.

[127] Divya Sharma, Kin Israel Notarte, Rey Arturo Fernandez, Giuseppe Lippi, M. Michael Gromiha, and Brandon M. Henry. In silico evaluation of the impact of omicron variant of concern sublineage BA.4 and BA.5 on the sensitivity of RT-qPCR assays for SARS-CoV-2 detection using whole genome sequencing. *Journal of Medical Virology*, 95(1), nov 2022. doi:10.1002/jmv.28241. URL https://doi.org/10.1002%2Fjmv.28241.

[128] Erik Volz, Swapnil Mishra, Meera Chand, Jeffrey C Barrett, Robert Johnson, Lily Geidelberg, Wes R Hinsley, Daniel J Laydon, Gavin Dabrera, Áine O'Toole, et al. Assessing transmissibility of SARS-CoV-2 lineage B. 1.1. 7 in England. *Nature*, 593(7858):266–269, 2021.

[129] Mónica L Acevedo, Aracelly Gaete-Argel, Luis Alonso-Palomares, Marco Montes de Oca, Andrés Bustamante, Aldo Gaggero, Fabio Paredes, Claudia P Cortes, Sergio

Pantano, Constanza Martínez-Valdebenito, et al. Differential neutralizing antibody responses elicited by CoronaVac and BNT162b2 against SARS-CoV-2 Lambda in Chile. *Nature Microbiology*, 7(4):524–529, 2022.

[130] 2021. https://www.who.int/en/activities/tracking-{SARS-CoV-2}-variants/.

[131] 2023. https://www.who.int/news/item/16-03-2023-statement-\on-the-update-of-who-s-working-definitions-and-tracking-system\-for-sars-cov-2-variants-of-concern-and-variants-of-interest.

[132] Chenyu Li, David Debruyne, Julia Spencer, Vidushi Kapoor, Lily Y. Liu, Bo Zhang, Lucie Lee, Rounak Feigelman, Grayson Burdon, Jeffrey Liu, Alejandra Oliva, Adam Borcherding, Jian Xu, Alexander E. Urban, Guoying Liu, and Zhitong Liu. High sensitivity detection of SARS-CoV-2 using multiplex PCR and a multiplex-PCR-based metagenomic method. *bioRxiv*, 2020. doi:10.1101/2020.03.12.988246. URL https://www.biorxiv.org/content/early/2020/03/14/2020.03.12.988246.

[133] Renfei Lu, Jian Wang, Min Li, Yaqi Wang, Jia Dong, and Weihua Cai. SARS-CoV-2 detection using digital PCR for COVID-19 diagnosis, treatment monitoring and criteria for discharge. *medRxiv*, 2020. doi:10.1101/2020.03.24.20042689. URL https://www.medrxiv.org/content/early/2020/03/30/2020.03.24.20042689.

[134] Matthew Frankel, Kenn Forberg, Kelly E Coller, Michael G Berg, John Hackett Jr, Gavin Cloherty, and George J Dawson. Development of a high-throughput multiplexed real time RT-PCR assay for detection of human pegivirus 1 and 2. *Journal of virological methods*, 241:34–40, 2017.

[135] Elham Sheikhzadeh, Shimaa Eissa, Aziah Ismail, and Mohammed Zourob. Diagnostic techniques for COVID-19 and new developments. *Talanta*, 220:121392, 2020. ISSN 0039-9140. doi:https://doi.org/10.1016/j.talanta.2020.121392. URL https://www.sciencedirect.com/science/article/pii/S0039914020306834.

[136] Ehsan Shabani, Sayeh Dowlatshahi, and Mohammad J Abdekhodaie. Laboratory detection methods for the human coronaviruses. *European Journal of Clinical Microbiology & Infectious Diseases*, pages 1–22, 2020.

[137] Alexey S. Chubarov, Igor P. Oscorbin, Maxim L. Filipenko, Alexander A. Lomzov, and Dmitrii V. Pyshnyi. Allele-specific PCR for KRAS mutation detection using phosphoryl guanidine modified primers. *Diagnostics*, 10(11), 2020. ISSN 2075-4418. doi:10.3390/diagnostics10110872. URL https://www.mdpi.com/2075-4418/10/11/872.

[138] Ruano G, Brash D, and Kidd K. PCR: The first few cycles. *Amplifications - a forum for PCR users*, 1991.

[139] Son V Chu, Son T Vu, Hang M Nguyen, Ngan T Le, Phuong T Truong, Van TT Vu, Thuy TB Phung, and Anh TV Nguyen. Fast and sensitive real-time PCR detection of major antiviral drug-resistance mutations in chronic hepatitis B patients using a predesigned panel of Locked-Nucleic-Acid TaqMan probes. *Journal of Clinical Microbiology*, pages JCM–00936, 2021. doi:10.1128/JCM.00936-21.

[140] Brenda Thornton and Chhandak Basu. Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education*, 39(2):145–154, 2011.

[141] Nicola Louise Jones. PCR: Principles, procedures, and parameters. *PCR Mutation Detection Protocols*, pages 37–46, 2002.

[142] Dandan Li, Jiawei Zhang, and Jinming Li. Primer design for quantitative real-time PCR for the emerging Coronavirus SARS-CoV-2. *Theranostics*, 10(16):7150, 2020.

[143] Anders Bergkvist, Malcolm Burns, Rebecca Sanders, Nicholas Redshaw, and Tim Wilkes. Good practice guide for the application of quantitative PCR (qPCR), 2013.

[144] Arundhati Rao, DY Goldstein, Donna M Wolk, and Leslie A Wolf. Development and evaluation of two SARS-CoV-2 RT-PCR laboratory developed tests on the ARIES® automated, sample-to-answer, real-time PCR system. *Accessed March*, 12, 2020.

[145] CW Dieffenbach, TM Lowe, and GS Dveksler. General concepts for PCR primer design. *PCR methods appl*, 3(3):S30–S37, 1993.

[146] W. Rychlik, W.J. Spencer, and R.E. Rhoads. Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Research*, 18(21):6409–6412, 11 1990. ISSN 0305-1048. doi:10.1093/nar/18.21.6409. URL https://doi.org/10.1093/nar/18.21.6409.

[147] Elena Kalle, Mikael Kubista, and Christopher Rensing. Multi-template polymerase chain reaction. *Biomolecular Detection and Quantification*, 2:11–29, 2014. ISSN 2214-7535. doi:https://doi.org/10.1016/j.bdq.2014.11.002. URL https://www.sciencedirect.com/science/article/pii/S2214753514000102.

[148] Myungsun Park, Joungha Won, Byung Yoon Choi, and C Justin Lee. Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Experimental & molecular medicine*, 52 (6):963–977, 2020.

[149] Elfath M. Elnifro, Ahmed M. Ashshi, Robert J. Cooper, and Paul E. Klapper. Multiplex PCR: Optimization and application in diagnostic virology. *Clinical Microbiology Reviews*, 13(4):559–570, 2000. doi:10.1128/CMR.13.4.559. URL https://journals.asm.org/doi/abs/10.1128/CMR.13.4.559.

[150] Kerstin Wernike, Markus Keller, Franz J. Conraths, Thomas C. Mettenleiter, Martin H. Groschup, and Martin Beer. Pitfalls in SARS-CoV-2 PCR diagnostics. *Transboundary and Emerging Diseases*, n/a(n/a), 2020. doi:https://doi.org/10.1111/tbed.13684. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/tbed.13684.

[151] Tao Suo, Xinjin Liu, Jiangpeng Feng, Ming Guo, Wenjia Hu, Dong Guo, Hafiz Ullah, Yang Yang, Qiuhan Zhang, Xin Wang, Muhanmmad Sajid, Zhixiang Huang, Liping Deng, Tielong Chen, Fang Liu, Ke Xu, Yuan Liu, Qi Zhang, Yingle Liu, Yong Xiong, Guozhong Chen, Ke Lan, and Yu Chen. ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens. *Emerging Microbes & Infections*, 9(1):1259–1268, 2020. doi:10.1080/22221751.2020.1772678. URL https://doi.org/10.1080/22221751.2020.1772678. PMID: 32438868.

[152] Kiran Munne, Venkanna Bhanothu, Vikrant Bhor, Vainav Patel, Smita D Mahale, and Shailesh Pande. Detection of SARS-CoV-2 infection by RT-PCR test: factors influencing interpretation of results. *VirusDisease*, pages 1–3, 2021. doi:https://doi.org/10.1007/s13337-021-00692-5.

[153] K. Nagamine, T. Hase, and T. Notomi. Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Molecular and Cellular Probes*, 16(3):223–229, 2002. ISSN 0890-8508. doi:https://doi.org/10.1006/mcpr.2002.0415. URL https://www.sciencedirect.com/science/article/pii/S0890850802904159.

[154] Lisa Becherer, Nadine Borst, Mohammed Bakheit, Sieghard Frischmann, Roland Zengerle, and Felix von Stetten. Loop-mediated isothermal amplification (LAMP)–review and classification of methods for sequence-specific detection. *Analytical Methods*, 12(6):717–746, 2020.

[155] Weihua Yang, Xiaofei Dang, Qingxi Wang, Mingjie Xu, Qianqian Zhao, Yunying Zhou, Huailong Zhao, Li Wang, Yihui Xu, Jun Wang, et al. Rapid and accurate detection of three genes in SARS-CoV-2 using RT-LAMP method. *Research Square*, 2020. doi:https://doi.org/10.21203/rs.3.rs-28070/v1.

[156] Krzysztof Pyrc, Aleksandra Milewska, and Jan Potempa. Development of loop-mediated isothermal amplification assay for detection of human coronavirus-NL63. *Journal of Virological Methods*, 175(1):133–136, 2011. ISSN 0166-0934. doi:https://doi.org/10.1016/j.jviromet.2011.04.024. URL https://www.sciencedirect.com/science/article/pii/S0166093411001625.

[157] Hong Thi Cam Thai, Mai Quynh Le, Cuong Duc Vuong, Manmohan Parida, Harumi Minekawa, Tsugunori Notomi, Futoshi Hasebe, and Kouichi Morita. Development and evaluation of a novel loop-mediated isothermal amplification method for rapid detection of severe acute respiratory syndrome coronavirus. *Journal of Clinical microbiology*, 42 (5):1956–1961, 2004.

[158] Kazuya Shirato, Takuya Yano, Syouhei Senba, Shigehiro Akachi, Takashi Kobayashi, Takamichi Nishinaka, Tsugunori Notomi, and Shutoku Matsuyama. Detection of Middle East respiratory syndrome coronavirus using reverse transcription loop-mediated isothermal amplification (RT-LAMP). *Virology journal*, 11:1–11, 2014. doi:10.1186/1743-422X-11-139.

[159] Minzhe Shen, Ying Zhou, Jiawei Ye, Abdu Ahmed Abdullah AL-maskri, Yu Kang, Su Zeng, and Sheng Cai. Recent advances and perspectives of nucleic acid detection for coronavirus. *Journal of Pharmaceutical Analysis*, 10(2):97–101, 2020. ISSN 2095-1779. doi:https://doi.org/10.1016/j.jpha.2020.02.010. URL https://www.sciencedirect.com/science/article/pii/S2095177920302082.

[160] 2021. https://international.neb.com/applications/dna-amplification-pcr-and-qpcr/isothermal-amplification/loop-mediated-isothermal-amplification-lamp.

[161] R. Lenarčič, D. Morisset, N. Mehle, and M. Ravnikar. Fast real-time detection of potato spindle tuber viroid by RT-LAMP. *Plant Pathology*, 62(5):1147–1156, 2013. doi:https://doi.org/10.1111/ppa.12017. URL https://bsppjournals.onlinelibrary.wiley.com/doi/abs/10.1111/ppa.12017.

[162] Gun-Soo Park, Keunbon Ku, Seung-Hwa Baek, Seong-Jun Kim, Seung Il Kim, Bum-Tae Kim, and Jin-Soo Maeng. Development of reverse transcription loop-mediated isothermal amplification (RT-LAMP) assays targeting SARS-CoV-2. *The Journal of Molecular Diagnostics*, 2020.

[163] Chitrangada Das Mukhopadhyay, Pramita Sharma, Koel Sinha, and Keshav Rajarshi. Recent trends in analytical and digital techniques for the detection of the SARS-CoV-2. *Biophysical Chemistry*, 270:106538, 2021. ISSN 0301-4622. doi:https://doi.org/10.1016/j.bpc.2020.106538. URL https://www.sciencedirect.com/science/article/pii/S0301462220302465.

[164] Laura E. Lamb, Sarah N. Bartolone, Elijah Ward, and Michael B. Chancellor. Rapid detection of novel coronavirus/Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by reverse transcription-loop-mediated isothermal amplification. *PLOS ONE*, 15:1–15, 06 2020. doi:10.1371/journal.pone.0234682. URL https://doi.org/10.1371/journal.pone.0234682.

[165] Ben Jia, Xueling Li, Wei Liu, Changde Lu, Xiaoting Lu, Liangxiao Ma, Yuan-Yuan Li, and Chaochun Wei. GLAPD: Whole genome based LAMP primer design for a set of target genomes. *Frontiers in Microbiology*, 10:2860, 2019. ISSN 1664-302X. doi:10.3389/fmicb.2019.02860. URL https://www.frontiersin.org/article/10.3389/fmicb.2019.02860.

[166] Michael D Buck, Enzo Z Poirier, Ana Cardoso, Bruno Frederico, Johnathan Canton, Sam Barrell, Rupert Beale, Richard Byrne, Simon Caidan, Margaret Crawford, et al. SARS-CoV-2 detection by a clinical diagnostic RT-LAMP assay. *Wellcome Open Research*, 6, 2021. doi:10.12688/wellcomeopenres.16517.2.

[167] Viet Loan Dao Thi, Konrad Herbst, Kathleen Boerner, Matthias Meurer, Lukas PM Kremer, Daniel Kirrmaier, Andrew Freistaedter, Dimitrios Papagiannidis, Carla Galmozzi, Megan L. Stanifer, Steeve Boulant, Steffen Klein, Petr Chlanda, Dina Khalid, Isabel Barreto Miranda, Paul Schnitzler, Hans-Georg Kräusslich, Michael Knop, and Simon Anders. A colorimetric RT-LAMP assay and LAMP-sequencing for detecting SARS-CoV-2 RNA in clinical samples. *Science Translational Medicine*, 12(556), 2020. ISSN 1946-6234. doi:10.1126/scitranslmed.abc7075. URL https://stm.sciencemag.org/content/12/556/eabc7075.

[168] Alisa Alekseenko, Donal Barrett, Yerma Pareja-Sanchez, Rebecca J Howard, Emilia Strandback, Henry Ampah-Korsah, Urška Rovšnik, Silvia Zuniga-Veliz, Alexander Klenov, Jayshna Malloo, et al. Direct detection of SARS-CoV-2 using non-commercial RT-LAMP reagents on heat-inactivated samples. *Scientific reports*, 11(1):1–10, 2021. doi:https://doi.org/10.1038/s41598-020-80352-8.

[169] Matthew A Lalli, Joshua S Langmade, Xuhua Chen, Catrina C Fronick, Christopher S Sawyer, Lauren C Burcea, Michael N Wilkinson, Robert S Fulton, Michael Heinz, William J Buchser, Richard D Head, Robi D Mitra, and Jeffrey Milbrandt. Rapid and extraction-free detection of SARS-CoV-2 from saliva by colorimetric reverse-transcription loop-mediated isothermal amplification. *Clinical Chemistry*, 67(2):415–424, 10 2020. ISSN 0009-9147. doi:10.1093/clinchem/hvaa267. URL https://doi.org/10.1093/clinchem/hvaa267.

[170] Bàrbara Reynés, Francisca Serra, and Andreu Palou. Rapid visual detection of SARS-CoV-2 by colorimetric loop-mediated isothermal amplification. *BioTechniques*, 70(4): 218–225, 2021. doi:10.2144/btn-2020-0159. URL https://doi.org/10.2144/btn-2020-0159. PMID: 33820475.

[171] Patrick Hardinge, Guy Kiddle, Laurence Tisi, and James AH Murray. Optimised LAMP allows single copy detection of 35Sp and NOSt in transgenic maize using bioluminescent assay in real time (BART). *Scientific reports*, 8(1):1–17, 2018.

[172] Olena Mayboroda, Ioanis Katakis, and Ciara K. O'Sullivan. Multiplexed isothermal nucleic acid amplification. *Analytical Biochemistry*, 545:20–30, 2018. ISSN 0003-2697. doi:https://doi.org/10.1016/j.ab.2018.01.005. URL https://www.sciencedirect.com/science/article/pii/S000326971830006X.

[173] Woong Sik Jang, Da Hye Lim, Jung Yoon, Ahran Kim, Minsup Lim, Jeonghun Nam, Richard Yanagihara, Sook-Won Ryu, Bo Kyeung Jung, Nam-Hee Ryoo, and Chae Seung Lim. Development of a multiplex loop-mediated isothermal amplification (LAMP) assay for on-site diagnosis of SARS CoV-2. *PLOS ONE*, 16(3):1–14, 03 2021. doi:10.1371/journal.pone.0248042. URL https://doi.org/10.1371/journal.pone.0248042.

[174] Marianna Soroka, Barbara Wasowicz, and Anna Rymaszewska. Loop-mediated isothermal amplification (LAMP): the better sibling of PCR? *Cells*, 10(8):1931, 2021.

[175] Yasuyoshi Mori, Hidetoshi Kanda, and Tsugunori Notomi. Loop-mediated isothermal amplification (LAMP): recent progress in research and development. *Journal of Infection and Chemotherapy*, 19(3):404–411, 2013. ISSN 1341-321X. doi:https://doi.org/10.1007/s10156-013-0590-0. URL https://www.sciencedirect.com/science/article/pii/S1341321X13701223.

[176] Meng Yee Lai, Choo Huck Ooi, and Yee Ling Lau. Validation of SYBR green I based closed-tube loop-mediated isothermal amplification (LAMP) assay for diagnosis of knowlesi malaria. *Malaria journal*, 20(1):1–6, 2021. doi:https://doi.org/10.1186/s12936-021-03707-0.

[177] Amanda E. Calvert, Brad J. Biggerstaff, Nathan A. Tanner, Molly Lauterbach, and Robert S. Lanciotti. Rapid colorimetric detection of Zika virus from serum and urine specimens by reverse transcription loop-mediated isothermal amplification (RT-LAMP). *PLOS ONE*, 12(9):1–16, 09 2017. doi:10.1371/journal.pone.0185340. URL https://doi.org/10.1371/journal.pone.0185340.

[178] Kayoko Ohtsuka, Keiko Yanagawa, Kosuke Takatori, and Yukiko Hara-Kudo. Detection of salmonella enterica in naturally contaminated liquid eggs by loop-mediated isothermal amplification, and characterization of salmonella isolates. *Applied and environmental microbiology*, 71(11):6730–6735, 2005. doi:10.1128/AEM.71.11.6730-6735.2005.

[179] Yee-Ling Lau, Meng-Yee Lai, Boon-Teong Teoh, Juraina Abd-Jamil, Jefree Johari, Sing-Sin Sam, Kim-Kee Tan, and Sazaly AbuBakar. Colorimetric detection of dengue by single tube reverse-transcription-loop-mediated isothermal amplification. *PLOS ONE*, 10(9):e0138694, sep 2015. doi:10.1371/journal.pone.0138694. URL https://doi.org/10.1371/journal.pone.0138694.

[180] Olamide K. Oloniniyi, Yohei Kurosaki, Hiroko Miyamoto, Ayato Takada, and Jiro Yasuda. Rapid detection of all known ebolavirus species by reverse transcription-loop-mediated isothermal amplification (RT-LAMP). *Journal of Virological Methods*, 246:8–14, 2017. ISSN 0166-0934. doi:https://doi.org/10.1016/j.jviromet.2017.03.011. URL https://www.sciencedirect.com/science/article/pii/S0166093416306231.

[181] Kelly A. Curtis, Donna L. Rudolph, and S. Michele Owen. Rapid detection of hiv-1 by reverse-transcription, loop-mediated isothermal amplification (RT-LAMP). *Journal of Virological Methods*, 151(2):264–270, 2008. ISSN 0166-0934. doi:https://doi.org/10.1016/j.jviromet.2008.04.011. URL https://www.sciencedirect.com/science/article/pii/S016609340800133X.

[182] Pei Huang, Hualei Wang, Zengguo Cao, Hongli Jin, Hang Chi, Jincun Zhao, Beibei Yu, Feihu Yan, Xingxing Hu, Fangfang Wu, Cuicui Jiao, Pengfei Hou, Shengnan Xu, Yongkun Zhao, Na Feng, Jianzhong Wang, Weiyang Sun, Tiecheng Wang, Yuwei Gao, Songtao Yang, and Xianzhu Xia. A rapid and specific assay for the detection of MERS-CoV. *Frontiers in Microbiology*, 9:1101, 2018. ISSN 1664-302X. doi:10.3389/fmicb.2018.01101. URL https://www.frontiersin.org/article/10.3389/fmicb.2018.01101.

[183] Wei E. Huang, Boon Lim, Chia-Chen Hsu, Dan Xiong, Wei Wu, Yejiong Yu, Huidong Jia, Yun Wang, Yida Zeng, Mengmeng Ji, Hong Chang, Xiuming Zhang, Hui Wang, and Zhanfeng Cui. RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2. *Microbial Biotechnology*, 13(4):950–961, 2020. doi:https://doi.org/10.1111/1751-7915.13586. URL https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1751-7915.13586.

[184] James P Broughton, Xianding Deng, Guixia Yu, Clare L Fasching, Jasmeet Singh, Jessica Streithorst, Andrea Granados, Alicia Sotomayor-Gonzalez, Kelsey Zorn, Allan Gopez, et al. Rapid detection of 2019 novel coronavirus SARS-CoV-2 using a CRISPR-based DETECTR lateral flow assay. *Nat. Biotechnol.*, 2020. doi:10.1038/s41587-020-0513-4.

[185] Alfredo Garcia-Venzor, Bertha Rueda-Zarazua, Eduardo Marquez-Garcia, Vilma Maldonado, Angelica Moncada-Morales, Hiram Olivera, Irma Lopez, Joaquin Zuñiga, and Jorge Melendez-Zajgla. SARS-CoV-2 direct detection without RNA isolation with loop-mediated isothermal amplification (LAMP) and CRISPR-Cas12. *Frontiers in Medicine*, 8:125, 2021. ISSN 2296-858X. doi:10.3389/fmed.2021.627679. URL https://www.frontiersin.org/article/10.3389/fmed.2021.627679.

