

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

B823i Brasil, Victor Henrique Felix.

Intersecções entre computação quântica e bioinformática: potenciais, desafios e aplicações práticas / Victor Henrique Felix Brasil. - João Pessoa, 2025.

23 f. : il.

Orientação: Thaís Gaudencio do Rêgo.  
TCC (Graduação) - UFPB/CI.

1. Computação quântica. 2. Bioinformática. 3. Pipeline híbrido. 4. GATK. 5. Identificação de variantes. 6. Subrotinas quânticas. I. Rêgo, Thaís Gaudencio do. II. Título.

UFPB/CI

CDU 004.8

# Intersecções entre Computação Quântica e Bioinformática: Potenciais, Desafios e Aplicações Práticas

Victor Henrique Felix Brasil<sup>1</sup>, Thaís Gaudencio do Rêgo<sup>1</sup>

<sup>1</sup>Centro de Informática – Universidade Federal da Paraíba (UFPB)  
João Pessoa - PB - Brasil

victorbrasil@cc.ci.ufpb.br, gaudenciothais@gmail.com

**Abstract.** *As the volume of biological data continues to increase rapidly, especially in genomics and molecular biology, traditional computational approaches have struggled to scale accordingly. This work presents a literature review focused on the intersection between quantum computing and bioinformatics, exploring the main concepts, methodologies, and applications that connect these two areas. The study maps how quantum algorithms and architectures are being considered to optimize classical bioinformatics tasks, and proposes a practical, incremental approach for integrating quantum computing into existing tools, such as the GATK, a toolkit widely used in bioinformatics for variant calling from sequencing data. Based on the analysis of recent scientific contributions and the evaluation of current challenges and opportunities, the study concludes that incorporating quantum subroutines into classical pipelines is conceptually viable, with the potential to accelerate specific stages, such as variant identification and prioritization, without requiring a complete restructuring of existing workflows. Furthermore, it proposes a conceptual foundation to guide future research efforts interested in applying quantum computing to diverse problems in bioinformatics.*

**Resumo.** *Com o crescimento rápido e constante dos dados biológicos, especialmente no campo da genômica e da biologia molecular, os métodos computacionais tradicionais enfrentam dificuldades para acompanhar essa escala. Este trabalho apresenta uma revisão da literatura sobre a interseção entre computação quântica e bioinformática, explorando os principais conceitos, metodologias e aplicações que conectam essas áreas. Além de mapear como algoritmos e arquiteturas quânticas vêm sendo considerados para otimizar tarefas clássicas da bioinformática, o estudo propõe uma forma prática e incremental de integrar a computação quântica a ferramentas já consolidadas, como o GATK, ferramenta amplamente utilizada na bioinformática para identificar variantes genéticas em dados de sequenciamento. A partir da análise de contribuições científicas recentes e da avaliação dos desafios e oportunidades atuais, o trabalho conclui que a inclusão de subrotinas quânticas em pipelines clássicos é conceitualmente viável, podendo acelerar etapas específicas, como a identificação e priorização de variantes, sem exigir a reestruturação completa dos fluxos existentes. Ademais, o estudo propõe uma base conceitual para futuras pesquisas interessadas em aplicar computação quântica a problemas da bioinformática distintos.*

# 1. Introdução

## 1.1. Contextualização

A bioinformática é um campo interdisciplinar que reúne conhecimentos da biologia, ciência da computação, matemática e estatística, com o objetivo de lidar com a complexidade dos dados de natureza genética, molecular e bioquímica [Prosdocimi 2007]. Desde meados da década de 1990, com a evolução das técnicas de sequenciamento de DNA, o volume de dados biológicos cresceu de forma exponencial, tornando-se cada vez mais presente em pesquisas científicas e bases de dados públicas. Esse alto crescimento tornou inviável qualquer tentativa de análise manual desses dados, o que transformou em necessidade o desenvolvimento de ferramentas e técnicas bioinformáticas. Porém, na computação, sabe-se que quanto maior a quantidade de dados a serem processados, maior também a demanda por poder computacional e, conseqüentemente, maior o tempo exigido para sua análise.

Buscando mitigar essas limitações de tempo e desempenho, surge como alternativa promissora a computação quântica. Por mais que seus dispositivos estejam em fase inicial de desenvolvimento, os princípios que servem como base para seu funcionamento permitem a representação e o processamento simultâneo de múltiplos estados de informação, o que oferece, em certas situações, um ganho computacional que os sistemas clássicos não conseguem entregar [Nałecz-Charkiewicz et al. 2024]. Ao contrário dos *bits* tradicionais, que só podem assumir o valor 0 ou 1, os *qubits* operam em superposição, abrindo caminho para a execução de algoritmos mais rápidos em problemas de alta complexidade. É nesse contexto que cresce o interesse pela interseção entre bioinformática e computação quântica, especialmente em áreas onde o volume de dados e a demanda computacional caminham lado a lado. Alguns desses desafios incluem a montagem de genomas completos, o alinhamento de seqüências em larga escala e a predição de estruturas de proteínas com alto grau de precisão, tarefas que muitas vezes impõem limitações altas aos métodos tradicionais [Nałecz-Charkiewicz et al. 2024, Pal et al. 2023]. Diante disso, começam a surgir classificações específicas dentro da chamada bioinformática quântica, como a *QCg-bioinformatics*, voltada ao uso de algoritmos quânticos sobre dados clássicos, e a *QCr-bioinformatics*, que propõe o uso direto de computadores quânticos nos fluxos bioinformáticos [Mokhtari et al. 2024].

Nesse contexto, a convergência entre as duas áreas deixa de ser apenas uma possibilidade remota e passa a ser encarada como uma transição necessária — ainda que gradual. A realidade é que, frente ao volume crescente de dados e às limitações dos modelos computacionais atuais, a integração de paradigmas alternativos tende a se tornar cada vez mais comum. A computação quântica, ao oferecer novas formas de abordar problemas que hoje são considerados gargalos, pode vir a desempenhar um papel estratégico nos próximos passos da bioinformática. Mesmo que não substitua imediatamente as técnicas atuais, seu uso complementar pode redefinir a maneira como lidamos com problemas de alto custo computacional.

## 1.2. Objetivos

O objetivo deste trabalho é realizar uma revisão da literatura científica existente sobre a interseção entre a bioinformática e a computação quântica, mapeando os conceitos fundamentais de ambas as áreas e investigando quais algoritmos, arquiteturas e metodologias

quânticas vêm sendo aplicadas a problemas clássicos da bioinformática. Ao fazer isso, pretende-se avaliar, com base em dados mais recentes, tanto os limites, quanto os potenciais dessa aproximação entre os dois campos. Além da revisão, o trabalho também propõe uma reflexão prática: como a computação quântica poderia ser incorporada, de maneira gradual e com menor custo, aos códigos e *pipelines* já consolidados na bioinformática, sem que seja necessário abandonar o que já funciona bem.

Ao reunir de maneira sistemática alguns pontos dos principais avanços publicados nos últimos anos, este trabalho não busca apenas apresentar um panorama do que tem sido feito, mas também oferecer referências conceituais e técnicas para outros pesquisadores que mostrarem interesse em explorar o uso da computação quântica em aplicações da bioinformática. A ideia é que essas informações não fiquem presas a um campo experimental ou apenas teórico, mas que sirvam de base para iniciativas futuras de integração entre os dois campos. Espera-se que ao final deste material fique claro o potencial e, de certa maneira, a necessidade dessa união.

## 2. Metodologia

Esta seção descreve a abordagem metodológica adotada para o desenvolvimento do trabalho, explicando o tipo de pesquisa, as estratégias de levantamento bibliográfico e os critérios utilizados na seleção e análise das fontes.

### 2.1. Abordagem

Este trabalho adota uma abordagem metodológica de natureza exploratória, baseada na análise de documentos e outros materiais já existentes, fazendo uma revisão da literatura, acompanhada de uma análise com foco em ideias e possibilidades para integrar Computação Quântica e Bioinformática. O objetivo não é apenas entender o que já foi produzido nesse campo, mas também apresentar caminhos e sugestões para o uso da computação quântica em ferramentas e algoritmos muito usados na bioinformática.

