

Dos Projetos Genoma à Covid-19: o papel da ciência aberta e colaborativa

Marimelia Porcionatto*

Resumo

Quase um século separa a pandemia da “gripe espanhola” (1918) da de Covid-19, infecção respiratória causada pelo vírus SARS-CoV-2, inicialmente denominado “novo coronavírus”. Apesar do aumento do conhecimento em áreas como virologia, imunologia, biologia molecular, patologia, ciências biomédicas e medicina, o novo coronavírus representa um desafio enfrentado apenas durante a pandemia de influenza. Milhares de cientistas trabalham para decifrar diversos aspectos desconhecidos dessa nova doença, da estrutura molecular viral à epidemiologia, dos mecanismos moleculares da infecção aos modelos matemáticos de transmissão. Cientistas de todo o mundo publicaram milhares de artigos científicos a respeito da Covid-19 em menos de um ano. No “I Ciclo Ciência Viva”, apresentei algumas ideias sobre como a comunidade científica global se adaptou de maneira rápida à nova realidade que a pandemia impôs e como os eventos científicos dos últimos 100 anos a preparam para isso.

Palavras-chave

Covid-19; Projeto Genoma; Ciência aberta; Ciência colaborativa

Abstract

Nearly one century separates the “Spanish flu” pandemic of 1918 from the Covid-19 pandemic. Despite the increase in knowledge on virology, immunology, molecular biology, pathology, and many others in biology and biomedical science and medicine, the “new coronavirus”, now identified as SARS-CoV-2, poses a challenge only seen during the “Spanish flu” pandemic. Hundreds of thousands of scientists work to decipher the many unknown aspects of the new coronavirus disease (Covid-19), from the viral molecular structure to epidemiology, from molecular mechanisms of infection to mathematical model of transmission. Scientists worldwide published thousands of scientific papers in less than a year. During the “I Ciclo Ciência Viva,” I presented some ideas on how the global scientific community quickly adapted itself to deal with the new reality the pandemic imposed and how the scientific events in the past one hundred years prepared it to do so.

Keywords

Covid19, Genome Project; Open Science; Collaborative science

* Departamento de Bioquímica, Escola Paulista de Medicina, UNIFESP. ✉ marimelia.porcionatto@unifesp.br.

Introdução

Estima-se que o número de mortos pela “gripe espanhola” na Inglaterra, entre 1918 e 1919, tenha sido de cerca de 200 mil pessoas. A pandemia de influenza se repetiu mais duas vezes ao longo do século XX, em 1957 e em 1968¹; estas ocorreram já na época da virologia moderna, o que permitiu a identificação dos agentes etiológicos da doença. Sabemos hoje que as três pandemias foram causadas por três subtipos diferentes do vírus A da influenza, sendo o subtipo H1N1 responsável pela pandemia de 1918, o H2N2, pela de 1957, e o H3N2, pela de 1968.²

No século XXI, a população mundial foi acometida por outras epidemias e pandemias causadas por diferentes vírus, como a pandemia de influenza causada pelo H1N1 em 2009. Embora existam relatos de contaminação humana com coronavírus aviário desde o final da década de 1960³, a primeira pandemia causada por um membro da família *Coronaviridae* ocorreu em 2002-2003. A Síndrome Respiratória Aguda Grave (SARS) foi causada pelo coronavírus SARS-CoV. Cerca de uma década depois, ocorreu a epidemia de Síndrome Respiratória do Oriente Médio, causada pelo MERS-CoV, que infecta dromedários.⁴

A mais recente pandemia, a de Covid-19, causada pelo SARS-CoV-2, representa um desafio para toda a população mundial, com repercussão na economia, na educação e nas relações humanas. A rápida disseminação do vírus e o aumento diário das mortes, agora contabilizadas e acompanhadas em tempo real, proporciona uma visão da dimensão do problema, algo impossível há algumas décadas. A pandemia de Covid-19 também alterou a maneira como os cientistas se relacionam e a forma como colaboram. As redes de colaboração entre cientistas vêm de longa data; porém, a Covid-19 mostrou-se tão desafiadora que a ciência aberta e colaborativa passou a ser realizada de maneira orgânica, quase natural, como forma de enfrentamento da doença.

Neste texto estão algumas das ideias apresentadas no “I Ciclo Ciência Viva” sobre a rápida adaptação da comunidade científica global à pandemia. Também discutimos como alguns projetos de grande porte, como os de sequenciamento de genomas completos, e os movimentos de ciência aberta e colaborativa contribuíram para o enfrentamento da pandemia de COVID-19.

¹ Edwin D. Kilbourne, "Influenza pandemics of the 20th century," *Emerging Infectious Diseases* 12 (2006): 9.

² Ibid.

