www.revistaft.com.br

Revista FT | ISSN 1678-0817 | v. 30, n. 164, 2026 | DOI: 10.69849/gqgs5a61

Identificação *in silico* de mutações das variantes de preocupação do SARS-CoV-2 em sítios de anelamento e redesenho dos *primers* e sondas de RT-qPCR afetados

In silico identification of SARS-CoV-2 variant-of-concern mutations in annealing sites and redesign of the affected RT-qPCR primers and probes

Gabriel Amorim Veloso[1]
Luis Augusto Peruzzo[2]
Samuel Bastos dos Santos[3]
Matheus Martins Daude[4]

RESUMO

A COVID-19, causada pelo SARS-CoV-2, apresenta elevada variabilidade genética, e mutações nas regiões reconhecidas por *primers* e sondas podem comprometer a confiabilidade dos testes de RT-qPCR. Este estudo avaliou o impacto das mutações das variantes de preocupação sobre os sítios de anelamento de ensaios diagnósticos e propôs conjuntos corrigidos para os ensaios afetados. Trata-se de estudo descritivo-analítico de abordagem *in silico*. Seis genomas completos, a referência Wuhan-Hu-1 e um representante de cada variante de preocupação (Alpha, Beta, Gamma, Delta e Omicron), foram obtidos no GenBank e alinhados com Clustal Omega. Quarenta oligonucleotídeos provenientes de cinco protocolos (CDC, Charité/OMS, China CDC, HKU e Institut Pasteur) foram localizados no genoma de referência. Após a consolidação de sequências idênticas entre protocolos, analisaram-se 37 oligonucleotídeos únicos organizados em 12 conjuntos. Trinta permaneceram íntegros nas cinco variantes e sete apresentaram ao menos um *mismatch*, totalizando 18 eventos, todos não sinônimos, distribuídos em 11 posições genômicas. O conjunto N do China CDC concentrou 12 desses eventos por sobrepor o agrupamento 28.881–28.883, origem das substituições R203K e G204R. Um evento foi classificado como de criticidade alta: a substituição na posição 15.451 (ORF1b:G662S, variante Delta), situada a uma base da extremidade 3′ do *primer* senso do conjunto RdRP_SARSr. Os novos conjuntos foram desenhados para os cinco ensaios afetados, restritos às 28.825 posições (96,4% do genoma) conservadas nas seis sequências, apresentando pareamento integral e parâmetros termodinâmicos adequados. A verificação de especificidade frente aos seis coronavírus humanos de referência não indicou reatividade esperada com MERS-CoV nem com os coronavírus sazonais, embora os conjuntos dirigidos a ORF1b e ao gene E mantenham o reconhecimento do SARS-CoV, propriedade herdada do desenho pan-*Sarbecovirus* original. Conclui-se que a conservação global do genoma não assegura a integridade dos sítios diagnósticos e que o monitoramento dessas regiões deve integrar a vigilância genômica de rotina.

Palavras-chave: COVID-19; Diagnóstico molecular; Mutações; *Primers*; RT-qPCR.

1 AFYA – Porto Nacional – Tocantins – Brasil.
2 AFYA – Porto Nacional – Tocantins – Brasil.
3 AFYA – Porto Nacional – Tocantins – Brasil.
4 AFYA – Porto Nacional – Tocantins – Brasil.

ABSTRACT

COVID-19, caused by SARS-CoV-2, exhibits high genetic variability, and mutations in regions recognized by primers and probes may compromise the reliability of RT-qPCR tests. This study assessed the impact of variant-of-concern mutations on the annealing sites of diagnostic assays and proposed corrected sets for the affected assays. This is a descriptive-analytical *in silico* study. Six complete genomes, the Wuhan-Hu-1 reference and one representative of each variant of concern (Alpha, Beta, Gamma, Delta and Omicron), were retrieved from GenBank and aligned with Clustal Omega. Forty oligonucleotides from five protocols (CDC, Charité/WHO, China CDC, HKU and Institut Pasteur) were mapped onto the reference genome. After consolidating sequences identical across protocols, 37 unique oligonucleotides organised into 12 sets were analysed. Thirty remained intact across the five variants and seven displayed at least one mismatch, totalling 18 events, all non-synonymous, distributed over 11 genomic positions. The China CDC N set concentrated 12 of these events by overlapping the 28,881–28,883 cluster, the origin of the R203K and G204R substitutions. One event was classified as highly critical: the substitution at position 15,451 (ORF1b:G662S, Delta variant), located one base from the 3′ end of the forward primer of the RdRP_SARSr set. New sets were designed for the five affected assays, restricted to the 28,825 positions (96.4% of the genome) conserved across all six sequences, showing complete complementarity and adequate thermodynamic parameters. Specificity screening against the six reference human coronaviruses indicated no expected reactivity with MERS-CoV or with the seasonal coronaviruses, although the sets targeting ORF1b and the E gene retain recognition of SARS-CoV, a property inherited from the original pan-*Sarbecovirus* design. Global genome conservation does not ensure the integrity of diagnostic sites, and monitoring these regions should be integrated into routine genomic surveillance.

Keywords: COVID-19; Molecular diagnosis; Mutations; Primers; RT-qPCR.

1 INTRODUÇÃO

A COVID-19, causada pelo coronavírus SARS-CoV-2, é uma infecção respiratória aguda de elevada transmissibilidade e alcance mundial. Identificado pela primeira vez em dezembro de 2019 na cidade de Wuhan, província de Hubei, China, o SARS-CoV-2 foi descrito como o sétimo coronavírus conhecido capaz de infectar seres humanos (Zhu *et al.*, 2020). Em 11 de março de 2020 a Organização Mundial da Saúde caracterizou a COVID-19 como pandemia, em razão de sua rápida disseminação geográfica (WHO, 2020). No Brasil, mais de 39 milhões de pessoas foram infectadas e aproximadamente 716 mil óbitos foram atribuídos a essa síndrome respiratória (Brasil, 2025). O impacto estendeu-se para além da esfera sanitária, afetando a organização dos serviços de saúde e o próprio trabalho em vigilância sanitária (Araújo *et al.*, 2023), e permanece pauta de pesquisa prioritária no país (Alves *et al.*, 2025).

Diante desse cenário, a detecção precoce da infecção tornou-se fundamental para o controle da pandemia. Os métodos de diagnóstico da COVID-19 abrangem testes moleculares, testes de antígeno e testes sorológicos, cada qual com indicações e limitações próprias, e a

Reação em Cadeia da Polimerase quantitativa com Transcrição Reversa (RT-qPCR) permanece o método molecular de escolha para o diagnóstico etiológico da infecção por SARS-CoV-2 (Tang *et al.*, 2020). A efetividade relativa das ferramentas de diagnóstico e manejo, incluindo os métodos de imagem, foi revisada por Alsharif e Qurashi (2021), e plataformas emergentes, como os nanobiossensores, vêm sendo investigadas como alternativas de resposta rápida (Mohammadi *et al.*, 2024). A sensibilidade da RT-qPCR, contudo, não é absoluta. Revisão sistemática com metanálise de 32 estudos, envolvendo mais de 18 mil pacientes, estimou taxa agregada de resultados falso-negativos de 12% no teste inicial (IC 95% de 10% a 14%, com certeza da evidência classificada como muito baixa), proporção que pode ser substancialmente maior conforme a prevalência da infecção na população testada (Pecoraro *et al.*, 2022). Essa limitação confere relevância prática à investigação dos fatores capazes de reduzir o desempenho do ensaio.

O RT-qPCR baseia-se na conversão do RNA viral em DNA complementar, seguida de amplificação cíclica catalisada por DNA polimerase (Mullis; Faloona, 1987) e monitorada em tempo real pela fluorescência proporcional ao produto formado (Kubista *et al*., 2006). A especificidade do ensaio decorre dos iniciadores (*primer*s), oligonucleotídeos curtos complementares a regiões definidas do genoma viral, e, nos formatos com sonda de hidrólise, de um terceiro oligonucleotídeo marcado que confere detecção específica do amplicon (Corman *et al*., 2020; Chu *et al*., 2020). O desempenho do método depende, portanto, da complementaridade entre esses oligonucleotídeos e a sequência-alvo.

Diversos fatores podem comprometer essa complementaridade ou a reação como um todo. Na etapa pré-analítica, o método de extração do RNA influencia a recuperação do material genético e, consequentemente, a sensibilidade do ensaio (Campos *et al.*, 2021), e a presença de inibidores de amplificação, provenientes da própria amostra clínica ou do processamento, pode reduzir a eficiência da reação e favorecer resultados falso-negativos (Fruehwirth *et al.*, 2020). Na etapa analítica, soma-se a esses fatores a variabilidade genética do próprio vírus: alterações nas regiões reconhecidas pelos oligonucleotídeos modificam as bases disponíveis para o pareamento e reduzem a estabilidade do híbrido formado.

O surgimento sucessivo de variantes de preocupação, classificadas e monitoradas pela Organização Mundial da Saúde segundo critérios de transmissibilidade, gravidade e escape imunológico (WHO, 2023), tornou essa possibilidade concreta. Diante disso, o presente trabalho teve por objetivo analisar, por meio de ferramentas de bioinformática, as regiões-alvo utilizadas em ensaios de RT-qPCR para detecção do SARS-CoV-2, identificando mutações capazes de comprometer a eficiência diagnóstica, caracterizando seu efeito sobre a proteína

codificada e sobre o anelamento, e propondo conjuntos *primer*/sonda corrigidos para os ensaios afetados.

2 MATERIAL E MÉTODOS

2.1 Desenho do estudo

Trata-se de estudo descritivo-analítico de abordagem *in silico* e caráter metodológico-comparativo, conduzido inteiramente em ambiente computacional a partir de dados genômicos de domínio público. A investigação compreendeu quatro etapas encadeadas: localização dos oligonucleotídeos diagnósticos no genoma de referência do SARS-CoV-2, identificação das divergências entre esses sítios e os genomas das variantes de preocupação, predição do efeito dessas divergências sobre a proteína codificada e sobre a amplificação, e proposição de novos conjuntos para os ensaios comprometidos. Por não envolver seres humanos, material biológico ou dados identificáveis, o estudo dispensa apreciação por Comitê de Ética em Pesquisa.

2.2 Obtenção das sequências genômicas

Foram selecionados seis genomas completos de SARS-CoV-2 depositados no *GenBank* (*National Center for Biotechnology Information*), compreendendo a cepa de referência Wuhan-Hu-1 e um representante de cada uma das cinco variantes de preocupação designadas pela Organização Mundial da Saúde (WHO, 2023), a saber Alpha, Beta, Gamma, Delta e Omicron (Quadro 1). Essa designação corresponde ao estatuto atribuído a cada variante durante o período em que circulou, classificação posteriormente revista pela própria instituição, que passou a listar parte delas como variantes de preocupação anteriormente circulantes. Adotaram-se como critérios de inclusão: genoma completo ou quase completo, atribuição explícita de linhagem Pango, disponibilidade de metadados de data e local de coleta, e citação do número de acesso em artigo revisado por pares. Foram excluídas sequências fragmentadas, com indícios de erro de montagem ou com elevada proporção de bases indeterminadas. O genoma NC_045512.2 (RefSeq) foi adotado como referência para todas as coordenadas apresentadas neste trabalho.