Para isso, foi realizado um levantamento bibliográfico voltado a identificar contribuições relevantes nas duas áreas e nas possíveis conexões entre elas. A ideia foi reunir conteúdos conceituais e técnicos que sustentassem as discussões feitas no decorrer do trabalho.

As buscas foram conduzidas em duas bases acadêmicas de amplamente reconhecidas: o PubMed, mantido pelo *National Center for Biotechnology Information* (NCBI), escolhido pela sua ampla coleção de materiais associados à biologia e medicina, e o Portal de Periódicos da CAPES, acessado por meio da Comunidade Acadêmica Federada (CAFe), com vínculo institucional pela UFPB. As buscas foram realizadas entre novembro de 2024 e abril de 2025.

Foram utilizados termos em português e inglês relacionados à área de estudo, como *Quantum*, *Quantum Computing*, *Quantum Bioinformatics*, *Bioinformatics*, *Bioinfo*, *Bioinformática*, *Computação Quântica*, entre outros. Também foram feitas combinações com nomes de ferramentas bioinformáticas clássicas, como *BLAST*, *GATK*, *SAMtools*, *BEDTools* e *Clustal*, associadas, utilizando o operador booleano "AND", a palavras como "results". Exemplos de combinações utilizadas incluem: "*GATK Bioinformatics + Results*", "*BLAST + Results*" e "*Quantum BLAST + results*".

Após as buscas iniciais, foram aplicados filtros para refinar os resultados com base nos seguintes critérios:

- Acesso aberto e/ou com disponibilidade institucional para a UFPB;
- Idioma: Português ou Inglês;
- Tipo de documento: artigos científicos;
- Recorte temporal: publicações da última década (2015–2025).

As bases PubMed e Portal de Periódicos da CAPES foram escolhidas por se alinharem diretamente aos temas centrais deste trabalho: bioinformática, genética, computação e aplicações médicas. O PubMed, por ser amplamente reconhecido em biociências, oferece acesso a uma enorme variedade de materiais relevantes, enquanto o Portal da CAPES garante acesso institucional a uma ampla variedade de periódicos científicos, incluindo áreas técnicas e multidisciplinares, o que se mostrou essencial para consultar estudos envolvendo tanto computação quântica quanto suas aplicações em biologia.

No geral, as buscas retornaram milhares de publicações em ambas as bases utilizadas. No Pubmed foram identificados cerca de 55.770 resultados, que após a aplicação dos filtros mencionados, foram reduzidos a, aproximadamente, 5.520. Já no Portal de Periódicos da CAPES, esse número de resultados foi também alto, principalmente para os termos gerais “*Bioinformática*” e “*Computação Quântica*” quando pesquisados no campo de título. Porém, devido às limitações da busca do Portal da CAPES, como, por exemplo, a impossibilidade de salvar ou rastrear os termos utilizados nas pesquisas, não foi possível quantificar e organizar os resultados de maneira melhor padronizada, como no PubMed. Assim, o fluxograma apresentado nesta seção descreve apenas o processo de triagem nesta última base. Além disso, observou-se uma escassez de estudos aplicados tratando diretamente da interseção entre computação quântica e bioinformática, principalmente publicações com foco prático. Isso reforçou a necessidade de aplicar filtros mais rigorosos, priorizando estudos com escopo aplicado ou com propostas claras de integração entre as áreas.

Outras bases acadêmicas, como IEEE Xplore e arXiv, chegaram a ser consideradas. No entanto, foram descartadas por diferentes motivos: o IEEE Xplore possui foco maior em materiais sobre engenharia elétrica e computação teórica e o arXiv, por mais que relevante para a obtenção de publicações, apresentou um grande volume de textos ainda não revisados por pares, o que dificultaria a triagem de materiais com aplicação prática direta nas áreas discutidas neste trabalho.

Após a aplicação dos filtros iniciais, foi conduzida uma etapa adicional de filtração dos materiais encontrados, com base nos títulos dos artigos, priorizando publicações que tratassem diretamente da interseção entre bioinformática e computação quântica ou de aplicações em comum. O foco maior foi dado a textos que mencionavam aplicações práticas, especialmente no campo da genética e no processamento de sequências biológicas. Essa etapa resultou em um conjunto de 111 publicações que estavam melhor alinhadas à proposta do trabalho.

Algumas publicações, por mais que estivessem com acesso livre para a UFPB, mostraram-se superficiais, duplicadas ou com conteúdo muito destoante da proposta do trabalho, sendo necessário seu descarte. E, ao final do processo, um total de 29 publicações, das 3 áreas abordadas, foram selecionadas para leitura integral, análise e discussão.

Agora, com os critérios para busca bibliográfica definidos, a próxima seção visa definir, utilizando os materiais filtrados como base, os fundamentos de cada área, assim como a maneira com que a intersecção entre ambas é desenvolvida dentro das publicações.

### 3. Fundamentos e Avanços Relacionados

Esta seção apresenta os conceitos fundamentais que sustentam a integração entre bioinformática e computação quântica, além de destacar avanços recentes relevantes para a proposta deste trabalho. Primeiramente, são abordadas as bases teóricas principais, em seguida, discute-se a aplicação prática de algoritmos e ferramentas, buscando contextualizar oportunidades e desafios na evolução da área.

#### 3.1. Fundamentos da Bioinformática

A bioinformática é uma área interdisciplinar que reúne conhecimentos da biologia, computação, estatística e matemática, voltada para a análise, interpretação e integração de dados biológicos, bem como para a modelagem de processos biomoleculares e a formulação de hipóteses científicas, especialmente em campos como a genômica, proteômica e transcriptômica [Prosdocimi 2007].

[Prosdocimi 2007] também propõe alguns conceitos úteis, e ainda muito válidos, para entender a bioinformática usando três metáforas distintas:

- A bioinformática “tijolo” refere-se à criação de ferramentas e à construção da infraestrutura digital para armazenar, processar e distribuir dados biológicos. Envolve a construção de bancos de dados e algoritmos que permitem organizar, processar e armazenar os dados. Dentre os exemplos, temos alguns softwares amplamente utilizados até hoje, como BLAST, usado para comparar sequências biológicas; o Clustal, voltado para o alinhamento múltiplo de sequências; e o GATK, utilizado na identificação e análise de variantes genéticas;
- A bioinformática “peneira” refere-se à filtragem e seleção dos dados mais relevantes em meio a grandes conjuntos de dados, como ocorre em estudos de expressão gênica, ou na identificação de variantes genéticas (do inglês, *variant calling*);
- A bioinformática “lupa” refere-se, no geral, ao uso desses dados com o objetivo de investigar hipóteses biológicas, construir modelos computacionais, ou realizar simulações que apoiem descobertas científicas.

Recentemente, uma das áreas que mais chamou a atenção dentro da bioinformática foi a bioinformática estrutural, cujo objetivo é entender e prever a estrutura tridimensional de biomoléculas, especialmente de proteínas. A estrutura de uma proteína determina diretamente sua função biológica, e prever essa estrutura a partir da sequência de aminoácidos foi, por muito tempo, um dos maiores desafios da ciência computacional e molecular. A dificuldade envolvia tanto a identificação do estado de menor energia, quanto a complexidade da amostragem, que envolvia escolher a conformação correta dentre várias possibilidades. O desenvolvimento do *AlphaFold* e outros sistemas similares que vieram depois dele, têm revolucionado a compreensão da estrutura das proteínas. O programa utiliza técnicas avançadas de aprendizagem profunda (do inglês, *Deep Learning* - DL) para prever essas estruturas com alta precisão, facilitando pesquisas em biologia estrutural e outras áreas [Marcu et al. 2022, Abramson et al. 2024].