[186] Yaqin Zhang, Minyan Chen, Chengrong Liu, Jiaqi Chen, Xinyi Luo, Yingying Xue, Qiming Liang, Li Zhou, Yu Tao, Mingqiang Li, Di Wang, Jianhua Zhou,

and Jiasi Wang. Sensitive and rapid on-site detection of SARS-CoV-2 using a gold nanoparticle-based high-throughput platform coupled with CRISPR/Cas12-assisted RT-LAMP. *Sensors and Actuators B: Chemical*, 345:130411, 2021. ISSN 0925-4005. doi:https://doi.org/10.1016/j.snb.2021.130411. URL https://www.sciencedirect.com/science/article/pii/S0925400521009795.

[187] Akansha Bhatt, Zeeshan Fatima, Munindra Ruwali, Chitra Seetharam Misra, Shyam Sunder Rangu, Devashish Rath, Ashok Rattan, and Saif Hameed. CLEVER assay: A visual and rapid RNA extraction-free detection of SARS-CoV-2 based on CRISPR--Cas integrated RT-LAMP technology. *Journal of Applied Microbiology*, apr 2022. doi:10.1111/jam.15571. URL https://doi.org/10.1111%2Fjam.15571.

[188] Tong Zhang, Wang Zhao, Xi Chen, Xinlian Zhang, Jinhui Zhu, Shenwei Li, Chuanyong Wu, Zhengan Tian, and Guodong Sui. Fully automated CRISPR-LAMP platform for SARS-CoV-2 delta and omicron variants. *Analytical Chemistry*, 94(44):15472–15480, oct 2022. doi:10.1021/acs.analchem.2c03607. URL https://doi.org/10.1021%2Facs.analchem.2c03607.

[189] Kasturi Selvam, Mohamad Ahmad Najib, Muhammad Fazli Khalid, Suharni Mohamad, Fahreddin Palaz, Mehmet Ozsoz, and Ismail Aziah. RT-LAMP CRISPR-Cas12/13-based SARS-CoV-2 detection methods. *Diagnostics*, 11(9):1646, 2021.

[190] Anetta Ptasinska, Celina Whalley, Andrew Bosworth, Charlotte Poxon, Claire Bryer, Nicholas Machin, Seden Grippon, Emma L. Wise, Bryony Armson, Emma L.A. Howson, Alice Goring, Gemma Snell, Jade Forster, Chris Mattocks, Sarah Frampton, Rebecca Anderson, David Cleary, Joe Parker, Konstantinos Boukas, Nichola Graham, Doriana Cellura, Emma Garratt, Rachel Skilton, Hana Sheldon, Alla Collins, Nusreen Ahmad, Simon Friar, Daniel Burns, Tim Williams, Keith M. Godfrey, Zandra Deans, Angela Douglas, Sue Hill, Michael Kidd, Deborah Porter, Stephen P. Kidd, Nicholas J. Cortes, Veronica Fowler, Tony Williams, Alex Richter, and Andrew D. Beggs. Diagnostic accuracy of loop-mediated isothermal amplification coupled to nanopore sequencing (LamPORE) for the detection of SARS-CoV-2 infection at scale in symptomatic and asymptomatic populations. *Clinical Microbiology and Infection*, 27(9):1348.e1–1348.e7, sep 2021. doi:10.1016/j.cmi.2021.04.008. URL https://doi.org/10.1016/j.cmi.2021.04.008.

[191] S.B. Needleman and C.D. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.*, 48(3):443–453, 1970.

[192] Temple F Smith, Michael S Waterman, et al. Identification of common molecular subsequences. *J. Mol. Biol.*, 147(1):195–197, 1981.

[193] Lars Kaderali and Alexander Schliep. Selecting signature oligonucleotides to identify organisms using DNA arrays. *Bioinformatics*, 18(10):1340–1349, 10 2002. ISSN 1367-4803. doi:10.1093/bioinformatics/18.10.1340. URL https://doi.org/10.1093/bioinformatics/18.10.1340.

[194] John SantaLucia, Jr., H T Allawi, and P A Seneviratne. Improved nearest-neighbour parameters for predicting DNA duplex stability. *Biochem.*, 35:3555–3562, 1996.

[195] M. Peyrard and A. R. Bishop. Statistical mechanics of a nonlinear model for DNA denaturation. *Phys. Rev. Lett.*, 62(23):2755–2757, 1989. doi:10.1103/PhysRevLett.62.2755.

[196] T. Dauxois, M. Peyrard, and A. R. Bishop. Entropy-driven DNA denaturation. *Phys. Rev. E*, 47(1):R44–R47, 1993. doi:10.1103/PhysRevE.47.R44.

[197] 2021. http://jae1001.user.srcf.net/ccp5_2007/.

[198] Ludmila V Yakushevich. *Nonlinear physics of DNA*. John Wiley & Sons, 2004.

[199] Philip M Morse. Diatomic molecules according to the wave mechanics. II. Vibrational levels. *Physical Review*, 34(1):57, 1929.

[200] M. Peyrard. Nonlinear dynamics and statistical physics of DNA. *Nonlinearity*, 17:R1, 2004.

[201] Gerald Weber. TfReg: Calculating DNA and RNA melting temperatures and opening profiles with mesoscopic models. *Bioinformatics*, 29:1345–1347, 2013. doi:10.1093/bioinformatics/btt133. URL http://bioinformatics.oxfordjournals.org/content/29/10/1345.

[202] Yong-Li Zhang, Wei-Mou Zheng, Ji-Xing Liu, and Y. Z. Chen. Theory of DNA melting based on the Peyrard-Bishop model. *Phys. Rev. E*, 56(6):7100–7115, 1997.

[203] Gerald Weber, Niall Haslam, Nava Whiteford, Adam Prügel-Bennett, Jonathan W. Essex, and Cameron Neylon. Thermal equivalence of DNA duplexes without melting temperature calculation. *Nat. Phys.*, 2:55–59, 2006. doi:10.1038/nphys189.

[204] Gerald Weber, Niall Haslam, Jonathan W. Essex, and Cameron Neylon. Thermal equivalence of DNA duplexes for probe design. *J. Phys.: Condens. Matter*, 21:034106, 2009. doi:10.1088/0953-8984/21/3/034106.

[205] Frances H Arnold, Steven Wolk, Phillip Cruz, and Ignacio Tinoco Jr. Structure, dynamics, and thermodynamics of mismatched DNA oligonucleotide duplexes d(CCCAGGG)$_2$ and d(CCCTGGG)$_2$. *Biochem.*, 26(13):4068–4075, 1987.

[206] Tom Brown, Gordon A Leonard, Ewan D Booth, and Jenny Chambers. Crystal structure and stability of a DNA duplex containing A(anti)·G(syn) base-pairs. *J. Mol. Biol.*, 207 (2):455–457, 1989.

[207] Yi-Gui Gao, Howard Robinson, Ruslan Sanishvili, Andrzej Joachimiak, and Andrew H-J Wang. Structure and recognition of sheared tandem G·A base pairs associated with human centromere DNA sequence at atomic resolution. *Biochem.*, 38(50):16452–16460, 1999.

[208] Frank Seela, Simone Budow, Khalil I Shaikh, and Anup M Jawalekar. Stabilization of tandem dG–dA base pairs in DNA-hairpins: replacement of the canonical bases by 7-deaza-7-propynylpurines. *Org. Biomol. Chem.*, 3(23):4221–4226, 2005.

[209] Ying Li, Gerald Zon, and W David Wilson. NMR and molecular modeling evidence for a GA mismatch base pair in a purine-rich DNA duplex. *Proc. Natl. Acad. Sci. USA*, 88 (1):26–30, 1991.

[210] 2020. https://www.cdc.gov/coronavirus/2019-ncov/lab/{RT-PCR}-panel-primer-probes.html.

[211] 2020. https://www.who.int/docs/default-source/coronaviruse/whoinhouseassays.pdf?sfvrsn=de3a76aa_2.

[212] 2020. https://www.captodayonline.com/LuminexSARSCoV219.pdf.

[213] Victor M Corman, Olfert Landt, Marco Kaiser, Richard Molenkamp, Adam Meijer, Daniel KW Chu, Tobias Bleicker, Sebastian Brünink, Julia Schneider, Marie Luisa Schmidt, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25(3):2000045, 2020.

[214] Jayeshkumar Narsibhai Davda, Keith Frank, Sivakumar Prakash, Gunjan Purohit, Devi Prasad Vijayashankar, Dhiviya Vedagiri, Karthik Bharadwaj Tallapaka, Krishnan Harinivas Harshan, Archana Bharadwaj Siva, Rakesh Kumar Mishra, Jyotsna Dhawan, and Imran Siddiqi. An inexpensive RT-PCR endpoint diagnostic assay for SARS-CoV-2 using nested PCR: Direct assessment of detection efficiency of RT-qPCR tests and suitability for surveillance. *bioRxiv*, 2020. doi:10.1101/2020.06.08.139477. URL https://www.biorxiv.org/content/early/2020/06/09/2020.06.08.139477.

[215] Paul R Grant, Melanie A Turner, Gee Yen Shin, Eleni Nastouli, and Lisa J Levett. Extraction-free COVID-19 (SARS-CoV-2) diagnosis by RT-PCR to increase capacity for national testing programmes during a pandemic. *bioRxiv*, 2020. doi:10.1101/2020.04.06.028316. URL https://www.biorxiv.org/content/early/2020/04/08/2020.04.06.028316.

[216] Yosuke Hirotsu, Hitoshi Mochizuki, and Masao Omata. Double-quencher probes improved the detection sensitivity of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by one-step RT-PCR. *medRxiv*, 2020. doi:10.1101/2020.03.17.20037903. URL https://www.medrxiv.org/content/early/2020/03/20/2020.03.17.20037903.

[217] Saakshi Jalali, Vikas K Patel, Venkatesh Prasad, Ajit Sapre, Santanu Dasgupta, and Bhaskar Bhadra. Detection of SARS-CoV-2 using high-throughput PCR. *J. Appl. Biol. Biotechnol*, 8(3):1–5, 2020.

[218] Daniel CF Lanza, João PMS Lima, and Selma MB Jerônimo. Design and in silico validation of polymerase chain reaction primers to detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). 2020. doi:10.21203/rs.3.rs-26306/v1.

[219] Roujian Lu, Xiang Zhao, Juan Li, Peihua Niu, Bo Yang, Honglong Wu, Wenling Wang, Hao Song, Baoying Huang, Na Zhu, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*, 395(10224):565–574, 2020.

[220] Bas B Oude Munnink, David F Nieuwenhuijse, Mart Stein, Áine O'Toole, Manon Haverkate, Madelief Mollers, Sandra K Kamga, Claudia Schapendonk, Mark Pronk, Pascal Lexmond, et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. *Nature medicine*, 26(9): 1405–1410, 2020.

[221] Arun K Nalla, Amanda M Casto, Meei-Li W Huang, Garrett A Perchetti, Reigran Sampoleo, Lasata Shrestha, Yulun Wei, Haiying Zhu, Keith R Jerome, and Alexander L Greninger. Comparative performance of SARS-CoV-2 detection assays using seven different primer-probe sets and one assay kit. *Journal of Clinical Microbiology*, 58(6): e00557–20, 2020.

[222] H Rahman, I Carter, K Basile, L Donovan, S Kumar, T Tran, D Ko, S Alderson, T Sivaruban, J-S Eden, et al. Interpret with caution: An evaluation of the commercial AusDiagnostics versus in-house developed assays for the detection of SARS-CoV-2 virus. *Journal of Clinical Virology*, page 104374, 2020.

[223] Tuna Toptan, Sebastian Hoehl, Sandra Westhaus, Denisa Bojkova, Annemarie Berger, Björn Rotter, Klaus Hoffmeier, Sandra Ciesek, and Marek Widera. Optimized qRT-PCR approach for the detection of intra- and extra-cellular SARS-CoV-2 RNAs. *bioRxiv*, 2020. doi:10.1101/2020.04.20.052258. URL https://www.biorxiv.org/content/early/2020/04/25/2020.04.20.052258.

[224] Chantal BF Vogels, Anderson F Brito, Anne L Wyllie, Joseph R Fauver, Isabel M Ott, Chaney C Kalinich, Mary E Petrone, Arnau Casanovas-Massana, M Catherine Muenker, Adam J Moore, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. *Nature Microbiology*, 5(10):1299–1305, 2020.

[225] Cyril Chik-Yan Yip, Chi-Chun Ho, Jasper Fuk-Woo Chan, Kelvin Kai-Wang To, Helen Shuk-Ying Chan, Sally Cheuk-Ying Wong, Kit-Hang Leung, Agnes Yim-Fong Fung, Anthony Chin-Ki Ng, Zijiao Zou, et al. Development of a novel, genome subtraction-derived, SARS-CoV-2-specific COVID-19-nsp2 real-time RT-PCR assay and its evaluation using clinical specimens. *International Journal of Molecular Sciences*, 21(7):2574, 2020.

[226] Juan GB Diego, Pedro Fernández-Soto, Marta Domínguez-Gil, Moncef Belhassen-García, Juan Luis Muñoz Bellido, and Antonio Muro. A simple, affordable, rapid, stabilized, colorimetric, versatile RT-LAMP assay to detect SARS-CoV-2. *Diagnostics*, 11(3), 2021. ISSN 2075-4418. doi:10.3390/diagnostics11030438. URL https://www.mdpi.com/2075-4418/11/3/438.

[227] Anurup Ganguli, Ariana Mostafa, Jacob Berger, Mehmet Y. Aydin, Fu Sun, Sarah A. Stewart de Ramirez, Enrique Valera, Brian T. Cunningham, William P. King, and Rashid Bashir. Rapid isothermal amplification and portable detection system for SARS-CoV-2. *Proceedings of the National Academy of Sciences*, 117(37):22727–22735, 2020. ISSN 0027-8424. doi:10.1073/pnas.2014739117. URL https://www.pnas.org/content/117/37/22727.

[228] Chao Ji, Shuxia Xue, Min Yu, Jinyu Liu, Qin Zhang, Feng Zuo, Qiuyue Zheng, Liangjuan Zhao, Hongwei Zhang, Jijuan Cao, Ke Wang, Wei Liu, and Wenjie Zheng. Rapid detection of SARS-CoV-2 virus using dual reverse transcriptional colorimetric loop-mediated isothermal amplification. *ACS Omega*, 6(13):8837–8849, 2021. doi:10.1021/acsomega.0c05781. URL https://doi.org/10.1021/acsomega.0c05781.

[229] Yee Ling Lau, Ilyiana Binti Ismail, Nur Izati Binti Mustapa, Meng Yee Lai, Tuan Suhaila Tuan Soh, Afifah Haji Hassan, Kalaiarasu M. Peariasamy, Yee Leng Lee, and Pik Pin Goh. A sensitive reverse transcription loop-mediated isothermal amplification assay for direct visual detection of SARS-CoV-2. *The American Journal of Tropical Medicine and Hygiene*, 103(6):2350 – 2352, 02 Dec. 2020. doi:10.4269/ajtmh.20-1079. URL https://www.ajtmh.org/view/journals/tpmd/103/6/article-p2350.xml.

[230] Zhen Luo, Chunhong Ye, Heng Xiao, Jialing Yin, Yicong Liang, Zhihui Ruan, DanjLuo, Daolong Gao, Qiuping Tan, Yongkui Li, Qiwei Zhang, Weiy-

ong Liu, and Jianguo Wu. Optimization of loop-mediated isothermal amplification (LAMP) assay for robust visualization in SARS-CoV-2 and emerging variants diagnosis. *Chemical Engineering Science*, 251:117430, 2022. ISSN 0009-2509. doi:https://doi.org/10.1016/j.ces.2022.117430. URL https://www.sciencedirect.com/science/article/pii/S0009250922000148.

[231] Lena Mautner, Christin-Kirsty Baillie, Heike Marie Herold, Wolfram Volkwein, Patrick Guertler, Ute Eberle, Nikolaus Ackermann, Andreas Sing, Melanie Pavlovic, Ottmar Goerlich, et al. Rapid point-of-care detection of SARS-CoV-2 using reverse transcription loop-mediated isothermal amplification (RT-LAMP). *Virology journal*, 17(1):1–14, 2020. doi:https://doi.org/10.1186/s12985-020-01435-6.

[232] Abu Naser Mohon, Lisa Oberding, Jana Hundt, Guido van Marle, Kanti Pabbaraju, Byron M. Berenger, Luiz Lisboa, Thomas Griener, Markus Czub, Cody Doolan, Venice Servellita, Charles Y. Chiu, Alexander L. Greninger, Keith R. Jerome, and Dylan R. Pillai. Optimization and clinical validation of dual-target RT-LAMP for SARS-CoV-2. *Journal of Virological Methods*, 286:113972, 2020. ISSN 0166-0934. doi:https://doi.org/10.1016/j.jviromet.2020.113972. URL https://www.sciencedirect.com/science/article/pii/S016609342030224X.

[233] Jesus Rodriguez-Manzano, Kenny Malpartida-Cardenas, Nicolas Moser, Ivana Pennisi, Matthew Cavuto, Luca Miglietta, Ahmad Moniri, Rebecca Penn, Giovanni Satta, Paul Randell, Frances Davies, Frances Bolt, Wendy Barclay, Alison Holmes, and Pantelis Georgiou. Handheld point-of-care system for rapid detection of SARS-CoV-2 extracted RNA in under 20 min. *ACS Central Science*, 7(2):307–317, 2021. doi:10.1021/acscentsci.0c01288. URL https://doi.org/10.1021/acscentsci.0c01288.

[234] C. Yan, J. Cui, L. Huang, B. Du, L. Chen, G. Xue, S. Li, W. Zhang, L. Zhao, Y. Sun, H. Yao, N. Li, H. Zhao, Y. Feng, S. Liu, Q. Zhang, D. Liu, and J. Yuan. Rapid and visual detection of 2019 novel coronavirus (SARS-CoV-2) by a reverse transcription loop-mediated isothermal amplification assay. *Clinical Microbiology and Infection*, 26(6):773–779, 2020. ISSN 1198-743X. doi:https://doi.org/10.1016/j.cmi.2020.04.001. URL https://www.sciencedirect.com/science/article/pii/S1198743X20301865.

[235] Qing Yang, Nicholas R Meyerson, Stephen K Clark, Camille L Paige, Will T Fattor, Alison R Gilchrist, Arturo Barbachano-Guerrero, Benjamin G Healy, Emma R Worden-Sapper, Sharon S Wu, Denise Muhlrad, Carolyn J Decker, Tassa K Saldi, Erika Lasda, Patrick Gonzales, Morgan R Fink, Kimngan L Tat, Cole R Hager, Jack C Davis, Christopher D Ozeroff, Gloria R Brisson, Matthew B McQueen, Leslie A Lein-

wand, Roy Parker, and Sara L Sawyer. Saliva TwoStep for rapid detection of asymptomatic SARS-CoV-2 carriers. *eLife*, 10, mar 2021. doi:10.7554/elife.65113. URL https://doi.org/10.7554%2Felife.65113.

[236] Rokusuke Yoshikawa, Haruka Abe, Yui Igasaki, Saeki Negishi, Hiroaki Goto, and Jiro Yasuda. Development and evaluation of a rapid and simple diagnostic assay for COVID-19 based on loop-mediated isothermal amplification. *PLOS Neglected Tropical Diseases*, 14(11):e0008855, nov 2020. doi:10.1371/journal.pntd.0008855. URL https://doi.org/10.1371%2Fjournal.pntd.0008855.

[237] 2020. https://www.ncbi.nlm.nih.gov/{SARS-CoV-2}/.

[238] 2020. https://www.ncbi.nlm.nih.gov/.

[239] 2021. https://www.gisaid.org/hcov19-variants/.

[240] 2021. https://bioinf.fisica.ufmg.br/software/.

[241] Antonio Martínez-Murcia, Adrián García-Sirera, Aaron Navarro, and Laura Pérez. Current RT-qPCR to detect SARS-CoV-2 may give positive results for related coronaviruses. *Archives of Microbiology*, 204(7), jun 2022. doi:10.1007/s00203-022-03029-y. URL https://doi.org/10.1007%2Fs00203-022-03029-y.

[242] Trestan Pillonel, Valentin Scherz, Katia Jaton, Gilbert Greub, and Claire Bertelli. Letter to the editor: SARS-CoV-2 detection by real-time RT-PCR. *Eurosurveillance*, 25(21): 2000880, 2020.

[243] Victor M Corman and Christian Drosten. Authors' response: SARS-CoV-2 detection by real-time RT-PCR. *Eurosurveillance*, 25(21):2001035, 2020.

[244] Yile Tao, Yang Yue, Guangyu Qiu, Zheng Ji, Martin Spillman, Zhibo Gai, Qingfa Chen, Michel Bielecki, Michael Huber, Alexandra Trkola, Qiyuan Wang, Junji Cao, and Jing Wang. Comparison of analytical sensitivity and efficiency for SARS-CoV-2 primer sets by TaqMan-based and SYBR green-based RT-qPCR. *Applied Microbiology and Biotechnology*, feb 2022. doi:10.1007/s00253-022-11822-4. URL https://doi.org/10.1007%2Fs00253-022-11822-4.

[245] Kazuya Shirato, Shutoku Matsuyama, and Makoto Takeda. Less frequent sequence mismatches in variants of concern (VOCs) of SARS-CoV-2 in the real-time RT-PCR assays developed by the National Institute of Infectious Diseases, Japan. *Jpn. J. Infect. Dis*, 75: 96–101, 2022.

[246] N Peyret, P A Seneviratne, H T Allawi, and John SantaLucia, Jr. Nearest-neighbour thermodynamics and NMR of DNA sequences with internal A·A, C·C G·G and T·T mismatches. *Biochem.*, 38(12):3468–3477, 1999. doi:10.1021/bi9825091.

[247] Anna Tikhomirova, Irina V Beletskaya, and Tigran V Chalikian. Stability of DNA duplexes containing GG, CC, AA, and TT mismatches. *Biochem.*, 45(35):10563–10571, 2006.

[248] Weihua Yang, Xiaofei Dang, Qingxi Wang, Mingjie Xu, Qianqian Zhao, Yunying Zhou, Huailong Zhao, Li Wang, Yihui Xu, Jun Wang, Shuyi Han, Min Wang, Fenyan Pei, and Yunshan Wan. Rapid detection of SARS-CoV-2 using reverse transcription RT-LAMP method. *medRxiv*, 2020. doi:10.1101/2020.03.02.20030130. URL https://www.medrxiv.org/content/early/2020/03/03/2020.03.02.20030130.

[249] 2021. https://primerexplorer.jp/e/v5_manual/pdf/PrimerExplorerV5_Manual_1.pdf.

