

INSaFLU – uma plataforma bioinformática *online* para análise e comparação do genoma completo do vírus influenza: um contributo pioneiro para o reforço da vigilância da gripe à escala global

INSaFLU – an online bioinformatics platform for genome-scale analysis of influenza virus: an innovative contribution for the worldwide flu surveillance

Vítor Borges¹, Miguel Pinheiro², Pedro Pechirra³, Raquel Guiomar³, João Paulo Gomes¹

vitor.borges@insa.min-saude.pt

(1) Núcleo de Bioinformática, Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

(2) Instituto de Biomedicina, Departamento de Ciências Médicas, Universidade de Aveiro, Aveiro, Portugal

(3) Laboratório Nacional de Referência para o Vírus da Gripe e outros Vírus Respiratórios, Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

_Resumo

A vigilância laboratorial da gripe entrou numa nova era marcada pela análise e caracterização do vírus da gripe à escala do seu genoma completo. Este artigo resume o contexto inerente ao desenvolvimento e à aplicação da plataforma bioinformática *online* INSaFLU ("INSide the FLU") (<https://insaflu.insa.pt/>), a qual consiste na primeira ferramenta bioinformática *online* gratuita especificamente criada para permitir que qualquer laboratório do mundo possa integrar facilmente a análise genoma do vírus na vigilância e investigação da gripe. Este avanço científico constitui um importante alicerce para a operacionalização de uma vigilância à escala global harmonizada pelo uso da sequência do genoma total do vírus da gripe.

_Abstract

A new era of laboratory flu surveillance has already started based on the genome-scale analysis and characterization of the causative agent, the influenza virus. This article summarizes the context behind the development and application of "INSaFLU" ("INSide the FLU") (<https://insaflu.insa.pt/>), which is the first influenza-oriented bioinformatics platform that allows any laboratory in the world to analyze the genome of the influenza virus in a user-friendly manner towards the generation core "genetic requests" for an effective and timely flu laboratory surveillance. INSaFLU supplies public health laboratories and influenza researchers with an open "one size fits all" framework, potentiating the operationalization of a harmonized multi-country WGS-based surveillance for influenza virus.

Comentário resumido do artigo original publicado: Borges V, Pinheiro M, Pechirra P, Guiomar R, Gomes JP. INSaFLU: an automated open web-based bioinformatics suite "from-reads" for influenza whole-genome-sequencing-based surveillance. *Genome Med.* 2018 Jun 29;10(1):46. doi: 10.1186/s13073-018-0555-0. (6)

_Introdução

A gripe representa uma grande preocupação em termos de saúde pública em todo o mundo, causando epidemias sazonais e ocasionalmente pandemias, sendo responsável por elevadas taxas de morbilidade e mortalidade (1, 2).

Iniciou-se recentemente uma nova era na vigilância laboratorial da gripe, a qual assenta na caracterização genética do vírus influenza à escala do seu genoma completo (3-5), o que será crucial para o aumento do conhecimento e inovação em diversas áreas, tais como: i) desenvolvimento de vacinas mais efetivas; ii) identificação de marcadores genéticos associados à resistência a fármacos antivirais; iii) identificação de características genéticas associadas a uma maior virulência e/ou capacidade de transmissão do vírus.

Embora esta transição tecnológica esteja a ser fortemente impulsionada pelas autoridades nacionais e internacionais de saúde, a falta de recursos humanos especializados e/ou de infraestruturas para a análise bioinformática necessária para lidar com a complexidade dos dados primários gerados pela tecnologia de sequenciação de nova geração (NGS, na sigla em inglês) tem adiado esta transição (3).

_Objetivo

Desenvolvimento de uma plataforma *online* de livre acesso e de fácil utilização para a integração da análise do genoma completo do vírus na vigilância da gripe, obedecendo às reco-

mendações das autoridades de saúde mundiais, no sentido de levar a cabo esta revolução tecnológica para o estudo da gripe.