Quadro 1 – Genomas de SARS-CoV-2 analisados. As posições cobertas referem-se às colunas do alinhamento efetivamente sequenciadas em cada genoma, excluídas as lacunas terminais. As bases ambíguas correspondem a posições indeterminadas (N) no depósito original.

Variante (OMS)	Linhagem Pango	Acesso GenBank	Procedência	Data da coleta	Posições cobertas	Bases ambíguas (N)
Referência	B	NC_045512.2	China (Wuhan)	dez/2019	1–29916	0
Alpha	B.1.1.7	MW487270.1	EUA (origem: Reino Unido)	dez/2020–jan/2021	52–29795	1
Beta	B.1.351	MZ443817.1	Malásia (origem: África do Sul)	2021	55–29849	612
Gamma	P.1	MW642248.1	Itália (origem: Brasil, Manaus)	jan/2021	31–29916	2
Delta	B.1.617.2	MZ208926.1	Marrocos (origem: Índia)	abr/2021	121–29860	0
Omicron	BA.1 (B.1.1.529)	OL672836.1	Bélgica (origem: África do Sul/Egito)	nov/2021	55–29781	0

2.3 Conjuntos de *primer*s e sondas avaliados

Foram compiladas as sequências oligonucleotídicas de cinco protocolos de RT-qPCR para detecção do SARS-CoV-2, extraídas dos documentos primários de cada instituição: o painel dos *Centers for Disease Control and Prevention* dos Estados Unidos (conjuntos N1, N2 e N3), o protocolo Charité/Organização Mundial da Saúde descrito por Corman *et al*. (2020), o protocolo do Centro Chinês de Controle e Prevenção de Doenças, o protocolo da Universidade de Hong Kong descrito por Chu *et al*. (2020) e o protocolo do Institut Pasteur de Paris. Os oligonucleotídeos dirigidos ao controle interno humano de RNase P foram excluídos por não se dirigirem ao genoma viral. Conjuntos cujas sequências se mostraram integralmente idênticas entre protocolos distintos foram consolidados em registro único, mantendo-se a indicação de ambas as fontes. A relação completa dos oligonucleotídeos analisados, com sequência e posição genômica, consta da Tabela 1.

2.4 Alinhamento múltiplo

Os seis genomas foram alinhados simultaneamente com o algoritmo Clustal Omega, executado no serviço web do European Bioinformatics Institute com parâmetros padrão. O alinhamento resultante contém 29.916 colunas, das quais 29.903 correspondem a posições do genoma de referência. Para cada genoma foram determinadas as lacunas terminais, decorrentes de regiões não sequenciadas e distintas de deleções verdadeiras, bem como as posições com bases ambíguas. Ambas as situações foram tratadas como dados ausentes e excluídas das comparações.

2.5 Localização dos oligonucleotídeos no genoma de referência

Cada oligonucleotídeo foi localizado no genoma de referência por varredura exaustiva de todas as posições possíveis, nas duas fitas, comparando-se tanto a sequência do oligonucleotídeo quanto o seu complemento reverso com a sequência genômica. A comparação respeitou o código de degenerescência IUPAC, de modo que uma base degenerada no oligonucleotídeo foi considerada compatível com qualquer base genômica contida em seu conjunto de possibilidades. Admitiu-se tolerância máxima de três divergências por oligonucleotídeo, a fim de permitir a localização de oligonucleotídeos pan-*Sarbecovirus*, originalmente concebidos para detectar também o SARS-CoV e, por isso, não integralmente complementares à referência. Para cada oligonucleotídeo foi retido o sítio de menor número de divergências, registrando-se a fita, as coordenadas inicial e final e o tamanho. As coordenadas obtidas foram confrontadas com as descritas nos protocolos de origem.

As coordenadas do genoma de referência foram convertidas em colunas do alinhamento, de modo que a comparação entre cada oligonucleotídeo e os genomas das variantes se desse posição a posição, sem necessidade de novo alinhamento local. Para oligonucleotídeos antissenso, as bases genômicas foram convertidas ao complemento reverso, de modo que todas as posições sejam reportadas na orientação $5' \rightarrow 3'$ do próprio oligonucleotídeo.

2.6 Identificação e classificação das divergências

Cada posição de cada oligonucleotídeo foi classificada em uma de cinco categorias: pareamento correto, divergência por substituição, divergência por deleção, base ambígua, quando o genoma apresenta base indeterminada na posição, e dado ausente, quando a posição recai em região não sequenciada daquele genoma. As duas últimas categorias foram tratadas como inconclusivas e não contabilizadas como divergência, evitando que falhas de cobertura fossem interpretadas como mutação.

As divergências já presentes no confronto entre o oligonucleotídeo e a própria sequência de referência foram registradas separadamente, sob a designação de divergências de desenho, por decorrerem da concepção original do oligonucleotídeo e não da evolução viral. Somente as divergências ausentes na referência e presentes em uma variante foram computadas como eventos de *mismatch* atribuíveis àquela variante.

2.7 Anotação gênica e predição do efeito sobre a proteína

As coordenadas das regiões codificantes foram obtidas da anotação RefSeq (NC_045512.2) e validadas computacionalmente: para cada uma das doze fases de leitura

verificou-se o comprimento múltiplo de três, a presença de códon de iniciação ATG e a presença de códon de terminação. O deslocamento de fase de leitura ribossomal de −1, que ocorre na posição 13.468 e separa as poliproteínas ORF1a e ORF1b, foi tratado explicitamente, adotando-se a numeração de resíduos própria de cada poliproteína. Os domínios de proteínas não estruturais (nsp) em ORF1ab foram igualmente anotados.

A predição do efeito de cada divergência sobre a proteína codificada foi efetuada pela leitura do códon completo da variante no alinhamento, e não pela substituição isolada de uma base no códon de referência. Esse procedimento é necessário porque mutações adjacentes dentro de um mesmo códon produzem, em conjunto, alteração de aminoácido distinta daquela que resultaria de cada substituição isolada. Códons contendo bases ambíguas, lacunas ou posições não sequenciadas foram considerados indetermináveis. O quadro de leitura empregado foi aferido contra posições de efeito conhecido e amplamente descrito na literatura, a saber 14.408 (ORF1b:P314), 23.403 (S:D614), 28.881 (N:R203), 26.270 (E:T9) e 3.037 (ORF1a:F924).

2.8 Classificação da criticidade

Cada evento de *mismatch* recebeu classificação de criticidade segundo critério posicional previamente definido. Atribuiu-se criticidade alta às divergências situadas nas cinco bases da extremidade 3′ de um *primer*, região em que o pareamento correto é requisito para o início da extensão pela DNA polimerase. Receberam criticidade moderada as deleções e as divergências localizadas no terço central de sondas, e criticidade baixa as demais posições. Por tratar-se de critério aplicado a cada evento isoladamente, registrou-se adicionalmente, para cada oligonucleotídeo e variante, o número total de divergências acumuladas, uma vez que divergências múltiplas reduzem a estabilidade do híbrido independentemente de sua distância à extremidade 3′ (Lefever *et al.*, 2013).

2.9 Análise de conservação e redesenho dos conjuntos

Para o redesenho dos conjuntos comprometidos, construiu-se previamente um mapa de conservação do genoma, no qual cada posição da referência foi classificada como conservada apenas quando as seis sequências apresentaram base idêntica, sem lacunas, sem bases ambíguas e sem recair em região não sequenciada de qualquer genoma. Esse critério conservador exclui do desenho toda posição cuja identidade não pôde ser confirmada nos seis genomas.

Os novos oligonucleotídeos foram desenhados com o programa Primer3, por meio da biblioteca primer3-py, restringindo-se a busca a janelas de 420 pares de base a montante e a

jusante de cada amplicon original e declarando-se como regiões excluídas, tanto para *primers* quanto para sondas, todas as posições não conservadas. Os parâmetros adotados foram os seguintes. Os *primers* foram restritos a 18 a 25 nucleotídeos, com temperatura de *melting* (Tm) entre 57 e 63 °C e conteúdo de guanina e citosina entre 30 e 65%. As sondas foram restritas a 18 a 32 nucleotídeos, com Tm alvo superior à dos *primers*. Os amplicons foram limitados a 70 a 150 pares de base, faixa ampliada progressivamente até 250 pares de base quando nenhum candidato era obtido, admitindo-se no máximo quatro nucleotídeos idênticos consecutivos. O cálculo termodinâmico adotou 50 mM de cátion monovalente, 3 mM de magnésio, 0,8 mM de desoxirribonucleotídeos e 250 nM de oligonucleotídeo. Para cada conjunto, a Tm mínima admitida para a sonda foi reduzida progressivamente a partir de 69 °C até que se obtivesse candidato válido, procedimento que assegura a maior Tm compatível com cada região-alvo.

2.10 Verificação *in silico* dos conjuntos propostos

Cada oligonucleotídeo proposto foi submetido a quatro verificações. A primeira consistiu na conferência do pareamento posição a posição contra os seis genomas alinhados, de modo a confirmar complementaridade integral. A segunda consistiu na busca de sítios secundários de anelamento no próprio genoma de referência, por varredura das duas fitas admitindo-se até três divergências, a fim de descartar a possibilidade de amplificação inespecífica intragenômica. A terceira compreendeu a avaliação das estruturas secundárias previstas, a saber grampo (*hairpin*), autodímero (*self-dimer*) e dímero entre os *primers* do par (*heterodimer*), expressas como Tm da estrutura e consideradas aceitáveis quando inferiores à temperatura de anelamento prevista para o ensaio. A quarta consistiu na avaliação da especificidade frente a outros coronavírus humanos, mediante varredura dos genomas de referência do SARS-CoV (NC_004718.3), do MERS-CoV (NC_019843.3) e dos coronavírus sazonais 229E (NC_002645.1), OC43 (NC_006213.1), NL63 (NC_005831.2) e HKU1 (NC_006577.2), obtidos do GenBank. Para cada oligonucleotídeo registrou-se, nas duas fitas, o menor número de divergências e a posição do melhor sítio, verificando-se em seguida se os *primers* de cada conjunto formavam geometria compatível com amplicon, isto é, *primer* senso na fita positiva a montante do *primer* antissenso na fita negativa, com produto inferior a 600 pares de base. O procedimento foi validado contra o próprio genoma de referência do SARS-CoV-2, no qual os quinze oligonucleotídeos propostos apresentaram pareamento integral.

2.11 Ferramentas computacionais

Todas as análises foram implementadas em Python 3.11, empregando-se as bibliotecas primer3-py 2.3.1 para o desenho de oligonucleotídeos e o cálculo termodinâmico, openpyxl 3.1.5 para a leitura e a geração das planilhas e matplotlib 3.10 para a construção das figuras. A localização dos oligonucleotídeos, a classificação das divergências, a anotação gênica e a predição do efeito proteico foram implementadas em rotinas próprias, reunidas em módulo único compartilhado por todas as figuras e tabelas, de modo que os valores apresentados nas diferentes peças derivem de uma única execução da análise.