[250] Letícia Trindade Almeida, Amanda Bonoto Gonçalves, Ana Paula Moreira Franco-Luiz, Thais Bárbara de Souza Silva, Pedro Augusto Alves, and Rubens Lima do Monte-Neto. Molecular detection of omicron SARS-CoV-2 variant is achieved by RT-LAMP despite genomic mutations. *Memórias do Instituto Oswaldo Cruz*, 117:e220050, 2022. doi:https://doi.org/10.1590/0074-02760220050.

[251] César MJA Metzger, Reto Lienhard, Helena MB Seth-Smith, Tim Roloff, Fanny Wegner, Jonas Sieber, Michael Bel, Gilbert Greub, and Adrian Egli. PCR performance in the SARS-CoV-2 omicron variant of concern? *Swiss Med Wkly*, 151(49):w30120, 2021. doi:10.4414/SMW.2021.w30120.

[252] Maria Artesi, Sébastien Bontems, Paul Göbbels, Marc Franckh, Piet Maes, Raphaël Boreux, Cécile Meex, Pierrette Melin, Marie-Pierre Hayette, Vincent Bours, et al. A recurrent mutation at position 26340 of SARS-CoV-2 is associated with failure of the E gene quantitative reverse transcription-PCR utilized in a commercial dual-target diagnostic assay. *Journal of clinical microbiology*, 58(10):10–1128, 2020.

[253] Katharina Ziegler, Philipp Steininger, Renate Ziegler, Jörg Steinmann, Klaus Korn, and Armin Ensser. SARS-CoV-2 samples may escape detection because of a single point mutation in the N gene. *Eurosurveillance*, 25(39):2001650, 2020.

[254] Mohammad Rubayet Hasan, Sathyavathi Sundararaju, Chidambaram Manickam, Faheem Mirza, Hamad Al-Hail, Stephan Lorenz, and Patrick Tang. A novel point mutation in the N gene of SARS-CoV-2 may affect the detection of the virus by reverse transcription-quantitative PCR. *Journal of Clinical Microbiology*, 59(4):10–1128, 2021.

[255] Mohammad Alkhatib, Maria Concetta Bellocchi, Greta Marchegiani, Sandro Grelli, Valeria Micheli, Daniele Stella, Bartolomeo Zerillo, Luca Carioti, Valentina Svicher, Paola

Rogliani, et al. First case of a COVID-19 patient infected by delta AY. 4 with a rare deletion leading to a N gene target failure by a specific real time PCR assay: novel Omicron VOC might be doing similar scenario? *Microorganisms*, 10(2):268, 2022.

[256] Xiaohua Huang, Ivan H El-Sayed, Wei Qian, and Mostafa A El-Sayed. Cancer cell imaging and photothermal therapy in the near-infrared region by using gold nanorods. *Journal of the American Chemical Society*, 128(6):2115–2120, 2006.

[257] Lin Wang, Yan Jin, Jing Deng, and Guozhen Chen. Gold nanorods-based FRET assay for sensitive detection of Pb 2+ using 8-17DNAzyme. *Analyst*, 136(24):5169–5174, 2011.

[258] PK Sudeep, ST Shibu Joseph, and K George Thomas. Selective detection of cysteine and glutathione using gold nanorods. *Journal of the American Chemical Society*, 127(18): 6516–6517, 2005.

[259] Jian Wang, Pu Zhang, Jing Yun Li, Li Qiang Chen, Cheng Zhi Huang, and Yuan Fang Li. Adenosine–aptamer recognition-induced assembly of gold nanorods and a highly sensitive plasmon resonance coupling assay of adenosine in the brain of model sd rat. *Analyst*, 135(11):2826–2831, 2010.

[260] Erik Dujardin, Long-Bao Hsin, CR Chris Wang, and Stephen Mann. DNA-driven self-assembly of gold nanorods. *Chemical Communications*, (14):1264–1265, 2001.

[261] Jing Deng, Yan Jin, Lin Wang, Guozhen Chen, and Chengxiao Zhang. Sensitive detection of endonuclease activity and inhibition using gold nanorods. *Biosensors and Bioelectronics*, 34(1):144–150, 2012.

[262] Wenhong Wang, Yina Zhao, and Yan Jin. Gold-nanorod-based colorimetric and fluorescent approach for sensitive and specific assay of disease-related gene and mutation. *ACS Applied Materials & Interfaces*, 5(22):11741–11746, 2013.

[263] Panagiotis G Georgiou, Collette S Guy, Muhammad Hasan, Ashfaq Ahmad, Sarah-Jane Richards, Alexander N Baker, Neer V Thakkar, Marc Walker, Sarojini Pandey, Neil R Anderson, et al. Plasmonic detection of sars-cov-2 spike protein with polymer-stabilized glycosylated gold nanorods. *ACS Macro Letters*, 11(3):317–322, 2022.

[264] 2021. http://rna.tbi.univie.ac.at/.

[265] Willi Quino, Diana Flores-León, Junior Caro-Castro, Carmen V Hurtado, Iris Silva, and Ronnie G Gavilan. Evaluation of reverse transcription-loop-mediated isothermal amplification for rapid detection of SARS-CoV-2. *Scientific Reports*, 11(1):24234, 2021.

[266] Jer-Horng Wu, Pei-Ying Hong, and Wen-Tso Liu. Quantitative effects of position and type of single mismatch on single base primer extension. *Journal of microbiological methods*, 77(3):267–275, 2009.

[267] 2021. https://www.idt{DNA}.com/pages/tools/oligoanalyzer.

[268] 2021. http://www.unafold.org/twostate.php.

[269] 2021. http://www.graphreader.com/.

[270] Alexei A Koshkin, Sanjay K Singh, Poul Nielsen, Vivek K Rajwanshi, Ravindra Kumar, Michael Meldgaard, Carl Erik Olsen, and Jesper Wengel. LNA (locked nucleic acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron*, 54(14):3607–3630, 1998.

[271] Birte Vester and Jesper Wengel. LNA (locked nucleic acid) high-affinity targeting of complementary RNA and DNA. *Biochem.*, 43(42):13233–13241, 2004. doi:10.1021/bi0485732.

[272] Michael Petersen, Christina B Nielsen, Katrine E Nielsen, Gitte A Jensen, Kent Bondensgaard, Sanjay K Singh, Vivek K Rajwanshi, Alexei A Koshkin, Britta M Dahl, Jesper Wengel, et al. The conformations of locked nucleic acids (LNA). *Journal of Molecular Recognition*, 13(1):44–53, 2000.

[273] Yong You, Bernardo G Moreira, Mark A Behlke, and Richard Owczarzy. Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Res.*, 34(8):e60, 2006. doi:10.1093/nar/gkl175.

[274] Patrick J Hrdlicka and Michael E Østergaard. Fluorophore-functionalised Locked Nucleic Acids (LNAs). *DNA Conjugates and Sensors; Fox, KR, Brown, T., Eds*, pages 1–33, 2012.

[275] Prabodhika Mallikaratchy, Jeffery Gardner, Lars Ulrik R Nordstrøm, Nicholas J Veomett, Michael R McDevitt, Mark L Heaney, and David A Scheinberg. A self-assembling short oligonucleotide duplex suitable for pretargeting. *Nucleic Acid Ther.*, 23(4):289–299, 2013. doi:10.1089/nat.2013.0425.

[276] Luca Morandi, Dario De Biase, Michela Visani, Valentina Cesari, Giovanna De Maglio, Stefano Pizzolitto, Annalisa Pession, and Giovanni Tallini. Allele specific locked nucleic acid quantitative PCR (ASLNAqPCR): an accurate and cost-effective assay to diagnose and quantify KRAS and BRAF mutation. *PLoS One*, 7(4):e36084, 2012.

[277] Irina V Astakhova, Alexey V Ustinov, Vladimir A Korshun, and Jesper Wengel. LNA for optimization of fluorescent oligonucleotide probes: improved spectral properties and target binding. *Bioconjugate Chemistry*, 22(4):533–539, 2011.

[278] Padmavathy Bakthavathsalam, Guillaume Longatte, Slade O Jensen, Mike Manefield, and J Justin Gooding. Locked nucleic acid molecular beacon for multiplex detection of loop mediated isothermal amplification. *Sensors and Actuators B: Chemical*, 268: 255–263, 2018.

[279] Kenji Rowel Q. Lim, Rika Maruyama, Yusuke Echigoya, Quynh Nguyen, Aiping Zhang, Hunain Khawaja, Sreetama Sen Chandra, Takako Jones, Peter Jones, Yi-Wen Chen, and Toshifumi Yokota. Inhibition of DUX4 expression with antisense LNA gapmers as a therapy for facioscapulohumeral muscular dystrophy. *Proc. Natl. Acad. Sci. USA*, 117(28):16509–16515, 2020. ISSN 0027-8424. doi:10.1073/pnas.1909649117. URL https://www.pnas.org/content/117/28/16509.

[280] Kathrin S Schmidt, Sandra Borkowski, Jens Kurreck, Andrew W Stephens, Rolf Bald, Maren Hecht, Matthias Friebe, Ludger Dinkelborg, and Volker A Erdmann. Application of locked nucleic acids to improve aptamer in vivo stability and targeting function. *Nucleic Acids Res.*, 32(19):5757–5765, 2004.

[281] Lin Wang, Chaoyong James Yang, Colin D Medley, Steven A Benner, and Weihong Tan. Locked nucleic acid molecular beacons. *J. Am. Chem. Soc.*, 127(45):15664–15665, 2005.

[282] Izabela Ferreira, Sofie Slott, Kira Astakhova, and Gerald Weber. Complete mesoscopic parametrization of single LNA modifications in DNA applied to oncogene probe design. *Journal of Chemical Information and Modeling*, 61:3615–3624, 2021. doi:10.1021/acs.jcim.1c00470.

[283] R. Owczarzy, Y. You, C.L. Groth, and A.V. Tataurov. Stability and mismatch discrimination of locked nucleic acid-DNA duplexes. *Biochemistry*, 50:9352–9367, 2011.

[284] Diethard Tautz. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. *DNA fingerprinting: State of the science*, pages 21–28, 1993.

[285] Udogadi Nwawuba Stanley, Abdullahi Mohammed Khadija, ADAMS TAJUDEEN Bukola, Imose Omusi Precious, and Esewi Ayevbuomwan Davidson. Forensic DNA profiling: autosomal short tandem repeat as a prominent marker in crime investigation. *The Malaysian journal of medical sciences: MJMS*, 27(4):22, 2020.

[286] Marzieh Mojbafan, Reza Bahmani, Samira Dabbagh Bagheri, Zohreh Sharifi, and Sirous Zeinali. Mutational spectrum of autosomal recessive limb-girdle muscular dystrophies in a cohort of 112 Iranian patients and reporting of a possible founder effect. *Orphanet Journal of Rare Diseases*, 15:1–10, 2020.

[287] Pierre Murat, Guillaume Guilbaud, and Julian E Sale. DNA polymerase stalling at structured DNA constrains the expansion of short tandem repeats. *Genome biology*, 21(1): 1–26, 2020.

[288] Walther Parson, Romana Kirchebner, Roswitha Mühlmann, Kathrin Renner, Anita Kofler, Stefan Schmidt, and Reinhard Kofler. Cancer cell line identification by short tandem repeat profiling: power and limitations. *The FASEB journal*, 19(3):1–18, 2005.

[289] Hao Fan and Jia-You Chu. A brief review of short tandem repeat mutation. *Genomics, proteomics & bioinformatics*, 5(1):7–14, 2007.

[290] Adrián Hernández Bustos, Elisa Martiny, Nadia Bom Pedersen, Rohith Pavan Parvathaneni, Jonas Hansen, Hanlee P. Ji, and Kira Astakhova. Short tandem repeat DNA profiling using perylene-oligonucleotide fluorescence assay. *Analytical Chemistry*, may 2023. doi:10.1021/acs.analchem.3c00063. URL https://doi.org/10.1021/acs.analchem.3c00063.

[291] John SantaLucia, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA*, 95(4):1460–1465, 1998. URL http://www.pnas.org/cgi/content/abstract/95/4/1460.

[292] Philip N Borer, Barbara Dengler, Ignacio Tinoco Jr, and Olke C Uhlenbeck. Stability of ribonucleic acid double-stranded helices. *Journal of Molecular Biology*, 86(4):843–853, 1974.

[293] K. J. Breslauer, R Frank, H Blocker, and L. A. Marky. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA*, 83(11):3746–3750, 1986.

[294] Wahyu W Hadiwikarta, J-C Walter, Jef Hooyberghs, and Enrico Carlon. Probing hybridization parameters from microarray experiments: nearest-neighbor model and beyond. *Nucleic Acids Res.*, 40(18):e138–e138, 2012. doi:10.1093/nar/gks475.

[295] Adrián Szobi, Katarína Buranovská, Nina Vojtaššáková, Daniel Lovíšek, Halil Ã–nder Ã–zbaşak, Sandra Szeibeczederová, Liudmyla Kapustian, Zuzana Hudáčová, Viera Kováčová, Diana Drobná, Piotr Putaj, Stanislava Bírová, Ivana Čirková, Martin Čarnecký, Peter Kilián, Peter Jurkáček, Viktória Čabanová, Kristína Boršová, Monika Sláviková, Veronika Vaňová, Boris Klempa, Pavol Čekan, and Evan D. Paul. Vivid COVID-19 LAMP is an ultrasensitive, quadruplexed test using LNA-modified primers and a zinc ion and 5-br-PAPS colorimetric detection system. *Communications Biology*, 6(1), mar 2023. doi:10.1038/s42003-023-04612-9. URL https://doi.org/10.1038/s42003-023-04612-9.

[296] Masahiro Itonaga, Ibu Matsuzaki, Kenji Warigaya, Takaaki Tamura, Yuki Shimizu, Masakazu Fujimoto, Fumiyoshi Kojima, Masao Ichinose, and Shin-ichi Murata. Novel methodology for rapid detection of KRAS mutation using PNA-LNA mediated loop-mediated isothermal amplification. *PLoS One*, 11(3):e0151654, 2016.

[297] Natali Vega-Magaña, Rocío Sánchez-Sánchez, Jorge Hernández-Bello, Alberto Antony Venancio-Landeros, Marcela Peña-Rodríguez, Rosa Alejandra Vega-Zepeda, Byron Galindo-Ornelas, Mauricio Díaz-Sánchez, Mariel García-Chagollán, Gabriela Macedo-Ojeda, et al. RT-qPCR assays for rapid detection of the N501Y, 69-70del, K417N, and E484K SARS-CoV-2 mutations: a screening strategy to identify variants with clinical impact. *Frontiers in Cellular and Infection Microbiology*, 11:672562, 2021.

[298] Viera Kováčová, Kristína Boršová, Evan D Paul, Monika Radvánszka, Roman Hajdu, Viktória Čabanová, Monika Sláviková, Martina Ličková, Ľubomíra Lukáčiková, Andrej Belák, Lucia Roussier, Michaela Kostičová, Anna Líšková, Lucia Maďarová, Mária Štefkovičová, Lenka Reizigová, Elena Nováková, Peter Sabaka, Alena Koščálová, Broňa Brejová, Tomáš Vinař, Jozef Nosek, Pavol Čekan, and Boris Klempa. A novel, multiplexed RT-qPCR assay to distinguish lineage B.1.1.7 from the remaining SARS-CoV-2 lineages. *medRxiv*, 2021. doi:10.1101/2021.02.09.21251168. URL https://www.medrxiv.org/content/early/2021/03/03/2021.02.09.21251168.

[299] Monika Radvánszka, Evan D. Paul, Roman Hajdu, Kristína Boršová, Viera Kováčová, Piotr Putaj, Stanislava Bírová, Ivana Čirková, Martin Čarnecký, Katarína Buranovská, Adrián Szobi, Nina Vojtaššáková, Diana Drobná, Viktória Čabanová, Monika Sláviková, Martina Ličková, Veronika Vaňová, Sabína Fumačová Havlíková, Ľubomíra Lukáčiková, Ivana Kajanová, Juraj Koči, Diana Rusňáková, Tatiana Sedláčková, Klaas E. A. Max, Thomas Tuschl, Tomáš Szemes, Boris Klempa, and Pavol Čekan. Sequential development of several RT-qPCR tests using LNA nucleotides and dual probe technology to differentiate SARS-CoV-2 from influenza a and b. *Microbial Biotechnology*, mar 2022. doi:10.1111/1751-7915.14031. URL https://doi.org/10.1111%2F1751-7915.14031.

[300] Esta Tamanaha, Yinhua Zhang, and Nathan A. Tanner. Profiling RT-LAMP tolerance of sequence variation for SARS-CoV-2 RNA detection. *PLOS ONE*, 17(3):e0259610, mar 2022. doi:10.1371/journal.pone.0259610. URL https://doi.org/10.1371%2Fjournal.pone.0259610.

[301] Scott Sherrill-Mix, Young Hwang, Aoife M Roche, Abigail Glascock, Susan R Weiss, Yize Li, Leila Haddad, Peter Deraska, Caitlin Monahan, Andrew Kromer, et al. Detection of SARS-CoV-2 RNA using RT-LAMP and molecular beacons. *Genome biology*, 22(1): 1–17, 2021. doi:https://doi.org/10.1186/s13059-021-02387-y.

# Appendix

Here we show all the support materials for the work. As the tables are too large to be viewed here, by clicking on their colored names you will be redirected to a web page of our group where the referended material is stored.

# A Primer/Probe Sets

Lists of PCR and LAMP primer sets collected and evaluated in this work.

## A.1 PCR Sets

Table 10. List of PCR primers and probes.

## A.2 LAMP Sets

Table 11. List of LAMP primers.

# B Genome Sets

In all tables of this session, contiguous genome codes, such as LR794478.1 to LR794483.1, are shown in compact notation LR7944(78–83).1.

## B.1 SARS-CoV-2 Wild Type

Table 12. List of SARS-CoV-2 genome accession codes.

## B.2 Alpha Variant (B.1.1.7)

Table 13. List of Alpha variant (B.1.1.7) genome accession codes.

## B.3 Beta Variant (B.1.351)

Table 14. List of Beta variant (B.1.351) genome accession codes.

## B.4 Gamma Variant (P.1)

Table 15. List of Gamma variant (P.1) genome accession codes.

## B.5 Delta Variant (B.1.617.2)

Table 16. List of Delta variant (B.1.617.2) genome accession codes.

## B.6 Lambda Variant (C.37)

Table 17. List of Lambda variant (C.37) genome accession codes.

## B.7 Mu Variant (B.1.621)

Table 18. List of Mu variant (B.1.621) genome accession codes.

## B.8 Omicron Variant (B.1.1.529)

Table 19. List of Omicron variant (B.1.1.529) genome accession codes.

## B.9 BA.2 Subvariant

Table 20. List of BA.2 subvariant genome accession codes.

## B.10 BA.3 Subvariant

Table 21. List of BA.3 subvariant genome accession codes.

## B.11 BA.4 Subvariant

Table 22. List of BA.4 subvariant genome accession codes.

## B.12 BA.5 Subvariant

Table 23. List of BA.5 subvariant genome accession codes.

## B.13 SARS-CoV-1

Table 24. List of SARS-CoV-1 genome accession codes.

## B.14 Other Coronaviruses (non-SARS)

Table 25. List of other coronaviruses genome accession codes.

# C PCR Primer/Probe Coverages

Strict and partial coverages for all PCR primers and probes to wild-type SARS-CoV-2 and its variants and subvariants, and the other six coronaviruses.

## C.1 SARS-CoV-2 and its variants

Table 26. Coverages for SARS-CoV-2 and its variants in PCR

## C.2 SARS-CoV-1 and non-SARS coronaviruses

Table 27. Coverages for SARS-CoV-1 and non-SARS in PCR.

# D LAMP Primer Coverages

Strict and partial coverages for all LAMP primers to wild-type SARS-CoV-2 and its variants and subvariants, and the other six coronaviruses.

## D.1 SARS-CoV-2 and its variants



## D.2 SARS-CoV-1 and non-SARS coronaviruses

# E Displacement Profiles

## E.1 PCR Profiles

Average opening for a few PCR primers/probes in Figure 46.

## E.2 LAMP Profiles

Average opening for a few LAMP primers in Figure 47.