³ A. Z. Kapikian *et al.*, "An epidemiologic study of altered clinical reactivity to respiratory syncytial (RS) virus infection in children previously vaccinated with an inactivated RS virus vaccine," *American Journal of Epidemiology* 89, n° 4 (abril 1969): 405–421,

⁴ Emmie de Wit *et al.*, "SARS and MERS: recent insights into emerging coronaviruses," *Nature Reviews in Microbiology* 14 (2016): 523.

Sequenciamento de genomas completos e as redes de pesquisadores

O Projeto Genoma Humano (PGH), iniciado oficialmente em 1990, foi um ambicioso projeto global, com sede nos Estados Unidos, que sequenciou pela primeira vez o genoma humano completo.^{5,6,7} A iniciativa contou com a participação de cientistas de outros cinco países: Alemanha, China, Japão, França e Reino Unido, totalizando 20 centros de sequenciamento.⁸ O PGH foi o maior projeto desenvolvido pelas ciências biológicas e médicas até então:

The Human Genome Project has been the first major foray of the biological and medical research communities into “big science”. [...] we present some of our experiences in organizing and managing such a complicated, publicly funded, international effort. We believe that many of the lessons we learned will be applicable to future large-scale projects in biology.⁹

Dentre as diversas contribuições do PGH, além do sequenciamento completo do genoma humano, destacam-se o trabalho em rede de colaboradores internacionais e a disponibilização das sequências em repositório público. Em reunião estratégica realizada na Ilhas Bermudas em 1996, pesquisadores do PGH estabeleceram os *Bermuda Principles*, conjunto de regras para o compartilhamento diário das sequências de DNA, obrigatório para todos os laboratórios financiados pelos Institutos Nacionais de Saúde (NIH) dos Estados Unidos, pelo Wellcome Trust, fundação de apoio à pesquisa biomédica sediada em Londres, no Reino Unido, e demais países participantes.¹⁰ Essa prática teria sido sugerida por dois cientistas da comunidade de pesquisadores que trabalhavam no estudo do verme *Caenorhabditis elegans*, um dos organismos-modelo da biologia, John Sulston e Robert Waterston, com o objetivo de servir como controle de qualidade e auxiliar na coordenação do projeto.

In the *C. elegans* community, and subsequently in the HGP, daily sharing served the pragmatic goals of quality control and project coordination. [...] in the HGP human genome, [...] the Bermuda Principles addressed concerns about gene patents impeding scientific advancement, and were aspirational and flexible in implementation and justification. They endured

⁵ Linda W. Engel, "The Human Genome Project. History, goals, and progress to date," *Archives of Pathol and Laboratory Medicine* 117 (1993): 459.

⁶ Margaret Everett, "The social life of genes: privacy, property and the new genetics," *Social Science & Medicine* 56, nº 1 (2003): 53.

⁷ Francis S. Collins *et al.*, "The Human Genome Project: lessons from large-scale biology," *Science* 300 (2003): 286.

⁸ *Ibid.*

⁹ *Ibid.*

¹⁰ Kathryn Maxson Jones *et al.*, "The Bermuda Triangle: The Pragmatics, Policies, and Principles for Data Sharing in the History of the Human Genome Project," *Journal of the History of Biology* 51 (2018): 693.

as an archetype for how rapid data sharing could be realized and rationalized, and permitted adaptation to the needs of various scientific communities.¹¹

Para além desses objetivos, a divulgação diária das sequências também foi vista como uma maneira de impedir que os genes fossem patenteados, algo muito debatido, e temido, à época do desenvolvimento do PGH.^{12, 13, 14, 15} Foi um período de grande preocupação com questões éticas relevantes a respeito de potenciais efeitos negativos que as patentes de genes poderiam ter na pesquisa e desenvolvimento, bem como na prática médica, além da sombria perspectiva de uso indevido em processos de eugenia baseada na composição genética dos indivíduos.¹⁶

O PGH alavancou a produção científica referente ao conhecimento do genoma humano, como pode ser observado pelo aumento do número de trabalhos publicados listados no principal indexador da área biológica e médica, o PubMed.¹⁷ De 1965 a 1989 foram publicados 4.387 trabalhos contendo o termo “*human genome*”. Esse número cresceu para 23.697 entre 1990 e 2001, ano da publicação do genoma humano nas revistas *Nature*¹⁸ e *Science*¹⁹. De 2002 até outubro de 2020, 147.923 trabalhos contendo o termo “*human genome*” foram listados no PubMed. O número de trabalhos publicados anualmente saltou de 24 (1965-1989) para mais de 8 mil no período após a publicação dos dois trabalhos com a sequência do genoma humano.