3 RESULTADOS E DISCUSSÃO

3.1 Ensaios avaliados e distribuição dos alvos no genoma

Foram avaliados 40 oligonucleotídeos dirigidos ao genoma do SARS-CoV-2, provenientes de cinco protocolos de RT-qPCR. Os três oligonucleotídeos do controle interno de RNase P do painel do CDC foram excluídos por se dirigirem ao genoma humano. Os conjuntos para o gene E descritos pela Charité e pelo Institut Pasteur mostraram-se idênticos em sequência. Consolidada essa duplicação, restaram 37 oligonucleotídeos únicos, organizados em 12 conjuntos *primer*/sonda.

Todos os 37 oligonucleotídeos foram localizados no genoma de referência NC_045512.2 (29.903 nt) mediante busca com tratamento das bases degeneradas IUPAC, e as coordenadas obtidas coincidiram com as descritas nos protocolos de origem (Tabela 1).

Tabela 1 – Conjuntos de *primers* e sondas de RT-qPCR avaliados, posição no genoma de referência NC_045512.2 e número de *mismatches* por variante. F indica *primer* senso, R *primer* antissenso e S sonda. Os valores por variante correspondem ao número de posições divergentes em relação à referência, desconsiderados os *mismatches* já presentes no próprio desenho do oligonucleotídeo.

Protocolo	Conjunto	Gene	Oligo	Sequência (5′→3′)	Posição	Alpha	Beta	Gamma	Delta	Omicron
CDC	2019-nCoV_N1	N	F	GACCCCAAAATCAGCGAAAT	28287–28306	0	1	0	0	0
CDC	2019-nCoV_N1	N	S	ACCCCGCATTACGTTTGGTGGACC	28309–28332	0	0	0	0	1
CDC	2019-nCoV_N1	N	R	TCTGGTTACTGCCAGTTGAATCTG	28335–28358	0	0	0	0	0
CDC	2019-nCoV_N3*	N	F	GGGAGCCTTGAATACACCAAAA	28681–28702	0	0	0	0	0
CDC	2019-nCoV_N3*	N	S	AYCACATTGGCACCCGCAATCCTG	28704–28727	0	0	0	0	0
CDC	2019-nCoV_N3*	N	R	TGTAGCACGATTGCAGCATTG	28732–28752	0	0	0	0	0
CDC	2019-nCoV_N2	N	F	TTACAAACATTGGCCGCAAA	29164–29183	0	0	0	0	0
CDC	2019-nCoV_N2	N	S	ACAATTTGCCCCCAGCGCTTCAG	29188–29210	0	0	0	0	0
CDC	2019-nCoV_N2	N	R	GCGCGACATTCCGAAGAA	29213–29230	0	0	0	0	0
Charité/OMS	RdRP_SARSr	ORF1b (nsp12)	F	GTGARATGGTCATGTGTGGCGG	15431–15452	0	0	0	1	0
Charité/OMS	RdRP_SARSr	ORF1b (nsp12)	S P2	CAGGTGGAACCTCATCAGGAGATGC	15470–15494	0	0	0	0	0
Charité/OMS	RdRP_SARSr	ORF1b (nsp12)	S P1	CCAGGTGGWACRTCATCMGGTGATGC	15469–15494	0	0	0	0	0
Charité/OMS	RdRP_SARSr	ORF1b (nsp12)	R	CARATGTTAAASACACTATTAGCATA	15505–15530	0	0	0	0	0
Charité/OMS e Inst. Pasteur	E_Sarbeco	E	F	ACAGGTACGTTAATAGTTAATAGCGT	26269–26294	0	0	0	0	1
Charité/OMS e Inst. Pasteur	E_Sarbeco	E	S	ACACTAGCCATCCTTACTGCGCTTCG	26332–26357	0	0	0	0	0
Charité/OMS e Inst. Pasteur	E_Sarbeco	E	R	ATATTGCAGCAGTACGCACACA	26360–26381	0	0	0	0	0
Charité/OMS	N_Sarbeco	N	F	CACATTGGCACCCGCAATC	28706–28724	0	0	0	0	0

Protocolo	Conjunto	Gene	Oligo	Sequência (5′→3′)	Posição	Alpha	Beta	Gamma	Delta	Omicron
Charité/OMS	N_Sarbeco	N	S	ACTTCCTCAAGGAACAACATTGCCA	28753–28777	0	0	0	0	0
Charité/OMS	N_Sarbeco	N	R	GAGGAACGAGAAGAGGCTTG	28814–28833	0	0	0	0	0
China CDC	ORF1ab	ORF1ab (nsp10/nsp12)	F	CCCTGTGGGTTTTACACTTAA	13342–13362	0	0	0	0	0
China CDC	ORF1ab	ORF1ab (nsp10/nsp12)	S	CCGTCTGCGGTATGTGGAAAGGTTATGG	13377–13404	0	0	0	0	0
China CDC	ORF1ab	ORF1ab (nsp10/nsp12)	R	ACGATTGTGCATCAGCTGA	13442–13460	0	0	0	0	0
China CDC	N	N	F	GGGGAACTTCTCCTGCTAGAAT	28881–28902	3	1	3	1	3
China CDC	N	N	S	TTGCTGCTGCTTGACAGATT	28934–28953	0	0	0	0	0
China CDC	N	N	R	CAGACATTTTGCTCTCAAGCTG	28958–28979	1	0	0	0	0
HKU	HKU-ORF1b-nsp14	ORF1b (nsp14)	F	TGGGGYTTTACRGGTAACCT	18778–18797	0	0	0	0	0
HKU	HKU-ORF1b-nsp14	ORF1b (nsp14)	S	TAGTTGTGATGCWATCATGACTAG	18849–18872	0	0	0	0	0
HKU	HKU-ORF1b-nsp14	ORF1b (nsp14)	R	AACRCGCTTAACAAAGCACTC	18889–18909	0	0	0	0	0
HKU	HKU-N	N	F	TAATCAGACAAGGAACTGATTA	29145–29166	0	0	0	0	0
HKU	HKU-N	N	S	GCAAATTGTGCAATTTGCGG	29177–29196	0	0	0	0	0
HKU	HKU-N	N	R	CGAAGGTGTGACTTCCATG	29236–29254	0	0	0	0	0
Inst. Pasteur	IP2	ORF1a (nsp9)	F	ATGAGCTTAGTCCTGTTG	12690–12707	0	0	0	0	0
Inst. Pasteur	IP2	ORF1a (nsp9)	S	AGATGTCTTGTGCTGCCGGTA	12717–12737	0	0	0	0	0
Inst. Pasteur	IP2	ORF1a (nsp9)	R	CTCCCTTTGTTGTGTTGT	12780–12797	0	0	0	0	0

Protocolo	Conjunto	Gene	Oligo	Sequência (5′→3′)	Posição	Alpha	Beta	Gamma	Delta	Omicron
Inst. Pasteur	IP4	ORF1b (nsp12)	F	GGTAACTGGTATGATTTCG	14080–14098	0	0	0	0	0
Inst. Pasteur	IP4	ORF1b (nsp12)	S	TCATACAAACCACGCCAGG	14105–14123	1	0	0	1	0
Inst. Pasteur	IP4	ORF1b (nsp12)	R	CTGGTCAAGGTTAATATAGG	14167–14186	0	0	0	0	0

Os alvos distribuem-se em três regiões: seis conjuntos no gene N, cinco em ORF1ab, sendo um em nsp9, um na transição nsp10/nsp12, dois em nsp12 (RNA-polimerase dependente de RNA) e um em nsp14, e um no gene E. Os *amplicons* variam de 67 a 132 pb (Figura 1a).

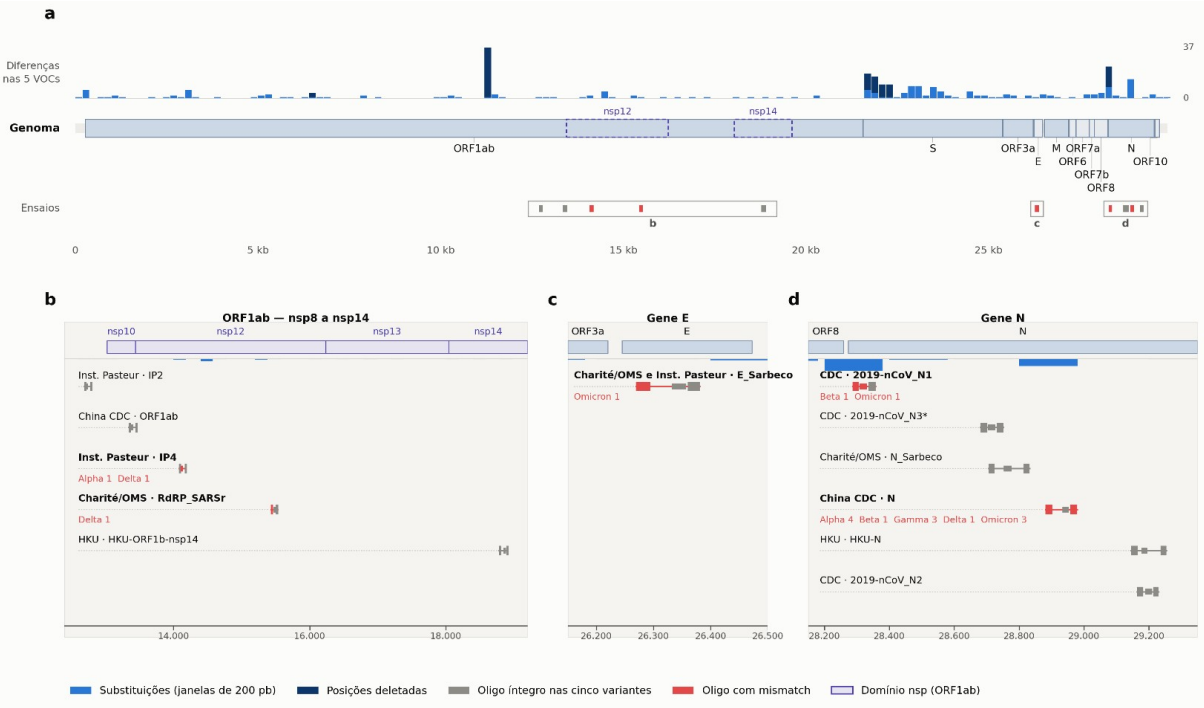


Figura 1 – Localização dos ensaios de RT-qPCR no genoma do SARS-CoV-2 e distribuição das diferenças genômicas nas variantes de preocupação. (a) Genoma completo da referência NC_045512.2 (Wuhan-Hu-1, 29.903 nt), com os genes anotados, as diferenças acumuladas das cinco variantes (Alpha, Beta, Gamma, Delta e Omicron), tanto substituições, agrupadas em janelas de 200 pb, quanto posições deletadas, e a posição dos doze conjuntos *primer*/sonda analisados. Os retângulos tracejados marcam os domínios nsp12 e nsp14 da ORF1ab, e os retângulos na faixa dos ensaios indicam as regiões ampliadas em (b), (c) e (d). (b) a (d) Detalhe das três regiões que concentram os alvos diagnósticos. Cada linha corresponde a um conjunto, com os blocos maiores representando os *primers* e o bloco menor, a sonda. Conjuntos em negrito e vermelho apresentam ao menos um *mismatch*, e os valores abaixo do nome indicam o número de posições divergentes em cada variante. Os protocolos Charité/OMS e Institut Pasteur empregam o mesmo conjunto para o gene E, representado em linha única. O conjunto 2019-nCoV_N3 (*) integrava a versão original do painel do CDC e foi posteriormente retirado do protocolo, sendo mantido na análise para fins de comparação.