Ferramentas como a BLAST (*Basic Local Alignment Search Tool*) também são importantes nesse contexto, visto que facilitam a comparação entre sequências e permitem inferências evolutivas e funcionais [Camacho et al. 2009, Cheng et al. 2022]. Já programas como o GATK (*Genome Analysis Toolkit*) e *DeepVariant*, são usados na identificação de variantes genéticas com precisão, sendo aplicados em estudos clínicos e populacionais [Lin et al. 2022, Lefouili and Nam 2022, Zhou et al. 2024].

Com o auxílio de métodos como o sequenciamento de RNA (*RNA-Seq*), a bioinformática também possibilita a investigação de redes de regulação gênica (do inglês, "Gene Regulatory Networks- GRNs) em diferentes organismos, como mostrado de forma geral nos estudos de [Huang et al. 2022], que analisam a interação entre RNAs em sistemas cardíacos.

Tudo isso ajuda a entender e evidenciar a bioinformática como uma área essencial na atualidade, sempre oferecendo novas metodologias para resolver problemas complexos em biologia, genética, medicina e outras áreas.

### **3.2. Fundamentos da Computação Quântica**

A computação quântica surge como uma nova maneira de lidar com problemas computacionais, funcionando com base em leis da física quântica. Princípios como a superposição, entrelaçamento e outros, conferem propriedades únicas aos sistemas quânticos, permitindo que certas tarefas sejam executadas de forma muito mais eficiente do que pelos métodos clássicos tradicionais [Wang 2012].

O funcionamento geral da computação quântica, como já foi dito, é baseado no uso dos *qubits* (*quantum bits*) que, diferente dos *bits* clássicos que só assumem valores 0 ou 1, podem existir em superposição desses estados ao mesmo tempo, mostrando uma melhora significativa na capacidade de representar e processar informações em paralelo [Marchetti et al. 2022]. Outra propriedade essencial é o entrelaçamento quântico (*quantum entanglement*), que ocorre quando dois ou mais *qubits* compartilham uma correlação entre si, de forma que qualquer alteração em um deles afeta instantaneamente os outros, mesmo à distância. Essa característica é única da computação quântica e é responsável por permitir altos níveis de paralelismo em algoritmos quânticos [Pal et al. 2023].

Dos algoritmos quânticos mais conhecidos, temos o de Grover, focado em buscas em grandes volumes de dados, e o de Shor, famoso por quebrar a fatoração de números primos, algo que até os melhores computadores clássicos têm dificuldade em fazer. Ambos mostram como os algoritmos quânticos podem trazer ganhos reais de desempenho, mostrando em alguns casos uma aceleração exponencial (Shor) ou quadrática (Grover). Esses algoritmos são implementados dentro do modelo conhecido como computação quântica baseada em portas (do inglês, *Gate-Based Quantum Computing*), no qual os *qubits* passam por portas lógicas quânticas que transformam seus estados com base em regras bem definidas. Parecidas com as portas lógicas da computação clássica, essas portas funcionam de maneira reversível e controlam transições entre os estados quânticos com alta precisão. Muitas dessas operações exploram interferências construtivas e destrutivas para aumentar as chances de obter uma resposta correta [Wang 2012].

Mesmo possuindo tanto potencial, a computação quântica ainda enfrenta desafios para se tornar amplamente aplicável. Um dos maiores obstáculos é o fenômeno da decoerência quântica, que ocorre quando os *qubits* interagem com o ambiente externo e perdem

sua estabilidade. Isso atrapalha a execução correta dos algoritmos, especialmente os maiores ou mais complexos. Para lidar com isso, pesquisadores vêm desenvolvendo técnicas de correção de erros, que tentam preservar o estado dos *qubits* durante o processamento [Pal et al. 2023].

Hoje em dia, a computação quântica se encontra na chamada era NISQ (*Noisy Intermediate-Scale Quantum*), marcada por ruídos e coerências limitadas, onde os dispositivos disponíveis possuem poucos *qubits* e ainda estão sujeitos a erros. Mesmo assim, já existem aplicações reais, no contexto da bioinformática, sendo testadas nessas máquinas, principalmente nas áreas de otimização e química quântica [Marchetti et al. 2022]. Publicações mais recentes, como [Mokhtari et al. 2024], evidenciam que esse cenário ainda perdura, com avanços contínuos e efetivos na área mesmo diante das limitações da era atual.

Mesmo com tantas limitações técnicas ainda presentes, é observável que a computação quântica continua avançando rapidamente e já começa a mostrar resultados significativos em vários casos.

### 3.3. Computação Quântica na Bioinformática

A combinação entre computação quântica e bioinformática vem acontecendo com cada vez mais frequência. Diversos estudos demonstram esse cruzamento ao utilizar algoritmos e sistemas quânticos para mitigar tarefas que demandam muito processamento e, portanto, podem ser demoradas na bioinformática tradicional [Mokhtari et al. 2024, Nałecz-Charkiewicz et al. 2024]. Essa união, convenientemente chamada de “bioinformática quântica”, engloba propriedades como superposição e entrelaçamento para tentar resolver, de maneira mais eficiente, problemas que exigem alto poder computacional [Hollenberg 2000, Pal et al. 2023]. Um bom exemplo disso é o trabalho de [Boev et al. 2021], que adaptou o problema de sobreposição de sequências genômicas para o formato de um problema de otimização do tipo QUBO (*Quadratic Unconstrained Binary Optimization*) e o resolveu usando técnicas de anelamento quântico (do inglês, *quantum annealing*), abordagem computacional que explora propriedades quânticas, como superposição e tunelamento, para encontrar a configuração de menor energia em uma paisagem de soluções possíveis. Mesmo lidando com sequências genômicas pequenas, os resultados obtidos no seu estudo mostraram que essa abordagem já é viável em dispositivos quânticos atuais e mostraram que, mesmo estando em estágios teoricamente iniciais, os modelos quânticos que já existem podem ser usados para testar soluções reais para alguns dos problemas da bioinformática.

Atualmente, em genômica, problemas como a identificação de variantes, que envolvem grandes volumes de dados, poderiam, idealmente, ser abordados por algoritmos quânticos com maior eficiência à medida que o *hardware* evolui [Pal et al. 2023]. Em muitos casos, os estudos apontam para os “algoritmos híbridos”, que fazem uso de computação quântica apenas em certas etapas do processamento de dados, como a solução mais viável no curto prazo. Trabalhos como o de [Nałecz-Charkiewicz et al. 2024] reforçam que esse pode ser o caminho mais acessível, até o momento, para que a integração entre as áreas ocorra. Neste mesmo quesito, é desenvolvido o conceito do termo *Q-bioinformatics*, que engloba desde os algoritmos híbridos até implementações reais em *hardwares* quânticos como *D-Wave* ou *IBM Q*. A categoria inclui desde abordagens diretamente inspiradas

na mecânica quântica, até o uso de *qubits* reais em etapas de processamento e análise de dados biomoleculares [Pal et al. 2023, Mokhtari et al. 2024]. Um exemplo importante dessa inspiração conceitual pode ser visto no trabalho de [Hollenberg 2000], que propôs o uso do algoritmo de Grover adaptado para bioinformática, utilizando a distância de Hamming como base para buscar alinhamentos exatos ou aproximados em bancos de dados proteicos.

Porém, mesmo com tantos avanços, as barreiras ainda são muitas, e compartilhadas com o campo geral da computação quântica em si. A quantidade de *qubits* disponível ainda é pequena para representar sequências muito grandes ou estruturas moleculares complexas. Muitos estudos se concentram em testes simulados ou em provas de conceito, o que dificulta a aplicação em cenários de uso real. Além disso, o custo do *hardware*, a escassez de profissionais qualificados, tanto em bioinformática, quanto em computação quântica e a falta de padronização de métricas de avaliação dificultam a adoção mais ampla dessas tecnologias [Marchetti et al. 2022, Khatami et al. 2023].