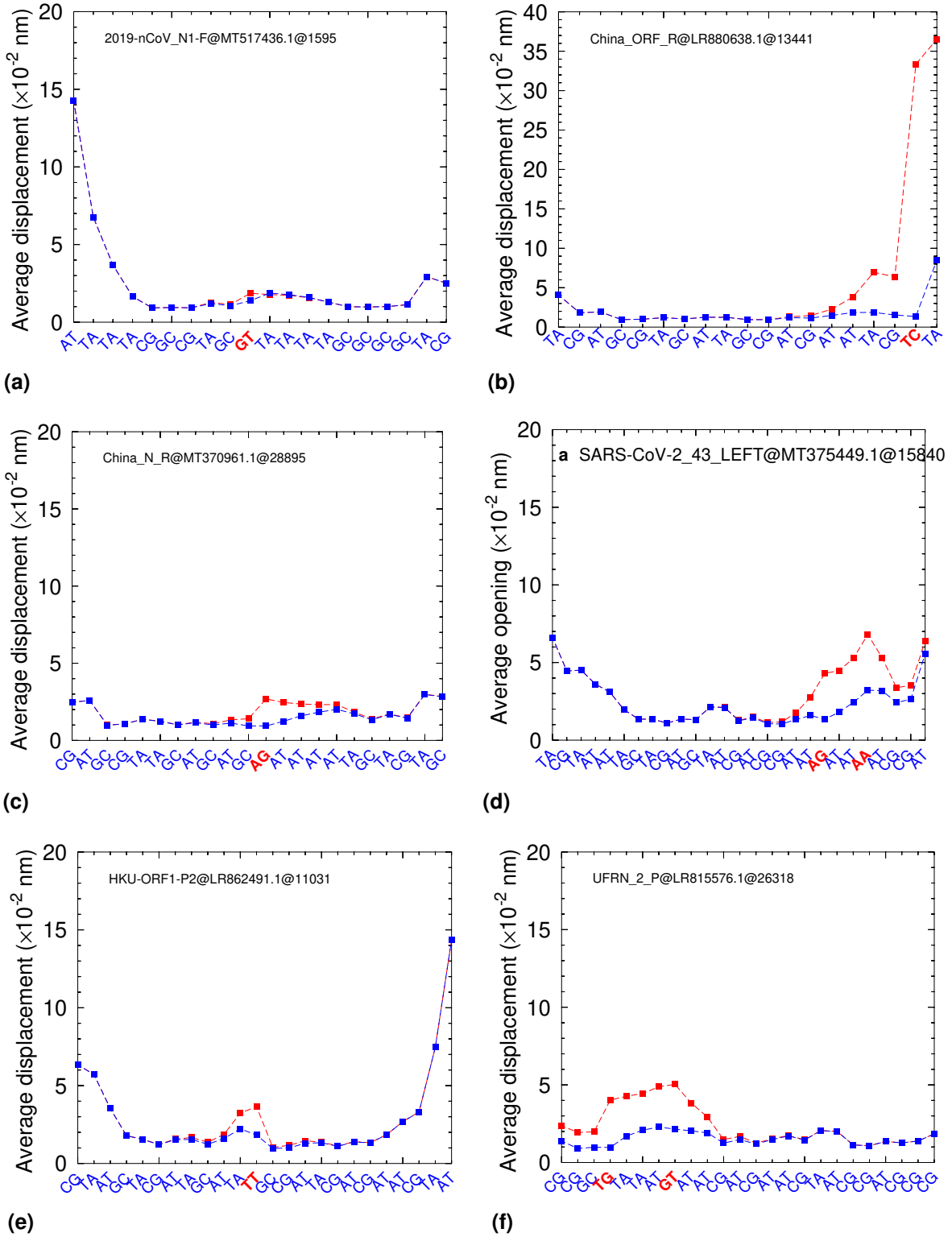



**Figure 46**

Displacement profiles for PCR primers. The blue line shows the displacement for matched aligment and the red one for mismatched alignment. The genome/primer hybridisation direction is $5' \longrightarrow 3'/3' \longrightarrow 5'$. (a) 2019-nCoV_N1-F primer [210] aligned to MT517436.1 genome at position 1595 and a GT mismatched pair ($T_{\text{ref.}} = 62.4°\text{C}$ and $T_{\text{MM}} = 56.5°\text{C}$); (b) China_ORF_R primer [212] aligned to LR880638.1 genome at position 13441 and a TC mismatched pair ($T_{\text{ref.}} = 61.5°\text{C}$ and $T_{\text{MM}} = 46.5°\text{C}$); (c) China_N_R primer [212] aligned to MT370961.1 genome at position 28895 and an AG mismatched pair ($T_{\text{ref.}} = 65.6°\text{C}$ and $T_{\text{MM}} = 56.1°\text{C}$); (d) SARS-CoV-2_43_LEFT primer [220] aligned to MT375449.1 genome at position 15840 and AG and AA mismatched pairs ($T_{\text{ref.}} = 69.1°\text{C}$ and $T_{\text{MM}} = 67.0°\text{C}$); (e) HKU_ORF1-P2 primer [224] aligned to LR862491.1 genome at position 11031 and a TT mismatched pair ($T_{\text{ref.}} = 63.6°\text{C}$ and $T_{\text{MM}} = 58.2°\text{C}$); (f) N_Set2_F1c primer [2] aligned to LR814012.2 genome at position 28341 and a GT mismatched pair ($T_{\text{ref.}} = 69.4°\text{C}$ and $T_{\text{MM}} = 69.0°\text{C}$).

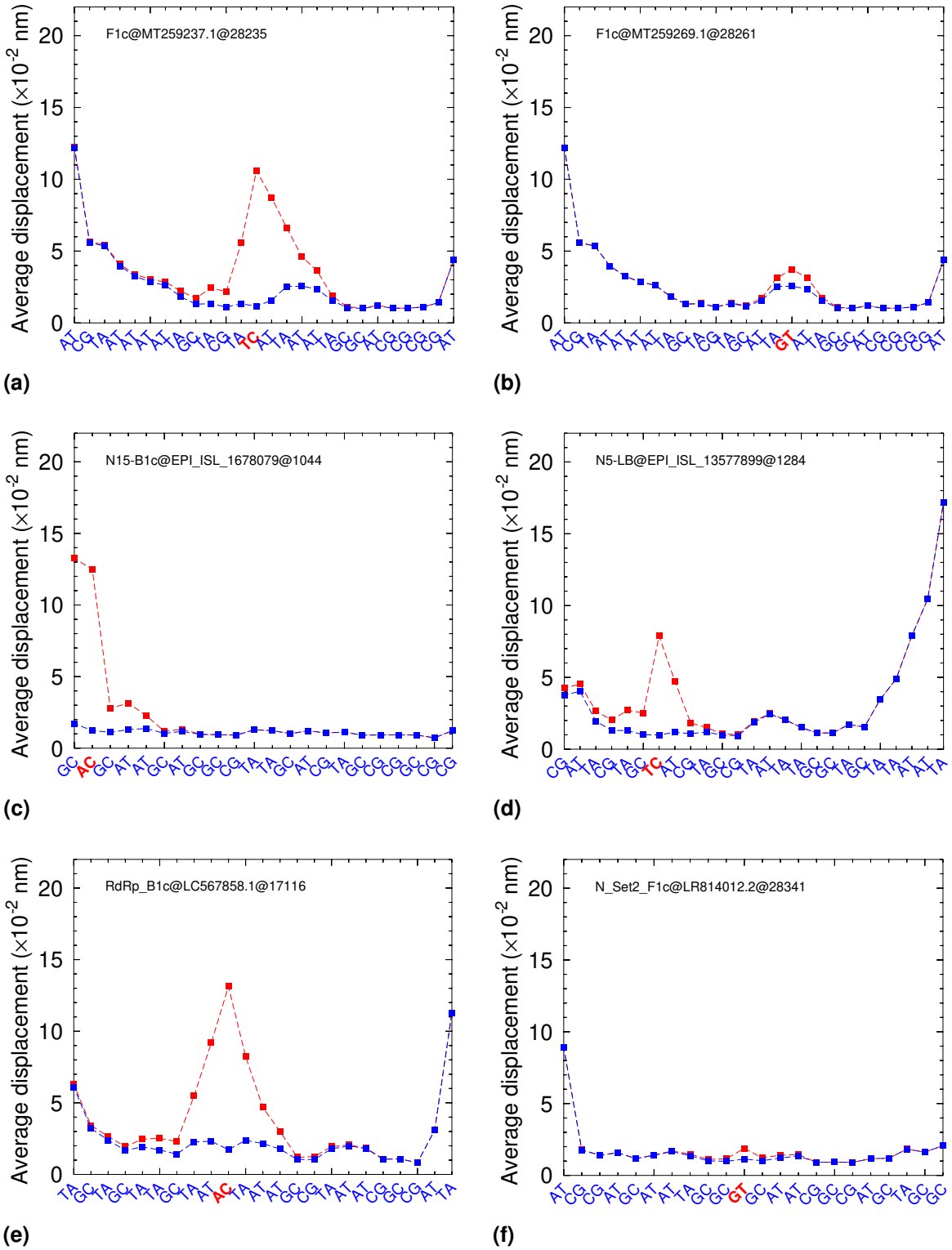



**Figure 47**

Displacement profiles for LAMP primers. The blue line shows the displacement for matched aligment and the red one for mismatched alignment. The genome/primer hybridisation direction is $5' \longrightarrow 3'/3' \longrightarrow 5'$. (a) F1c primer [229] aligned to MT259237.1 genome at position 28235 and a TC mismatched pair ($T_{\text{ref.}} = 69.7°\text{C}$ and $T_{\text{MM}} = 64.8°\text{C}$); (b) F1c primer [229] aligned to MT259269.1 genome at position 28261 and a GT mismatched pair ($T_{\text{ref.}} = 69.6°\text{C}$ and $T_{\text{MM}} = 72.0°\text{C}$); (c) N15-B1c primer [183] aligned to EPI_ISL_1678079 genome at position 1044 and an AC mismatched pair ($T_{\text{ref.}} = 76.6°\text{C}$ and $T_{\text{MM}} = 77.8°\text{C}$); (d) N5-LB primer [226] aligned to EPI_ISL_13453963 genome at position 1247 and a TC mismatched pair ($T_{\text{ref.}} = 67.6°\text{C}$ and $T_{\text{MM}} = 68.5°\text{C}$); (e) RdRp_B1c primer [2] aligned to LC567858.1 genome at position 17116 and a AC mismatched pair ($T_{\text{ref.}} = 63.4°\text{C}$ and $T_{\text{MM}} = 54.4°\text{C}$); (f) N_Set2_F1c primer [2] aligned to LR814012.2 genome at position 28341 and a GT mismatched pair ($T_{\text{ref.}} = 69.4°\text{C}$ and $T_{\text{MM}} = 69.0°\text{C}$).

# F Drop-out LAMP primers

All possible drop-out LAMP primers for wild-type SARS-CoV-2 and its variants and subvariants.

## F.1 SARS-CoV-2

Table 32. Potential drop-out primers for SARS-CoV-2.

## F.2 Alpha Variant

Table 33. Potential drop-out primers for Alpha variant.

## F.3 Beta Variant

Table 34. Potential drop-out primers for Beta variant.

## F.4 Gamma Variant

Table 35. Potential drop-out primers for Gamma variant.

## F.5 Delta Variant

Table 36. Potential drop-out primers for Delta variant.

## F.6 Lambda Variant

Table 37. Potential drop-out primers for Lambda variant.

## F.7 Mu Variant

Table 38. Potential drop-out primers for Mu variant.

## F.8 Omicron Variant

Table 39. Potential drop-out primers for Omicron variant.

## F.9 BA.2 Subvariant

Table 40. Potential drop-out primers for BA.2 subvariant.

## F.10 BA.3 Subvariant

Table 41. Potential drop-out primers for BA.3 subvariant.

## F.11 BA.4 Subvariant

Table 42. Potential drop-out primers for BA.4 subvariant.

## F.12 BA.5 Subvariant

Table 43. Potential drop-out primers for BA.5 subvariant.

# G FIP and BIP LAMP Primers – Terminal Mismatches

FIP and BIP LAMP primers with terminal mismatches that have an increase in their coverage.

## G.1 SARS-CoV-2

Table 44. Mismatched position in FIP and BIP primers for SARS-CoV-2.

## G.2 Alpha Variant

Table 45. Mismatched position in FIP and BIP primers for Alpha variant.

## G.3 Beta Variant

Table 46. Mismatched position in FIP and BIP primers for Beta variant

## G.4 Gamma Variant

Table 47. Mismatched position in FIP and BIP primers for Gamma variant

## G.5 Delta Variant

Table 48. Mismatched position in FIP and BIP primers for Delta variant

## G.6 Lambda Variant

Table 49. Mismatched position in FIP and BIP primers for Lambda variant

## G.7 Mu Variant

Table 50. Mismatched position in FIP and BIP primers for Mu variant

## G.8 Omicron Variant

Table 51. Mismatched position in FIP and BIP primers for Omicron variant

## G.9 BA.2 Subvariant

Table 52. Mismatched position in FIP and BIP primers for BA.2 subvariant.

## G.10 BA.3 Subvariant

Table 53. Mismatched position in FIP and BIP primers for BA.3 subvariant.

## G.11 BA.4 Subvariant

Table 54. Mismatched position in FIP and BIP primers for BA.4 subvariant.

## G.12 BA.5 Subvariant

Table 55. Mismatched position in FIP and BIP primers for BA.5 subvariant.

# H CRISPR LAMP Results

## H.1 Primers

Table 56. List of CRISPR LAMP primers from Ref. 60.

## H.2 Coverages

Table 57. Coverages for SARS-CoV-2 and Alpha, Beta, Gamma and Delta variants.

# I Mismatch Position Primers

## I.1 Primer Sequences

Table 58. List of primers from Ref. 266.

## I.2 Correlations

Correlation between average opening and extension efficiency for set1 and set2 in Figures 48 and 49, respectively.

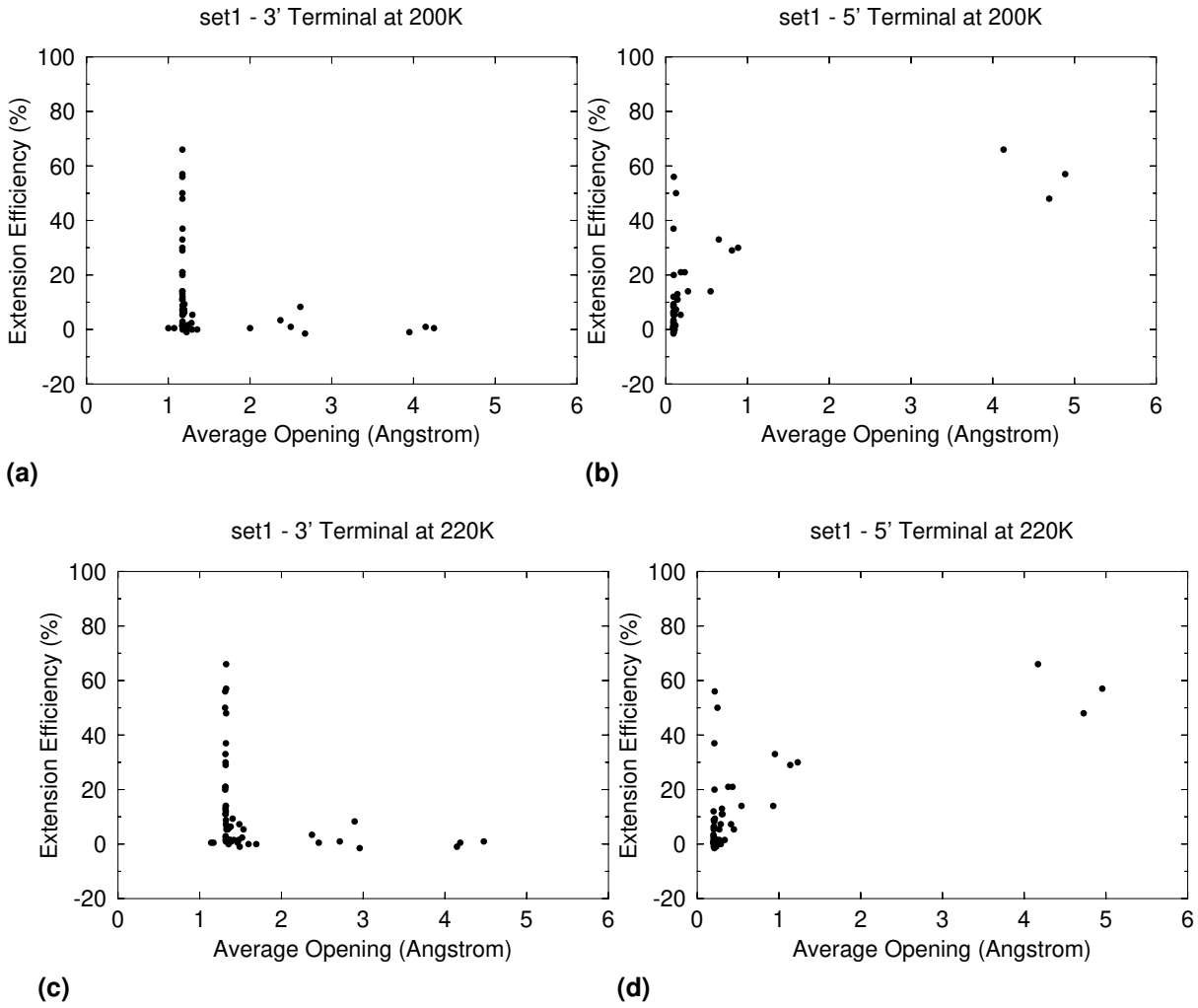



**Figure 48**
Correlation between average opening and extension efficiency for set1. In a temperature of calculation at 200 K to both 3′ (a) and 5′ (b) terminals, and in a temperature of calculation at 220 K in both 3′ (c) and 5′ (d) terminals.

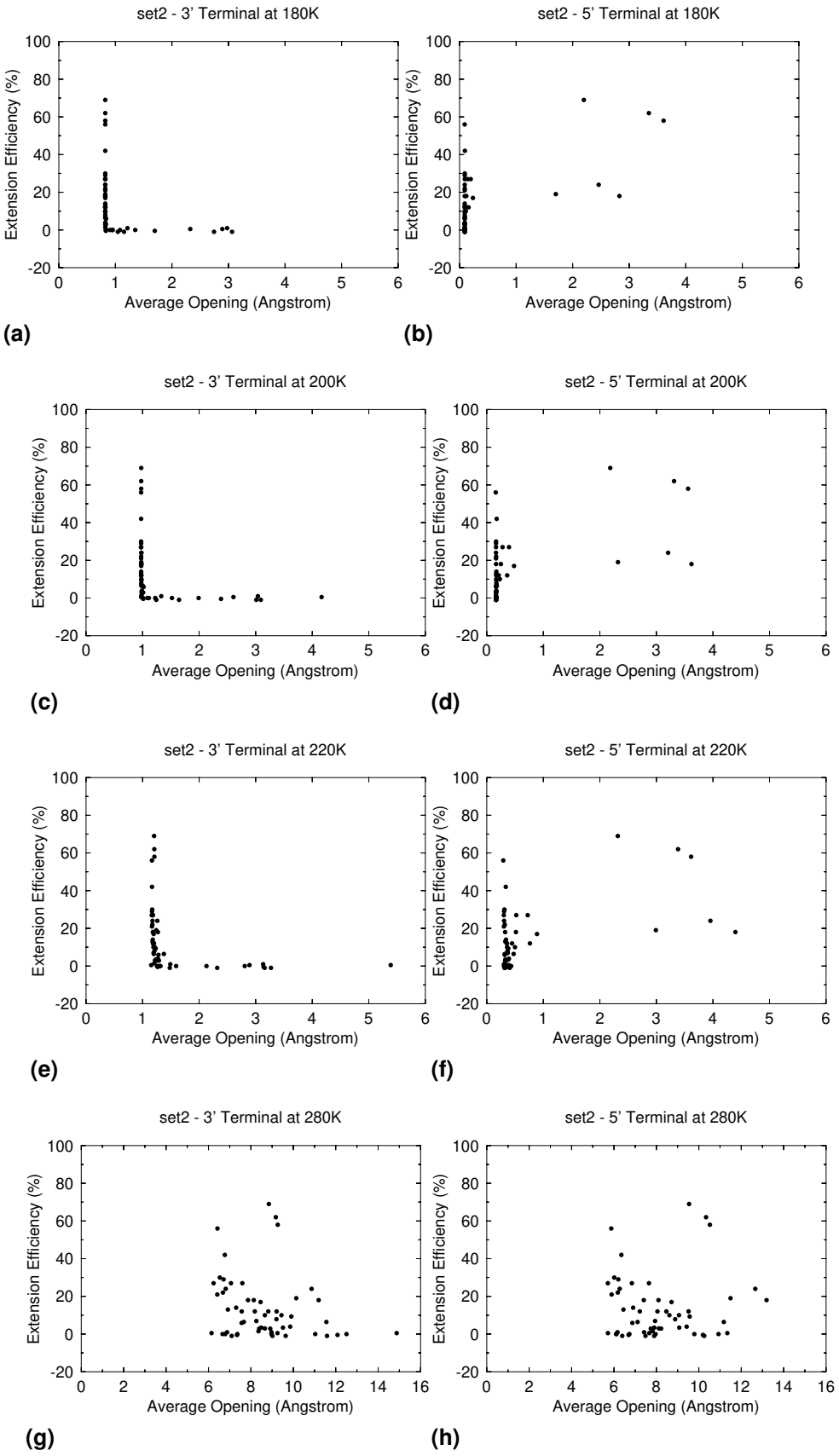



**Figure 49**

Correlation between average opening and extension efficiency for set2. In a temperature of calculation at 180 K to both 3′ (a) and 5′ (b) terminals; 200 K for 3′ (c) and 5′ (d) terminals; 220 K for 3′ (e) and 5′ (f) terminals, and 280 K for 3′ (g) and 5′ (h) terminals.

# J LNA-modified LAMP Primers

Table 59. List of LAMP primers from Ref. 295.

# K Published Articles

Articles published during this doctoral project. The first paper presents the evaluation of PCR primers/probes for wild-type SARS-CoV-2 and other six coronaviruses. The second, the evaluation of PCR primers/probes for SARS-CoV-2 variants of concern at the time. The third, the assessment of LAMP primers for wild-type and mutated SARS-CoV-2 genomes.

Molecular and Cellular Probes 56 (2021) 101707

Contents lists available at ScienceDirect

Molecular and Cellular Probes

journal homepage: www.elsevier.com/locate/ymcpr

Short communication

# Thermodynamic evaluation of the impact of DNA mismatches in PCR-type SARS-CoV-2 primers and probes

Pâmella Miranda [a,b], Gerald Weber [a,*]

[a] Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil
[b] Programa Interunidades de Pós-Graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

ARTICLE INFO

*Keywords:*
DNA stability
DNA mismatches
Mesoscopic models
PCR primer design
SARS-CoV-2

ABSTRACT

*Background:* DNA mismatches can affect the efficiency of PCR techniques if the intended target has mismatches in primer or probe regions. The accepted rule is that mismatches are detrimental as they reduce the hybridization temperatures, yet a more quantitative assessment is rarely performed.
*Methods:* We calculate the hybridization temperatures of primer/probe sets after aligning to SARS-CoV-2, SARS-CoV-1 and non-SARS genomes, considering all possible combinations of single, double and triple consecutive mismatches. We consider the mismatched hybridization temperature within a range of 5 °C to the fully matched reference temperature.
*Results:* We obtained the alignments of 19 PCR primers sets that were recently reported for the detection of SARS-CoV-2 and to 21665 SARS-CoV-2 genomes as well as 323 genomes of other viruses of the coronavirus family of which 10 are SARS-CoV-1. We find that many incompletely aligned primers become fully aligned to most of the SARS-CoV-2 when mismatches are considered. However, we also found that many cross-align to SARS-CoV-1 and non-SARS genomes.
*Conclusions:* Some primer/probe sets only align substantially to most SARS-CoV-2 genomes if mismatches are taken into account. Unfortunately, by the same mechanism, almost 75% of these sets also align to some SARS-CoV-1 and non-SARS viruses. It is therefore recommended to consider mismatch hybridization for the design of primers whenever possible, especially to avoid undesired cross-reactivity.

## 1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has caused a flurry of activity regarding the detection of SARS-CoV-2, in particular a substantial amount of new RT-PCR primers were developed for this specific purpose [1–19]. A number of factors can influence the reliability of the PCR detection, such as sample contamination [20], cross-reactivity with other viruses [14], contamination of reagents [21], non-specific annealing [22] and poor amplification efficiency [23]. A crucial primer design factor is its hybridization melting temperature [24] that is related to the annealing of oligonucleotides. A set of primers with close melting temperatures and in the ideal range for primer extension usually ensures good PCR performance [25].

