

INFRAESTRUTURA DIGITAL · UFC · 29 DE DEZEMBRO DE 2025

SNAP: Streamlined Nextflow Analysis Pipeline for Immunoprecipitation-Based Epigenomic Profiling of Circulating Chromatin

Por Ze Zhang, Paulo Da Silva Cordeiro, Surya B. Chhetri, Brad Fortunato, Zhenjie Jin, Razane El Hajj Chehade, Karl Semaan, Gunsagar S. Gulati, Garyoung Gary Lee, Christopher Hemauer, Weiwei Bian, Shahabeddin Sotudian, Ziwei Zhang, David O. Osei-Hwedieh, Tanya Heim, Corrie Painter, Rashad Nawfal, Marc Eid, Damien Vasseur e John Canniff

Fonte: OpenAlex · UFC · [Publicação original](#)

INOVAÇÃO	APLICABILIDADE	POTENCIAL ECONÔMICO	MATURIDADE
9/10	9/10	9/10	Média

SNAP é uma pipeline de bioinformática inovadora e escalável para o perfil epigenômico de cromatina circulante, abordando a lacuna de ferramentas especializadas. Ele otimiza a análise de dados de biópsia líquida para detecção e monitoramento de doenças, especialmente câncer, melhorando a qualidade dos dados e a precisão diagnóstica através de métricas de controle de qualidade otimizadas e estimativa direta de ctDNA. Sua arquitetura modular e compatibilidade com ambientes de computação de alto desempenho o tornam uma ferramenta robusta para pesquisa e aplicação clínica.

NÚCLEO DA ANÁLISE

Ideia de startup ou produto

Uma startup chamada 'EpiFlow Diagnostics' que oferece uma plataforma SaaS (Software as a Service) baseada em nuvem, segura e amigável, alimentada pelo SNAP. Esta plataforma permitiria que laboratórios clínicos e pesquisadores processassem e analisassem dados epigenômicos de cromatina livre de células de forma eficiente, fornecendo insights acionáveis para diagnóstico e monitoramento de doenças. Os serviços poderiam incluir análise de dados, interpretação de resultados e integração com sistemas de informação laboratorial (LIS).

Aplicações práticas

Detecção precoce de câncer e outras doenças, monitoramento da progressão da doença e resposta ao tratamento, desenvolvimento de diagnósticos não invasivos baseados em biópsia líquida, pesquisa de biomarcadores epigenéticos e medicina personalizada.

Potencial de mercado

ALTO. O mercado global de biópsia líquida está em rápida expansão, impulsionado pela demanda por diagnósticos não invasivos. SNAP preenche uma lacuna crítica na análise de dados epigenômicos, oferecendo uma ferramenta valiosa para laboratórios de diagnóstico clínico, empresas farmacêuticas (P&D), centros de oncologia e instituições de pesquisa. O potencial de mercado inclui diagnósticos de câncer, monitoramento de doenças crônicas e desenvolvimento de terapias personalizadas.

FUNDAMENTO CIENTÍFICO

Problema abordado

A ausência de pipelines de bioinformática especializados para as características únicas da cromatina livre de células circulantes, o que dificulta a análise precisa e eficiente do perfil epigenômico para detecção e monitoramento de doenças de forma minimamente invasiva.

Metodologia

SNAP é uma pipeline baseada em Nextflow, reproduzível, escalável e modular, projetada especificamente para métodos baseados em imunoprecipitação para perfil de cromatina livre de células. Incorpora métricas de controle de qualidade otimizadas (escore de enriquecimento, limiares de contagem de fragmentos), estimativa direta do conteúdo de ctDNA a partir de distribuições de comprimento de fragmentos e impressão digital de SNP para verificação de identidade da amostra.

Principais descobertas

Os filtros de qualidade do SNAP melhoraram significativamente o desempenho da classificação e mantiveram alta retenção de dados em dados de cfChIP-seq e cfMeDIP-seq em múltiplos tipos de câncer. A validação independente em pacientes com osteossarcoma confirmou a detecção de assinaturas epigenômicas associadas ao tumor que se correlacionaram com os níveis de ctDNA e refletiram a biologia da doença. Sua arquitetura modular permite fácil extensão a outros ensaios baseados em imunoprecipitação de células livres.