### 3.4. Trabalhos Relacionados

Esta seção apresenta uma análise estruturada de algumas das produções filtradas que sustentam a interseção entre computação quântica e bioinformática, buscando identificar como e onde as tecnologias quânticas vêm sendo exploradas nesse campo. Para isso, foi elaborada a Tabela 1, contendo 13 das principais publicações selecionadas entre os 29 artigos analisados integralmente. Foram incluídos na tabela apenas aqueles que apresentavam, de forma mais clara e objetiva, alguma proposta de aplicação ou simulação computacional envolvendo diretamente a computação quântica no contexto da bioinformática. A seleção foi feita com base na leitura crítica dos materiais, filtrando os que não discutiam conceitos de ambas as áreas de maneira integrada, embora esses outros artigos ainda sejam retomados em seções posteriores. Cada entrada da tabela está organizada em três categorias: **Título & Autor**, **Contexto da Interseção** (descrição breve do escopo do trabalho e da conexão entre as áreas) e **Uso de Computação Quântica** (como, caso exista, essa aplicação é abordada no material).

Tabela 1: **Trabalhos que abordam o uso de Computação Quântica na Bioinformática**

<b>Título &amp; Autor</b>	<b>Contexto da Interseção</b>	<b>Uso de Computação Quântica</b>
<i>Fast quantum search algorithms in protein sequence comparisons: Quantum bioinformatics</i> [Hollenberg 2000]	Aplicação de algoritmos de busca quântica para alinhamento de sequências de proteínas, utilizando a aplicação teórica dos algoritmos de Grover e BBHT para localizar subsequências em bancos de dados com base na distância de Hamming.	Não há implementação ou validação do uso, visto que se trata de uma análise conceitual sobre o potencial de ganho de eficiência na fase de comparação entre sequências do <i>pipeline</i> de ferramentas como, nesse caso, a BLAST.

<p><i>The prospects of quantum computing in computational molecular biology</i> [Outeiral et al. 2020]</p>	<p>Avalia o potencial da computação quântica para resolver desafios computacionais em biologia molecular, com foco em problemas como predição estrutural e otimização de moléculas.</p>	<p>Não há implementação prática ou validação experimental, visto que suas propostas são apresentadas em termos de possíveis ganhos futuros, especialmente em tarefas de simulação molecular e modelagem de interações proteína-ligante.</p>
<p><i>Biology begins to tangle with quantum computing</i> [Marx 2021]</p>	<p>Explora a relação emergente entre computação quântica e biologia, destacando iniciativas, colaborações e aplicações promissoras. Fala sobre projetos de <i>design</i> de peptídeos com o uso de subrotinas quânticas implementadas em hardware da <i>D-Wave</i> e cita avanços em algoritmos e <i>hardware</i> aplicáveis a áreas como bioinformática, neurociência e biologia estrutural.</p>	<p>Há uso prático com uma proposta híbrida na etapa do processamento das cadeias laterais de aminoácidos, dentro do <i>pipeline</i> de modelagem de estrutura proteica.</p>
<p><i>Genome assembly using quantum and quantum-inspired annealing</i> [Boev et al. 2021]</p>	<p>Apresenta uma abordagem experimental para a montagem genômica <i>de novo</i> como um problema de otimização do tipo QUBO, viabilizando sua resolução por meio de anelamento quântico e algoritmos inspirados. A abordagem é aplicada tanto em dados sintéticos, quanto reais, com foco em sequências curtas, demonstrando sua viabilidade no curto prazo.</p>	<p>Há uso experimental de computação quântica aplicada à bioinformática. É proposta uma abordagem híbrida na etapa central do <i>pipeline</i> de montagem genômica responsável por reconstruir a sequência original completa.</p>

<p><i>Investigating the potential for a limited quantum speedup on protein lattice problems</i> [Outeiral et al. 2021]</p>	<p>Analisa a aplicação do anelamento quântico em problemas de enovelamento de proteínas modelados em redes, focando na viabilidade de melhorias de desempenho em comparação a métodos clássicos.</p>	<p>Há uso teórico com a proposta de uma abordagem híbrida na etapa de predição estrutural de proteínas, parte do <i>pipeline</i> de modelagem de estrutura terciária de proteínas.</p>
<p><i>Quantum computing algorithms: getting closer to critical problems in computational biology</i> [Marchetti et al. 2022]</p>	<p>Discussão de aplicações teóricas de algoritmos quânticos como VQE, QAOA e QiBAM em tarefas críticas da biologia computacional. Propõe o uso híbrido dessas abordagens em problemas de modelagem multiescala, <i>docking</i> molecular, análise genômica e interpretação de dados biomoleculares de larga escala.</p>	<p>Há uso teórico de computação quântica aplicada à bioinformática. É proposta uma abordagem híbrida na etapa de análise e interpretação de dados ômicos, parte do <i>pipeline</i> de análise de padrões biológicos em larga escala.</p>
<p><i>Gate-based quantum computing for protein design</i> [Khatami et al. 2023]</p>	<p>Apresenta três modelos baseados no algoritmo de Grover para o <i>design</i> de proteínas, com diferentes níveis de complexidade. As soluções são testadas em simuladores e, parcialmente, em <i>hardware</i> real da IBM, comparando-se aos métodos clássicos quanto à escalabilidade e ao custo computacional.</p>	<p>Há uso direto com uma proposta de uma abordagem híbrida na etapa de <i>design</i> de proteínas, parte do <i>pipeline</i> de engenharia biomolecular, que visa identificar sequências de menor energia.</p>

<p><i>Quantum Computing for Molecular Biology</i> [Bairdi et al. 2023]</p>	<p>Uma revisão ampla sobre como a computação quântica pode superar as limitações dos modelos clássicos em simulações moleculares, com foco em problemas-chave da biologia molecular, como enovelamento de proteínas. O trabalho também propõe um roteiro para se alcançar vantagem na implementação da computação quântica no contexto da bioinformática.</p>	<p>Há uso direto da computação quântica em diversos elementos do <i>pipeline</i> da biologia molecular, identificando quais componentes já são viáveis nos dispositivos NISQ atuais e onde se espera obter vantagem nos próximos anos.</p>
<p><i>Quantum Computing in the Next-Generation Computational Biology Landscape: From Protein Folding to Molecular Dynamics</i> [Pal et al. 2023]</p>	<p>Revisão sobre os fundamentos da computação quântica e seu potencial em bioinformática. São discutidas aplicações em áreas como predição de estrutura de proteínas, modelagem de redes de regulação gênica, montagem genômica e simulações moleculares, com destaque para algoritmos e arquiteturas quânticas em ascensão.</p>	<p>Há uso conceitual de computação quântica aplicada à bioinformática, tratando-se de uma abordagem de propostas apenas teóricas. São discutidas múltiplas áreas do <i>pipeline</i> bioinformático que poderiam ser beneficiadas, como predição estrutural de proteínas, montagem genômica e análise de redes de regulação gênicas.</p>

<p><i>Quantum computing on nucleic acid research : Approaching towards next-generation computing</i> [Chakraborty et al. 2023]</p>	<p>Revisão das aplicações emergentes da computação quântica em pesquisas com ácidos nucleicos. São discutidos problemas como montagem genômica, predição de regiões de ligação de fatores de transcrição, reconstrução de sequências e predição de estrutura secundária do RNA.</p>	<p>Há uso conceitual de computação quântica aplicada à bioinformática. São apontadas diversas etapas do <i>pipeline</i> genômico nas quais a computação quântica pode ser integrada, como montagem genômica, análise de ligação DNA-proteína e predição de estruturas secundárias do RNA. Não há proposta ou implementação específica, sendo uma pesquisa teórica baseada em trabalhos anteriores da área.</p>
<p><i>Sampling a rare protein transition with a hybrid classical-quantum computing algorithm</i> [Ghamari et al. 2023]</p>	<p>Descrição de um método híbrido que combina aprendizado de máquina, dinâmica molecular e computação quântica para simular em proteínas. Utilizando uma máquina DWAVE, os autores realizaram simulações em nível atômico, com tempos de transição de milissegundos, o que normalmente exige supercomputadores especializados.</p>	<p>Há uso prático de computação quântica aplicada à bioinformática, por meio de um algoritmo híbrido que integra aprendizado de máquina, simulações clássicas e computação quântica. Trata-se de uma implementação real, mas ainda experimental, que explora o potencial da computação quântica em dinâmica molecular.</p>