_Resultados

Neste projeto inovador, desenvolvemos e implementámos a plataforma INSaFLU ("INSide the FLU") (<https://insaflu.insa.pt/>) (6), a qual consiste na primeira ferramenta bioinformática *online* gratuita especificamente criada para processar de um modo automático dados primários ("reads") de NGS, convertendo-os nos *outputs* que são atualmente os principais requisitos genéticos para uma eficaz e atempada vigilância laboratorial da gripe (por exemplo, determinação do tipo e subtipo do vírus influenza, descodificação da sequência do genoma viral, identificação de fenómenos de recombinação, anotações das variantes genéticas e construção de árvores filogenéticas) (figura 1).

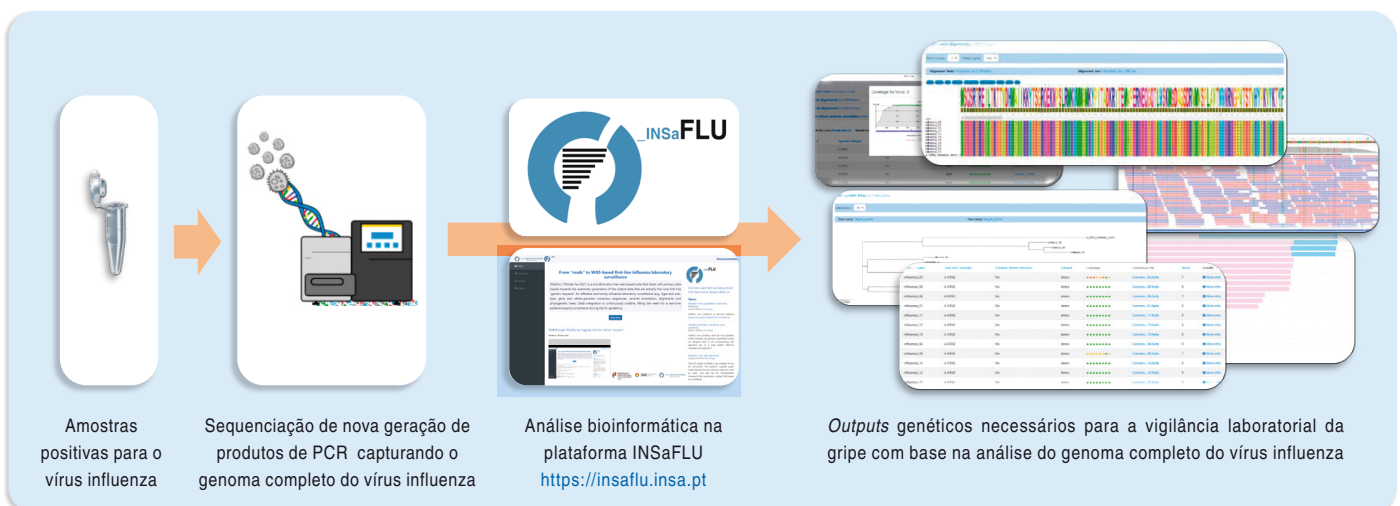
A plataforma INSaFLU, apesar da complexidade inerente à sua construção e implementação, permite que qualquer laboratório a nível mundial realize, de uma forma rápida e simples, análises computacionais complexas, sem necessidade de estudos avançados em bioinformática. Os dados são processados de uma forma confidencial e podem ser integrados de uma forma cumulativa e escalável, o que vem ao encontro da necessidade de uma vigilância epidemiológica contínua durante as epi-

demias e pandemias de gripe. Os vários *outputs* são gerados de forma padronizada e compatível com outros *software* para análises adicionais, facilitando a análise integrada dos perfis genéticos do vírus com os dados epidemiológicos e clínicos associados aos casos de gripe.