LEITURA PARA GESTÃO PÚBLICA

Governos e agências de saúde pública poderiam promover a padronização de pipelines de bioinformática como o SNAP para análises de biópsia líquida em diretrizes clínicas e programas de pesquisa. Isso facilitaria a adoção generalizada de diagnósticos não invasivos, melhoraria a precisão diagnóstica em saúde pública e aceleraria a descoberta de biomarcadores, potencialmente reduzindo custos de saúde a longo prazo.

PARCERIA UNIVERSIDADE × EMPRESA

Parcerias estratégicas entre a UFC (desenvolvedores do SNAP) e grandes empresas de diagnóstico in vitro (ex: Illumina, Guardant Health, Natera) para integrar o SNAP em suas plataformas de biópsia líquida ou desenvolver novos kits de diagnóstico. Colaborações com empresas farmacêuticas para o desenvolvimento de diagnósticos complementares e descoberta de biomarcadores epigenéticos para ensaios clínicos.

OPORTUNIDADES DERIVADAS

4 direções estratégicas identificadas

STARTUP · IMPACTO ALTO · HEALTHTECH

Plataforma EpiFlow Diagnostics para Análise Epigenômica

Desenvolver uma plataforma comercial baseada em nuvem que utilize o SNAP para oferecer serviços de análise epigenômica de cromatina circulante a laboratórios clínicos e empresas farmacêuticas, focando em diagnóstico e monitoramento de câncer.

PARCERIA · IMPACTO ALTO · HEALTHTECH

Integração do SNAP em Plataformas de Biópsia Líquida Existentes

Colaborar com empresas líderes em biópsia líquida para integrar o SNAP como o motor de bioinformática padrão para o perfil epigenômico, expandindo a capacidade e precisão de seus produtos atuais.

POLÍTICA PÚBLICA · IMPACTO MÉDIO · GOVTECH

Padronização Nacional de Bioinformática para Biópsia Líquida

Influenciar políticas públicas para a adoção de pipelines padronizados como o SNAP em diretrizes de saúde, promovendo o uso de diagnósticos não invasivos e a pesquisa em epigenômica no sistema de saúde.

PRODUTO CORPORATIVO · IMPACTO ALTO · CIÊNCIA DE DADOS

Ferramenta de Descoberta de Biomarcadores para P&D Farmacêutico

Licenciar o SNAP para empresas farmacêuticas como uma ferramenta especializada para a descoberta e validação de biomarcadores epigenéticos em suas fases de pesquisa e desenvolvimento de medicamentos, especialmente em oncologia.

ABSTRACT ORIGINAL

Epigenomic profiling of circulating chromatin is a powerful and minimally invasive approach for detecting and monitoring disease, but there are no bioinformatics pipelines tailored to the unique characteristics of cell-free chromatin. We present SNAP (Streamlined Nextflow Analysis Pipeline), a reproducible, scalable, and modular workflow specifically designed for immunoprecipitation-based methods for profiling cell-free chromatin. SNAP incorporates quality control metrics optimized for circulating chromatin, including enrichment score and fragment count thresholds, as well as direct estimation of circulating tumor DNA (ctDNA) content from fragment length distributions. It also includes SNP fingerprinting to enable sample identity verification. When applied to cfChIP-seq and cfMeDIP-seq data across multiple cancer types, SNAP's quality filters significantly improved classification performance while maintaining high data retention. Independent validation using plasma from patients with osteosarcoma confirmed the detection of tumor-associated epigenomic signatures that correlated with ctDNA levels and reflected disease biology. SNAP's modular architecture enables straightforward extension to additional cell-free immunoprecipitation-based assays, providing a robust framework to support studies of circulating chromatin broadly. SNAP is compatible with cloud and high-performance computing environments and is publicly available at <https://github.com/prc992/SNAP/>.

MATÉRIA PARA LEIGOS

Para leigos: Nova Ferramenta da UFC Acelera a Detecção de Doenças por Análise de DNA no Sangue

O cenário atual

Acompanhar e detectar doenças usando apenas uma amostra de sangue, de forma pouco invasiva, é uma área de pesquisa muito promissora. Isso é feito analisando o DNA que circula livremente em nosso corpo, conhecido como cromatina circulante. Essa análise, chamada de perfil epigenômico, pode revelar informações importantes sobre a presença e o desenvolvimento de diversas enfermidades.

No entanto, o desafio é que não existiam ferramentas de bioinformática (programas de computador para analisar dados biológicos) feitas sob medida para as características únicas dessa cromatina circulante. Isso dificultava que os pesquisadores pudessem processar e interpretar esses dados de maneira eficiente e confiável.