<p><i>New classifications for quantum bioinformatics: Q-bioinformatics, QCt-bioinformatics, QCg-bioinformatics, and QCr-bioinformatics</i> [Mokhtari et al. 2024]</p>	<p>Revisão que propõe uma nova classificação para as subáreas da bioinformática quântica, definindo quatro categorias principais (Q-B, QCt-B, QCg-B, QCr-B). Ele discute os conceitos fundamentais da mecânica quântica aplicados à biologia, as diferenças entre bioinformática quântica e biologia quântica, e os desafios que envolvem essa interseção. O objetivo central é fomentar colaborações interdisciplinares e propor terminologias mais claras para o avanço do campo.</p>	<p>Não há aplicação prática. A proposta do trabalho é sistematizar a área de bioinformática quântica, classificando possíveis tipos de aplicação de computação quântica com base no tipo de dado (biológico clássico ou quântico), na plataforma utilizada (clássica ou quântica) e na natureza do processo computacional (extração de dados, análise, modelagem, etc.).</p>
<p><i>Quantum computing in bioinformatics: a systematic review mapping</i> [Nałecz-Charkiewicz et al. 2024]</p>	<p>Realiza o primeiro mapeamento sistemático dedicado exclusivamente ao uso da computação quântica na bioinformática, levantando, classificando e organizando produções científicas relevantes. O artigo analisa abordagens quânticas aplicadas em diferentes áreas da bioinformática, discute tendências emergentes, destaca obstáculos ainda presentes e propõe direções futuras para a área.</p>	<p>Não há uso direto, mas cataloga e classifica os estudos que o fazem. Por mais que não haja aplicação prática direta ou código, o levantamento evidencia os pontos do <i>pipeline</i> bioinformático nos quais há maior interesse na aplicação de computação quântica, como montagem de genomas, enovelamento de proteínas e análise de sequências.</p>

As publicações listadas na Tabela 1 deixam claro que a inserção da computação quântica em bioinformática já vem sendo explorada, e alguns deles inspiram diretamente a proposta deste trabalho. Enquanto algumas publicações, como [Hollenberg 2000] ou [Outeiral et al. 2020], desenvolveram projeções conceituais sem implementação prática, estudos mais recentes apresentam abordagens concretas de integração híbrida. Em especial, [Boev et al. 2021] e [Marx 2021] servem de referência por demonstrar que é possível encaixar uma subrotina quântica em um *pipeline* bioinformático tradicional, alavancando ganhos de desempenho em um ponto específico do processo sem, como já dito em seções anteriores, refazer todo o fluxo de análise. [Boev et al. 2021], por exemplo, propuseram

uma solução na qual a etapa central da montagem genômica foi reformulada como um problema adaptado à computação quântica (um problema QUBO resolvido via anelamento quântico em *hardware* D-Wave). Essa abordagem híbrida permitiu acelerar a montagem de genomas ao atacar diretamente um subproblema crítico do *pipeline*, mas manteve as demais fases (pré-processamento de sequências, montagem inicial, pós-validação, etc.) clássicas. Do mesmo jeito, [Marx 2021] descreve o uso de rotinas quânticas em um ponto específico do *pipeline* de modelagem de proteínas, focando na otimização das cadeias laterais de aminoácidos na construção de estruturas proteicas. Nesse caso, um módulo quântico experimental foi inserido para resolver a busca pelo estado de menor energia dentro do fluxo de design de proteínas, enquanto o restante do procedimento permanece clássico. Esses dois exemplos práticos, em genômica e em biologia estrutural, mostram, na prática, as vantagens de realizar a integração híbrida entre a computação quântica e *pipelines* bioinformáticos.

[Chakraborty et al. 2023], por exemplo, mapeiam diversas fases do *pipeline* genômico nas quais subrotinas quânticas poderiam ser integradas com benefício, incluindo montagem de genomas, análise de ligação DNA-proteína e predição de estrutura de RNA. Tudo isso deixando claro que a integração pontual tende a ser mais eficaz no curto prazo do que uma substituição integral de ferramentas como um todo. No fim, começa a se formar um consenso de que, na era atual, a NISQ, a melhor estratégia é usar a computação quântica como aceleradora em problemas isolados, dentro de um conjunto híbrido onde o restante do processamento continua dentro da computação clássica.

Seguindo essa discussão, este trabalho defende ser aplicável e promissor, como será discutido nas seções à seguir, inserir uma subrotina quântica no módulo de identificação de variantes do GATK, aproveitando a computação quântica de forma otimizada para potencializar a eficiência dessa fase crucial, sem destoar das práticas atuais que foram discutidas nem ignorar as restrições dos sistemas quânticos disponíveis atualmente.

## 4. Análise e Discussão

Esta seção apresenta uma análise dos avanços relevantes na integração entre computação quântica e bioinformática, bem como dos principais desafios técnicos e limitações ainda existentes na área. Inicialmente, são sintetizados os avanços recentes (4.1), seguidos pela identificação de barreiras tecnológicas e questões práticas (4.2). Por fim, é proposta uma discussão conceitual sobre a implementação de soluções híbridas no contexto da bioinformática quântica, dando destaque para um caso teórico envolvendo a identificação de variantes genômicas (4.3).

### 4.1. Síntese dos Avanços Relevantes

Como discutido nos tópicos anteriores, o aumento do volume e da complexidade dos dados biológicos fez com que se tornassem necessárias soluções computacionais cada vez mais robustas. É nesse contexto que a computação quântica mostra-se como ideal para lidar com as limitações da computação clássica. Já na bioinformática, esse potencial começa a aparecer em aplicações reais, onde algoritmos quânticos voltados para busca e otimização podem ser aplicados à varredura em bancos de dados de sequências biológicas, especialmente em cenários de *big data*. Da mesma maneira, problemas tradicionalmente custosos, como o alinhamento múltiplo de sequências e a predição da estrutura de proteí-

nas, também aparecem como bons candidatos para abordagens baseadas no uso de *qubits* [Marx 2021].

Outro avanço importante está relacionado à simulação de sistemas biomoleculares. A computação quântica, por lidar com fenômenos em escala atômica, tende a oferecer vantagens teóricas relevantes nesse tipo de tarefa. Por mais que os testes atuais ainda estejam limitados a poucos *qubits*, já foi demonstrado que, em interações mais complexas, como ligações entre proteínas e ligantes, podem ser modeladas com maior precisão ao serem utilizadas abordagens quânticas [Marx 2021, Huang et al. 2022]. Trabalhos mais recentes, como os de [Baiardi et al. 2023] e [Chakrabarti et al. 2023], reforçam que essas aplicações seguem em crescimento, destacando tanto o potencial prático da computação quântica em simulações moleculares quanto os desafios computacionais e físicos que ainda existem na área.

Reiterando que esses avanços da área ainda estão em constante desenvolvimento, o que é evidenciado pela alta presença de simulações e provas de conceito, como é discutido por [Nałecz-Charkiewicz et al. 2024] e [Pal et al. 2023], a maioria das implementações quânticas descritas na literatura ainda são apenas protótipos, provas de conceito ou simulações. Além disso, adaptar algoritmos bioinformáticos para a realidade quântica exige mais que trocar linguagens ou ajustar códigos, é necessário repensar, como um todo, como os dados são estruturados, representados e analisados [Wang 2012]. É justamente por isso que alguns pesquisadores defendem que, ao menos inicialmente, o uso da computação quântica será exclusivamente híbrido: com módulos quânticos integrados a fluxos clássicos de análise [Marx 2021].