No Brasil, mais especificamente no estado de São Paulo, a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) financiou um grande projeto em rede nos moldes do PGH, em menor escala e tendo como alvo um organismo mais simples: *Xylella fastidiosa*, bactéria causadora da clorose variegada dos citros (CVC), mais conhecida como praga do amarelinho. O Projeto Genoma da *Xylella* (PGX) foi realizado entre 1997 e 1999 por 35 laboratórios e cerca de 200 pesquisadores.²⁰ Os laboratórios e pesquisadores formaram a Rede de Organização para o Sequenciamento e Análise de Nucleotídeos (Onsa, no acrônimo em inglês), em uma alusão ao *The Institute for Genomic Research* (TIGR), criado pelo bioquímico norte-americano Craig Venter em 1992 para realizar o sequenciamento do genoma de diversos eucariotos e procariotos. Os Estados Unidos tinham o TIGR; o estado de São Paulo, a Onsa.

¹¹ Ibid.

¹² Christopher Anderson, "US patent application stirs up gene hunters," *Nature* 353 (1991): 485.

¹³ David J. Aston, "Patenting of genes," *Nature* 355 (1992): 104.

¹⁴ Mark Blaxter, "Patenting of genes," *Nature* 355 (1992): 104.

¹⁵ Colin Macilwain, "OTA panel opens inquiry into patenting of genes," *Nature* 362 (1993): 386.

¹⁶ Jon F. Merz & Mildred K. Cho, "What are gene patents and why are people worried about them?," *Community Genet* 8 (2005): 203.

¹⁷ National Center for Biotechnology Information, "PubMed.gov," National Library of Medicine, <https://pubmed.ncbi.nlm.nih.gov/>. Endereço permanente do PubMed.

¹⁸ Eric S. Lander et al., "Initial sequencing and analysis of the human genome," *Nature* 409 (2001): 860.

¹⁹ John C. Venter, "The sequence of the human genome," *Science* 291 (2001): 1304.

²⁰ Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), "Linha do Tempo," Centro de Documentação e Informação da FAPESP, <https://bv.fapesp.br/linha-do-tempo/1206/genoma-xylella/>. Linha do tempo do Projeto Genoma da *Xylella*.

Ciência colaborativa na era da Covid-19

Evidentemente, as redes de colaboração entre pesquisadores não são uma novidade da era da Covid-19, assim como não foram no desenvolvimento dos projetos genoma. Uma análise realizada em 2007 que avaliou cerca de 20 milhões de trabalhos publicados desde 1955 e mais de 2 milhões de patentes mostrou que publicações com apenas um autor eram as mais citadas no início da década de 1950. No entanto, esse padrão mudou nas últimas décadas.²¹ Os autores verificaram que trabalhos com múltiplos autores são mais citados em diversas áreas, não sendo exclusividade das ciências biológicas, biomédicas ou médicas, com maior tradição em publicações com múltiplos autores.

Research is increasingly done in teams across nearly all fields. Teams typically produce more frequently cited research than individuals do, and this advantage has been increasing over time. Teams now also produce the exceptionally high-impact research, even where that distinction was once the domain of solo authors. These results are detailed for sciences and engineering, social sciences, arts and humanities, and patents, suggesting that the process of knowledge creation has fundamentally changed.²²

Uma das explicações para o aumento da colaboração entre pesquisadores e grupos de pesquisa seria o avanço tecnológico que, em áreas do conhecimento que necessitam de equipamentos para realizar experimentos, demanda pessoas altamente especializadas, além de equipamentos caros para serem adquiridos por pesquisadores individualmente. Esse foi o caso dos projetos genoma. Quando o PGH teve início, o sequenciamento de DNA demandava muito tempo, o que estimulou o desenvolvimento de um sequenciador de DNA automático de primeira geração por uma empresa sediada nos Estados Unidos. Para o desenvolvimento do PGX, a FAPESP investiu na compra de sequenciadores de DNA automáticos e equipou os laboratórios participantes para concluírem o sequenciamento dos 2,7 milhões de pares de bases do genoma da *X. fastidiosa*. Além da aquisição de equipamentos e reagentes, o PGX também possibilitou o treinamento de pesquisadores na área de genômica, criando expertise que foi fundamental para outros avanços na ciência no país.

O movimento de Ciência Aberta (*Open Science*), com vistas a tornar públicos e gratuitos os resultados de pesquisas bem como dados brutos depositados em repositórios acessíveis, também parece contribuir de maneira importante para as atuais redes colaborativas. Uma coalisão formada por instituições internacionais propôs, em 2002, a

²¹ Stefan B. Wuchty *et al.*, "The increasing dominance of teams in production of knowledge," *Science* 316 (2007): 1036.

²² *Ibid.*

Budapest Open Access Initiative (BOAI), com o objetivo de tornar público e gratuito o acesso a resultados de pesquisa para o avanço na ciência, medicina e saúde.²³ Nos anos seguintes, o movimento continuou crescendo e, além de acesso a trabalhos publicados, passou a propor também o compartilhamento de dados brutos de pesquisa. Revistas científicas e agências de fomento ao redor do mundo já exigem que os cientistas tenham os dados gerados nas pesquisas depositados em repositórios.