A factor that may interfere with the hybridization temperatures are the presence of mismatches, that is non-Watson-Crick base pairs, between the primer and the DNA target. This affects the stability of the duplex, usually leading to a decrease in the hybridization temperature [26,27]. As a result, the presence of mismatches may influence the performance of primers restraining the amplification of DNA target. New mismatches arise due to mutations in primer regions of the target DNA, and may lead to false-negative results [20,28–30]. This is of special concern for the case of RNA viruses that have a high mutation rates [19,30]. Mutations that occur in the SARS-CoV-2 genome [31,32] imply that the presence of mismatches between primer/probe and the template eventually become inevitable. On the other hand, it is known that mismatch presence may affect only the first few cycles of PCR [33,34] and with proper design may even be advantageous [35]. Therefore, as a rule of thumb, the occurrence of single mismatches are admitted in the hope that they may not affect the detection of the target and its amplification [36]. Unfortunately, the thermodynamic instability caused by the presence of mismatches is rarely quantified in primer design for a number of reasons. One of which is that the prediction of hybridization temperatures involving mismatches carries large uncertainties. Unlike Watson-Crick complementary base pairs, AT and CG, the hydrogen

* Corresponding author.
*E-mail address:* gweber@ufmg.br (G. Weber).

https://doi.org/10.1016/j.mcp.2021.101707
Received 26 December 2020; Received in revised form 4 February 2021; Accepted 9 February 2021
Available online 17 February 2021

bonding and stacking interactions of mismatches are strongly dependent on the adjacent base pairs. Temperature predictions rely on experimental melting temperature data which typically do not cover the full combinatorial spectrum of mismatches and were carried out under high sodium buffer conditions [37,38]. However, this has now changed. A recent development from our group has reworked the parametrization for a comprehensive set of 4032 sequences containing up to three consecutive mismatches [39]. This now enables the analysis in unprecedented detail of the effect of mismatches in primer/probe hybridization. A key finding of this work was that up to 15% of all mismatch contexts may result in larger melting temperatures than the least stable canonical AT base pair, which means that the occurrence of mismatches may in some cases result in increased stabilization [39].

Here, we analyse how and if mismatches do influence the melting temperatures of primer/probehybridised to SARS-CoV-2 genomes. We collected 19 PCR primer/probe sets (297 primers and 43 probes) which cover seven different gene regions of SARS-CoV-2 genome (N, E, S, M, ORF1ab, RdRp and nsp2 genes) [1–19]. These primer/probes were aligned to 21665 genomes of SARS-CoV-2 and 323 genomes of other coronaviruses. Melting temperatures are calculated with a mesoscopic model using the newly developed parameters for up to three consecutive mismatches [39]. Using the mesoscopic model for the calculation of mismatches has an important advantage over nearest-neighbour models [37] as it naturally accounts for end effects, that is, mismatches located near the primer end may have different hybridization temperatures than those that are centred which reflects experimental observations on PCR efficiencies [34]. Despite this, in some cases, the presence of mismatches in the 3' end of the primer contributes to the amplification efficiency [40] and may avoid false-priming [24,25].

## 2. Materials and methods

### 2.1. Genomes and primer sets

We collected $N_G = 21665$ genomes of SARS-CoV-2 at NCBI [41], in 8 October 2020, and ensured that all were at least 25000 bp in size. The accession codes of these genomes are shown in Supplementary Table S1. To verify cross-reactivity we also performed the same analysis for $N_{\text{h.c.}} = 323$ human coronaviruses (229E, NL63, OC43, HKU1, MERS), including SARS-CoV-1, and their accession codes are shown in Supplementary Tables S2 and S3.

A total of 19 different primer/probe sets for RT-PCR were obtained from Refs. [1–19], their full details are shown in Supplementary Table S4. Note that several publications include primers from earlier reports. In particular, CDC primers [2] are included in several publications. Therefore, for each set we only considered those that were not repeated from other publications. Note that some primers and probes were designed for both SARS-CoV-1 and SARS-CoV-2 [4,14].

### 2.2. Primer and genome alignment

Each primer or probe sequence is aligned against a given genome using a Smith-Waterman algorithm [42], where matching base pairs AT and CG were given score 2, mismatches score −1, and no gaps were considered. Alignments were carried out in two strand configurations, first for the genome sequences as obtained from the database and taking the primer/probe sequence as complementary strand

$$5' - (\text{unmodified target genome sequence}) - 3'$$
$$3' - (\text{primer/probe}) - 5'$$

and next by taking the complementary of the genome sequence

$$5' - (\text{complementary target genome sequence}) - 3'$$
$$3' - (\text{primer/probe}) - 5'$$

These alignments are carried out regardless if the primer was identified as forward or reverse. In all cases the nominal directions of the primers were identified correctly.

A primer/probe that was completely aligned to a target genome, without mismatches, was termed as strictly matched. If there were up to three contiguous mismatches in the alignment it was called as partially matched. The limit of three contiguous mismatches relates to the available melting temperature parameters. Alignments with four or more contiguous mismatches were considered as not aligned.

As an example of partial alignment, we show the RdRp_SARSr-R1 primer (bottom strand) in the MT457390 genome

$$5' - \text{TATGCTAATAGTGTT}\underline{\text{T}}\text{TTAACATTTG} - 3'$$
$$3' - \text{ATACGATTATCACA}\underline{\text{C}}\text{AAATTGTAAAC} - 5'$$

where the mismatched site is underlined.

### 2.3. Calculation of melting temperatures

Hybridization temperatures $T_m$ are calculated from

$$T_m = a_0 + a_1\tau, \tag{1}$$

where $\tau$ is a statistical index calculated from the classical partition function of a model Hamiltonian, and $a_0$ and $a_1$ are regression coefficients obtained from a set of 4096 experimental melting temperatures of which 4032 are from sequences containing up to three consecutive mismatched base pairs [39]. The buffer conditions for these parameters are 50 mM sodium chloride, 10 mM sodium phosphate, pH 7.4, and total strand concentration 1.0 μM. For a complete description of the melting temperature calculation and experimental conditions see Refs. [39,43]. The index $\tau$ was calculated for each primer/probe, using the parameters reported in Ref. [39], after aligning against a reference genome. The calculation of $\tau$ also yields the average displacement profile which shows the expected base-pair opening along the oligonucleotide duplex, for details see Eq. (5) from Ref. 39.

### 2.4. Coverage evaluation

We calculated the hybridization temperatures from Eq. (1) for each primer/probe assuming a complete Watson-Crick complementarity which we called the reference temperature $T_{\text{ref.}}$, which are shown in Supplementary Table S5.

All 19 primer/probe sets were aligned against $N_G$ genomes and we kept only those alignments with up to three consecutive mismatches. The coverage for a strictly non-mismatched alignment $C_{\text{strict}}$ was calculated as

$$C_{\text{strict}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{MM}}}{N_G} \tag{2}$$

where $N_G$ is the total number of genomes which are at least 25000 bp in size, $N_{\text{n.a.}}$ is the number of genomes for which no alignment was found, and $N_{\text{MM}}$ is the number of genomes for which a partial alignment with up to three consecutive mismatches was found.

Next, for each of the $N_{\text{MM}}$ partial alignments we calculated the hybridization temperatures $T_{\text{MM}}$ from Eq. (1) taking into account the mismatches, and the difference to the reference temperature $T_{\text{ref.}}$ is

$$\Delta T_{\text{MM}} = T_{\text{ref.}} - T_{\text{MM}} \tag{3}$$

$T_{\text{MM}}$ is usually, but not always, lower than $T_{\text{ref.}}$ [39]. We will consider the partially mismatched coverage $C_{\text{part.}}$ as

$$C_{\text{part.}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{low}}(\Delta T_{\text{lim.}})}{N_G} \tag{4}$$

where $N_{\text{low}}$ is the number of primers/probes satisfying

$$\Delta T_{\text{MM}} \leq \Delta T_{\text{lim.}} \tag{5}$$

**Table 1**
Summary of the results for all primer/probe sets. Shown are the number of primers/probes $N_{pp}$ for each set, the range of reference temperatures $T_{ref.}$, the range of strict and partially mismatched coverages, for SARS-CoV-2, SARS-CoV-1 and non-SARS genomes. Detailed results for each primer are shown in Supplementary Tables S5, S6 and S7.

| Name of Set | $N_{pp}$ | $T_{ref.}$ (°C) | SARS-CoV-2 | | SARS-CoV-1 | | non-SARS | |
|---|---|---|---|---|---|---|---|---|
| | | | $C_{strict}$ (%) | $C_{part.}$ (%) | $C_{strict}$ (%) | $C_{part.}$ (%) | $C_{strict}$ (%) | $C_{part.}$ (%) |
| CDC [2] | 6 | 61.1–75.5 | 98.4–99.3 | 99.1–99.4 | 0.0–80.0 | 0.0–80.0 | 0.0 | 0.0 |
| WHO [1] | 21 | 51.3–70.3 | 97.9–99.4 | 97.9–99.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| Luminex [3] | 6 | 59.4–81.0 | 64.9–99.5 | 65.1–99.5 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Jalali et al [8]. | 6 | 61.3–64.2 | 98.7–99.4 | 98.7–99.4 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Corman et al [4]. | 21 | 61.7–81.7 | 0.0–99.4 | 0.0–99.4 | 0.0–100 | 0.0–100 | 0.0 | 0.0–19.2 |
| Davda et al [5]. | 16 | 56.5–70.3 | 95.3–99.4 | 95.3–99.4 | 0.0–80.0 | 0.0–90 | 0.0 | 0.0 |
| Grant et al [6]. | 2 | 62.4–79.9 | 97.5–99.3 | 97.5–99.5 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Hirotsu et al [7]. | 3 | 60.6–68.8 | 0.0–99.4 | 0.0–99.4 | 0.0 | 0.0 | 0.0 | 0.0 |
| Lanza et al [9]. | 27 | 60.5–75.0 | 98.5–99.4 | 98.8–99.5 | 0.0 | 0.0–90.0 | 0.0 | 0.0 |
| Li et al [10]. | 2 | 67.5–70.4 | 98.2–99.2 | 98.3–99.2 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Lu et al [11]. | 3 | 64.0–74.7 | 98.6–99.4 | 99.3–99.4 | 0.0 | 0.0 | 0.0 | 0.0 |
| Munnink et al [12]. | 171 | 65.4–74.9 | 45.9–99.5 | 46.0–99.7 | 0.0–100 | 0.0–100 | 0.0 | 0.0–0.639 |
| Nalla et al [13]. | 4 | 51.7–68.0 | 0.0–99.5 | 0.0–99.5 | 0.0–90.0 | 0.0–100 | 0.0–68.7 | 0.0–68.7 |
| Niu et al [14]. | 6 | 59.2–84.2 | 0.0–99.4 | 0.0–99.4 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Park et al [15]. | 20 | 59.3–65.4 | 94.9–99.5 | 94.9–99.5 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Rahman et al [16]. | 6 | 64.2–76.1 | 97.6–99.4 | 97.6–99.4 | 0.0 | 0.0–100 | 0.0 | 0.0 |
| Toptan et al [17]. | 6 | 62.2–65.4 | 98.8–99.5 | 98.8–99.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| Vogels et al [18]. | 11 | 58.6–65.5 | 0.0–99.3 | 0.0–99.3 | 0.0–100 | 0.0–100 | 0.0 | 0.0 |
| Yip et al [19]. | 2 | 61.4–63.4 | 99.1–99.2 | 99.1–99.3 | 0.0 | 0.0 | 0.0 | 0.0 |

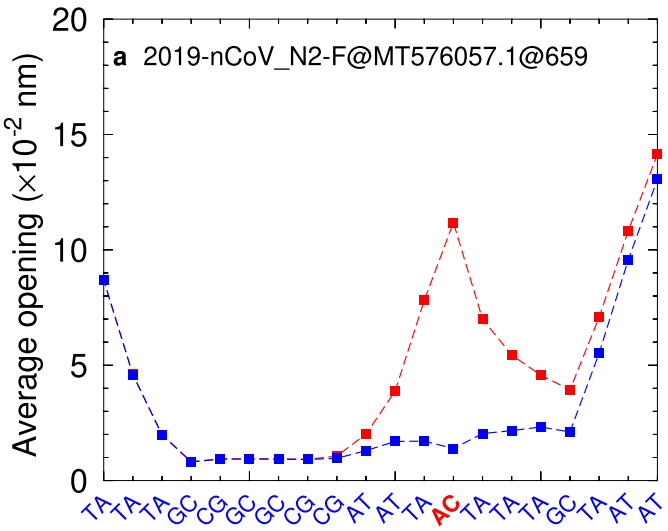



**Fig. 1.** Displacement profile for CDC 2019-nCoV_N2–F when aligned to SARS-COV-2 MT576057.1 at position 659 has a mismatch AC (red symbols) instead of CG (blue symbols).

Here, we will use $\Delta T_{lim.} = 5$ °C, that is, we will consider that mismatched primers/probes with $T_{MM}$ no more than 5 °C below the reference temperature $T_{ref.}$ are still acceptable.

### 2.5. Availability

The software packages used for this work are freely available, please see supplementary section S1 for step-by-step instructions on how to download, install, run and perform the analysis described here. The authors will consider requests for primer analysis, please see contact details.

## 3. Results and discussion

After aligning all primer/probe sets to all genomes we calculated their hybridization temperatures taking into account up to three consecutive mismatches, as detailed in the methods section. Table 1 summarises all sets analysed, their range of reference hybridization temperatures, strict and partial coverage for SARS-CoV-2 and for non-SARS-CoV-2 viruses. The detailed results for each primer are shown in Supplementary Table S5 for SARS-CoV-2, including the amount of mismatched alignments in the last three positions related to both terminals. If the PCR primers can in principle bind to non-SARS-CoV-2 targets then this set may not specific [22]. Considering this, we analysed the 19 primer/probe sets in relation to genomes of other coronaviruses (non-SARS) as well as SARS-CoV-1 to verify if there was some cross-reactivity. The detailed results are shown in Supplementary Table S6 for SARS-COV-1, and S7 for non-SARS viruses.

The typical design rules for PCR primers and probes recommend that the range of hybridization temperatures in a given set should be narrow, of the order of 10 °C [25,44]. Several authors even suggest that the range for primer pair should be no more than 5°C [22,45] or even less than 1°C [46]. However, it is evident, from Table 1, that very few sets have temperature ranges below 10 °C, while some even exceed 20 °C. For example, the Luminex set, which includes the primer/probe set of China CDC, shows differences in the primer temperatures up to 21.6 °C. However, when mismatches are considered the hybridization temperatures may go far below the design range. In Fig. 1, we show an example of a displacement profile where a single AC mismatch completely disturbs the surrounding AT base pairs and the hybridization temperature drops to $T_{MM} = 48.5$ °C, down from a reference temperature of $T_{ref.} = 61.1$ °. However, a presence of one or more mismatches does not necessarily imply in a reduction of hybridization temperature. For example, the SARS-CoV-2_89_RIGHT primer when aligned to MT259228.1 has two consecutive mismatches towards the 5' end, see Supplementary Fig. S1. Even though these mismatches induce a small end fraying, it has a calculated temperature of $T_{MM} = 68.7$ °C which is even somewhat higher than its reference temperature $T_{ref.} = 68.3$ °C. This stability is caused by an increased stacking interaction between the GA and AA mismatches [39].

In terms of SARS-CoV-2 strict coverage, most sets have $C_{strict}$ typically above 90%, which is expected as the primer design is guided by existing genomes. However, a number of specific probes, such as RdRp_SARSr-P1-1 from Corman et al. [4] go from 0 to 99.4% only if mismatches are taken into account. Indeed, as pointed out by Pillonel et al. [47] several of the probes from Corman et al. [4] do not fully match the available SARS-CoV-2 genomes (1623 at the time [47]). However, it was also observed that the mismatches had little effect on their efficiencies [13,28] which is consistent with our calculations. The reason

for the presence of mismatches in this case, as explained in Corman and Drosten [28], was the incomplete genomic information available at the time when this set was designed. It is worth noticing that while accounting for mismatches increases the coverage of this particular primer/probe set, it also increases the coverage for SARS-CoV-1 and even non-SARS as shown in Table 1.

We observed that in some cases the presence of few mismatches substantially decrease the hybridization temperature, leading to a complete absence of coverage. For example, four primers from Vogels et al. [18] do not align with any genome at all, not even when considering the mismatches as their hybridization temperatures $T_{MM}$ are too low in comparison to reference temperature $T_{ref.}$. In contrast, for several cases when mismatches are taken into account the coverage becomes almost complete. A special example is probe 2019-nCoV_N1–P from CDC set that had 223 further mismatch alignments increasing the strict coverage of 98.4% to partial coverage of 99.4%. Similar findings were observed for SARS-CoV-2_6_LEFT [12] and NIID_WH-1_F501 of WHO [1].

The cross-reactivity, that is, the coverage of SARS-Cov-1 and non-SARS, appear in most primer/probe sets when mismatches are taken into account. Of the 19 primer/probe sets, we found only 5 sets that do not present cross-reactivity at all, see Table 1.

## 4. Conclusion

We evaluated the impact of mismatches in the hybridization of primers and probes for the detection of SARS-CoV-2 and other genomes. We have shown that the effect of mismatches on the probe/primer hybridization is not straightforward and can only be fully evaluated with a detailed calculation with up-to-date model parameters. In particular, our calculations showed that a substantial amount of the existing primers/probes may cross-react to SARS-CoV-1 and non-SARS genomes, which further highlights the need for taking mismatch hybridization into account.

## Funding statement

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes/Ação Emergencial, Brazil, Finance Code 001).

## Author statement

**Pâmella Miranda**: Conceptualization, Methodology, Formal analysis, Investigation, Writing -Original Draft, Visualization.

**Gerald Weber**: Conceptualization, Methodology, Supervision, Writing - Original Draft, Visualization, Project administration

## Declaration of competing interest

None.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mcp.2021.101707.

## References

[1] https://www.who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of–sars-cov-2-institut-pasteur-paris.pdf?sfvrsn=3662fcb6_2.
[2] https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html, 2020.
[3] https://www.captodayonline.com/LuminexSARSCoV219.pdf, 2020.
[4] V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D.K. Chu, T. Bleicker, S. Brünink, J. Schneider, M.L. Schmidt, et al., Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR, Euro Surveill. 25 (3) (2020) 2000045.
[5] J.N. Davda, K. Frank, S. Prakash, G. Purohit, D.P. Vijayashankar, D. Vedagiri, K. B. Tallapaka, K.H. Harshan, A.B. Siva, R.K. Mishra, et al., An Inexpensive RT-PCR Endpoint Diagnostic Assay for SARS-CoV-2 Using Nested PCR: Direct Assessment of Detection Efficiency of RT-qPCR Tests and Suitability for Surveillance, bioRxiv (2020), 2020.06.08.139477.
[6] P.R. Grant, M.A. Turner, G.Y. Shin, E. Nastouli, L.J. Levett, Extraction-free COVID-19 (SARS-CoV-2) Diagnosis by RT-PCR to Increase Capacity for National Testing Programmes during a Pandemic, BioRxiv (2020), 2020.04.06.028316.
[7] Y. Hirotsu, H. Mochizuki, M. Omata, Double-Quencher Probes Improved the Detection Sensitivity of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by One-step RT-PCR, medRxiv (2020), 2020.03.17.20037903.
[8] S. Jalali, V.K. Patel, V. Prasad, S. Dasgupta, A. Sapre, B. Bhadra, Detection of SARS-CoV-2 using high-throughput PCR, J. Appl. Biol. Biotechnol. 8 (3) (2020) 1–5.
[9] D. C. Lanza, J. P. Lima, S. M. Jerônimo, Design and in Silico Validation of Polymerase Chain Reaction Primers to Detect Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) . https://www.researchsquare.com/article/rs-26306/v1.
[10] C. Li, D. Debruyne, J. Spencer, V. Kapoor, L.Y. Liu, B. Zhang, L. Lee, R. Feigelman, G. Burdon, J. Liu, et al., High Sensitivity Detection of SARS-CoV-2 Using Multiplex PCR and a Multiplex-PCR-Based Metagenomic Method, bioRxiv (2020), 2020.03.12.988246.
[11] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet 395 (10224) (2020) 565–574.
[12] B.B.O. Munnink, D.F. Nieuwenhuijse, M. Stein, A. O'Toole, M. Haverkarte, M. Mollers, S.K. Kamga, C. Schapendonk, P. Lexmond, M. Pronk, et al., Rapid SARS-CoV-2 Whole Genome Sequencing for Informed Public Health Decision Making in the Netherlands, Nature Medicine 26 (9) (2020) 1405–1410.
[13] A.K. Nalla, A.M. Casto, M.-L.W. Huang, G.A. Perchetti, R. Sampoleo, L. Shrestha, Y. Wei, H. Zhu, K.R. Jerome, A.L. Greninger, Comparative performance of SARS-CoV-2 detection assays using seven different primer-probe sets and one assay kit, J. Clin. Microbiol. 58 (6) (2020) e00557–20.
[14] P. Niu, R. Lu, L. Zhao, H. Wang, B. Huang, F. Ye, W. Wang, W. Tan, Three Novel Real-Time RT-PCR Assays for Detection of COVID-19 Virus, China CDC Weekly, 2020, pp. 1–5.
[15] M. Park, J. Won, B.Y. Choi, C.J. Lee, Optimization of primer sets and detection protocols for SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR, Exp. Mol. Med. 52 (6) (2020) 963–977.
[16] H. Rahman, I. Carter, K. Basile, L. Donovan, S. Kumar, T. Tran, D. Ko, S. Alderson, T. Sivaruban, J.-S. Eden, et al., Interpret with caution: an evaluation of the commercial AusDiagnostics versus in-house developed assays for the detection of SARS-CoV-2 virus, J. Clin. Virol. (2020) 104374.
[17] T. Toptan, S. Hoehl, S. Westhaus, D. Bojkova, A. Berger, B. Rotter, K. Hoffmeier, S. Ciesek, M. Widera, Optimized qRT-PCR approach for the detection of intra- and extra-cellular SARS-CoV-2 RNAs, bioRxiv URL https://www.biorxiv.org/content/early/2020/04/25/2020.04.20.052258.
[18] C.B. Vogels, A.F. Brito, A.L. Wyllie, J.R. Fauver, I.M. Ott, C.C. Kalinich, M. E. Petrone, A. Casanovas-Massana, M.C. Muenker, A.J. Moore, et al., Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets, Nat. Microbiol. 5 (2020) 1299–1305, https://doi.org/10.1038/s41564-020-0761-6.
[19] C.C.-Y. Yip, C.-C. Ho, J.F.-W. Chan, K.K.-W. To, H.S.-Y. Chan, S.C.-Y. Wong, K.-H. Leung, A.Y.-F. Fung, A.C.-K. Ng, Z. Zou, et al., Development of a novel, genome subtraction-derived, SARS-CoV-2-specific COVID-19-nsp2 real-time RT-PCR assay and its evaluation using clinical specimens, Int. J. Mol. Sci. 21 (7) (2020) 2574.
[20] G. Lippi, A.-M. Simundic, M. Plebani, Potential Preanalytical and Analytical Vulnerabilities in the Laboratory Diagnosis of Coronavirus Disease 2019 (COVID-19), Clinical Chemistry and Laboratory Medicine 58 (7) (2020) 1070–1076.
[21] K. Wernike, M. Keller, F.J. Conraths, T.C. Mettenleiter, M.H. Groschup, M. Beer, Pitfalls in SARS-CoV-2 PCR Diagnostics, Transboundary and Emerging Diseases (2020). https://onlinelibrary.wiley.com/doi/abs/10.1111/tbed.13684.
[22] D. Li, J. Zhang, J. Li, Primer design for quantitative real-time PCR for the emerging Coronavirus SARS-CoV-2, Theranostics 10 (16) (2020) 7150.
[23] T. Suo, X. Liu, J. Feng, M. Guo, W. Hu, D. Guo, H. Ullah, Y. Yang, Q. Zhang, X. Wang, et al., ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens, Emerg. Microb. Infect. 9 (2020) 1259–1268, https://doi.org/10.1080/22221751.2020.1772678.
[24] M. Mitsuhashi, Technical report: Part 1. Basic requirements for designing optimal oligonucleotide probe sequences, J. Clin. Lab. Anal. 10 (5) (1996) 277–284.
[25] M. Mitsuhashi, Technical report: Part 2. Basic requirements for designing optimal PCR primers, J. Clin. Lab. Anal. 10 (5) (1996) 285–293.
[26] D. de Silva, C.T. Wittwer, Monitoring hybridization during polymerase chain reaction, J. Chromatogr. B Biomed. Sci. Appl. 741 (1) (2000) 3–13.
[27] R. Stadhouders, S.D. Pas, J. Anber, J. Voermans, T.H. Mes, M. Schutten, The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay, J. Mol. Diagn. 12 (1) (2010) 109–117.
[28] V.M. Corman, C. Drosten, Authors' response: SARS-CoV-2 detection by real-time RT-PCR, Euro Surveill. 25 (21) (2020) 2001035.
[29] C.R. Carpenter, P.A. Mudd, C.P. West, E. Wilber, S.T. Wilber, Diagnosing COVID-19 in the Emergency Department: A Scoping Review of Clinical Examinations, Laboratory Tests, Imaging Accuracy, and Biases, Academic Emergency Medicine 27 (8) (2020) 653–670.