Dessa maneira, à medida que os dispositivos evoluírem e mais ferramentas forem sendo desenvolvidas, espera-se que soluções quânticas passem a integrar fluxos reais de análise, inclusive em tarefas mais críticas como montagem de genomas, identificação de variantes, ou simulações moleculares, consolidando de vez a Bioinformática Quântica.

## 4.2. Limitações e Desafios da Área

Apesar da notável expansão dos campos da bioinformática quântica, ainda existem muitas barreiras técnicas específicas que impedem o avanço integral das aplicações práticas dessas tecnologias. Um dos principais desafios está na própria natureza dos dispositivos quânticos atuais da era NISQ, caracterizada por um número limitado de *qubits*, sensibilidade a ruídos e instabilidade. Como discutido por [Marchetti et al. 2022], e reafirmado indiretamente por [Hegde et al. 2024], esses *hardwares* ainda estão sujeitos a instabilidades significativas, e as aplicações práticas permanecem restritas a simulações, provas de conceito e implementações em pequena escala. Além disso tudo [Nałecz-Charkiewicz et al. 2024] também apontam a dificuldade de fazer uma comparação entre desempenho quântico e clássico, devido à ausência de métricas padronizadas. Por mais que, teoricamente, fosse possível utilizar o mesmo dado de entrada para comparar métodos quânticos e clássicos, na prática isso se torna difícil: os testes quânticos, em sua maioria, são realizados em ambientes simulados ou com dados artificiais, enquanto os métodos tradicionais operam sobre dados reais de larga escala [Nałecz-Charkiewicz et al. 2024]. Ou seja, não é apenas o dado que muda: as condições de execução, o nível de ruído e a robustez dos dispositivos também são muito diferentes, o que dificulta comparações diretas de desempenho. Ainda assim, mesmo quando um algoritmo quântico apresenta bons resultados em problemas simulados ou de pequena escala,

isso não garante que o mesmo desempenho será mantido em aplicações maiores e reais, como em genomas completos ou grandes bancos de dados de expressão gênica. A escalabilidade depende de fatores importantes, como o número de *qubits* físicos disponíveis, o controle de erros e a estabilidade dos dispositivos, que ainda apresentam limitações na era NISQ. Fenômenos como o aumento do ruído e da instabilidade tendem a se intensificar com o crescimento da escala do problema.

Também existe uma escassez de profissionais qualificados para trabalhar com essa combinação de áreas. A maioria dos profissionais conhece bem só um dos lados: ou a biologia e os dados genéticos, ou a física quântica e a programação. Para que novas soluções funcionem integralmente, é preciso juntar esses conhecimentos em equipes que trabalhem em conjunto.

Por fim, ao observar tudo isto, é possível afirmar que boa parte destas limitações analisadas não são impossíveis de serem mitigadas. Por estarmos tratando de uma ciência nova e em constante adaptação, todos os limites aqui mencionados estão sempre próximos de serem superados, pois é justamente com o mapeamento deles, que é aberto espaço para discutir como subrotinas quânticas, por mais que ainda não possam ser integralmente aplicadas, podem solucionar cada um dos problemas que forem sendo encontrados no caminho para o desenvolvimento da área.

### **4.3. Uma proposta de aplicação de computação quântica em uma ferramenta clássica da bioinformática**

Esta seção apresenta uma proposta conceitual de aplicação da computação quântica em uma ferramenta de bioinformática já consolidada, com o objetivo de explorar melhorias em etapas específicas de processamento de dados. Primeiramente, discute-se a motivação para a escolha do *pipeline*, seguida de uma descrição teórica do formato híbrido proposto. Por fim, são analisados os benefícios potenciais, as considerações técnicas e as perspectivas futuras relacionadas a essa integração.

#### **4.3.1. Proposta de Aplicação de Computação Quântica no *Pipeline* do GATK**

Uma das ferramentas mais utilizadas na bioinformática para identificação de variantes genômicas é o GATK. Seu *pipeline* envolve diversas etapas de pré-processamento, alinhamento, identificação de variantes e filtragem. Por mais que a evolução do GATK e de ferramentas relacionadas já tenha reduzido o tempo de execução em muitas situações, o aumento constante dos dados genômicos segue desafiando a sua escalabilidade. Por isso, a incorporação de subrotinas quânticas em alguns passos críticos do GATK pode acabar abrindo caminhos para ganhos otimistas de desempenho.

No *pipeline* clássico, a identificação de variantes costuma ser um dos pontos de maior custo computacional, pois requer análises intensivas de sequências múltiplas alinhadas para identificar Polimorfismo de Nucleotídeo Único (do inglês, *Single Nucleotide Polymorphisms* - SNPs), pequenas inserções/deleções e outras variações [Cheng et al. 2022, Lin et al. 2022].

Pensando em possíveis ganhos de desempenho através do uso da computação quântica, uma alternativa seria focar especificamente na etapa de identificação de variantes, ou em subetapas que envolvem buscas e comparações extensas dentro de grandes

volumes de dados. Nesse caso, um algoritmo quântico poderia ser usado como uma subrotina, gerando candidatos a variantes genéticas (do inglês, *candidate variants*) com base em estratégias de busca ou otimização.

#### 4.3.2. Formato Híbrido: *Pipeline* Clássico + Subrotina Quântica

No *pipeline* clássico do GATK, o processo de identificação de variantes segue uma sequência consolidada de etapas, iniciando com o alinhamento das sequências de um genoma de referência, seguido pela remoção de duplicatas e filtragem de dados com ruído. Depois disso, são aplicados procedimentos como a recalibração da qualidade das sequências e a detecção de variantes genéticas, com base em heurísticas e modelos estatísticos tradicionais. Esse fluxo, embora bastante eficiente, ainda enfrenta desafios relacionados ao aumento contínuo do volume de dados genômicos, impactando diretamente no tempo de execução e na escalabilidade das análises [Cheng et al. 2022, Lin et al. 2022].

Nesse contexto, destaca-se a etapa de identificação de variantes genéticas (do inglês, *variant calling*), que envolve a análise e comparação de inúmeras combinações possíveis. Essa característica torna essa fase particularmente adequada para o uso de uma subrotina quântica, visto que problemas de busca e otimização, em que múltiplas soluções precisam ser avaliadas em paralelo, são áreas onde a computação quântica pode oferecer ganhos promissores. Desta maneira, a inserção de um módulo quântico para apoiar ou substituir parcialmente essa etapa representa uma alternativa favorável para melhorar a eficiência e a escalabilidade do *pipeline*.

A seguir está uma visão geral simplificada do funcionamento de um *pipeline* híbrido teórico, combinando etapas tradicionais com a inserção de um módulo quântico para filtragem de variante, como também mostrado pela Figura 1 em seguida.

##### 1. Pré-processamento

- Essa fase segue o padrão já consolidado no uso do GATK. As leituras são alinhadas a um genoma de referência, duplicatas são removidas e dados com baixa qualidade são filtrados. Esses procedimentos são bem estabelecidos e funcionam de forma eficiente em arquiteturas/computação de alto desempenho (do inglês, *High Performance Computing* - HPC).