Para além dos tradicionais requisitos genéticos acima citados, esta plataforma possibilita também a identificação de potenciais infeções mistas, bem como a deteção de variantes genéticas do vírus dentro do mesmo paciente. Esta dupla abordagem tem o potencial de fortalecer a capacidade dos laboratórios não só para detetar a emergência de variantes antigénicas ou associadas à resistência a fármacos antivirais, como também para descodificar as complexas vias de adaptação do vírus e cadeias de transmissão. É ainda expetável que, num cenário de futura pandemia de gripe, esta ferramenta para análise do genoma completo do vírus desempenhe um papel importante para a identificação dos fenómenos genéticos subjacentes à origem de uma nova variante do vírus influenza.

A plataforma INSaFLU foi apresentada publicamente à comunidade científica internacional, nomeadamente em encontros científicos promovidos pelo projecto iMOVE+ (*Integrated Monitoring of Vaccines in Europe*), e pelo Centro

Figura 1: Esquema simplificado do enquadramento da plataforma INSaFLU no *workflow* da vigilância laboratorial do vírus da gripe na era da sequenciação e análise do seu genoma completo.



Europeu de Prevenção e Controlo das Doenças (ECDC, na sigla em inglês) e Organização Mundial da Saúde (OMS) (http://www.euro.who.int/__data/assets/pdf_file/0008/385802/who-euro-ecdc-influenza-meeting-rep-eng.pdf?ua=1). Cerca de um ano após a sua disponibilização, a plataforma INSaFLU conta já com utilizadores de instituições académicas e Institutos de Saúde Pública de diversos países, constituindo também a principal ferramenta bioinformática usada pelo Laboratório Nacional de Referência para o Vírus da Gripe e outros Vírus Respiratórios do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) para caracterização das variantes genéticas do vírus da gripe circulante em Portugal.

Importa acrescentar que, face à sua versatilidade para análise de dados de NGS obtidos a partir de quaisquer produtos de PCR (“amplificações”), o INSA está já a testar e/ou aplicar esta plataforma bioinformática para investigação e/ou vigilância de outros agentes patogénicos.

_Conclusão

A plataforma INSaFLU encontra-se disponível para todos os laboratórios à escala mundial, constituindo um importante alicerce para a operacionalização de uma vigilância harmonizada pelo uso da sequência do genoma completo do vírus da gripe.

O artigo científico associado à plataforma INSaFLU foi publicado na revista *Genome Medicine* (6) e pode ser lido em: <https://doi.org/10.1186/s13073-018-0555-0>.

Agradecimento

À Doutora Cristina Furtado pela revisão do artigo.

Referências bibliográficas:

- (1) Stöhr K. Influenza--WHO cares. *Lancet Infect Dis.* 2002;2(9):517.
- (2) Iuliano AD, Roguski KM, Chang HH, et al; Global Seasonal Influenza-associated Mortality Collaborator Network. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet.* 2018;391(10127):1285-1300. doi: 10.1016/S0140-6736(17)33293-2. Epub 2017 Dec 14.
- (3) Ali R, Blackburn RM, Kozlakidis Z. Next-Generation Sequencing and Influenza Virus: a short review of the published implementation attempts. *HAYATI J Biosci.* 2016;23(4):155-9. <https://doi.org/10.1016/j.hjb.2016.12.007>
- (4) Revez J, Espinosa L, Albiger B, et al.; ECDC National Microbiology Focal Points and Experts Group. Survey on the Use of Whole-Genome Sequencing for Infectious Diseases Surveillance: rapid expansion of european national capacities, 2015-2016. *Front Public Health.* 2017;5:347. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5741818/>
- (5) Goldstein EJ, Harvey WT, Wilkie GS, et al. Integrating patient and whole-genome sequencing data to provide insights into the epidemiology of seasonal influenza A(H3N2) viruses. *Microb Genom.* 2018;4(1). Epub 2017 Dec 21. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5857367/>
- (6) Borges V, Pinheiro M, Pechirra P, et al. INSaFLU: an automated open web-based bioinformatics suite “from-reads” for influenza whole-genome-sequencing-based surveillance. *Genome Med.* 2018;10(1):46. doi: 10.1186/s13073-018-0555-0. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6027769/>