[30] C. Koo, S. Kaur, Z.-Y. Teh, H. Xu, A. Nasir, Y.-L. Lai, E. Khan, L.-C. Ng, H. C. Hapuarachchi, Genetic variability in probe binding regions explains false negative results of a molecular assay for the detection of dengue virus, Vector Borne Zoonotic Dis. 16 (7) (2016) 489–495.
[31] K.A. Khan, P. Cheung, Presence of mismatches between diagnostic PCR assays and coronavirus SARS-CoV-2 genome, Royal Soc. Open Sci. 7 (6) (2020) 200636.
[32] J. Carneiro, C. Gomes, C. Couto, F.CoV2ID Pereira, Detection and Therapeutics Oligo Database for SARS-CoV-2, bioRxiv (2020), 2020.04.19.048991.
[33] S. Lefever, F. Pattyn, J. Hellemans, J. Vandesompele, Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays, Clin. Chem. 59 (10) (2013) 1470–1480.
[34] J.-H. Wu, P.-Y. Hong, W.-T. Liu, Quantitative effects of position and type of single mismatch on single base primer extension, J. Microbiol. Methods 77 (3) (2009) 267–275.
[35] Y. Higashi, H. Niimi, I. Sakamaki, Y. Yamamoto, I. Kitajima, Rapid identification of Candida species in candidemia directly from blood samples using imperfect match probes, Sci. Rep. 10 (1) (2020) 1–9.
[36] K. Warton, Y. Xu, C. E. Ford, Target sequence heterogeneity causes the 'hook effect'in fluorescent dye-based quantitative PCR, Biotechniques 69 (0).
[37] N. Peyret, P.A. Seneviratne, H.T. Allawi, J. SantaLucia Jr., Nearest-Neighbour thermodynamics and NMR of DNA sequences with internal A A, C C G G and T T mismatches, Biochemistry 38 (12) (1999) 3468–3477.
[38] H. Allawi, J. SantaLucia, J, Thermodynamics of internal C.T mismatches in DNA, Nucleic Acids Res. 26 (11) (1998) 2694–2701. http://nar.oupjournals.org/cgi/content/abstract/26/11/2694.
[39] L.M. Oliveira, A.S. Long, T. Brown, K.R. Fox, G. Weber, Melting temperature measurement and mesoscopic evaluation of single, double and triple DNA mismatches, Chem. Sci. 11 (2020) 8273–8287, https://doi.org/10.1039/d0sc01700k. https://pubs.rsc.org/en/content/articlelanding/2020/SC/D0SC01700K.
[40] S. Kwok, D. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson, J. Sninsky, Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies, Nucleic Acids Res. 18 (4) (1990) 999–1005.
[41] https://www.ncbi.nlm.nih.gov/sars-cov-2/, 2020.
[42] T.F. Smith, M.S. Waterman, et al., Identification of common molecular subsequences, J. Mol. Biol. 147 (1) (1981) 195–197.
[43] G. Weber, N. Haslam, J.W. Essex, C. Neylon, Thermal equivalence of DNA duplexes for probe design, J. Phys. Condens. Matter 21 (2009), 034106, https://doi.org/10.1088/0953-8984/21/3/034106.
[44] C. Dieffenbach, T. Lowe, G. Dveksler, General concepts for PCR primer design, PCR Methods Appl. 3 (3) (1993) S30–S37.
[45] E. van Pelt-Verkuil, A. Van Belkum, J.P. Hays, Principles and Technical Aspects of PCR Amplification, Springer Science & Business Media, 2008.
[46] B. Thornton, C. Basu, Real-time PCR (qPCR) primer design using free online software, Biochem. Mol. Biol. Educ. 39 (2) (2011) 145–154.
[47] T. Pillonel, V. Scherz, K. Jaton, G. Greub, C. Bertelli, Letter to the editor: SARS-CoV-2 detection by real-time RT-PCR, Euro Surveill. 25 (21) (2020) 2000880.

# Mesoscopic Evaluation of DNA Mismatches in PCR Primer-Target Hybridisation to Detect SARS-CoV-2 Variants of Concern

Pâmella Miranda[1,2(✉)], Vivianne Basílio Barbosa[1], and Gerald Weber[1]

[1] Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
{pamella-fisica,gweber}@ufmg.br
[2] Programa Interunidades de Pós-Graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

**Abstract.** Mismatches are any type of base-pairs other than AT and CG. They are an expected occurrence in PCR primer-target hybridisation and may interfere with the amplification and in some cases even prevent the detection of viruses and other types of target. Given the natural occurrence of mutations it is expected that the number of primer-target mismatches increases which may result in a larger number of false-negative PCR diagnostics. However, mismatches may equally improve the primer-target hybridisation since some types of mismatches may stabilize the helix. Only very recently have thermodynamic parameters become available that would allow the prediction of mismatch effects at buffer conditions similar to that of PCR. Here we collected primers from WHO recommendation and aligned them to the genomes of the current variants of concern (VOC): Alpha, Beta, Gamma and Delta variants. We calculated the hybridisation temperatures taking into account up to three consecutive mismatches with the new parameters. We assumed that hybridisation temperatures to mismatched alignments within a range of 5 °C of the non-mismatched temperature to still result in functional primers. In addition, we calculated strict and partial coverages for complete and mismatched alignments considering only single, double and triple consecutive mismatches. We found that if mismatches are taken into account, the coverage of WHO primers actually increase for VOCs and for the Delta variant it becomes 100%. This suggest that, at least for the moment, these primers should continue to be effective for the detection of VOCs.

**Keywords:** DNA mismatches · PCR primers · Mesoscopic models

Supported by organization Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes/Ação Emergencial, Brazil, Finance Code 001).


P. F. Stadler et al. (Eds.): BSB 2021, LNBI 13063, pp. 145–150, 2021.
https://doi.org/10.1007/978-3-030-91814-9_15

## 1 Introduction

The emergence of the pandemic of COVID-19 required the deployment of large-scale testing to control and monitor the disease. For this purpose, several protocols of PCR-based methods, mainly RT-PCR, were developed. Although RT-PCR is the gold standard molecular diagnostic, a few factors can interfere with its accuracy and performance such as sample quality and low amplification efficiency [9]. PCR efficiency in particular may be affected by destabilizing mismatches in primer-target. They may affect the ability of primers hybridise to the target, which may lead to non-amplification and, consequently, to non-detection. The influence caused in the hybridisation by mismatches depends on their length, sequential environment, position and number [4]. Even so, mismatches in primer-target duplex impact only the first few cycles of the PCR reaction [9]. They also affect the melting temperature, which is an important parameter to the primer design and is related to their stability and performance.

Here, we describe the evaluation of 21 primers and probes for RT-PCR recommended by WHO [1] in early 2020 to be applied to the detection of "original" SARS-CoV-2, which was evaluated in a previous work [5]. We collected those from Institut Pasteur, Department of Medical Sciences (Thailand) and National Institute of Infectious Diseases (Japan). We applied a mesoscopic model to calculate the hybridisation temperatures of alignments using a newly developed parameters for up to three consecutive mismatches [7]. The primers/probes were analysed regarding to SARS-CoV-2 variants of concern (VOC) classified so far: B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma) and B.1.617.2 (Delta) variants.

## 2 Materials and Methods

**Primer/Genome Sets.** We collected 21 primers and probes from the summary of protocols recommended by WHO [1]. Regarding the genomes, we collected from GISAID [2] 7247 genomes of Alpha, 7497 of Beta and 2308 of Gamma variants in 7 April 2021, and 7943 genomes of Delta variant in 5 June 2021.

**Primer/Genome Alignments.** Primers and probes were aligned against each genome using Smith-Waterman algorithm [8], where AT and CG base pairs were given score 2, mismatches score $-1$, and no gaps were considered. Alignments were carried out regarding two strand configurations. The genome sequence as obtained from the database

$$\begin{array}{c} 5'-(\text{unmodified target genome sequence})-3' \\ 3'-(\text{primer/probe sequence})-5' \end{array}$$

and its complementary counterpart

$$\begin{array}{c} 5'-(\text{complementary target genome sequence})-3' \\ 3'-(\text{primer/probe sequence})-5' \end{array}$$

The alignments without mismatches were termed as strictly matched, those which contained up to three consecutive mismatches as partially matched and alignments with four or more consecutive mismatches were considered as not aligned. The limit of three consecutive mismatches is due to the available melting temperature parameters.

**Calculating Hybridisation Temperatures.** Hybridisation temperatures were calculated using

$$T_m = a_0 + a_1\tau, \tag{1}$$

where $\tau$ is a statistical index, which is calculated from the classical partition function of a model Hamiltonian, and $a_0$ and $a_1$ are regression coefficients obtained from a set of sequences containing up to three contiguous mismatched base pairs [7]. Moreover, the calculation of $\tau$ also generate the average displacement profile which shows the expected base-pair opening along the primer-target duplex. For a complete description of this calculation see Ref. [7].

**Calculating Strict and Partial Coverages.** We calculated the melting temperatures for the 21 primers/probes assuming a perfect hybridisation, which we called the reference temperature $T_{\text{ref.}}$, see Table 1. Alignments were carried out between primer and genomes of VOCs and kept only those with up to three contiguous mismatches. The coverage for strictly matched alignments $C_{\text{strict}}$ was calculated as

$$C_{\text{strict}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{MM}}}{N_G} \tag{2}$$

where $N_G$ is the total number of genomes which are at least 25000 bp in size, $N_{\text{n.a.}}$ the number of genomes for which no alignment was found, and $N_{\text{MM}}$ the number of genomes for which a partial alignment with up to three contiguous mismatches was found.

For partially matched alignments, we calculated the melting temperature $T_{\text{MM}}$ taking into account the mismatches, and assumed the difference to the reference temperature $T_{\text{ref.}}$

$$\Delta T_{\text{MM}} = T_{\text{ref.}} - T_{\text{MM}} \tag{3}$$

Then, we calculated the partially coverage $C_{\text{part.}}$ as

$$C_{\text{part.}} = \frac{N_G - N_{\text{n.a.}} - N_{\text{low}}(\Delta T_{\text{lim.}})}{N_G} \tag{4}$$

where $N_{\text{low}}$ is the number of primers satisfying

$$\Delta T_{\text{MM}} \leq \Delta T_{\text{lim.}} \tag{5}$$

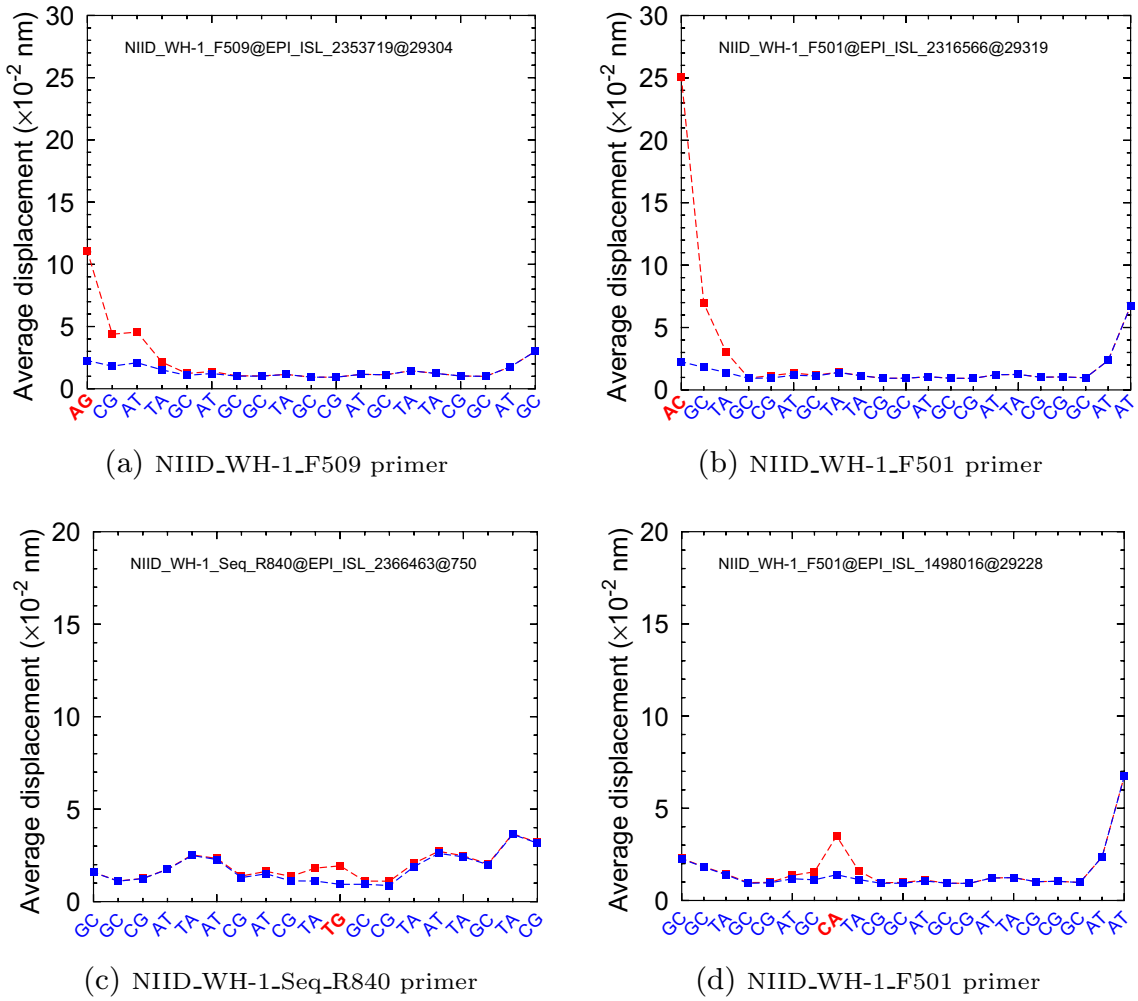
Here, we use $\Delta T_{\text{lim.}} = 5\ ^\circ\text{C}$, that is, we consider that primers with up to three consecutive mismatches with $T_{\text{MM}}$ no more than 5 °C below the reference temperature $T_{\text{ref.}}$ are acceptable.

**Availability.** The software packages used for this work are freely available and can be found in https://bioinf.fisica.ufmg.br/software/analyse_mismatch_primers.tar.gz.

## 3 Results and Discussion

Mismatches in primer-template duplex may avoid the amplification and turn the PCR reaction non-functional [4]. However, in some cases, mismatches may contribute to stabilize the duplex, even the hybridisation may be greater considering mismatches in comparison to AT-rich primers with no mismatches [6]. Mismatches in the direction of 3′ end are more detrimental to PCR reaction [3], yet at and near 3′ end they may prevent false priming [6].

In Fig. 1a, a single AG mismatch is shown located at 3′ end, which yields a small surrouding perturbation (red line). Its temperature $T_{\mathrm{MM}} = 55.3$ °C is




(a) NIID_WH-1_F509 primer (b) NIID_WH-1_F501 primer

(c) NIID_WH-1_Seq_R840 primer (d) NIID_WH-1_F501 primer

**Fig. 1.** Average displacement profiles. The blue line shows the displacement for full matched aligment and the red one for mismatched alignment. In each figure, the label shows the primer name, genome accession code and position, separated by the symbol @. (Color figure online)

out of the limit in relation to reference temperature $T_{\text{ref.}} = 63.6\ ^{\circ}\text{C}$. In contrast, Fig. 1b shows a single AC mismatch at $3'$ end, which yields a large end fraying and may impact in the DNA polymerase action, leading to a non-amplification. Nevertheless, its temperature $T_{\text{MM}} = 69.0\ ^{\circ}\text{C}$ is slightly lower than the reference temperature $T_{\text{ref.}} = 70.3\ ^{\circ}\text{C}$, which indicates a feasible effective hybridisation. In Figs. 1c and 1d, we show single mismatches in the middle of the alignment, TG and CA pairs, respectively. Both single mismatches display a small perturbation to duplex in comparison to matched reference (blue line). However, TG sequence hybridises at a temperature of $T_{\text{MM}} = 45.4\ ^{\circ}\text{C}$, considerable lower than its reference temperature $T_{\text{ref.}} = 60.2\ ^{\circ}\text{C}$, whereas CA sequence hybridises at $T_{\text{MM}} = 71.0\ ^{\circ}\text{C}$, which is slightly over to reference temperature $T_{\text{ref.}} = 70.3\ ^{\circ}\text{C}$. The latter shows a feasible contribution of a single mismatch, which could stabilise the primer-target duplex without impact in the amplification.

In Table 1, we show both strict and partial coverages for the four variants of concern. In a considerable number of cases, the coverage increases considering mismatches and, in special for Delta variant, it increases to 100%.

**Table 1.** Results for 21 PCR primers and probes from WHO [1] recommendation. Shown are the reference temperatures $T_{\text{ref.}}$ and the range of strict and partially coverages for Alpha, Beta, Gamma and Delta variants of concern (VOC) genomes, respectively.

| Primer/Probe | $T_{\text{ref.}}$ (°C) | Alpha variant | | Beta variant | | Gamma variant | | Delta variant | |
|---|---|---|---|---|---|---|---|---|---|
| | | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) | $C_{\text{strict}}$ (%) | $C_{\text{part.}}$ (%) |
| NIID_WH-1_F24381 | 61.2 | 99.1 | 99.1 | 98.3 | 98.3 | 98.9 | 98.9 | 99.8 | 99.8 |
| NIID_WH-1_F501 | 70.3 | 99.5 | 99.7 | 99.3 | 99.5 | 99.6 | 99.7 | 99.7 | 100 |
| NIID_WH-1_F509 | 63.3 | 99.5 | 99.5 | 99.0 | 99.1 | 99.3 | 99.3 | 99.3 | 99.3 |
| NIID_WH-1_R24873 | 61.5 | 99.7 | 99.7 | 99.3 | 99.3 | 99.4 | 99.4 | 100 | 100 |
| NIID_WH-1_R854 | 61.7 | 99.1 | 99.1 | 98.3 | 98.3 | 99.6 | 99.6 | 99.4 | 99.4 |
| NIID_WH-1_R913 | 69.2 | 99.3 | 99.8 | 98.1 | 98.6 | 99.7 | 99.9 | 99.8 | 100 |
| NIID_WH-1_Seq_F24383 | 60.4 | 99.1 | 99.1 | 98.3 | 98.3 | 98.9 | 98.9 | 99.8 | 99.8 |
| NIID_WH-1_Seq_F519 | 58.8 | 99.3 | 99.3 | 98.4 | 98.4 | 99.1 | 99.1 | 98.2 | 98.2 |
| NIID_WH-1_Seq_R24865 | 60.1 | 99.7 | 99.7 | 99.3 | 99.3 | 99.3 | 99.3 | 100 | 100 |
| NIID_WH-1_Seq_R840 | 60.2 | 98.9 | 98.9 | 98.1 | 98.1 | 99.6 | 99.6 | 99.5 | 99.5 |
| WH-NICN-F | 64.4 | 99.8 | 99.8 | 99.1 | 99.1 | 99.0 | 99.0 | 98.8 | 98.8 |
| WH-NICN-P | 51.3 | 99.9 | 99.9 | 98.6 | 98.6 | 99.2 | 99.2 | 99.9 | 99.9 |
| WH-NICN-R | 64.1 | 99.8 | 99.9 | 98.6 | 98.7 | 98.9 | 99.0 | 99.8 | 99.8 |
| WuhanCoV-spk1-f | 65.4 | 99.4 | 99.5 | 98.4 | 98.6 | 98.9 | 98.9 | 99.8 | 99.9 |
| WuhanCoV-spk2-r | 64.6 | 99.9 | 99.9 | 99.4 | 99.4 | 99.4 | 99.4 | 99.7 | 99.7 |
| nCoV_IP2-12669Fw | 54.3 | 99.8 | 99.8 | 99.3 | 99.3 | 99.2 | 99.2 | 100 | 100 |
| nCoV_IP2-12696bProbe | 67.0 | 99.7 | 99.8 | 99.4 | 99.4 | 99.9 | 99.9 | 98.9 | 100 |
| nCoV_IP2-12759Rv | 53.7 | 99.6 | 99.6 | 98.6 | 98.6 | 99.4 | 99.4 | 99.8 | 99.8 |
| nCoV_IP4-14059Fw | 54.8 | 99.9 | 99.9 | 99.5 | 99.5 | 100 | 100 | 100 | 100 |
| nCoV_IP4-14084Probe | 61.3 | 90.1 | 90.1 | 99.5 | 99.5 | 99.1 | 99.1 | 99.7 | 99.7 |
| nCoV_IP4-14146Rv | 54.8 | 99.5 | 99.5 | 99.5 | 99.5 | 99.8 | 99.8 | 98.0 | 98.0 |

## 4 Conclusion

We evaluated DNA mismatches in PCR-type primers/probes recommended by WHO to the detection of SARS-CoV-2 virus. We carried it out regarding the

variants of concern classified so far. The impact caused by mismatches are not straightforward and a full evaluation can be carried out with a detailed calculation and up-to-date model parameters. Nevertheless, we showed that these primers are able to align to VOCs genomes in a high coverage and it is feasible a contribution of mismatches to primer-target hybridisation.