##### 2. Módulo Quântico de Seleção de Candidatos

- Nesta etapa, um algoritmo quântico recebe fragmentos de leituras e regiões do genoma consideradas incertas.
- O objetivo é, a partir de certos critérios de pontuação (qualidade de mapeamento, evidências de polimorfismos), encontrar rapidamente as combinações de variantes candidatas mais prováveis.
- Dependendo do tipo de abordagem adotada, o funcionamento pode variar. No formato *gate-based*, é possível representar cada posição do genoma como uma combinação de estados quânticos, cada um indicando uma possível variante. Já no modelo de anelamento quântico, o problema é tratado como uma tarefa de otimização, onde cada variante recebe um peso proporcional à sua qualidade de evidência, avaliada por

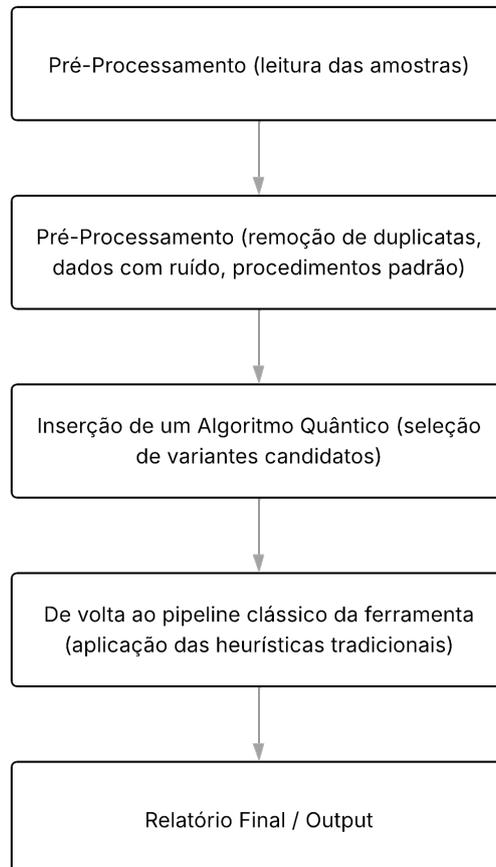
critérios como a confiabilidade do mapeamento, a profundidade de cobertura das leituras e a detecção consistente de polimorfismos relevantes [Cheng et al. 2022, Lin et al. 2022].

### 3. Reintegração ao GATK

- As variantes candidatas selecionadas pelo módulo quântico voltam para o *pipeline* clássico para verificação, refinamento e recalibração da qualidade.
- Nesta etapa, o GATK aplica suas heurísticas tradicionais, mas agora com um número menor de variantes candidatas para explorar, o que reduz a carga de processamento e acelera o andamento do processo como um todo.

### 4. Filtragem Final e *Output*

- As variantes que passam pelos critérios de qualidade são registradas como resultado final.
- Todo o restante do *pipeline* (anotação, formatação dos arquivos de saída, etc.) continua no formato clássico, pois as ferramentas já consolidadas fornecem suporte adequado para essas etapas.



**Figura 1. fluxograma simplificado demonstrando o *pipeline* do GATK com uma subrotina quântica inserida na sua etapa de filtragem de variantes**

### 4.3.3. Benefícios Potenciais, Desafios e Considerações

Um dos possíveis ganhos de desempenho consiste na otimização proposta da busca: com a superposição quântica, é possível testar vários candidatos ao mesmo tempo, reduzindo o número total de combinações que precisam ser verificadas. Isso pode levar a uma redução significativa no tempo de execução, especialmente em regiões do genoma mais complexas, ou em análises com grande volume de dados.

Outro ganho é a possibilidade de uma integração gradual. Em vez de substituir todo o *pipeline*, basta aplicar a computação quântica em uma ou duas subrotinas específicas, o que torna a adoção mais realista e acessível. Assim, ferramentas como o GATK poderiam continuar sendo usadas, mas com o suporte de módulos quânticos em pontos estratégicos.

Por mais que haja potencial, ainda existem limitações técnicas importantes. A escalabilidade é uma delas: o número de qubits disponível e a sensibilidade dos dispositivos NISQ ainda limitam os tipos de problema que podem ser tratados com eficiência. Além do mais, quando se escolhe a implementação via anelamento quântico, é necessário converter o processo de filtragem de variantes para um modelo de otimização binária, onde as possíveis soluções são representadas como variáveis que assumem valores 0 ou 1. Essa adaptação é fundamental, pois o anelamento quântico resolve problemas de minimização sobre esse tipo de espaço binário, e pode requerer ajustes no formato original dos dados para se adequar a essa estrutura. Já no caso das arquiteturas *gate-based*, é preciso garantir que as leituras genômicas estejam representadas em estados quânticos compatíveis com circuitos de portas lógicas, o que implica organizar as leituras de forma que possam ser codificadas como combinações de múltiplos *qubits*, respeitando as restrições físicas e lógicas desses dispositivos, o que pode acabar se tornando extremamente custoso.

Outro ponto que deve ser levado em conta é o custo de comunicação, pois, para que uma subrotina quântica funcione bem, os dados precisam ser transferidos com eficiência entre os sistemas clássicos e quânticos. Isso pode gerar um tempo extra que, se não for compensado pelos ganhos em desempenho, torna a solução pouco vantajosa. Estratégias como compressão de dados ou uso de lotes (do inglês, *batching*) podem ser alternativas viáveis para mitigar os custos de comunicação e transferência de dados entre sistemas clássicos e quânticos, conforme discutido por [Hollenberg 2000] e [Nałęcz-Charkiewicz et al. 2024]. Com tudo isso em mente, fica visível que por mais que a proposta de inserção de subrotinas quânticas em *pipelines* clássicos apresente benefícios conceituais e práticos, é importante reconhecer que há cenários em que essa integração pode não ser vantajosa ou até prejudicial ao desempenho geral do sistema. Um exemplo é o custo de comunicação entre os módulos clássico e quântico. *Pipelines* como o GATK, por exemplo, trabalham com grandes volumes de dados e requerem fluxos contínuos e estáveis. Inserir uma subrotina quântica que dependa de transferência e conversão constante de dados para formatos compatíveis pode introduzir atrasos e gargalos que anulam os ganhos computacionais esperados. Além disso, os dispositivos quânticos atuais ainda enfrentam limitações de escala, tempo de coerência e filas de acesso, o que pode comprometer a eficiência em cenários que exigem resposta em tempo real. Assim, a escolha de uma abordagem híbrida precisa considerar cuidadosamente o perfil do problema como um todo, sempre observando o risco de gerar resultados menos eficientes que os métodos

clássicos já consolidados.

Com isso, torna-se importante destacar que qualquer solução quântica precisa passar por uma etapa de validação cuidadosa. Ruídos, erros de medição e falhas no processo de calibração ainda são obstáculos frequentes que podem comprometer a confiabilidade dos resultados [Marchetti et al. 2022, Nałęcz-Charkiewicz et al. 2024].

## 5. Considerações Finais

Este trabalho demonstrou, teoricamente, que a integração de computação quântica em ferramentas clássicas da bioinformática é conceitualmente viável e pode trazer benefícios reais para a área. Através da proposta teórica do uso de uma subrotina quântica inserida no *pipeline* do GATK, caracterizando uma hibridização do mesmo, foi possível evidenciar ganhos potenciais de desempenho em etapas específicas, como a filtragem de variantes, sem que seja necessário reescrever completamente os fluxos que já existem, caracterizando uma hibridização dos processos realizado pela ferramenta. Em resumo, é possível afirmar que a adoção de forma incremental de módulos quânticos mostra-se como uma alternativa válida e promissora para acelerar tarefas essenciais dos *pipelines* bioinformáticos ao mesmo tempo que continua aproveitando a robustez dos métodos já existentes.

Mesmo com tantos avanços conceituais comprovados, as limitações técnicas atuais ainda são desafios significativos no que tange a integração entre a computação quântica e a bioinformática. A capacidade de *hardware* quântico disponível (como o número ainda diminuto de *qubits* na era NISQ), problemas de ruído e erros de calibração limitam muito a confiabilidade das aplicações práticas. Além disso, há um custo de comunicação e adaptação entre os sistemas, pois transferir dados entre plataformas clássicas e quânticas pode acabar gerando um custo de tempo maior que o de um processamento feito integralmente por métodos clássicos. Se esses custos não forem compensados pelos ganhos computacionais quânticos, a vantagem de se utilizar uma abordagem híbrida diminui significativamente. Assim, no estágio atual, qualquer solução quântica híbrida precisa passar por validações rigorosas para garantir que os resultados permaneçam corretos e passivos de reprodução, mesmo diante das instabilidades observadas nos dispositivos quânticos.