A distribuição das diferenças em relação à referência, calculada em janelas de 200 pb, mostrou-se heterogênea ao longo do genoma. O maior valor (37 posições) situa-se em 11.201–11.400, em ORF1ab, e resulta quase integralmente de uma deleção de nove nucleotídeos em 11.288–11.296, no domínio nsp6, compartilhada por Alpha, Beta, Gamma e Omicron. Apenas uma substituição ocorre nessa janela. Considerando exclusivamente substituições, o valor máximo desloca-se para a janela 28.801–29.000, no gene N, com 14 substituições, que é a mesma região ocupada pelo conjunto do China CDC (Figura 1d).

A concentração de divergências nessa janela não é um acidente estatístico. A deleção de nove nucleotídeos em nsp6, que responde por quase toda a diferença observada, mostrou-se discriminante o bastante entre linhagens para ser convertida em alvo diagnóstico. Vogels *et al.* (2021) adotaram a deleção ORF1a Δ3675-3677 como alvo primário de um ensaio multiplex de detecção de variantes de preocupação, no qual a própria falha de anelamento da sonda sinaliza a presença da deleção. Convém registrar, porém, que a identidade do evento não é a mesma em todas as linhagens. Zhou *et al.* (2023) descrevem ΔSGF, correspondente aos resíduos 106 a 108 de nsp6, em Alpha, Beta e Gamma, e mostram que a linhagem BA.1 de Omicron carrega majoritariamente ΔLSG, nos resíduos 105 a 107, com ΔSGF voltando a predominar a partir de BA.2. Como as duas deleções recobrem extensões quase coincidentes, as coordenadas aqui reportadas devem ser lidas como a posição atribuída pelo alinhamento, e não como a delimitação do códon removido em cada linhagem.

3.2 Mutações nos sítios de anelamento

Dos 37 oligonucleotídeos avaliados, 30 permaneceram íntegros nas cinco variantes e 7 apresentaram ao menos um *mismatch* (Figura 2).

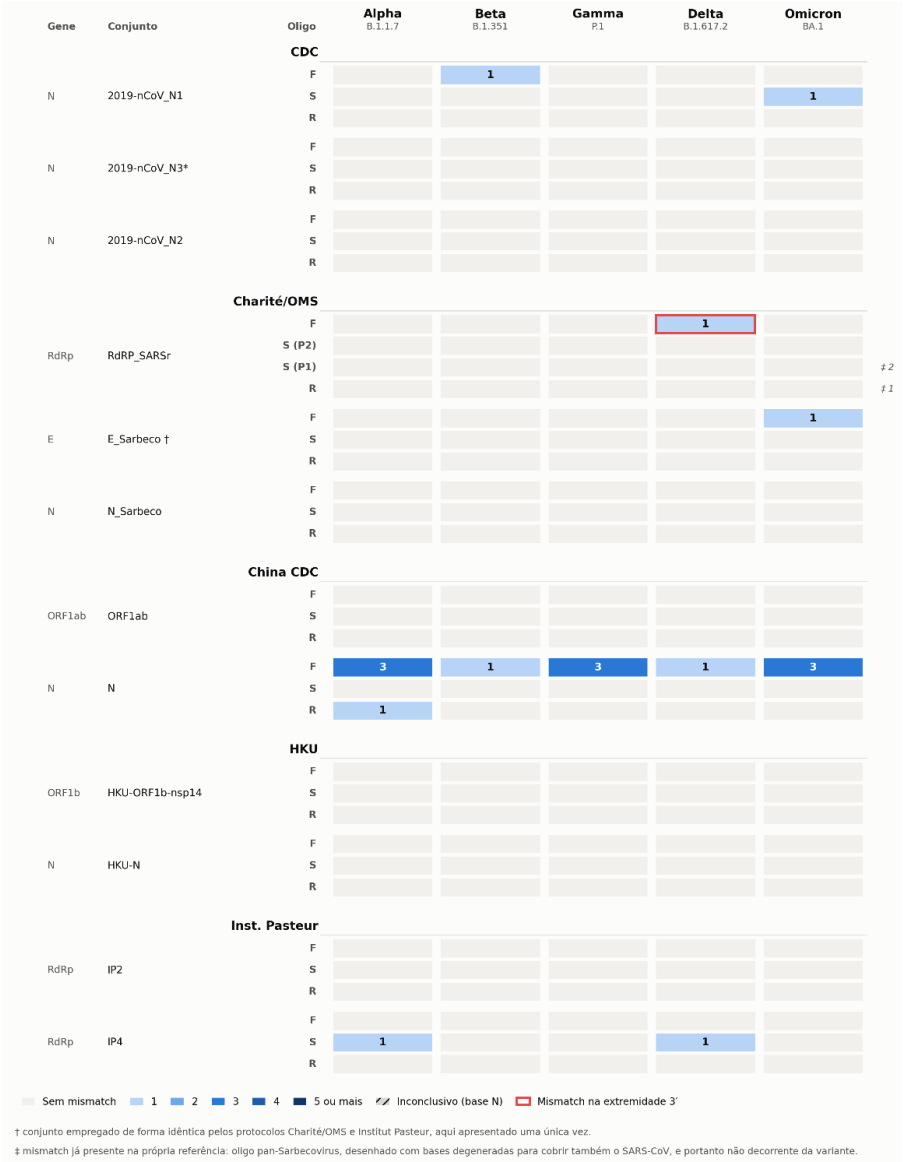


Figura 2 – *Mismatches* dos *primers* e sondas de RT-qPCR nas variantes de preocupação. Número de posições divergentes em cada um dos 37 oligonucleotídeos, por variante, em relação à referência NC_045512.2 (Wuhan-Hu-1). Células em cinza claro, sem número, indicam sítio de anelamento íntegro. F indica *primer* senso, R *primer* antissenso e S sonda. O conjunto RdRP_SARSr possui duas sondas, P1 e P2. O contorno vermelho assinala *mismatch* nas cinco bases da extremidade 3′ do *primer*, região crítica para a extensão pela polimerase. * Conjunto que integrava a versão original do painel do CDC e foi posteriormente retirado do protocolo, mantido na análise para fins de comparação. † Conjunto empregado de forma idêntica pelos protocolos Charité/OMS e Institut Pasteur, apresentado uma única vez. ‡ *Mismatch* já presente na própria referência, decorrente do desenho pan-*Sarbecovirus* com bases degeneradas e, portanto, não atribuível às variantes. O número ao lado indica a quantidade de posições.

Em termos de conjuntos, 7 dos 12 não exibiram qualquer divergência: ORF1ab (China CDC), IP2 (Institut Pasteur), HKU-ORF1b-nsp14 e HKU-N (HKU), 2019-nCoV_N2 e 2019-nCoV_N3 (CDC) e N_Sarbeco (Charité/OMS). Foram contabilizados 18 eventos de *mismatch*,

distribuídos entre Alpha (5), Omicron (5), Gamma (3), Delta (3) e Beta (2). Todos correspondem a substituições de base, não tendo sido observada qualquer deleção sobre sítio de anelamento. Os eventos concentram-se em 11 posições genômicas distintas: sete no gene N, três em ORF1ab e uma no gene E.

A carga é acentuadamente desigual entre os conjuntos. O conjunto N do China CDC concentra 12 dos 18 eventos, seguido por IP4 e 2019-nCoV_N1, com dois cada, e por RdRP_SARSr e E_Sarbeco, com um cada (Tabela 2).

Tabela 2 – *Mismatches* identificados nos sítios de anelamento, com localização gênica, efeito sobre a proteína codificada e classificação de criticidade. Considerou-se de criticidade alta o *mismatch* situado nas cinco bases da extremidade 3′ de um *primer*, de criticidade moderada a deleção ou o *mismatch* no terço central da sonda, e de criticidade baixa as demais posições. A coluna "Carga" indica o número total de *mismatches* daquele oligonucleotídeo na mesma variante. ORF1a e ORF1b seguem a numeração de resíduos de cada poliproteína, separadas pelo deslocamento de fase −1 na posição 13.468.

Variante	Conjunto	Oligo	Posição	Gene/domínio	Troca	Efeito	Pos. no oligo	Dist. 3′	Carga	Criticidade
Alpha	Inst. Pasteur · IP4	S	14120	ORF1b (nsp12)	C → T	P218L	16	3	1	Baixa
Delta	Inst. Pasteur · IP4	S	14122	ORF1b (nsp12)	G → A	G219S	18	1	1	Baixa
Delta	Charité/OMS · RdRP_SARSr	F	15451	ORF1b (nsp12)	G → A	G662S	21	1	1	Alta
Omicron	Charité/OMS · E_Sarbeco	F	26270	E	C → T	T9I	2	24	1	Baixa
Beta	CDC · 2019-nCoV_N1	F	28300	N	G → T	Q9H	14	6	1	Baixa
Omicron	CDC · 2019-nCoV_N1	S	28311	N	C → T	P13L	3	21	1	Baixa
Alpha	China CDC · N	F	28881	N	G → A	R203K	1	21	3	Baixa
Gamma	China CDC · N	F	28881	N	G → A	R203K	1	21	3	Baixa
Delta	China CDC · N	F	28881	N	G → T	R203M	1	21	1	Baixa
Omicron	China CDC · N	F	28881	N	G → A	R203K	1	21	3	Baixa
Alpha	China CDC · N	F	28882	N	G → A	R203K	2	20	3	Baixa
Gamma	China CDC · N	F	28882	N	G → A	R203K	2	20	3	Baixa
Omicron	China CDC · N	F	28882	N	G → A	R203K	2	20	3	Baixa
Alpha	China CDC · N	F	28883	N	G → C	G204R	3	19	3	Baixa
Gamma	China CDC · N	F	28883	N	G → C	G204R	3	19	3	Baixa
Omicron	China CDC · N	F	28883	N	G → C	G204R	3	19	3	Baixa

Variante	Conjunto	Oligo	Posição	Gene/domínio	Troca	Efeito	Pos. no oligo	Dist. 3′	Carga	Criticidade
Beta	China CDC · N	F	28887	N	C → T	T205I	7	15	1	Baixa
Alpha	China CDC · N	R	28977	N	C → T	S235F	3	19	1	Baixa

Esse padrão mostra que a conservação global do genoma do SARS-CoV-2 não garante a conservação dos sítios efetivamente usados no diagnóstico molecular. Ainda que a maioria dos oligonucleotídeos tenha permanecido íntegra, a distribuição acentuadamente desigual dos eventos indica que a vulnerabilidade de um ensaio depende menos da taxa média de substituição do vírus e mais da posição específica escolhida para o anelamento. Observação convergente foi relatada por Vogels *et al.* (2020), que, ao compararem a sensibilidade analítica de quatro protocolos amplamente utilizados, atribuíram o desempenho inferior do conjunto RdRP_SARSr a um *mismatch* no *primer* antissenso, e não a diferenças de eficiência intrínseca entre os ensaios, uma vez que todos apresentaram eficiência de amplificação próxima ou superior a 90%.