## References

1. https://www.who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf?sfvrsn=3662fcb6_2
2. https://www.gisaid.org/hcov19-variants/
3. Bru, D., Martin-Laurent, F., Philippot, L.: Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16s RRNA gene as an example. Appl. Environ. Microbiol. **74**(5), 1660–1663 (2008). https://doi.org/10.1128/AEM.02403-07
4. Elaswad, A., Fawzy, M.: Mutations in animal SARS-CoV-2 induce mismatches with the diagnostic PCR assays. Pathogens **10**(3) (2021). https://doi.org/10.3390/pathogens10030371
5. Miranda, P., Weber, G.: Thermodynamic evaluation of the impact of DNA mismatches in PCR-type SARS-CoV-2 primers and probes. Mol. Cell. Probes **56**, 101707 (2021). https://doi.org/10.1016/j.mcp.2021.101707
6. Mitsuhashi, M.: Technical report: Part 1. basic requirements for designing optimal oligonucleotide probe sequences. J. Clin. Lab. Anal. **10**(5), 277–284 (1996)
7. Oliveira, L.M., Long, A.S., Brown, T., Fox, K.R., Weber, G.: Melting temperature measurement and mesoscopic evaluation of single, double and triple DNA mismatches. Chem. Sci. **11**, 8273–8287 (2020). https://doi.org/10.1039/d0sc01700k, https://pubs.rsc.org/en/content/articlelanding/2020/SC/D0SC01700K
8. Smith, T.F., Waterman, M.S., et al.: Identification of common molecular subsequences. J. Mol. Biol. **147**(1), 195–197 (1981)
9. Suo, T., et al.: ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens. Emerg. Microbes Infect. **9**(1), 1259–1268 (2020). https://doi.org/10.1080/22221751.2020.1772678, pMID: 32438868

**The Open COVID Journal**
*DOI: 10.2174/0126669587279780240130063422, 2024, 4, e26669587279780*
**ISSN: 2666-9587**
*1*

**RESEARCH ARTICLE** **OPEN ACCESS**

# *In silico* Thermodynamic Evaluation of the Effectiveness of RT-LAMP Primers for SARS-CoV-2 Variants Detection

Pâmella Miranda[1,2,*], Pedro A. Alves[3,4], Rubens L. do Monte-Neto[3] and Gerald Weber[1]

[1]*Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil*
[2]*Programa Interunidades de Pós-Graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil*
[3]*Instituto René Rachou – Fundação Oswaldo Cruz, Belo Horizonte-MG, Brazil*
[4]*Centro de Tecnologia em Vacinas, Belo Horizonte-MG, Brazil*

**Abstract:**

***Background:*** Viral mutations are the primary cause of mismatches in primer-target hybridisation, affecting the sensibility of molecular techniques, and potentially leading to detection dropouts. Despite its importance, little is known about the quantitative effect of mismatches in primer-target hybridisation. We have used up-to-date and highly detailed thermodynamic model parameters of DNA mismatches to evaluate the sensibility to variants of SARS-CoV-2 RT-LAMP primers.

***Methods:*** We aligned 18 RT-LAMP primer sets, which underwent clinical validation, to the genomes of the wild-type strain (ws), 7 variants and 4 subvariants, and calculated hybridisation temperatures allowing up to three consecutive mismatches. We calculated the coverage when the mismatched melting temperature fell by more than 5°C in comparison to the matched alignments. If no mismatches were considered, the average coverage found was 94% for ws, falling to the lowest value for Omicron, *i.e.*, 84%.

***Results:*** However, considering mismatches, the coverage was much higher, *i.e.*, 97% (ws) to 88% (Omicron). Stabilizing mismatches (higher melting temperatures) accounted for roughly 1/3 of this increase. The number of primer dropouts increased for each new variant; however, the effect was much less severe if mismatches were considered.

***Conclusion:*** We suggest using melting temperature calculations to continuously assess the trend of primer dropouts.

**Keywords:** DNA mismatches, Diagnosis, LAMP primer design, SARS-CoV-2, DNA thermodynamic models, Melting temperature calculations.



*Address correspondence to this author at the Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil; E-mail: pamella-fisica@ufmg.br

*Cite as:* Miranda P, Alves P, do Monte-Neto R, Weber G. *In silico* Thermodynamic Evaluation of the Effectiveness of RT-LAMP Primers for SARS-CoV-2 Variants Detection. , 2024; 4: e26669587279780.
http://dx.doi.org/10.2174/0126669587279780240130063422

Received: October 05, 2023
Revised: December 22, 2023
Accepted: January 16, 2024
Published: ?? ??, 2024



## 1. INTRODUCTION

There are eight possible mismatched (MM) base pairs in DNA: AA, AC, AG, CC, CT, GG, GT, and TT. They may arise from DNA replication [1], genetic recombination [2], and primer-template hybridisation in PCR reactions [3], which may lead to false-negative results [4]. Their presence may influence the stability and structural properties of DNA duplex, changing hydrogen bonds conformation and stacking interactions. However, some mismatches show a similar overall shape to a canonical

pair and a relatively stable configuration, *e.g.*, a GT-mismatched pair [5, 6]. MM pairs may be found in anti-syn or syn-anti conformations differently from DNA pairs, which are naturally in an anti-anti conformation. Mismatch impact varies from weakly bound (CC pair) to strongly bound (GG pair) in a local conformation, while molecular dynamics and NMR experiments have shown no impact on a global conformation, such as for AA and TT pairs [1, 7]. Internal and terminal mismatches influence primer-target hybridisation in different ways [8]. Mismatches located far from the 3' end have a moderate effect without influencing PCR performance [3]. On the other hand, those near the 3' terminal are critical and may lead to non-amplification of the target [8, 9]. Nevertheless, mismatches either near or at the 3' terminal may avoid false priming unlike the 5' terminal and internal mismatches [10]. Although it is known that mismatches typically destabilise the primer-target duplex, some types of mismatches are more stable than others, and some even more than AT base pairs, which may contribute towards the stability of the duplex [1, 11, 12]. A few mismatches in PCR primers may contribute to the design of antisense oligonucleotides [13], SNP [14], and allele-specific identification [12].

The isothermal PCR known as RT-LAMP (reverse transcription loop-mediated isothermal amplification) is a robust, fast, and inexpensive molecular technique, and can be carried out in less than an hour [15, 16]. It has been used as a molecular diagnostic test for several diseases, such as ebola [17], zika [18], HIV [19], SARS [20], MERS-CoV [21], and SARS-CoV-2 [22], the causative agent of COVID-19. To detect these diseases, it is necessary to design specific primers to identify the target agent. Unlike PCR, which usually uses a single pair of primers, RT-LAMP uses 2 or 3 pairs: F3 and B3 (outer primers), FIP and BIP (inner primers), and LF and LB (loop primers). The outer and inner primers act at the beginning of the reaction, but just the inner ones act in later cycles. FIP and BIP primers are long primers that contain two parts: F1c and F2 for FIP and B1c and B2 for BIP, which correspond to sense and antisense sequences of the target [23]. Finally, the loop primers are included to accelerate the reaction [24].

For LAMP, in the same way as for PCR [25], mismatches may appear between target and primers due to mutations potentially causing false-negative results [26, 27]. Yet, mismatches may enhance the technique's performance. In fact, SARS-CoV-2 PCR primers designed by Corman *et al*. [27] during the earlier stages of the pandemic had mismatches that did not hinder the detection of the coronavirus [28]. A few Cas12 enzymes in CRISPR assays have shown mismatch tolerance [29], and resistant mutants may be detected after an antibiotic administration when a mismatch is incorporated at the 3' terminal [30]. Recent PCR-based methods have used mismatches either at or near the 3' terminal to detect Delta variant [31] and Omicron subvariants [32].

In a previous study [33], we evaluated the impact of mismatches on RT-PCR primers and probes, where we showed that the mismatches do not always have a negative impact on thermodynamic stability. The reason for this is that there are a number of mismatch configurations that may actually increase the melting temperatures. This was confirmed recently by Scapaticci *et al*. [34], who found that mutations may have higher melting temperatures and suggested that the melting temperature analysis could be used to detect specific variants. As for PCR, it is expected that mismatches in primer-target hybridisation may appear for LAMP primers, especially for both FIP and BIP primers in which mismatched base pairs in either 5' or 3' ends may prevent the elongation by Bst DNA polymerase, leading to a low amplification efficiency [35]. Although one or two mismatches have been shown to be tolerable for LAMP [30, 36], studies with three or more consecutive mismatches are, to our knowledge, not available.

Here, we have shown the evaluation of DNA mismatches in 18 RT-LAMP primer sets [37-54], which were designed for wild-type SARS-CoV-2 genomes. One of those sets [38] was previously successfully evaluated by us for a few variants and now for amplified genome sets. We applied a previous workflow [33] to analyse those primers for the detection of SARS-CoV-2 variants as Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Lambda (C.37), Mu (B.1.621), and Omicron (B.1.1.529) variants, and BA.2 to BA.5 subvariants. The outcomes show if those primers may still be effective in detecting the variants and how the presence of mismatches may contribute to covering more genomes, consequently detecting the coronavirus. Furthermore, we reinforce the fact that a continuous evaluation of RT-LAMP primer sets is needed to cover variants that may arise, as already suggested [48].

## 2. MATERIALS AND METHODS

### 2.1. Genome Sets

We randomly collected 21665 genomes of original SARS-CoV-2 (wild type strain) on 8$^{th}$ October, 2020, at NCBI [55]; 7247 genomes of the Alpha variant, 7497 of the Beta variant, and 2308 of the Gamma variant on 7$^{th}$ April, 2021; 7943 of Delta variant on 5$^{th}$ June, 2021; 7029 of Omicron variant on 16$^{th}$ December, 2021; 6610 of Mu variant, 9340 of Lambda variant, and 7393 and 348 of Omicron subvariants BA.2 and BA.3 on 11$^{th}$ February 2022; and 629 and 1231 of Omicron subvariants BA.4 and BA.5 on 19$^{th}$ September 2022, at GISAID [56].

### 2.2. Primer Sets

We collected 18 different RT-LAMP primer sets designed for SARS-CoV-2 original genomes that underwent clinical validation [37-54], resulting in a total of 436 primers. Their details are shown in Table **S1**. FIP and BIP primers were divided in F1c/F2 and B1c/B2 primers, respectively, except those from three sets [38, 40, 51], for which the division of primers was already given. We found all possible combinations of primer pairs and selected those according to the temperatures of the same type of pair from the three sets just mentioned.

### 2.3. Evaluation Workflow

All primers were aligned to each genome set using a Smith-Waterman algorithm, as described earlier [33]. Fully matched alignments were called strictly matched and those with single, double, and triple consecutive mismatches were termed partially matched. Alignments with four or more consecutive mismatches were considered as not aligned. The limit of the maximal number of consecutive mismatches is be due to the fact that the available parameter only covers up to three contiguous mismatches [11]. In addition, it is very likely that four or more mismatches will destabilize the primers far beyond the limits considered here. Also, deletions in the viral genome, as in the Omicron variant [57, 58], may lead to no alignment of the primers.

Hybridisation temperatures for matched ($T_{ref.}$) and mismatched ($T_{MM}$) alignments were calculated from a mesoscopic model with the parameters obtained from a previous work [11].

$$T_m = a_0 + a_1 \tau, \quad (1)$$

The reference hybridisation temperature $T_{ref.}$ for each primer is shown in Supplementary Table **S1**. It should be noted that the parameters [11] are for a sodium buffer, which is different from those typically used in PCR reactions that contain $Mg^+$. Therefore, the absolute temperatures $T_{ref.}$ may be different from the actual melting temperatures of the primers. However, since our analysis deals with temperature differences, which are not strongly buffer-dependent, we expect them to be sufficiently accurate for our purposes.

We define a strictly matched (AT and CG only) alignment coverage for each primer as follows:

$$C_{strict} = \frac{N_G - N_{n.a.} - N_{MM}}{N_G} \quad (2)$$

Where, $N_G$ is the total number of genomes, $N_{n.a.}$ is the number of genomes for which no alignment was found, and $N_{MM}$ is the number of genomes for which a partial alignment containing mismatches was found.

The difference between reference hybridisation temperature $T_{ref.}$ and mismatched alignments $T_{MM}$ is defined as follows:

$$\Delta T_{MM} = T_{ref.} - T_{MM} \quad (3)$$

Where, $T_{MM}$ is usually lower than $T_{ref.}$ [11]. The partial coverage for alignments with up to three contiguous mismatches is defined below:

$$C_{part.}(\Delta T_{lim.}) = \frac{N_G - N_{n.a.} - N_{low}(T_{MM} < T_{ref.} - \Delta T_{lim.})}{N_G} \quad (4)$$

Where, $N_{low}$ is the number of alignments where the mismatched melting temperature $T_{MM}$ is lower by $\Delta T_{lim.}$ than the reference $T_{ref.}$. It should be noted that as there are many mismatch configurations that have an increased $T_{MM}$, that is, there are situations where $C_{part.} > C_{strict}$ even for $\Delta T_{lim.} = 0$. A previous work has provided additional details of this workflow [33].

All 18 primer sets were aligned against the genomes of SARS-CoV-2 variants. We calculated the hybridisation temperatures and coverages for both matched and mismatched alignments considering single, double, and triple consecutive mismatches. The complete evaluation was carried out in approximately 120 h computing time.

### 2.4. Availability

The software packages used to carry out this work are freely available and can be found at https://bioinf.fisica.ufmg.br/software/, in the *analyse primer lamp.tar.gz* package.

## 3. RESULTS AND DISCUSSION

We assessed 18 clinical validated RT-LAMP primer sets [37-54]. They showed high strict and partial coverages for wild-type SARS-CoV-2 and its variants. Even for variants and subvariants, a few primers achieved more than 90% coverage. A considerable number of primers showed high coverages only when mismatches were taken into account. Also, primers utilized by Alves *et al*. [38] achieved high coverages for wild-type SARS-CoV-2 and Gamma variant, in agreement with the experimental results. Furthermore, Almeida *et al*. [59] showed E1 and N2 primer subsets to be able to identify the Omicron variant target despite the presence of only a single mismatch. All strict and partial coverages are shown in Tables **S2**-**S55**.

Given the continuous mutation of the SARS-CoV-2 genomes, it is expected that over time, mismatches should increasingly occur within the primer regions. Fig. (**1**), where we show the $C_{strict}$ averaged over all 436 primers, illustrates this decreasing coverage as variants appear. In comparison to the wild-type strain (ws) coverage, all variants decreased their coverage. When we considered partial coverages in the presence of mismatches with $\Delta T_{lim} = 0°C$, that is, primers with $T_{MM} \geq T_{ref.}$, the curve uniformly shifted upwards. For $\Delta T_{lim} = 5°C$, Beta, Gamma, Delta, and Mu partial coverage became slightly higher than the ws strict coverage. However, the rate of decrease was not uniform, and some variants had higher coverage than their presumed predecessor variants. For the Omicron variant, which had a larger number of mutations [60], we have observed a sharp drop in the coverage. However, for the subsequent subvariant, the picture has been mixed; the BA.3 subvariant shared the low coverage, but BA.2, BA.4, and BA.5 have shown a higher coverage. The reason for this oscillation was not clear.

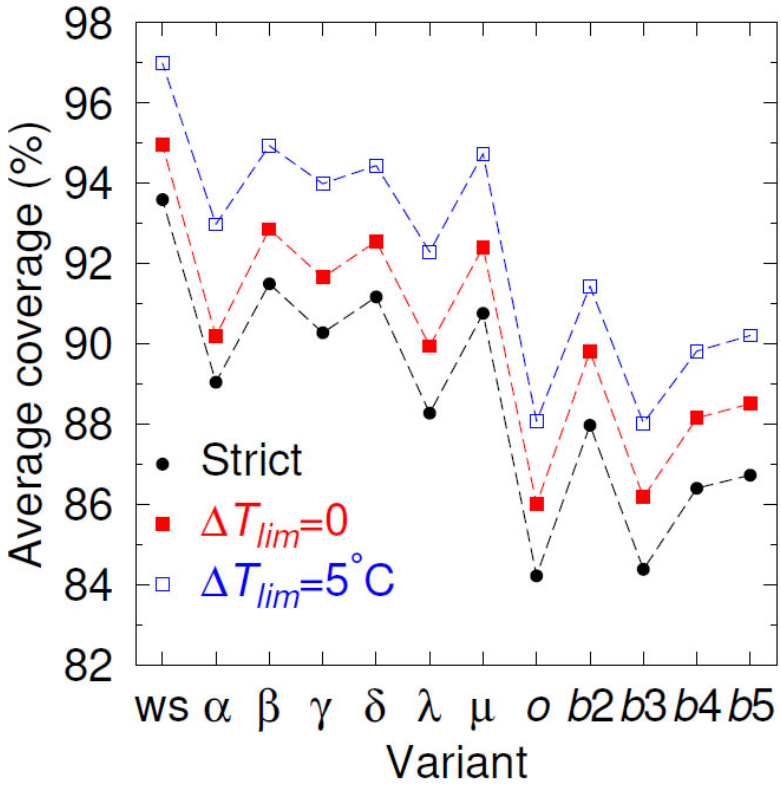



**Fig. (1).** Coverage averaged over all primers as for wild-type SARS-CoV-2 genomes (ws), Alpha (α), Beta (β), Gamma (γ), Delta (δ), Lambda (λ), Mu (μ), and Omicron (*o*) variants, and BA.2 (b2), BA.3 (b3), BA.4 (b4) and BA.5 (b5) subvariants. Black bullets are for $C_{\text{strict}}$ and red (blue) boxes are for $\Delta T_{\text{lim}}$ = 0°C (5°C). The dashed line connecting the data point is only intended as a guide to the eye.

**Table 1. Examples of primer coverages with stabilizing mismatches, $T_{\text{lim.}}$ = 0°C, which have $C_{\text{part.}}$(0°C) > 90% while having $C_{\text{strict}}$ = 0. Only those primers having stabilizing mismatches for the given variant are shown.**

| | $C_{\text{part.}}$(0°C) (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Primer** | **ws** | **Alpha** | **Beta** | **Gamma** | **Delta** | **Lambda** | **Mu** | **Omicron** | **BA.2** | **BA.3** | **BA.4** | **BA.5** |
| As1e F1c [28] | 99.2 | 99.4 | 98.7 | 99.7 | 99.2 | 99.0 | 99.7 | 98.5 | 100 | 98.6 | 99.5 | 99.7 |
| iLACO-F1c [28] | 99.2 | - | 99.1 | 99.5 | 99.3 | 99.8 | 99.8 | 99.5 | 99.8 | 99.7 | 100 | 99.9 |
| N15-B1c [33] | 97.7 | 98.8 | 96.5 | 98.4 | 99.9 | 99.4 | 99.6 | 99.2 | 99.9 | 99.7 | 98.9 | 98.9 |
| N1-B1c [35] | 98.4 | 99.5 | 97.1 | 99.6 | 99.6 | 99.7 | 99.3 | 93.8 | 99.8 | 94.3 | 99.0 | 98.7 |
| N1-F1c [35] | - | - | - | - | - | - | - | 92.7 | 97.5 | 94.0 | - | - |
| N2-F1c [35] | 99.1 | 99.1 | 96.6 | 99.1 | 99.9 | 98.7 | 99.7 | 99.3 | 99.9 | 99.7 | 99.0 | 99.0 |
| NEB orf1a-A-F1c [37] | 98.9 | 99.5 | 99.2 | 99.5 | 99.8 | 99.6 | 99.1 | 99.4 | 99.7 | 99.1 | 99.2 | 96.9 |
| F1c [38] | - | - | - | - | - | 98.9 | 99.2 | 93.9 | 99.5 | 93.1 | - | - |

Similar to what we have seen for RT-PCR [61], many alignments that would result in a null strict coverage achieve partial coverage beyond 99% if mismatches are considered. In some cases, a large partial coverage is already obtained for $\Delta T_{\text{lim}}$ = 0°C, that is, if we consider only mismatches that do not destabilize the duplex.

In Table 1 we show a few examples of primers that have zero strict coverage but go beyond 90% if stabilizing mismatches are considered. It is somewhat surprising that some primers achieved high coverages for Omicron only and not for the other variants, despite the fact that all were designed for the wild-type strain. While this seems to be an opposite trend to the overall decline for Omicron, one should note that a higher coverage for Omicron was rather exceptional and only occurred for very few primers. On the other hand, this quite clearly highlights that the assessment of mismatch influence is far from trivial.