O trabalho também evidencia estratégias para contornar algumas dessas limitações, dando ênfase às vantagens de uma abordagem híbrida. Ao aplicar computação quântica apenas em subrotinas especializadas, como na filtragem de variantes genômicas, e manter grande parte do *pipeline* no domínio clássico, é possível mitigar algumas das restrições de *hardware* quântico presentes hoje. Essa divisão entre as tarefas busca aproveitar o melhor dos dois mundos: módulos quânticos que lidam com a etapa mais custosa de maneira otimizada, enquanto o restante do fluxo permanece com a infraestrutura clássica e eficiente. Mesmo com os obstáculos técnicos discutidos, essa estratégia híbrida mostra um caminho viável de pesquisa, unindo uma ferramenta da bioinformática já estabelecida com os avanços contínuos da computação quântica. Dessa maneira, a análise genômica é acelerada sem que seja necessário reinventar completamente os processos já consolidados, representando um cenário promissor para a próxima geração de *pipelines* da biologia computacional.

Falando em cenários, quanto às perspectivas futuras, espera-se que, além de que seja feita a validação das propostas feitas aqui neste trabalho, a evolução do *hardware*

quântico e a melhoria das arquiteturas permitam aumentar de forma gradativa o papel das subrotinas quânticas. Com dispositivos mais estáveis e potentes, a hibridização poderá abranger conjuntos de dados ainda maiores, assim como mais contextos da bioinformática, até que, em algum momento, tenhamos *pipelines* bioinformáticos integralmente quânticos, conforme o amadurecimento das áreas. Além disso, a mesma lógica apresentada na "identificação quântica" de variantes pode ser adaptada para outras tarefas complexas de outros *pipelines* da bioinformática, como a análise de expressão gênica, ou predição de estruturas proteicas. Esses cenários oferecem diversas oportunidades para ganhos de desempenho similares ao proposto neste trabalho, desde que as subrotinas quânticas sejam ajustadas às características específicas de cada problema.

Por fim, os resultados teorizados neste trabalho confirmam que a conciliação entre algoritmos quânticos e ferramentas bioinformáticas não é apenas possível, como também pode trazer vantagens reais e imediatas. Embora diversos desafios técnicos persistam, as evidências discutidas em seções anteriores apontam que integrar subrotinas quânticas em fluxos clássicos, como, por exemplo, o *pipeline* do GATK, é um caminho promissor para impulsionar a bioinformática como um todo. Essa união incremental representa uma ponte sólida entre o presente das análises genômicas e um futuro possivelmente revolucionado pela computação quântica.

## Referências

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A. I., Cowie, A., Figurnov, M., Fuchs, F. B., Gladman, H., Jain, R., Khan, Y. A., Low, C. M. R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E. D., Zielinski, M., Židek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., and Jumper, J. M. (2024). Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature*, 630:493–500. Open Access.
- Baiardi, A., Christandl, M., and Reiher, M. (2023). Quantum computing for molecular biology. *ChemBioChem*, 24(13):e202300120. Open Access.
- Boev, A. S., Rakitko, A. S., Usmanov, S. R., Kobzeva, A. N., Popov, I. V., Ilinsky, V. V., Kiktenko, E. O., and Fedorov, A. K. (2021). Genome Assembly Using Quantum and Quantum-Inspired Annealing. *Scientific Reports*, 11(13183). Open Access.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, 10(421). Open Access.
- Chakrabarti, B. K., Leschke, H., Ray, P., Shirai, T., and Tanaka, S. (2023). Quantum annealing and computation: challenges and perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 381(2241):20210419. Open Access.

- Chakraborty, C., Bhattacharya, M., Dhama, K., and Lee, S.-S. (2023). Quantum Computing on Nucleic Acid Research: Approaching Towards Next-Generation Computing. *Molecular Therapy – Nucleic Acids*, 33:53–56. Open Access.
- Cheng, T., Chin, P.-J., Cha, K., Petrick, N., and Mikailov, M. (2022). Profiling the BLAST bioinformatics application for load balancing on high-performance computing clusters. *BMC Bioinformatics*, 23(544). Open Access.
- Hegde, P. R., Kyriienko, O., Heimonen, H., Tolia, P., Netzer, G., Barkoutsos, P., Vinuesa, R., Peng, I., and Markidis, S. (2024). Beyond the buzz: Strategic paths for enabling useful nisq applications.
- Hollenberg, L. C. L. (2000). Fast quantum search algorithms in protein sequence comparisons: Quantum bioinformatics. *Phys. Rev. E*, 62:7532–7535.
- Huang, K., Wu, L., Gao, Y., Li, Q., Wu, H., Liu, X., and Han, L. (2022). Transcriptome Sequencing Data Reveal LncRNA-miRNA-mRNA Regulatory Network in Calcified Aortic Valve Disease. *Frontiers in Cardiovascular Medicine*, 9(886995). Open Access.
- Khatami, M. H., Mendes, U. C., Wiebe, N., and Kim, P. M. (2023). Gate-based quantum computing for protein design. *PLOS Computational Biology*, 19(4):1–20. Open Access.
- Lefouili, M. and Nam, K. (2022). The evaluation of Bcftools mpileup and GATK HaplotypeCaller for variant calling in non-human species. *Scientific Reports*, 12(11331). Open Access.
- Lin, Y.-L., Chang, P.-C., Hsu, C., Hung, M.-Z., Chien, Y.-H., Hwu, W.-L., Lai, F., and Lee, N.-C. (2022). Comparison of GATK and DeepVariant by trio sequencing. *Scientific Reports*, 12(1809). Open Access.
- Marchetti, L., Nifosì, R., Martelli, P. L., Da Pozzo, E., Cappello, V., Banterle, F., Trincavelli, M. L., Martini, C., and D’Elia, M. (2022). Quantum computing algorithms: getting closer to critical problems in computational biology. *Briefings in Bioinformatics*, 23(6):bbac437. Open Access.
- Marcu, -B., Tăbîrcă, S., and Tangney, M. (2022). An Overview of AlphaFold’s Breakthrough. *Frontiers in Artificial Intelligence*, 5(875587). Open Access.
- Marx, V. (2021). Biology Begins to Tangle with Quantum Computing. *Nature Methods*, 18:715–719. Open Access.
- Mokhtari, M., Khoshbakht, S., Ziyaei, K., Akbari, M. E., and Moravveji, S. S. (2024). New classifications for quantum bioinformatics: Q-bioinformatics, QCt-bioinformatics, QCg-bioinformatics, and QCr-bioinformatics. *Briefings in Bioinformatics*, 25(2):bbae074.
- Nałęcz-Charkiewicz, K., Charkiewicz, K., and Nowak, R. M. (2024). Quantum computing in bioinformatics: a systematic review mapping. *Briefings in Bioinformatics*, 25(5):bbae391.
- Outeiral, C., Strahm, M., Shi, J., Morris, G. M., Benjamin, S. C., and Deane, C. M. (2020). The prospects of quantum computing in computational molecular biology. *WIREs Computational Molecular Science*, 11(1):e1481. Open Access.

- Pal, S., Bhattacharya, M., Lee, S.-S., and Chakraborty, C. (2023). Quantum Computing in the Next-Generation Computational Biology Landscape: From Protein Folding to Molecular Dynamics. *Molecular Biotechnology*, 65(7):679–701. Open Access.
- Prosdocimi, F. (2007). *Curso Online: Introdução à Bioinformática*. Publicação independente. Publicação independente e disponível gratuitamente online.
- Wang, Y. (2012). Quantum Computation and Quantum Information. *Statistical Science*, 27(3):373 – 394. Open Access.
- Zhou, Y., Kathiresan, N., Yu, Z., Rivera, L. F., Yang, Y., Thimma, M., Manickam, K., Chebotarov, D., Mauleon, R., Chougule, K., Wei, S., Gao, T., Green, C. D., Zuccolo, A., Xie, W., Ware, D., Zhang, J., McNally, K. L., and Wing, R. A. (2024). A High-Performance Computational Workflow to Accelerate GATK SNP Detection Across a 25-Genome Dataset. *BMC Biology*, 22(13). Open Access.