A concentração de divergências no conjunto N do China CDC explica-se pela sua geometria: o limite 5′ do amplicon coincide exatamente com o agrupamento 28.881–28.883, origem das substituições R203K e G204R. Esse mesmo agrupamento já foi associado experimentalmente a falhas de detecção do gene N. Bozidis *et al*. (2022) descreveram atraso ou ausência de amplificação do alvo N em três sistemas comerciais distintos, com diferenças médias de pelo menos 14 ciclos em relação aos demais alvos, em isolados que carregavam R203K, G204R, A208G e M234I. Zimmermann *et al*. (2021), ao investigarem o deslocamento de Ct observado com ensaio comercial em amostras da linhagem B.1.1.7, identificaram *mismatch* de três nucleotídeos no *primer* senso do alvo N. A convergência entre esses relatos experimentais e o presente achado *in silico* reforça que a região 28.881–28.883 constitui ponto crítico para ensaios dirigidos ao gene N, ainda que os conjuntos envolvidos não sejam os mesmos.

A vulnerabilidade do gene N não se restringe a esse agrupamento. Wollschläger *et al.* (2021) recorreram a experimentos de genética reversa com plasmídeos contendo a sequência codificante de N com e sem a mutação D3L e obtiveram deslocamento de Ct de 7,1 a 10,1 ciclos atribuível a essa única alteração, próxima à extremidade 5′ do gene. Nas amostras clínicas da linhagem B.1.1.7, o deslocamento foi de 6 a 10 ciclos e o alvo N deixou de ser detectado naquelas com Ct acima de 29. O *primer* senso do conjunto 2019-nCoV_N1, posicionado em 28.287–28.306, situa-se imediatamente a jusante dessa região e apresentou *mismatch* na variante Beta, o que indica que também esse trecho do gene N acumula variação relevante para o diagnóstico.

Cabe notar que nenhum dos 18 eventos corresponde a deleção, o que é favorável. A deleção sobre um sítio de anelamento produz a forma mais severa de falha, a ausência completa

de sinal do alvo, e não a simples elevação do Ct. O exemplo canônico é a deleção Δ69-70 do gene S, que suprime a detecção do alvo correspondente no ensaio TaqPath e deu origem ao fenômeno conhecido como falha do alvo S. McMillen *et al.* (2022) verificaram que essa falha identifica a variante Alpha com sensibilidade de 100% e especificidade de 99,6%, o que ilustra tanto a magnitude do efeito quanto a possibilidade de explorá-lo como assinatura. A ausência de deleções sobre os doze conjuntos avaliados significa que o modo de falha mais grave não se materializou em nenhum deles, embora deleções estejam presentes no genoma analisado, como mostra a própria janela de nsp6.

3.3 Localização gênica e efeito sobre a proteína codificada

A tradução do códon correspondente em cada variante, efetuada sobre a sequência alinhada de modo a tratar corretamente mutações adjacentes, indicou que os 18 eventos são não sinônimos. Correspondem a 11 alterações de aminoácido distintas, sendo sete no gene N (R203K, G204R, R203M, T205I, S235F, P13L e Q9H), três em ORF1b (P218L, G219S e G662S) e uma no gene E (T9I), conforme a Tabela 2. O quadro de leitura empregado foi validado contra posições de efeito conhecido na literatura, a saber 14.408 (ORF1b:P314), 23.403 (S:D614), 28.881 (N:R203), 26.270 (E:T9) e 3.037 (ORF1a:F924), todas reproduzidas corretamente.

O caráter integralmente não sinônimo dos eventos merece registro. Em genomas de RNA, parte expressiva das substituições fixadas é silenciosa, de modo que a concentração de trocas com efeito sobre a proteína em um número reduzido de posições sugere que esses sítios não são neutros. O agrupamento 28.881–28.883 é o caso mais bem documentado. Wu *et al.* (2021), trabalhando com vírus autêntico, demonstraram que as substituições R203K e G204R aumentam a replicação, a infectividade e a virulência do SARS-CoV-2, e identificaram assinatura de seleção positiva nessas posições. A implicação para o diagnóstico é desfavorável, porque sítios que conferem vantagem replicativa tendem a fixar-se e a reaparecer em linhagens independentes. Um oligonucleotídeo desenhado sobre eles acumula o risco de perder desempenho repetidas vezes, e não apenas uma.

A classificação de criticidade, fundamentada na posição do *mismatch* no oligonucleotídeo, identificou um evento de criticidade alta e 17 de criticidade baixa. O evento de alta criticidade é a substituição G→A na posição 15.451 (ORF1b:G662S, variante Delta), situada a uma base da extremidade 3′ do *primer* senso do conjunto RdRP_SARSr.

Esse critério, contudo, é aplicado evento a evento e não capta o efeito cumulativo: nove dos 18 eventos ocorrem em oligonucleotídeos que acumulam mais de um *mismatch* na mesma

variante, condição que reduz a estabilidade do híbrido independentemente da distância à extremidade 3′.

A distinção entre número e posição das divergências é central para a interpretação desses achados. Lefever *et al.* (2013), avaliando sistematicamente o efeito de *mismatches* em sítios de anelamento, demonstraram que o impacto sobre o ciclo de quantificação é determinado principalmente pelo número de nucleotídeos divergentes e pela distância à extremidade 3′ do *primer*, e que *mismatches* isolados situados a mais de cinco pares de base dessa extremidade exercem efeito moderado e podem ser tolerados. Os mesmos autores verificaram que o bloqueio completo da reação exige pelo menos quatro divergências em um único *primer*, ou três em um *primer* e duas no outro. Aplicado aos presentes dados, esse critério sustenta interpretação consistente. As três divergências do *primer* senso do China CDC, situadas a 19 a 21 nucleotídeos da extremidade 3′, encontram-se em faixa intermediária, compatível com elevação do valor de Ct e perda de sensibilidade analítica, mas não com falha completa de amplificação. É esse o comportamento descrito por Bozidis *et al.* (2022) e por Wollschläger *et al.* (2021), nos quais predominou a elevação do Ct e a ausência completa de sinal ficou restrita às amostras de menor carga viral.

Além do número e da posição, o tipo de par desemparelhado modula o efeito. Stadhouders *et al.* (2010), ao quantificarem o impacto de *mismatches* na região 3′ de *primers* em ensaios de nuclease 5′, observaram variação de menos de 1,5 a mais de 7,0 ciclos conforme o par formado, com impacto reduzido para combinações como A·C, C·A, T·G e G·T e impacto severo para A·A, G·A, A·G e C·C. Essa gradação ajuda a explicar por que divergências numericamente equivalentes podem produzir consequências distintas e reforça que a simples contagem de *mismatches* é insuficiente para prever o desempenho de um ensaio.

O único evento de criticidade alta aqui identificado insere-se nesse mesmo quadro interpretativo. Cao *et al.* (2022) demonstraram que mutações situadas nas sequências correspondentes às sondas e à extremidade 3′ dos *primers* exercem impacto significativo sobre a eficiência de detecção por qPCR. No caso do mutante G28916T do gene N, o limite de detecção de dois conjuntos comerciais passou de 1 para 10 cópias por microlitro, o que confere plausibilidade à classificação atribuída à substituição na posição 15.451 do conjunto RdRP_SARSr na variante Delta.

3.4 Distribuição dos *mismatch*es ao longo dos oligonucleotídeos

O detalhamento base a base dos sete oligonucleotídeos afetados está apresentado na Figura 3.



Figura 3 – Sítios de anelamento dos oligonucleotídeos afetados por mutações nas variantes de preocupação. Dos 37 oligonucleotídeos avaliados, são apresentados os sete que exibem ao menos um *mismatch*, identificados pelo protocolo, conjunto e tipo de oligo, com o gene e as coordenadas na referência NC_045512.2 indicados no canto superior direito de cada painel. A primeira linha traz a sequência do oligonucleotídeo no sentido 5′→3′, e a linha “Referência” assinala as divergências já presentes na própria referência. Pontos indicam base idêntica à do oligonucleotídeo e bases coloridas assinalam as posições divergentes, com a alteração de aminoácido correspondente à direita. A base destacada em lilás no painel (b) corresponde a uma posição degenerada (código IUPAC) do oligo. A faixa rosa destaca as cinco bases da extremidade 3′ dos *primers*. A numeração abaixo das sequências refere-se à posição dentro do oligonucleotídeo, no sentido 5′→3′.

O *primer* senso do conjunto N do China CDC é o mais comprometido. Em Alpha, Gamma e Omicron as posições 1 a 3 divergem simultaneamente, correspondendo ao bloco GGG→AAC em 28.881–28.883 e às substituições R203K e G204R. Em Delta diverge isoladamente a posição 1 (R203M) e, em Beta, a posição 7 (T205I). O *primer* antissenso do mesmo conjunto apresenta *mismatch* adicional em Alpha, na posição 3 (S235F).

Os demais oligonucleotídeos afetados apresentam um único *mismatch* por variante: a sonda do IP4 nas posições 16 (Alpha) e 18 (Delta), o *primer* senso do RdRP_SARSr na posição 21 (Delta), o *primer* senso do E_Sarbeco na posição 2 (Omicron), o *primer* senso do 2019-nCoV_N1 na posição 14 (Beta) e a sonda do mesmo conjunto na posição 3 (Omicron).

Dois desses achados encontram confirmação experimental direta na literatura. Liu *et al.* (2024) avaliaram, por RT-qPCR e por PCR digital em gotículas, o efeito de cinco variantes de nucleotídeo único situadas nas regiões de anelamento dos conjuntos 2019-nCoV_N1 do CDC e E_Sarbeco do protocolo Charité. Entre elas estão C28311T, na sonda do 2019-nCoV_N1, e C26270T, no *primer* senso do E_Sarbeco, que são precisamente as duas posições aqui identificadas como divergentes na variante Omicron, correspondentes a P13L e T9I. Os autores verificaram que a presença de uma ou duas dessas substituições no *primer* ou na sonda afetou a sensibilidade de ambas as plataformas em extensões variáveis e recomendaram a atualização periódica dos conjuntos. Essa correspondência entre a predição computacional e a medida experimental confere sustentação externa ao método aqui empregado.

O conjunto E_Sarbeco merece registro à parte por já haver sido objeto de falha documentada. Artesi *et al.* (2020) descreveram a ausência de amplificação do alvo E do ensaio comercial cobas SARS-CoV-2 em oito pacientes portadores de uma transição C→U na posição 26.340, mutação que surgiu de forma independente ao menos quatro vezes segundo os genomas então depositados. A agência regulatória norte-americana mantém, desde então, lista pública de ensaios cujo desempenho pode ser afetado por mutações virais, recomendando o acompanhamento contínuo dessa interação (FDA, 2023). A divergência aqui observada nesse mesmo conjunto reforça a necessidade de vigilância sobre alvos já reconhecidamente suscetíveis.