O que os pesquisadores fizeram

Diante dessa lacuna, pesquisadores desenvolveram uma nova ferramenta chamada SNAP (Streamlined Nextflow Analysis Pipeline), que pode ser traduzida como "Pipeline de Análise Nextflow Simplificado". O SNAP é um fluxo de trabalho (uma sequência de etapas automatizadas) que foi criado especificamente para analisar dados de métodos baseados em imunoprecipitação, usados para estudar a cromatina livre que circula no sangue. O principal objetivo do SNAP é ser reprodutível (ou seja, os resultados podem ser obtidos novamente por outros pesquisadores), escalável (funciona bem com grandes volumes de dados) e modular (permite adicionar ou remover partes facilmente).

Como funciona na prática

O SNAP incorpora uma série de recursos importantes para garantir a qualidade e a precisão das análises:

- **Controle de Qualidade Otimizado:** Ele usa métricas de controle de qualidade específicas para a cromatina circulante, como pontuações de enriquecimento e limites de contagem de fragmentos, que ajudam a filtrar dados de baixa qualidade.
- **Deteção de DNA Tumoral:** A ferramenta consegue estimar diretamente a quantidade de DNA tumoral circulante (ctDNA) a partir da distribuição do tamanho dos fragmentos de DNA, o que é crucial para a detecção de câncer.
- **Verificação de Identidade:** Inclui um método de "impressão digital" de SNPs (variações genéticas) para verificar a identidade das amostras, garantindo que os dados correspondam ao paciente correto.

Sua arquitetura modular permite que o SNAP seja facilmente expandido para outros tipos de testes baseados em imunoprecipitação para cromatina livre. Além disso, ele é compatível com ambientes de computação em nuvem e de alto desempenho, tornando-o acessível e poderoso para diversos laboratórios. O SNAP está disponível publicamente para uso no endereço: <https://github.com/prc992/SNAP/>.

Resultados e evidência

Para testar a eficácia do SNAP, os pesquisadores o aplicaram a dados de diferentes tipos de câncer, usando técnicas como cfChIP-seq e cfMeDIP-seq, que analisam a cromatina circulante. Os resultados mostraram que os filtros de qualidade do SNAP melhoraram significativamente a capacidade de classificar as amostras, ao mesmo tempo em que mantiveram uma alta retenção de dados úteis.

Uma validação independente foi realizada usando amostras de plasma de pacientes com osteossarcoma (um tipo de câncer ósseo). Essa validação confirmou que o SNAP foi capaz de detectar assinaturas epigenômicas associadas ao tumor, e essas assinaturas se correlacionaram com os níveis de DNA tumoral circulante (ctDNA), refletindo a biologia da doença.

Implicações práticas

O desenvolvimento do SNAP oferece uma estrutura robusta para apoiar amplamente os estudos sobre a cromatina circulante. Isso significa que a ferramenta pode facilitar a pesquisa e o desenvolvimento de novos métodos minimamente invasivos para detectar e monitorar doenças, potencialmente levando a diagnósticos mais precoces e acompanhamentos mais eficazes para pacientes.

Limitações e próximos passos

O artigo não detalha explicitamente as limitações específicas da ferramenta SNAP, nem os próximos passos de pesquisa para o projeto. No entanto, a arquitetura modular do SNAP sugere que ele pode ser estendido para outros ensaios baseados em imunoprecipitação com cromatina livre, o que indica um caminho para futuras aplicações e adaptações.

QUEM SÃO OS PESQUISADORES

Quem são os pesquisadores

Os autores deste trabalho são Ze Zhang, Paulo Da Silva Cordeiro, Surya B. Chhetri, Brad Fortunato, Zhenjie Jin, Razane El Hajj Chehade, Karl Semaan, Gunsagar S. Gulati, Garyoung Gary Lee, Christopher Hemauer, Weiwei Bian, Shahabeddin Sotudian, Ziwei Zhang, David O. Osei-Hwedie, Tanya Heim, Corrie Painter, Rashad Nawfal, Marc Eid, Damien Vasseur e John Canniff. O trabalho está associado à Universidade Federal do Ceará (UFC).

O paper não detalha as afiliações específicas de cada pesquisador, seus papéis individuais no estudo, sua formação acadêmica, trajetória profissional, ou quaisquer outras informações biográficas além de seus nomes e a universidade associada ao estudo.

PALAVRAS-CHAVE

Epigenomics · Profiling (computer programming) · Workflow · Modular design · Pipeline (software) · Chromatin