**Table 2. Sets having at least one potential drop-out primer for any of the variants. Only the reference number is shown for each set. Drop-out primers are considered as those with a partial coverage ($\Delta T_{lim.}$ = 5°C) below 5%, $N_{drop}$, for at least one variant. $N_{primers}$ is the number of separate primers for each set.**

| Ref. | $N_{primers}$ | $N_{drop}$ | ws | Alpha | Beta | Gamma | Delta | Lambda | Mu | Omicron | BA.2 | BA.3 | BA.4 | BA.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [29] | 32 | 6 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 3 | 3 | 3 | 4 | 3 |
| [30] | 40 | 7 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 3 | 3 | 3 | 3 | 3 |
| [31] | 81 | 28 | 4 | 11 | 5 | 5 | 5 | 6 | 12 | 9 | 5 | 10 | 7 | 6 |
| [32] | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| [33] | 32 | 6 | 0 | 0 | 0 | 1 | 0 | 3 | 1 | 3 | 2 | 3 | 2 | 2 |
| [34] | 15 | 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 2 |
| [35] | 32 | 5 | 1 | 3 | 1 | 3 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 2 |
| [36] | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| [37] | 32 | 5 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 2 | 2 | 4 | 2 |
| [38] | 10 | 6 | 0 | 2 | 1 | 0 | 2 | 3 | 0 | 3 | 3 | 3 | 3 | 3 |
| [39] | 48 | 12 | 0 | 3 | 0 | 5 | 3 | 6 | 0 | 7 | 7 | 7 | 8 | 7 |
| [40] | 8 | 5 | 0 | 2 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| [41] | 16 | 5 | 0 | 1 | 2 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| [42] | 8 | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| [43] | 15 | 3 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| [44] | 15 | 2 | 0 | 1 | 0 | 1 | 1 | 2 | 0 | 1 | 1 | 1 | 1 | 1 |
| [45] | 16 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

**Table 3. Primers obtained from the work of Ji *et al*. [44] with single, double, and triple contiguous mismatches.**

| Primers | Mismatches | $T_{ref.}$ (°C) | $T_{MM}$ (°C) |
|---|---|---|---|
| N1-B1c | GT | 73.9 | 80.6 |
| N2-B1c | TT/CT | 69.0 | 66.6 |
| ORFlab-1-F1c | GT/TT | 71.0 | 70.1 |
| N2-F1c | GT/CT/TT | 75.6 | 79.9 |

While most primers had large coverages, an important amount of primers failed to achieve significant coverage for at least one variant, and may represent a potential dropout. A summary of the amount of primers that could potentially represent dropouts is provided in Table **2**. Here, we have considered a very stringent threshold of 5% at $\Delta T_{lim}$ = 5°C, that is, primers where even considering a maximal 5°C melting temperature below the reference temperature covered less than 5% of the available genomes for a given variant. Only the set proposed by Alekseenko *et al*. [37] had no potential dropout primers at all. The complete list of potential dropout primers for each variant is shown in Tables **S56**-**S67**.

Mismatched pairs in 5' or 3' terminals of FIP and BIP primers may hamper the amplification by Bst DNA polymerase. However, we found a few alignments with either 5' and 3' terminal mismatches to have a hybridisation temperature within the threshold and contribute to the increase in the coverage when mismatches are taken into account. Clearly, in some cases, mismatches in both terminals reduced the temperature. An interesting case was found for four primers from the work of Ji *et al*. [44], which showed single, double, and triple contiguous mismatches at the 3’ terminal (Table **3**). We observed that only the double mismatched pair cases decreased the temperature. On the other hand, the single and triple mismatched pairs increased the temperature. Perhaps, due to the GT mismatched pair has been reported as a strong pair [2, 11, 62, 63]. FIP and BIP primers with terminal mismatches that had an increase in their coverage are shown in Tables **S68**-**S79** for each genome set. It should be noted that FIP and BIP primers were divided into F1c/F2 and B1c/B2, respectively, and as such treated individually. With respect to the LAMP technique, the F1c and B1c depend on their respective F2 and B2 complements, and the dropout may in practice be higher.

## CONCLUSION

In this work, we have evaluated the coverage of 18 RT-LAMP primer sets considering single, double, and triple mismatches in primer-target hybridisation to SARS-CoV-2 variants. In general, the average coverage of these primer sets decreased for the new variants, when compared to the wild-type strain. Overall, the coverage was lowest for the Omicron and BA.3 variants. However, a clear monotonic decrease in the coverage was not observed; instead, for some variants, the coverage increased when compared to its putative predecessor, as exemplified most notably by the Mu variant, which showed one of the highest coverages. Coverage uniformly increased if

mismatches were taken into account, while not enough to completely compensate for the loss in comparison to the wild-type strain, as is shifted the worst case from 84% to 88%. Similarly, the number of potential dropout primers increased with each new variant, and only one out of 18 sets showed no potential primer drop-out. We suggest the use of the methodology described here to continuously evaluate the effectiveness of RT-LAMP primer as new variants emerge. Furthermore, our method can be applied to the detection of other infectious diseases.

## LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| α | = | Alpha |
| β | = | Beta |
| γ | = | Gamma |
| δ | = | Delta |
| λ | = | Lambda |
| μ | = | Mu |
| *o* | = | Omicron |
| b2 | = | BA.2 |
| b3 | = | BA.3 |
| b4 | = | BA.4 |
| b5 | = | BA.5 |

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

## HUMAN AND ANIMAL RIGHTS

No humans/animals were used for studies that are the basis of this research.

## CONSENT FOR PUBLICATION

Not applicable.

## AVAILABILITY OF DATA AND MATERIALS

The data and supportive information are available within the article.

## FUNDING

PM is supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes/Ação Emergencial, Brazil, Finance Code 001). PA is supported by the Brazilian Ministry of Science, Technology and Innovation, through the “Rede Virus” (MCTI, FINEP grant number 01.2N5.00). RMN is supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig grant number PPM-00699-18). RMN and GW are research fellows from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq grant numbers RMN 312965/2020-6, GW 307538/2019-2).

## CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

## ACKNOWLEDGEMENTS

Declared none.

## SUPPLEMENTARY MATERIAL

Table **S1** shows all primers used and their reference temperatures. Tables **S2**-**S55** show the strict and partial coverages of all primers. Tables **S56**-**S67** show the potential drop-out primers for each variant. Tables **S68**-**S79** show the terminal mismatched positions of FIP and BIP primers for each genome.

## REFERENCES

[1] Rossetti G, Dans PD, Gomez-Pinto I, Ivani I, Gonzalez C, Orozco M. The structural impact of DNA mismatches. Nucleic Acids Res 2015; 43(8): 4309-21. http://dx.doi.org/10.1093/nar/gkv254 PMID: 25820425

[2] Ke S-H, Wartell RM. Influence of nearest neighbor sequence on the stability of base pair mismatches in long DNA: determination by temperature-gradient gel electrophoresis. Nucleic Acids Res 1993; 21(22): 5137-43.

[3] Khan KA, Cheung P. Presence of mismatches between diagnostic PCR assays and coronavirus SARS-CoV-2 genome. Royal Society Publishing 2020.

[4] Kuchinski KS, Jassem AN, Prystajecky NA. Assessing oligonucleotide designs from early lab developed PCR diagnostic tests for SARS-CoV-2 using the PCR strainer pipeline. J Clin Virol 2020; 131: 104581.

[5] Allison L A. Fundamental Molecular Biology. Blackwell Publishing Ltd. 2007.

[6] Johnson SJ, Beese LS. Structures of mismatch replication errors observed in a DNA polymerase. Cell 2004; 116(6): 803-16. http://dx.doi.org/10.1016/S0092-8674(04)00252-1 PMID: 15035983

[7] Gervais V, Cognet JAH, Bret M, Sowers LC, Fazakerley GV. Solution structure of two mismatches A.A and T.T in the K-ras gene context by nuclear magnetic resonance and molecular dynamics. Eur J Biochem 1995; 228(2): 279-90. http://dx.doi.org/10.1111/j.1432-1033.1995.00279.x PMID: 7705340

[8] Hendling M, Barišić I. In-silico design of DNA oligonucleotides: Challenges and approaches. Comput Struct Biotechnol J 2019; 17: 1056-65. http://dx.doi.org/10.1016/j.csbj.2019.07.008 PMID: 31452858

[9] Elaswad A, Fawzy M. Mutations in animal SARS-CoV-2 induce mismatches with the diagnostic PCR assays. Pathogens 2021; 10(3): 371. http://dx.doi.org/10.3390/pathogens10030371 PMID: 33808783

[10] Mitsuhashi M. Technical report: Part 1. Basic requirements for designing optimal oligonucleotide probe sequences. J Clin Lab Anal 1996; 10(5): 277-84. http://dx.doi.org/10.1002/(SICI)1098-2825(1996)10:5<277::AID-JCLA8>3.0.CO;2-5 PMID: 8887007

[11] Oliveira LM, Long AS, Brown T, Fox KR, Weber G. Melting temperature measurement and mesoscopic evaluation of single, double and triple DNA mismatches. Chem Sci 2020; 11(31): 8273-87. http://dx.doi.org/10.1039/D0SC01700K PMID: 34094181

[12] Stadhouders R, Pas SD, Anber J, Voermans J, Mes THM, Schutten M. The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5′ nuclease assay. J Mol Diagn 2010; 12(1): 109-17. http://dx.doi.org/10.2353/jmoldx.2010.090035 PMID: 19948821

[13] Sugimoto N, Nakano M, Nakano S. Thermodynamics-structure relationship of single mismatches in RNA/DNA duplexes. Biochemistry 2000; 39(37): 11270-81. http://dx.doi.org/10.1021/bi000819p PMID: 10985772

[14] Gibson NJ. The use of real-time PCR methods in DNA sequence variation analysis. Clin Chim Acta 2006; 363(1-2): 32-47. http://dx.doi.org/10.1016/j.cccn.2005.06.022 PMID: 16182268
[15] Becherer L, Borst N, Bakheit M, Frischmann S, Zengerle R, von Stetten F. Loop-mediated isothermal amplification (LAMP) – review and classification of methods for sequence-specific detection. Anal Methods 2020; 12(6): 717-46. http://dx.doi.org/10.1039/C9AY02246E
[16] Park G-S, Ku K, Baek S-H, *et al*. Development of reverse transcription loop-mediated isothermal amplification (RT-LAMP) assays targeting SARS-CoV-2. J Mol Diagn 2020; 22: 729-35. http://dx.doi.org/10.1016/j.jmoldx.2020.03.006 PMID: 32276051
[17] Oloniniyi OK, Kurosaki Y, Miyamoto H, Takada A, Yasuda J. Rapid detection of all known ebolavirus species by reverse transcription-loop-mediated isothermal amplification (RT-LAMP). J Virol Methods 2017; 246: 8-14. http://dx.doi.org/10.1016/j.jviromet.2017.03.011 PMID: 28356221
[18] Calvert AE. Rapid colorimetric detection of Zika virus from serum and urine specimens by reverse transcription loop-mediated isothermal amplification (RT-LAMP). PLoS One 2017; 12: 1-16.
[19] Curtis KA, Rudolph DL, Owen SM. Rapid detection of HIV-1 by reverse-transcription, loop-mediated isothermal ampli- fication (RT-LAMP). J Virol Methods 2008; 151: 264-70.
[20] Hong TC, Mai QL, Cuong DV, *et al*. Development and evaluation of a novel loop-mediated isothermal amplification method for rapid detection of severe acute respiratory syndrome coronavirus. J Clin Microbiol 2004; 42(5): 1956-61. http://dx.doi.org/10.1128/JCM.42.5.1956\OT1\textendash1961.2004 PMID: 15131154
[21] Huang P, Wang H, Cao Z, *et al*. A rapid and specific assay for the detection of MERS-CoV. Front Microbiol 2018; 9: 1101. http://dx.doi.org/10.3389/fmicb.2018.01101 PMID: 29896174
[22] da Costa VD, Santos AC, da Silva LL, *et al*. RT-LAMP multicenter study for SARS-CoV-2 genome molecular detection in brazilian swab and saliva samples. Diagnostics 2023; 13(2): 210. http://dx.doi.org/10.3390/diagnostics13020210 PMID: 36673025
[23] Notomi T, Okayama H, Masubuchi H, *et al*. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res 2000; 28(12): E63. PMID: 10871386
[24] Nagamine K, Hase T, Notomi T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. Mol Cell Probes 2002; 16(3): 223-9. http://dx.doi.org/10. 1006/mcpr.2002.0415 PMID: 12144774
[25] Laine PN, Mustanoja E, Lyyski A, *et al*. SARS-CoV-2 variant with mutations in n gene affecting detection by widely used PCR primers. J Med Virol 2021; 94: 1227-31. http://dx.doi.org/10.1002/jmv.27418
[26] Lippi G, Simundic A-M, Plebani M. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). Clin Chem Lab Med 2020; 58: 1070-6. http://dx.doi.org/10.1515/cclm-2020-0285
[27] Corman VM, Landt O, Kaiser M, *et al*. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance 2020; 25: 2000045.
[28] Bustin S, Kirvell S, Huggett JF, Nolan T. RT-qPCR diagnostics: The "Drosten" SARS-CoV-2 assay paradigm. Int J Mol Sci 2021; 22(16): 8702. http://dx.doi.org/10.3390/ijms22168702 PMID: 34445406
[29] Ooi KH, Liu MM, Tay JWD, *et al*. An engineered CRISPR-Cas12a variant and DNA-RNA hybrid guides enable robust and rapid COVID-19 testing. Nat Commun 2021; 12(1): 1739. http://dx.doi.org/10.1038/s41467-021-21996-6 PMID: 33741959
[30] Wang D. Effect of internal primer–template mismatches on loop-mediated isothermal amplification. Biotechnol Biotechnol Equip 2016; 30(2): 314-8.
[31] Garson JA, Badru S, Parker E, Tedder RS, McClure MO. Highly sensitive and specific detection of the SARS-CoV-2 Delta variant by double-mismatch allele-specific real time reverse transcription PCR. J Clin Virol 2022; 146: 105049. http://dx.doi.org/10.1016/j.jcv.2021.105049 PMID: 34871906
[32] Bozidis P, Petridi E, Gartzonika K. An ARMS-multiplex PCR targeting SARS-CoV-2 omicron sub-variants. Pathogens 2023; 12(8): 1017. http://dx.doi.org/10.3390/pathogens12081017 PMID: 37623977
[33] Miranda P, Weber G. Thermodynamic evaluation of the impact of DNA mismatches in PCR-type SARS-CoV-2 primers and probes. Mol Cell Probes 2021; 56: 101707. http://dx.doi.org/10.1016/j.mcp.2021.101707 PMID: 33609730
[34] Scapaticci M, Bartolini A, Vitone F, *et al*. Detection of a characteristic melting profile of a SARS-CoV-2 Kappa variant in Italy using the SARS-CoV-2 Variants ELITe MGB® Kit. J Virol Methods 2022; 301: 114458. http://dx.doi.org/10.1016/j.jviromet.2022.114458 PMID: 35026304
[35] Zhou Y, Wan Z, Yang S, *et al*. A mismatch-tolerant reverse transcription loop-mediated isothermal amplification method and its application on simultaneous detection of all four serotype of dengue viruses. Front Microbiol 2019.
[36] Peyrefitte C N, Boubis L, Coudrier D, *et al*. Real-time reverse-transcription loop-mediated isothermal amplification for rapid detection of rift valley fever virus. J Clin Microbiol 2008; 46: 3653-9.
[37] Alekseenko A, Barrett D, Pareja-Sanchez Y, *et al*. Direct detection of SARS-CoV-2 using non-commercial RT-LAMP reagents on heat-inactivated samples. Sci Rep 2021; 11(1): 1820. http://dx.doi.org/10.1038/s41598-020-80352-8 PMID: 33469065
[38] Alves PA, de Oliveira EG, Franco-Luiz APM, *et al*. Optimization and clinical validation of colorimetric reverse transcription loop-mediated isothermal amplification, a fast, highly sensitive and specific COVID- 19 molecular diagnostic tool that is robust to detect SARS-CoV-2 variants of concern. Front Microbiol 2021; 12: 713713. http://dx.doi.org/10.3389/fmicb.2021.713713 PMID: 34867841
[39] García-Bernalt Diego J, Fernández-Soto P, Domínguez-Gil M, Belhassen-García M, Bellido JLM, Muro A. A simple, affordable, rapid, stabilized, colorimetric, versatile RT-LAMP assay to detect SARS-CoV-2. Diagnostics 2021; 11(3): 438. http://dx.doi.org/10.3390/diagnostics11030438
[40] Ganguli A, Mostafa A, Berger J, *et al*. Rapid isothermal amplification and portable detection system for SARS-CoV-2. bioRxiv 2020; 2020.05.21.108381. http://dx.doi.org/10.1101/2020.05.21.108381
[41] Garcia-Venzor A, Rueda-Zarazua B, Marquez-Garcia E, *et al*. SARS-CoV-2 direct detection without RNA isolation with loop-mediated isothermal amplification (LAMP) and CRISPR-Cas12. Front Med 2021; 8: 627679. http://dx.doi.org/10.3389/fmed.2021.627679 PMID: 33681254
[42] Huang WE, Lim B, Hsu C-C, *et al*. RT-LAMP for rapid diagnosis of coronavirus SARS-CoV-2. Microb Biotechnol 2020; 13(4): 950-61.
[43] Jang WS, Lim DH, Yoon J, *et al*. Development of a multiplex loop-mediated isothermal amplification (LAMP) assay for on-site diagnosis of SARS CoV-2. PLoS One 2021; 16: 1-14. http://dx.doi.org/10.1371/journal.pone.0248042
[44] Ji C, Xue S, Yu M, *et al*. Rapid detection of SARS-CoV-2 virus using dual reverse transcriptional colorimetric loop-mediated isothermal amplification. ACS Omega 2021; 6(13): 8837-49.
[45] Jiang M, Pan W, Arasthfer A, *et al*. C. Lass-Fl o¨rl, Development and validation of a rapid, single-step reverse transcriptase loop-mediated isothermal amplification (RT-LAMP) system potentially to be used for reliable and high-throughput screening of COVID-19. Front Cell Infect Microbiol 2020; 10: 331. http://dx.doi.org/10.3389/fcimb.2020.00331 PMID: 32626666
[46] Lalli MA, Langmade JS, Chen X, *et al*. Rapid and extraction-free detection of SARS-CoV-2 from saliva by colorimetric reverse-transcription loop-mediated isothermal amplification. Clin Chem 2020; 67(2): 415-24.
[47] Lau YL, Ismail IB, Izati Binti Mustapa N, *et al*. A sensitive reverse transcription loop-mediated isothermal amplification assay for direct visual detection of SARS-CoV-2. Am J Trop Med Hyg 2020;

103(6): 2350-2.
http://dx.doi.org/10.4269/ajtmh.20-1079 PMID: 33098286

[48] Luo Z, Ye C, Xiao H, Yin J, Liang Y, Ruan Z. Optimization of loop-mediated isothermal amplification (LAMP) assay for robust visualization in SARS-CoV-2 and emerging variants diagnosis. Chem Eng Sci 2022; 251: 117430.

[49] Mautner L, Baillie CK, Herold HM, *et al*. Rapid point-of-care detection of SARS-CoV-2 using reverse transcription loop-mediated isothermal amplification (RT-LAMP). Virol J 2020; 17(1): 160.
http://dx.doi.org/10.1186/s12985-020-01435-6 PMID: 33087160

[50] Mohon AN, Oberding L, Hundt J, *et al*. Optimization and clinical validation of dual-target RT-LAMP for SARS-CoV-2. J Virol Methods 2020; 286: 113972.
http://dx.doi.org/10.1016/j.jviromet.2020.113972 PMID: 32941977

[51] Rodriguez-Manzano J, Malpartida-Cardenas K, Moser N, *et al*. Handheld point-of-care system for rapid detection of SARS-CoV-2 extracted RNA in under 20 min. ACS Cent Sci 2021; 7(2): 307-17.
http://dx.doi.org/10.1101/2020.06.29.20142349

[52] Yan C, Cui J, Huang L, *et al*. Rapid and visual detection of 2019 novel coronavirus (SARS-CoV-2) by a reverse transcription loop-mediated isothermal amplification assay. Clin Microbiol Infect 2020; 26: 773-9.
http://dx.doi.org/10.1016/j.cmi.2020.04.001

[53] Yang Q, Meyerson NR, Clark SK, *et al*. Saliva TwoStep for rapid detection of asymptomatic SARS-CoV-2 carriers. eLife 2021.

[54] Yoshikawa R, Abe H, Igasaki Y, Negishi S, Goto H, Yasuda J. Development and evaluation of a rapid and simple diagnostic assay for COVID-19 based on loop-mediated isothermal amplification. PLOS Negl Trop Dis 2020; 14: e0008855.

[55] Severe acute respiratory syndrome coronavirus 2. 2020. Available from: https://www.ncbi.nlm.nih.gov/SARS-CoV-2

[56] Map of tracked variant occurrence. 2021. Available from: https://www.gisaid.org/hcov19-variants/

[57] Sharma D, Ye C, Lippi G, *et al*. *In silico* evaluation of the impact of the omicron variant on the sensitivity of RT-qPCR assays for SARS-CoV-2 detection using whole genome sequencing Research Square 2022.
http://dx.doi.org/10.21203/rs.3.rs-1220446/v1

[58] Sharma D, Notarte KI, Fernandez RA, Lippi G, Gromiha MM, Henry BM. *In silico* evaluation of the impact of omicron variant of concern sublineage BA.4 and BA.5 on the sensitivity of RT-qPCR assays for SARS-CoV-2 detection using whole genome sequencing. J Med Virol 2022; 95(1): e28241.

[59] Almeida LT, Gonçalves AB, Franco-Luiz APM, Silva TBS, Alves PA, Monte-Neto RL. Molecular detection of omicron SARS-CoV-2 variant is achieved by RT-LAMP despite genomic mutations. Mem Inst Oswaldo Cruz 2022; 117: e220050.
http://dx.doi.org/10.1590/0074-02760220050 PMID: 35766650

[60] Metzger CMJA, Lienhard R, Seth-Smith HMB, *et al*. PCR performance in the SARS-CoV-2 Omicron variant of concern? Swiss Med Wkly 2021; 151(4950): w30120.
http://dx.doi.org/10.4414/SMW.2021.w30120 PMID: 34909869

[61] Miranda P, Barbosa VB, Weber G. Mesoscopic evaluation of DNA mismatches in PCR primer-target hybridisation to detect SARS-CoV-2 variants of concern Brazilian Symposium on Bioinformatics. 145-50.
http://dx.doi.org/10.1007/978-3-030-91814-9_15

[62] Peyret N, Seneviratne PA, Allawi HT, SantaLucia J Jr. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G, and T·T mismatches. Biochemistry 1999; 38(12): 3468-77.
http://dx.doi.org/10.1021/bi9825091 PMID: 10090733

[63] Aboul-ela F, Koh D, Tinoco I Jr, Martin FH. Base-base mismatches. Thermodynamics of double helix formation for dCA $_3$ XA $_3$ G + dCT $_3$ YT $_3$ G (X, Y = A,C,G,D. Nucleic Acids Res 1985; 13(13): 4811-24.
http://dx.doi.org/10.1093/nar/13.13.4811 PMID: 4022774