A pandemia de Covid-19 acontece em um cenário em que os movimentos de ciência aberta e colaborativa ocorrem há vários anos, favorecendo o a internacionalização da produção científica. Cientistas de, virtualmente, todos os países e de diversas áreas do conhecimento usam ferramentas de compartilhamento de dados e interação via internet para colaborar tendo como objetivo principal o avanço do conhecimento e a solução de problemas em comum. Chamados a enfrentar os desafios da Covid-19, cientistas do mundo se uniram para buscar compreender a doença e suas consequências. Dos mecanismos celulares e moleculares da infecção ao uso de modelos matemáticos preditivos da disseminação do vírus, milhares de cientistas trabalham em uníssono para solucionar um problema que, em maior ou menor grau, atinge a todos.

Os primeiros trabalhos no PubMed contendo o termo “*coronavirus*” datam da década de 1950 e, de lá até 2019, estão registrados pouco mais de 15 mil trabalhos contendo o termo. Somente de janeiro a outubro de 2020, mais de 48 mil trabalhos contêm termo “*coronavirus*”. Desses, cerca de 36 mil também contêm o termo “*SARS-CoV-2*”, coronavírus responsável pela Covid-19, e pouco mais de 17 mil contêm também o termo “*SARS-CoV*”, coronavírus responsável pela Sars, no início dos anos 2000.

Os cinco países que mais publicaram em Covid-19 nos primeiros meses de 2020 (janeiro a maio) foram, nesta ordem: China, Estados Unidos, Itália, Reino Unido e Índia.²⁴ O Brasil aparece entre os 25 países que mais publicaram trabalhos científicos nesse período. Lee e Haupt analisaram o nível de colaboração internacional entre janeiro e maio de 2020, compararam com um período anterior à pandemia (2015-2019) e observaram crescimento das colaborações internacionais que passaram de 23,4% (2015-2019) para 33,6%.²⁵ Embora o nível de colaboração internacional geral tenha aumentado, na China, país que mais publicou em Covid-19 em 2020, o índice de colaboração internacional caiu de 22,5% de 2015-2019 para 20,0% em 2020. Os Estados Unidos, segundo país com maior número de publicações no período permaneceu estatisticamente inalterado, com 39,9% de publicações com colaborações internacionais em 2015-2019 e 42,9% em 2020.²⁶

Nessa análise, o Brasil passou de 32,1% de trabalhos com colaboração internacional publicados entre 2015-2019 para 46,2% entre janeiro e maio de 2020. Para exemplificar, o grupo do Instituto de Medicina Tropical da Universidade de São Paulo que sequenciou o SARS-CoV-2 isolado do primeiro paciente brasileiro, publicou trabalhos com diversos

²³ Budapest Open Access Initiative, “Budapest Open Access Initiative,” BOAI, <https://www.budapestopenaccessinitiative.org/>.

²⁴ Jenny J. Lee & John P. Haupt, “Scientific globalism during a global crisis: research collaboration and open access publications on Covid-19,” *Higher Education* (2020): 1.

²⁵ Ibid.

²⁶ Ibid.

colaboradores da Bélgica, Canadá, Estados Unidos, França e Reino Unido.²⁷ Não está no escopo desta discussão analisar se haviam colaborações prévias entre estes grupos ou se elas foram estabelecidas para solucionar questões referentes à Covid-19. Também será interessante avaliar se essas colaborações se manterão no período pós-pandemia.

Conclusões

Os movimentos de ciência aberta e colaborativa, além de experiências anteriores de grandes projetos internacionais, parecem ter contribuído para os avanços na compreensão de diversos aspectos da pandemia de Covid-19. Diferentemente de projetos de grande porte realizados anteriormente, não existe uma coordenação central dessas colaborações. Fica também a verificar se os grupos internacionais se organizaram a partir de colaborações anteriores e, talvez mais relevante para o futuro, seja acompanhar o amadurecimento desse formato de fazer ciência, de maneira aberta e colaborativa.

Agradecimentos

A autora agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Bolsa de Produtividade em Pesquisa (Proc. 309679/2018-4), e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo Auxílio à Pesquisa Regular (Proc. 2020/05289-2).

²⁷ Darlan S. Candido *et al.*, "Evolution and epidemic spread of SARS-CoV-2 in Brazil," *Science* 369 (2020): 1255 ; Darlan S. Candido *et al.*, "Routes for Covid-19 importation in Brazil," *Journal of Travel Medicine* 27 (2020): taaa042 ; Jaqueline G. Jesus *et al.*, "Importation and early local transmission of Covid-19 in Brazil, 2020," *Revista do Instituto de Medicina Tropical de São Paulo* 62 (2020): e30.