O conjunto IP4 também possui histórico próprio. Gdoura *et al.* (2022), ao compararem quatro protocolos aprovados pela Organização Mundial da Saúde em amostras clínicas, registraram ausência de sinal do alvo IP4 em trinta casos nos quais os demais alvos amplificaram com valores baixos de Ct, e relataram que o centro nacional de referência francês havia comunicado, em março de 2021, perda de sensibilidade desse alvo associada à mutação C14050T da variante Alpha. Essa posição situa-se a montante do amplicon aqui analisado, que

se estende de 14.080 a 14.186, e não corresponde, portanto, aos eventos identificados neste trabalho, ambos sobre a sonda, nas posições 14.120 em Alpha e 14.122 em Delta. Os dois conjuntos de achados convergem, contudo, no essencial, pois indicam que a região de nsp12 interrogada pelo IP4 acumulou variação relevante já nas primeiras variantes de preocupação. Em levantamento *in silico* conduzido sobre 177 genomas brasileiros depositados até meados de 2020, o IP4 figurava entre os conjuntos sem qualquer *mismatch* (Santos *et al.*, 2020), o que delimita no tempo o surgimento dessa vulnerabilidade.

Dois oligonucleotídeos do conjunto RdRP_SARSr, a sonda pan-*Sarbecovirus* P1 e o *primer* antissenso R1, apresentam divergências em relação à própria sequência de referência, em duas e uma posição, respectivamente. Tais divergências decorrem do desenho original, concebido com bases degeneradas para detectar igualmente o SARS-CoV (Corman *et al*., 2020), e não constituem efeito das variantes analisadas. Zimmermann *et al*. (2021) quantificaram o custo analítico dessa escolha, demonstrando que a remoção das ambiguidades do conjunto RdRP do protocolo Charité reduz em aproximadamente três ciclos o valor de Ct obtido. Trata-se, portanto, de limitação inerente ao ensaio, que se soma ao efeito das mutações das variantes sem com ele se confundir.

3.5 Posição do *mismatch* e consequência prevista para a amplificação

Os achados configuram dois modos distintos de comprometimento (Figura 4).

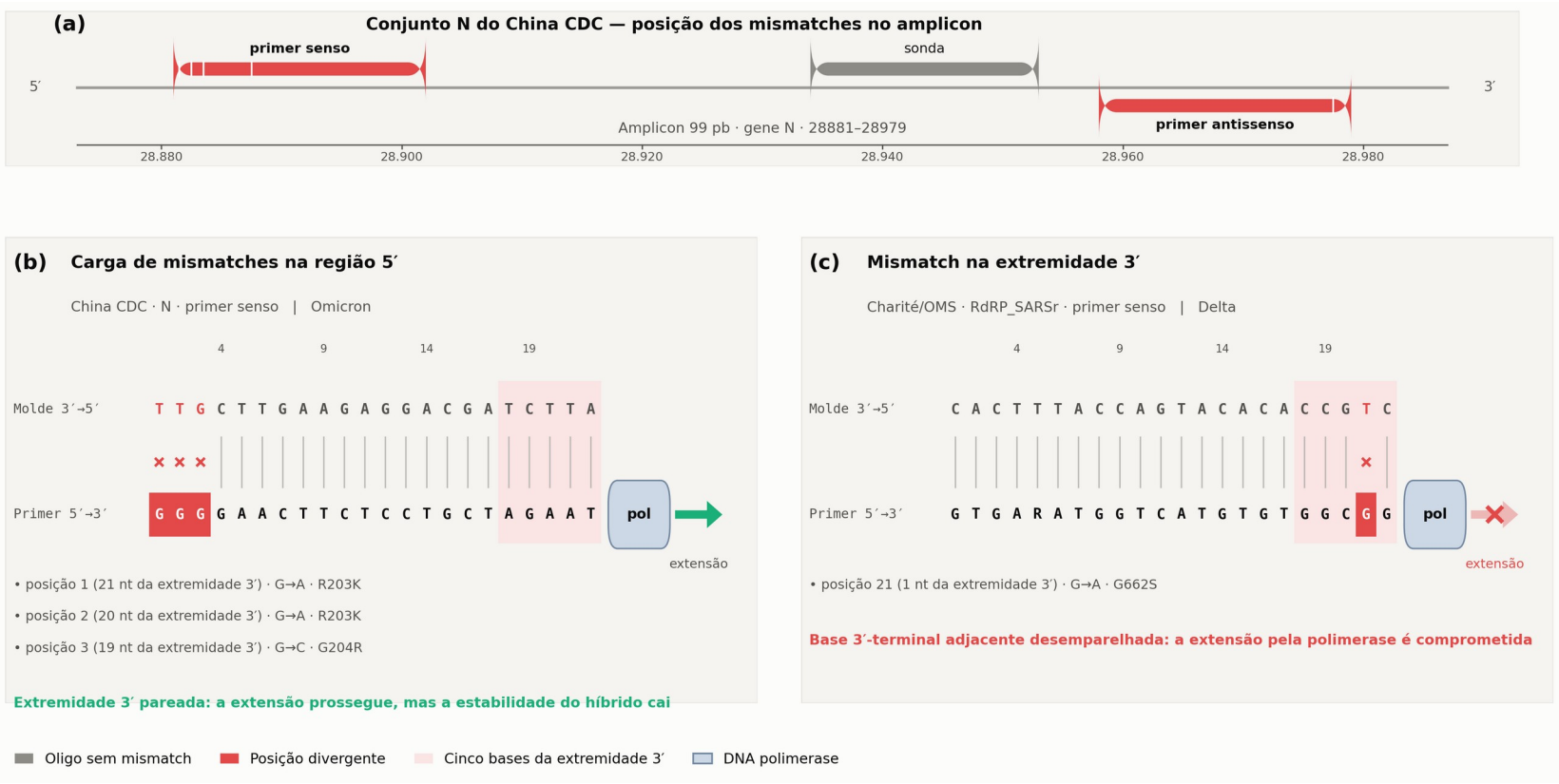


Figura 4 – Posição do *mismatch* no híbrido *primer*-molde e consequência prevista para a amplificação. (a) Organização do conjunto N do China CDC sobre o amplicon de 99 pb (posições 28.881 a 28.979 da referência NC_045512.2). Os *primers*, em vermelho, apresentam *mismatch* e a sonda, em cinza, permanece íntegra. Traços brancos marcam as posições divergentes em pelo menos uma variante. (b, c) Híbrido entre o *primer* senso e a fita-molde nos dois modos distintos de falha. Em (b), *primer* senso do China CDC · N na variante Omicron, os três *mismatches* concentram-se na região 5′ e a extremidade 3′ permanece pareada, de modo que a extensão prossegue, embora com menor estabilidade do híbrido. Em (c), *primer* senso do Charité/OMS · RdRP_SARSr na variante Delta, o *mismatch* ocupa a penúltima base, adjacente ao nucleotídeo 3′-terminal, o que compromete o início da extensão pela polimerase. Traços verticais representam pareamento Watson-Crick e o símbolo × assinala as posições desemparelhadas. A faixa rosa destaca as cinco bases da extremidade 3′, e a seta indica a extensão pela DNA polimerase (pol), prevista em (b) e comprometida em (c).

No conjunto N do China CDC, os *mismatches* localizam-se a 19 a 21 nucleotídeos da extremidade 3′ do *primer* senso. O terminal 3′ permanece pareado e a extensão pela polimerase não é diretamente impedida, sendo esperada perda de estabilidade do híbrido. No conjunto RdRP_SARSr, ao contrário, o único *mismatch* situa-se a uma base da extremidade 3′, região em que o pareamento correto constitui requisito para o início da extensão.

A coexistência desses dois modos de comprometimento tem consequência prática direta: ensaios diagnósticos para SARS-CoV-2 devem contemplar múltiplos alvos genômicos independentes. No relato de Artesi *et al.* (2020), os pacientes acometidos pela falha do alvo E ainda assim foram corretamente identificados porque o ensaio interrogava duas regiões distintas do genoma, e os autores argumentaram explicitamente que dois alvos devem ser empregados para evitar resultados falso-negativos. Tombuloglu *et al.* (2022) aplicaram esse princípio ao desenvolverem ensaio multiplex dirigido simultaneamente aos genes N e RdRp, com controle interno humano. O presente levantamento oferece subsídio adicional a essa estratégia ao demonstrar que, entre os doze conjuntos avaliados, sete permaneceram íntegros em todas as variantes, de modo que a redundância entre alvos tende a converter a vulnerabilidade localizada de um ensaio em perda parcial de sensibilidade, e não em falha diagnóstica.

A mesma lógica aplica-se à vigilância. Vogels *et al.* (2021) converteram a variação nos sítios de anelamento em ferramenta epidemiológica ao desenvolverem ensaio multiplex que detecta variantes de preocupação a partir de diferenças de amplificação, diferenciando B.1.1.7 do par formado por B.1.351 e P.1. Peterson *et al.* (2021) empregaram princípio análogo em águas residuárias, detectando mutações-assinatura de Alpha, Beta e Gamma em 261 amostras provenientes de cinco cidades canadenses. O mesmo fenômeno que compromete um ensaio diagnóstico pode ser explorado como sinal de triagem populacional, desde que a relação entre mutação e desempenho do ensaio seja previamente caracterizada.

3.6 Conjuntos *primer*/sonda redesenhados

A comparação das seis sequências identificou 28.825 posições idênticas em todas elas (96,4% do genoma), descontadas as regiões não sequenciadas e as bases ambíguas. Restringindo o desenho a essas regiões, foram propostos novos conjuntos para os cinco ensaios afetados (Tabela 3).

Tabela 3 – Conjuntos de *primers* e sondas propostos para os ensaios afetados, com parâmetros termodinâmicos. Tm calculada pelo modelo de vizinho mais próximo para 50 mM de cátion monovalente, 3 mM de Mg^{2+}, 0,8 mM de dNTP e 250 nM de oligonucleotídeo. Grampo e autodímero expressos como Tm da estrutura secundária prevista.

Conjunto proposto	Substitui	Oligo	Sequência (5′→3′)	nt	Posição	Tm (°C)	GC (%)	Grampo (°C)	Autodímero (°C)
IP4-v2	Inst. Pasteur · IP4	F	GAAAACCCAGATATATTACGCG	22	13939–13960	58.8	40.9	0.0	10.0
IP4-v2	Inst. Pasteur · IP4	S	ACGCCAACTTAGGTGAACGTGTACGCCAAGCT	32	13964–13995	66.6	53.1	37.4	9.9
IP4-v2	Inst. Pasteur · IP4	R	CAACAATACCAGCATTTCGC	20	14028–14047	59.8	45.0	42.9	0.0
RdRP_SARSr-v2	Charité/OMS · RdRP_SARSr	F	ACTCAAATGAATCTTAAGTATGCC	24	15058–15081	58.9	33.3	0.0	6.9
RdRP_SARSr-v2	Charité/OMS · RdRP_SARSr	S	GTGCAAAGAATAGAGCTCGCACCGTAGCTGGT	32	15086–15117	65.2	53.1	43.8	16.5
RdRP_SARSr-v2	Charité/OMS · RdRP_SARSr	R	TAGTGGCGGCTATTGATTTC	20	15168–15187	59.5	45.0	0.0	16.8
E_Sarbeco-v2	Charité/OMS · E_Sarbeco	F	TGCTTTCGTGGTATTCTTGC	20	26307–26326	60.6	45.0	0.0	0.0
E_Sarbeco-v2	Charité/OMS · E_Sarbeco	S	ACTGCGCTTCGATTGTGTGCGTACTGCTGCA	31	26347–26377	67.5	54.8	44.2	7.7
E_Sarbeco-v2	Charité/OMS · E_Sarbeco	R	AGAAGGTTTTACAAGACTCACG	22	26388–26409	60.1	40.9	0.0	0.0
2019-nCoV_N1-v2	CDC · 2019-nCoV_N1	F	AGATTCAACTGGCAGTAACC	20	28336–28355	59.3	45.0	0.0	28.2
2019-nCoV_N1-v2	CDC · 2019-nCoV_N1	S	GGGCGCGATCAAAACAACGTCGGCCCCAAGG	31	28374–28404	69.9	64.5	44.0	18.4
2019-nCoV_N1-v2	CDC · 2019-nCoV_N1	R	CTTGCCATGTTGAGTGAGAG	20	28437–28456	60.3	50.0	0.0	0.0
N-v2	China CDC · N	F	TCTCGTTCCTCATCACGTAG	20	28823–28842	60.4	50.0	38.5	0.0
N-v2	China CDC · N	S	TGGCTGGCAATGGCGGTGATGCTGCTCTTGC	31	28902–28932	69.6	61.3	45.8	13.5
N-v2	China CDC · N	R	TCAAGCTGGTTCAATCTGTC	20	28946–28965	59.6	45.0	0.0	0.0

Os 15 oligonucleotídeos propostos apresentam pareamento perfeito nas seis sequências analisadas e nenhuma ocorrência secundária no próprio genoma de referência, admitindo-se até três *mismatches*. Os amplicons variam de 103 a 143 pb. Os *primers* apresentam Tm entre 58,8 e 60,6 °C e as sondas entre 65,2 e 69,9 °C, de modo que cada sonda supera em 5,7 a 10,6 °C os *primers* do respectivo conjunto, faixa próxima à diferença de 5 a 10 °C recomendada para ensaios com sonda de hidrólise (Bio-Rad, 2006). As Tm das estruturas secundárias previstas mantiveram-se abaixo da temperatura de anelamento, com valores máximos de 45,8 °C para grampo, 28,2 °C para autodímero e 21,7 °C para dímero entre *primers*. As sondas resultaram longas, com 31 a 32 nucleotídeos, consequência do baixo conteúdo de guanina e citosina das regiões conservadas disponíveis, evidente nos *primers* senso dos alvos situados em ORF1b, com 33,3% e 40,9%.

O conjunto proposto para o gene N posiciona o *primer* senso em 28.823–28.842 e a sonda em 28.902–28.932, de maneira que o agrupamento 28.881–28.887, origem das substituições R203K, G204R, R203M e T205I, permanece no interior do amplicon, sem sobreposição a qualquer oligonucleotídeo. A região continua a ser amplificada, mas deixa de interferir no anelamento. Essa configuração preserva o alvo gênico consagrado sem herdar a sua vulnerabilidade.

A exigência de sondas longas não vem do procedimento de desenho, mas da composição das regiões conservadas disponíveis. Radvánszka *et al*. (2022) enfrentaram limitação análoga ao redesenharem os alvos RdRp e E do protocolo Charité e a contornaram mediante incorporação de nucleotídeos de ácido nucleico travado (LNA) e tecnologia de sonda dupla, obtendo ganho de sensibilidade e especificidade. Estratégia equivalente, ou o emprego de sondas com ligante de sulco menor, permitiria reduzir o comprimento dos oligonucleotídeos aqui propostos preservando suas propriedades térmicas.

Convém lembrar que o desempenho de um ensaio de RT-qPCR não depende apenas da complementaridade dos oligonucleotídeos. Silva *et al.* (2023), ao avaliarem o efeito da variante Omicron sobre a extração de RNA e a detecção do SARS-CoV-2 em laboratório brasileiro, atribuíram a falha do gene controle ao aumento da viscosidade das amostras associado ao tropismo dessa variante pelas vias aéreas superiores, falha corrigida por modificação na etapa de extração. Fatores pré-analíticos, portanto, interagem com as características biológicas do vírus na determinação do resultado final. A otimização dos sítios de anelamento, objeto deste trabalho, é condição necessária, mas não suficiente, para a confiabilidade diagnóstica.

3.7 Especificidade frente a outros coronavírus humanos

Os quinze oligonucleotídeos propostos foram confrontados com os genomas de referência do SARS-CoV, do MERS-CoV e dos quatro coronavírus sazonais 229E, OC43, NL63 e HKU1. A varredura foi previamente validada contra o genoma de referência do SARS-CoV-2, no qual os quinze apresentaram pareamento integral, o que confirma a adequação do procedimento. Os resultados por oligonucleotídeo constam da Tabela 4.

Tabela 4 – Menor número de divergências entre cada oligonucleotídeo proposto e os genomas de referência de outros coronavírus humanos. SARS-CoV (NC_004718.3), MERS-CoV (NC_019843.3), HCoV-229E (NC_002645.1), HCoV-OC43 (NC_006213.1), HCoV-NL63 (NC_005831.2) e HCoV-HKU1 (NC_006577.2). Os valores correspondem ao menor número de posições divergentes encontrado nas duas fitas de cada genoma. F indica *primer* senso, R *primer* antissenso e S sonda. Contra o genoma de referência do SARS-CoV-2 (NC_045512.2) todos os oligonucleotídeos apresentam zero divergências.

Conjunto proposto	Oligo	SARS-CoV	MERS-CoV	229E	OC43	NL63	HKU1
IP4-v2	F	5	8	8	7	6	7
	S	5	8	14	14	14	14
	R	5	5	5	6	4	6
RdRP_SARSr-v2	F	0	3	5	4	4	4
	S	0	5	13	4	12	3
	R	1	6	5	6	7	6
E_Sarbeco-v2	F	0	6	5	5	5	7
	S	0	12	12	12	13	13
	R	5	7	6	6	7	6
2019-nCoV_N1-v2	F	2	6	6	7	6	6
	S	7	13	13	13	14	14
	R	1	6	6	6	5	7
N-v2	F	1	6	5	6	6	5
	S	9	11	12	10	12	11
	R	0	6	6	6	6	6

Nos coronavírus sazonais e no MERS-CoV, o menor número de divergências observado em qualquer oligonucleotídeo foi três, e a maior parte dos valores situou-se entre cinco e quatorze. Duas combinações de *primers* chegaram a formar geometria compatível com amplicon, o IP4-v2 no MERS-CoV e o RdRP_SARSr-v2 no 229E, porém com cinco a oito divergências por *primer*, de modo que não é esperada amplificação em nenhum dos dois casos.

O SARS-CoV comportou-se de maneira distinta. O conjunto RdRP_SARSr-v2 apresentou pareamento integral do *primer* senso e da sonda, uma única divergência no *primer* antissenso e amplicon de 130 pb. O conjunto E_Sarbeco-v2 apresentou *primer* senso e sonda íntegros, amplicon de 103 pb e cinco divergências no *primer* antissenso, todas situadas fora das cinco bases da extremidade 3′. Os conjuntos 2019-nCoV_N1-v2 e N-v2 formaram amplicons de

121 e 143 pb, mas com sete e nove divergências nas respectivas sondas. O IP4-v2 não formou amplicon.

A reatividade cruzada dos conjuntos RdRP_SARSr-v2 e E_Sarbeco-v2 com o SARS-CoV não constitui falha do redesenho, e sim consequência da escolha das regiões-alvo. Os conjuntos originais da Charité foram concebidos como pan-*Sarbecovirus*, com bases degeneradas destinadas a detectar igualmente o SARS-CoV (Corman *et al.*, 2020), e as regiões conservadas entre as variantes de preocupação do SARS-CoV-2 são, em boa medida, as mesmas conservadas naquela espécie. Como o SARS-CoV não circula desde 2004, a consequência prática para o diagnóstico é desprezível, mas o comportamento precisa ser declarado, porque dois dos cinco conjuntos propostos não discriminam entre as duas espécies. Os conjuntos dirigidos ao gene N comportaram-se de modo oposto e mostraram-se específicos para o SARS-CoV-2, uma vez que a ausência de hibridação da sonda impede a geração de sinal ainda que o amplicon se forme.

Esse resultado tem uma leitura metodológica que extrapola os conjuntos aqui propostos. Restringir o desenho às posições conservadas entre variantes de uma espécie tende a selecionar justamente as regiões sob maior restrição funcional, que são também as mais conservadas entre espécies próximas. Ganha-se robustez frente à evolução intraespecífica e perde-se, na mesma medida, poder de discriminação entre espécies do mesmo subgênero. Quando a especificidade de espécie for requisito, o critério de conservação precisa ser combinado a um critério explícito de exclusão, o que o presente desenho não incorporou.

3.8 Limitações do estudo

A análise baseou-se em um genoma representativo por variante, o que permite identificar sítios candidatos a comprometimento, mas não autoriza qualquer estimativa de prevalência das alterações observadas nas populações virais circulantes. A confirmação dessa dimensão exigiria amostragem ampliada. A abordagem é inteiramente computacional e, portanto, prevê o comportamento esperado do híbrido *primer*-molde sem medi-lo: os efeitos aqui discutidos sobre eficiência de amplificação e valores de Ct derivam da literatura experimental citada, e não de ensaios conduzidos neste trabalho. A verificação de especificidade restringiu-se aos seis coronavírus humanos de referência, não tendo sido conduzida busca contra bases completas de sequências nucleotídicas, que abrangeria também o transcriptoma humano e outros agentes respiratórios. Por fim, a validação laboratorial dos oligonucleotídeos propostos, com determinação de eficiência de amplificação, limite de

detecção e especificidade analítica, constitui requisito indispensável antes de qualquer aplicação diagnóstica.

No conjunto, os achados sustentam que o monitoramento das regiões-alvo deve integrar-se à vigilância genômica de rotina, e não ser conduzido apenas em resposta a falhas já observadas. A análise comparativa entre sequências de variantes emergentes e os conjuntos *primer*/sonda em uso permite antecipar perdas de desempenho e orientar a revisão dos protocolos antes que o comprometimento se manifeste na prática laboratorial. Essa necessidade foi apontada de modo explícito por Liu *et al*. (2024) e por Cao *et al*. (2022), e o presente estudo mostra que atendê-la é exequível com ferramentas computacionais de acesso livre.

4 CONCLUSÃO

A análise *in silico* de 37 oligonucleotídeos provenientes de cinco protocolos de RT-qPCR demonstrou que a elevada conservação global do genoma do SARS-CoV-2 não assegura a integridade dos sítios efetivamente utilizados no diagnóstico molecular. Trinta oligonucleotídeos permaneceram íntegros nas cinco variantes de preocupação avaliadas, enquanto os 18 eventos de *mismatch* identificados concentraram-se em sete oligonucleotídeos e, de forma acentuada, em um único conjunto.

Os achados evidenciam que a avaliação de um ensaio frente a variantes emergentes exige considerar, além do número de divergências, a posição que ocupam no oligonucleotídeo. O conjunto N do China CDC acumulou três divergências na região 5′ do *primer* senso, condição associada, na literatura experimental, a perda de estabilidade do híbrido e elevação do valor de Ct. O conjunto RdRP_SARSr apresentou divergência única, porém situada a uma base da extremidade 3′, posição em que o pareamento é requisito para o início da extensão. Dois modos distintos de comprometimento, portanto, que a simples contagem de *mismatches* não distingue.

Os conjuntos redesenhados, restritos às regiões conservadas nas seis sequências analisadas, apresentaram complementaridade integral e parâmetros termodinâmicos adequados, com destaque para a reformulação do alvo N, que deslocou o agrupamento mutacional para o interior do amplicon. A avaliação computacional de especificidade não indicou reatividade esperada com o MERS-CoV nem com os coronavírus sazonais, embora os conjuntos dirigidos a ORF1b e ao gene E preservem o reconhecimento do SARS-CoV. Tais propostas constituem resultado *in silico* e requerem validação laboratorial antes de qualquer emprego diagnóstico.

Conclui-se que o monitoramento das regiões-alvo de ensaios moleculares deve integrar a vigilância genômica de rotina. A abordagem computacional aqui empregada mostrou-se capaz

de identificar sítios vulneráveis e orientar a revisão de protocolos de forma antecipada, contribuindo para a manutenção da acurácia diagnóstica diante da evolução viral.

REFERÊNCIAS

ALSHARIF, W.; QURASHI, A. Effectiveness of COVID-19 diagnosis and management tools: a review. **Radiography**, v. 27, n. 2, p. 682-687, 2021. DOI: https://doi.org/10.1016/j.radi.2020.09.010.

ALVES, N. S. *et al.* Agenda de prioridades de pesquisa em COVID-19 e pós-COVID no Brasil: resultados de consulta ampla e consenso Delphi, 2022-2023. **Epidemiologia e Serviços de Saúde**, Brasília, v. 34, e20240623, 2025. DOI: https://doi.org/10.1590/S2237-96222025v34e20240623.

ARAÚJO, P. S. *et al.* Efeitos da pandemia de COVID-19 no trabalho em vigilância sanitária. **Ciência & Saúde Coletiva**, v. 28, n. 5, p. 1365-1376, 2023. DOI: https://doi.org/10.1590/1413-81232023285.11852022.

ARTESI, M. *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**, v. 58, n. 10, e01598-20, 2020. DOI: https://doi.org/10.1128/JCM.01598-20.

BIO-RAD LABORATORIES. **Real-time PCR applications guide**. Hercules: Bio-Rad Laboratories, 2006.

BOZIDIS, P. *et al.* Unusual N gene dropout and Ct value shift in commercial multiplex PCR assays caused by mutated SARS-CoV-2 strain. **Diagnostics**, v. 12, n. 4, art. 973, 2022. DOI: https://doi.org/10.3390/diagnostics12040973.

BRASIL. Ministério da Saúde. **COVID-19 no Brasil**: casos e óbitos. Disponível em: https://infoms.saude.gov.br/extensions/covid-19_html/covid-19_html.html. Acesso em: 18 ago. 2025.

CAMPOS, K. R. *et al.* COVID-19 laboratory diagnosis: comparative analysis of different RNA extraction methods for SARS-CoV-2 detection by two amplification protocols. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 63, e52, 2021. DOI: https://doi.org/10.1590/S1678-9946202163052.

CAO, L. *et al.* The impact of accumulated mutations in SARS-CoV-2 variants on the qPCR detection efficiency. **Frontiers in Cellular and Infection Microbiology**, v. 12, art. 823306, 2022. DOI: https://doi.org/10.3389/fcimb.2022.823306.

CHU, D. K. W. *et al.* Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. **Clinical Chemistry**, v. 66, n. 4, p. 549-555, 2020. DOI: https://doi.org/10.1093/clinchem/hvaa029.

CORMAN, V. M. *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. **Eurosurveillance**, v. 25, n. 3, art. 2000045, 2020. DOI: https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045.

FDA - U.S. FOOD AND DRUG ADMINISTRATION. **SARS-CoV-2 viral mutations**: impact on COVID-19 tests. Silver Spring: FDA, 2023. Disponível em:

https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/sars-cov-2-viral-mutations-impact-covid-19-tests. Acesso em: 24 set. 2025.

FRUEHWIRTH, M. *et al.* False-negative result in molecular diagnosis of SARS-CoV-2 in samples with amplification inhibitors. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 56, 2020. DOI: https://doi.org/10.5935/1676-2444.20200060.

GDOURA, M. *et al.* SARS-CoV2 RT-PCR assays: *in vitro* comparison of 4 WHO approved protocols on clinical specimens and its implications for real laboratory practice through variant emergence. **Virology Journal**, v. 19, n. 1, art. 54, 2022. DOI: https://doi.org/10.1186/s12985-022-01784-4.

KUBISTA, M. *et al.* The real-time polymerase chain reaction. **Molecular Aspects of Medicine**, v. 27, n. 2-3, p. 95-125, 2006. DOI: https://doi.org/10.1016/j.mam.2005.12.007.

LEFEVER, S. *et al.* Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. **Clinical Chemistry**, v. 59, n. 10, p. 1470-1480, 2013. DOI: https://doi.org/10.1373/clinchem.2013.203653.

LIU, S. *et al.* Sensitivity analysis of RT-qPCR and RT-ddPCR for SARS-CoV-2 detection with mutations on N1 and E primer-probe region. **Microbiology Spectrum**, v. 12, n. 8, e0429223, 2024. DOI: https://doi.org/10.1128/spectrum.04292-23.

McMILLEN, T. *et al.* The spike gene target failure (SGTF) genomic signature is highly accurate for the identification of Alpha and Omicron SARS-CoV-2 variants. **Scientific Reports**, v. 12, art. 18968, 2022. DOI: https://doi.org/10.1038/s41598-022-21564-y.

MOHAMMADI, M. *et al.* COVID-19 diagnosis on the basis of nanobiosensors' prompt interactivity: a holistic review. **Pathology - Research and Practice**, v. 262, art. 155565, 2024. DOI: https://doi.org/10.1016/j.prp.2024.155565.

MULLIS, K. B.; FALOONA, F. A. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. **Methods in Enzymology**, v. 155, p. 335-350, 1987. DOI: https://doi.org/10.1016/0076-6879(87)55023-6.

PECORARO, V. *et al.* Estimate false-negative RT-PCR rates for SARS-CoV-2. A systematic review and meta-analysis. **European Journal of Clinical Investigation**, v. 52, n. 2, e13706, 2022. DOI: https://doi.org/10.1111/eci.13706.

PETERSON, S. W. *et al.* RT-qPCR detection of SARS-CoV-2 mutations S 69-70 del, S N501Y and N D3L associated with variants of concern in Canadian wastewater samples. **Science of the Total Environment**, v. 810, art. 151283, 2021. DOI: https://doi.org/10.1016/j.scitotenv.2021.151283.

RADVÁNSZKA, M. *et al.* 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**, v. 15, n. 7, p. 1995-2021, 2022. DOI: https://doi.org/10.1111/1751-7915.14031.

SANTOS, R. S. *et al. In silico* analysis of mismatches in RT-qPCR assays of 177 SARS-CoV-2 sequences from Brazil. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 53, e20200657, 2020. DOI: https://doi.org/10.1590/0037-8682-0657-2020.

SILVA, L. M. *et al.* The influence of the Omicron variant on RNA extraction and RT-qPCR detection of SARS-CoV-2 in a laboratory in Brazil. **Viruses**, v. 15, n. 8, art. 1690, 2023. DOI: https://doi.org/10.3390/v15081690.

STADHOUDERS, R. *et al.* The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5′ nuclease assay. **The Journal of Molecular Diagnostics**, v. 12, n. 1, p. 109-117, 2010. DOI: https://doi.org/10.2353/jmoldx.2010.090035.

TANG, Y.-W. *et al.* Laboratory diagnosis of COVID-19: current issues and challenges. **Journal of Clinical Microbiology**, v. 58, n. 6, e00512-20, 2020. DOI: https://doi.org/10.1128/JCM.00512-20.

TOMBULOGLU, H. *et al.* Multiplex real-time RT-PCR method for the diagnosis of SARS-CoV-2 by targeting viral N, RdRP and human RP genes. **Scientific Reports**, v. 12, art. 2853, 2022. DOI: https://doi.org/10.1038/s41598-022-06977-z.

VOGELS, C. B. F. *et al.* Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. **Nature Microbiology**, v. 5, p. 1299-1305, 2020. DOI: https://doi.org/10.1038/s41564-020-0761-6.

VOGELS, C. B. F. *et al.* Multiplex qPCR discriminates variants of concern to enhance global surveillance of SARS-CoV-2. **PLOS Biology**, v. 19, n. 5, e3001236, 2021. DOI: https://doi.org/10.1371/journal.pbio.3001236.

WHO - WORLD HEALTH ORGANIZATION. **WHO Director-General's opening remarks at the media briefing on COVID-19, 11 March 2020**. Geneva: WHO, 2020. Disponível em: https://www.who.int/news-room/speeches/item/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Acesso em: 31 out. 2026.

WHO - WORLD HEALTH ORGANIZATION. **Tracking SARS-CoV-2 variants**. Geneva: WHO, 2023. Disponível em: https://www.who.int/activities/tracking-SARS-CoV-2-variants. Acesso em: 24 set. 2025.

WOLLSCHLÄGER, P. *et al.* SARS-CoV-2 N gene dropout and N gene Ct value shift as indicator for the presence of B.1.1.7 lineage in a commercial multiplex PCR assay. **Clinical Microbiology and Infection**, v. 27, n. 9, p. 1353.e1-1353.e5, 2021. DOI: https://doi.org/10.1016/j.cmi.2021.05.025.

WU, H. *et al.* Nucleocapsid mutations R203K/G204R increase the infectivity, fitness, and virulence of SARS-CoV-2. **Cell Host & Microbe**, v. 29, n. 12, p. 1788-1801.e6, 2021. DOI: https://doi.org/10.1016/j.chom.2021.11.005.

ZHOU, Y. *et al.* SARS-CoV-2 nonstructural protein 6 from Alpha to Omicron: evolution of a transmembrane protein. **mBio**, v. 14, n. 4, e00688-23, 2023. DOI: https://doi.org/10.1128/mbio.00688-23.

ZHU, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. **The New England Journal of Medicine**, v. 382, n. 8, p. 727-733, 2020. DOI: https://doi.org/10.1056/NEJMoa2001017.

ZIMMERMANN, F. *et al. In vitro* evaluation of the effect of mutations in primer binding sites on detection of SARS-CoV-2 by RT-qPCR. **Journal of Virological Methods**, v. 299, art. 114352, 2021. DOI: https://doi.org/10.1016/j.jviromet.2021.114352.